Lecture Notes in Networks and Systems 909

Radek Silhavy
Petr Silhavy *Editors*

Software Engineering Methods in Systems and Network Systems

Proceedings of 7th Computational Methods in Systems and Software 2023, Vol. 1

Springer

Lecture Notes in Networks and Systems 909

Series Editor

Janusz Kacprzyk, *Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Advisory Editors

Fernando Gomide, *Department of Computer Engineering and Automation—DCA, School of Electrical and Computer Engineering—FEEC, University of Campinas—UNICAMP, São Paulo, Brazil*
Okyay Kaynak, *Department of Electrical and Electronic Engineering, Bogazici University, Istanbul, Türkiye*
Derong Liu, *Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, USA*
Institute of Automation, Chinese Academy of Sciences, Beijing, China
Witold Pedrycz, *Department of Electrical and Computer Engineering, University of Alberta, Alberta, Canada*
Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
Marios M. Polycarpou, *Department of Electrical and Computer Engineering, KIOS Research Center for Intelligent Systems and Networks, University of Cyprus, Nicosia, Cyprus*
Imre J. Rudas, *Óbuda University, Budapest, Hungary*
Jun Wang, *Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong*

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Radek Silhavy · Petr Silhavy
Editors

Software Engineering Methods in Systems and Network Systems

Proceedings of 7th Computational Methods in Systems and Software 2023, Vol. 1

Springer

Editors
Radek Silhavy
Faculty of Applied Informatics
Tomas Bata University in Zlin
Zlin, Czech Republic

Petr Silhavy
Faculty of Applied Informatics
Tomas Bata University in Zlin
Zlin, Czech Republic

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-3-031-53548-2 ISBN 978-3-031-53549-9 (eBook)
https://doi.org/10.1007/978-3-031-53549-9

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

Welcome to Volume 1 of the conference proceedings for the esteemed Computational Methods in Systems and Software 2023 (CoMeSySo). This volume, titled "Software Engineering Methods in Systems and Network Systems," encapsulates the innovative strides and groundbreaking research presented by experts, scholars, and professionals from around the globe.

In today's digital age, the role of software engineering in shaping the future of systems and network systems cannot be understated. The papers and articles contained within this volume delve deep into the methodologies, practices, and tools that are at the forefront of this dynamic field. From novel approaches to software development to the optimization of network systems, the breadth and depth of topics covered here are a testament to the vibrant and evolving nature of software engineering.

The CoMeSySo conference has always been a melting pot of ideas, fostering collaborations, and discussions that push the boundaries of what's possible in computational methods. This year, we were privileged to witness a confluence of minds, all dedicated to advancing the state of the art in software engineering for systems and network systems.

We want to extend our heartfelt gratitude to all the authors, reviewers, and members of the organizing committee. Their dedication, hard work, and passion have made this volume not just a collection of papers but a beacon for future research and development.

To our readers, we hope this volume serves as both an inspiration and a resource. Whether you are a seasoned professional, a budding researcher, or a curious enthusiast, the insights and knowledge shared within these pages will enrich your understanding and fuel your passion for software engineering.

Thank you for being a part of this journey. We look forward to the continued growth and evolution of the CoMeSySo community and to the innovations that the future holds.

Radek Silhavy
Petr Silhavy

Organization

Program Committee

Program Committee Chairs

Petr Silhavy	Tomas Bata University in Zlin, Faculty of Applied Informatics
Radek Silhavy	Tomas Bata University in Zlin, Faculty of Applied Informatics
Zdenka Prokopova	Tomas Bata University in Zlin, Faculty of Applied Informatics
Roman Senkerik	Tomas Bata University in Zlin, Faculty of Applied Informatics
Roman Prokop	Tomas Bata University in Zlin, Faculty of Applied Informatics
Viacheslav Zelentsov	Doctor of Engineering Sciences, Chief Researcher of St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS)
Roman Tsarev	Department of Information Technology, International Academy of Science and Technologies, Moscow, Russia
Stefano Cirillo	Department of Computer Science, University of Salerno, Fisciano (SA), Italy

Program Committee Members

Juraj Dudak	Faculty of Materials Science and Technology in Trnava, Slovak University of Technology, Bratislava, Slovak Republic
Gabriel Gaspar	Research Centre, University of Zilina, Zilina, Slovak Republic
Boguslaw Cyganek	Department of Computer Science, University of Science and Technology, Krakow, Poland
Krzysztof Okarma	Faculty of Electrical Engineering, West Pomeranian University of Technology, Szczecin, Poland

Monika Bakosova	Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology, Bratislava, Slovak Republic
Pavel Vaclavek	Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic
Miroslaw Ochodek	Faculty of Computing, Poznan University of Technology, Poznan, Poland
Olga Brovkina	Global Change Research Centre Academy of Science of the Czech Republic, Brno, Czech Republic & Mendel University of Brno, Czech Republic
Elarbi Badidi	College of Information Technology, United Arab Emirates University, Al Ain, United Arab Emirates
Luis Alberto Morales Rosales	Head of the Master Program in Computer Science, Superior Technological Institute of Misantla, Mexico
Mariana Lobato Baes	Research-Professor, Superior Technological of Libres, Mexico
Abdessattar Chaâri	Laboratory of Sciences and Techniques of Automatic Control & Computer Engineering, University of Sfax, Tunisian Republic
Gopal Sakarkar	Shri. Ramdeobaba College of Engineering and Management, Republic of India
V. V. Krishna Maddinala	GD Rungta College of Engineering & Technology, Republic of India
Anand N. Khobragade (Scientist)	Maharashtra Remote Sensing Applications Centre, Republic of India
Abdallah Handoura	Computer and Communication Laboratory, Telecom Bretagne, France
Almaz Mobil Mehdiyeva	Department of Electronics and Automation, Azerbaijan State Oil and Industry University, Azerbaijan

Technical Program Committee Members

Ivo Bukovsky, Czech Republic
Maciej Majewski, Poland
Miroslaw Ochodek, Poland
Bronislav Chramcov, Czech Republic
Eric Afful Dazie, Ghana
Michal Bliznak, Czech Republic

Donald Davendra, Czech Republic
Radim Farana, Czech Republic
Martin Kotyrba, Czech Republic
Erik Kral, Czech Republic
David Malanik, Czech Republic
Michal Pluhacek, Czech Republic
Zdenka Prokopova, Czech Republic
Martin Sysel, Czech Republic
Roman Senkerik, Czech Republic
Petr Silhavy, Czech Republic
Radek Silhavy, Czech Republic
Jiri Vojtesek, Czech Republic
Eva Volna, Czech Republic
Janez Brest, Slovenia
Ales Zamuda, Slovenia
Roman Prokop, Czech Republic
Boguslaw Cyganek, Poland
Krzysztof Okarma, Poland
Monika Bakosova, Slovak Republic
Pavel Vaclavek, Czech Republic
Olga Brovkina, Czech Republic
Elarbi Badidi, United Arab Emirates

Organizing Committee Chair

Radek Silhavy — Tomas Bata University in Zlin, Faculty of Applied Informatics
email: comesyso@openpublish.eu

Conference Organizer (Production)

Silhavy s.r.o.
Web: https://comesyso.openpublish.eu
Email: comesyso@openpublish.eu

Conference Website, Call for Papers

https://comesyso.openpublish.eu

Contents

An Artificial Intelligence-Based Concept and Survey for Personalized Learning in Schools in the Republic of Bulgaria 1
Ekaterina Ivanova and Pavel Zlatarov

Correlation Analysis and Predictive Factors for Building a Mathematical Model 14
V. A. Nelyub, V. S. Tynchenko, A. P. Gantimurov, K. V. Degtyareva, and O. I. Kukartseva

Orca Predator Algorithm for Feature Selection 26
Camilo Ravelo, Sebastian Medina, and Rodrigo Olivares

Skeleton Detection Using MediaPipe as a Tool for Musculoskeletal Disorders Analysis 35
Josef Böhm, Taotao Chen, Karel Štícha, Jan Kohout, and Jan Mareš

Model Design Research Plan for Warehouse Barcode Image Recognition in Smart Systems 51
Jan Tyrychtr, Shady Aly, Adéla Hamplová, and Tomáš Benda

Robust H_∞ Controller Design for Satellite Systems with Uncertain Inertia Matrix: A Linear Matrix Inequality Approach 58
Ibrahim Shaikh, Samuel Emebu, and Radek Matušů

Pipeline Leakage Detection via Extreme Seeking Entropy 67
Jakub Steinbach, Jakub Seiner, and Jan Vrba

Holographic Discs-Volumetric Media and Quality Data Storage Media 75
Salahaddin Yusifov, Elnare Firdus, Durdana Rustamova, Veyis Aliyev, Sabina Sharifli, Rauf Mayilov, and Almaz Mehdiyeva

Ensuring Operational Continuity and Safety in Refineries Through a Robust Emergency Shutdown System Design and Implementation 82
Salahaddin Yusifov, Rauf Mayilov, Abdulaga Gurbanov, Ijabika Sardarova, Kerim Bagirzade, Mehriban Mammadova, Sevil Ahmadova, and Zarifa Mahmudova

Cyber-Physical Fire Detection and Recognition System with Smart Glasses 93
Nikolay Gospodinov and Georgi Krastev

Trends and Challenges in Surveillance - A Systematic Review of Camera Systems Implementing Artificial Intelligence 103
Boyana Ivanova, Kamelia Shoilekova, and Rumen Rusev

Usage of the Summary Model DELIS-CH for Starting the Design Process of an Educational Video Game for Cultural Heritage 113
Yavor Dankov and Andjela Dankova

User-Oriented Dashboard Design Process for the DIZU-EVG Instrument for Visualizing Results from Educational Video Games 121
Yavor Dankov

Heuristic Approaches to Delivering Cloud Resources at Minimal Cost in IT Infrastructure Management 130
Georgi Shipkovenski and Oleg Asenov

Vocal Folds Image Segmentation Based on YOLO Network 141
Jakub Steinbach, Zuzana Urbániová, and Jan Vrba

Knowledge Discovery Systems: An Overview 150
Serafeim A. Triantafyllou

Novel Radio Scheduling Framework for Optimal Energy Efficiency in Wireless Sensor Network 160
K. Deepa Mathew and T. Anita Jones Mary Pushpa

A Novel Approach of Intrusion Detection System for IoT Against Modern Attacks Using Deep Learning 172
A. Durga Bhavani and Neha Mangla

Predictive Classification Framework for Software Demand Using Ensembled Machine Learning 183
Salma Firdose and Burhan Ul Islam Khan

Secure Data Transmission Scheme in Wireless Sensor Network Resisting Unknown Lethal Threats 196
Chaya Puttaswamy and Nandini Prasad Kanakapura Shivaprasad

An IoT-Based Cloud Data Platform with Real-Time Connecting Maritime Autonomous Surface Ships 208
Hyoseong Hwang and Inwhee Joe

Deep Learning-Based Tag Mapping Automation of Ship Data Models with Natural Language Processing 221
Jiawei Huang, Hyoseong Hwang, and Inwhee Joe

Deep Reinforcement Learning-Based Task Offloading in Multi-access Edge Computing for Marine IoT 233
Ducsun Lim and Dongkyun Lim

Neural Network Development for Quality Analysis of ERP Systems 245
A. D. Selyutin, V. A. Kushnikov, A. S. Bogomolov, A. F. Rezchikov, V. A. Ivashchenko, E. V. Berdnova, T. V. Pakhomova, O. I. Dranko, I. A. Stepanovskaya, A. A. Kositzyn, and A. A. Dnekeshev

Designing an Algorithm for Recognizing the Kazakh-Latin Alphabet in an Image 254
Zhumazhan Kulmagambetova, Damir Murzagulov, Ulmeken Smailova, Gulmira Shangytbayeva, and Bazargul Kulzhagarova

Influence of TMDC Layers on the Optical Properties of Silicon Nanoparticles 263
Denis Kislov and Vjaceslavs Bobrovs

Flexible GaNP Nanowire-Based Platform: Optical Studies 271
Alina Kurinnaya, Olga Koval, Alex Serov, Vjaceslavs Bobrovs, Igor Shtrom, and Alexey Bolshakov

Transverse Kerker Effects in All-Dielectric Conical Nanoparticles 278
Alexey V. Kuznetsov and Vjaceslavs Bobrovs

Generalized Kerker Effects in All-Dielectric Conical Nanoparticles 283
Alexey V. Kuznetsov and Vjaceslavs Bobrovs

Ultrashort Pulse Generation in Spaser Through Nonlinear Regime 288
Morteza A. Sharif, Mehdi Borjkhani, and Vjaceslavs Bobrovs

Numerical Solution of Mass Transfer Resistances Problem in an Electrolysis Process 292
Ever Peralta-Reyes, Iris C. Valdez-Dominguez, Alejandro Regalado-Méndez, Reyna Natividad, Edson E. Robles-Gómez, Hugo Pérez-Pastenes, and Rubi Romero

Trust in Electronic Record Management System: Insights from Islamic-Based Professional and Moral Engagement-Based Digital Archive 303
Miftachul Huda, Reda Owis Hassan Serour, Mukhamad Hadi Musolin, Mohd Azman, Andi Muhammad Yauri, Abu Bakar, Muhammad Zuhri, Mujahidin, and Uswatun Hasanah

Digital Record Management in Islamic Education Institution: Current Trends on Enhancing Process and Effectiveness Through Learning Technology 316
Miftachul Huda, Mukhamad Hadi Musolin, Reda Owis Hassan Serour, Mohd Azman, Andi Muhammad Yauri, Abu Bakar, Muhammad Zuhri, Mujahidin, and Uswatun Hasanah

Integrating Machine Learning Algorithms with EEG Signals to Identify Emotions Among University Students 334
Mohd Fahmi Mohamad Amran, Venothanee Sundra Mohan, Nurhafizah Moziyana Mohd Yusop, Yuhanim Hani Yahaya, Muhammad Fairuz Abd Rauf, Noor Afiza Mat Razali, Fazilatulaili Ali, and Sharifah Aishah Syed Ali

Comparison of Game Development Framework and Model for Parkinson Disease Rehabilitation 343
Muhammad Fairuz Abd Rauf, Saliyah Kahar, Mohd Fahmi Mohamad Amran, Suziyanti Marjudi, Zuraidy Adnan, and Rita Wong

Information and Communication Skills for Higher Learners Competence Model 357
Muhammad Hasbi Abd Rahman, Jazurainifariza Jaafar, and Miftachul Huda

Twitter Sentiment Analysis with Machine Learning for Political Approval Rating 377
Rodrigo Loayza Abal, Juan J. Soria, and Lidia Segura Peña

Evaluation of Multiplatfom Component for Biometric Authentication in Low-Code Programming Platform – Case Study 398
Zdzisław Sroczyński

Leveraging Machine Learning and Raspberry Pi for Enhanced Wildlife Remote Monitoring and Localization 410
Fabrice Manzi, Emmanuel Tuyishime, Antoine Hitayezu, Gedeon Muhawenayo, Philibert Nsengiyumva, and Kayalvizhi Jayavel

A Software Library for Managing Groups of Collector Motors in Robotics 424
Rinat Galin, Daniiar Volf, Saniya Galina, and Mark Mamchenko

EdApp as a Tool to Intensify Foreign Language Professional Training in the Digitalization of the Educational Environment 433
Dmitrii Burylin, Damir Ibraimov, Nadezhda Chernova, Natalia Katakhova, Irina Osliakova, Tatiana Kudinova, and Svetlana Katahova

Virtual Reality as a Toolkit in the Professional Training of Students 439
Tatiana A. Shchuchka, Natalia A. Gnezdilova, Nataliya V. Chernousova, and Pavel V. Pankin

Methods for the Formation of an Automated Distribution of Pursuers in Group Pursuit 451
Alexander Dubanov

Author Index 461

An Artificial Intelligence-Based Concept and Survey for Personalized Learning in Schools in the Republic of Bulgaria

Ekaterina Ivanova[1] and Pavel Zlatarov[2(✉)]

[1] Department of Pedagogy, University of Ruse, Ruse 7017, Bulgaria
eivanova@uni-ruse.bg
[2] Department of Computing, University of Ruse, Ruse 7017, Bulgaria
pzlatarov@uni-ruse.bg

Abstract. The modern generation of learners, also known as the 'Digital Generation', needs adequate training, tailored to the ways of perceiving and learning given information. This fact leads to the need to study the possibilities of familiarizing primary school-age students with their surrounding world, by including in the educational process technical devices (smartphones, tablets, televisions, computers), the Internet, activities such as games, animation, videos, music, photography, and more. Furthermore, with the rising popularity of artificial intelligence-based systems, students can use technology to benefit from a highly personalized experience in and outside of the classroom.

This document discusses benefits of introducing AI-based technology in the educational process, describes the most popular online-based educational tools for teaching primary school-aged students in the Republic of Bulgaria, and how they can be used together with artificial intelligence to make personalized recommendations of educational content. Furthermore, the document presents the results of a public survey regarding the use of artificial intelligence for education personalization.

Keywords: Artificial intelligence · personalized education · primary education · online based tools · teachers · pupils

1 Introduction

It has been three years since the spread of COVID-19, which evolved into a global pandemic. These three years imposed a new way of life on society in several areas. It was necessary to observe social distance and suspend all classes in educational institutions and organizations, schools, and universities, and this led to the introduction of a distance form of education.

The COVID-19 pandemic has required the educational system in Bulgaria (and those worldwide) to face difficulties and challenges, forcing changes regarding the organization and course of the educational process, ensuring appropriate means and conditions

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 1–13, 2024.
https://doi.org/10.1007/978-3-031-53549-9_1

for learning and teaching at home, the implementation of communication between teachers and students, etc. [1]. This difficult situation placed the teaching professionals in a position of perceived additional stress related to their professional lives. The challenge they had to face was to immediately switch to a distance form of learning, with no time to adapt to the new role of distance educators. Although in 2018, the Ministry of Education and Science of the Republic of Bulgaria approved the national program 'ICT in the System of Preschool and School Education', which states, 'Access to ICT for today's children is an integral and increasingly important part of access to education. The electronification of the learning process is a key element of the modern school, and the introduction of ICT-based innovations into the education system optimizes the learning process and increases its effectiveness' [2] Distance learning appeared to be something new for the teachers, although some of them had previously used online-based educational resources [3].

The new demands of the pandemic on education have opened up new opportunities for teaching and challenged educators to transcend themselves, break out of the rut of everyday life, to discover new ways of inspiration, as well as unique and irreplaceable learning opportunities that can only exist in an online environment [4].

The need for this is inextricably linked with the modern generation of learners, also known as the 'digital generation', which needs adequate training, tailored to the ways of perceiving and learning given information.

This fact leads to the need to study the possibilities of familiarizing primary school-age students with their surroundings, by including in the educational process technical devices (smartphones, tablets, televisions, computers), the Internet, activities such as games, animation, videos, music, photography, and more.

The introduction of so many technological advancements in the classroom pave the way for other advancements, such as personalized learning, which can be used to deliver content better suited to the learner's individual needs, and artificial intelligence (computer systems that can simulate the human intelligence process and capable of reasoning and/or problem-solving), which has not only gained significant popularity in recent years, but also has huge potential to revolutionize many fields and industries, one of them being education.

2 Artificial Intelligence and Personalized Education

In the recent years, organizations working on artificial intelligence (AI) have demonstrated incredible developments in multiple domains. Most notably, natural language processing models, such as GPT-3 and GPT-4 (both powering the immensely popular ChatGPT) have developed into indispensable tools for language processing and text generation, powering numerous applications across many industries [5].

Artificial intelligence has also advanced in the field of computer vision and image generation – models like DALL-E and others, developed by researchers, are capable of generating images based on natural language description [6], and various models have long been used in fields such as manufacturing and surveillance to tell certain objects apart. Developments have been made in fields such as healthcare, physics and biology; AI models such as AlphaFold, developed by DeepMind, have shown great accuracy and performance in highly complex tasks, such as protein prediction [7].

Of course, artificial intelligence is also useful for augmenting everyday tasks, such as vehicle autonomy and driver assistance, voice assistants in smartphone and smart home applications, translation engines and recommendation engines.

Personalized education has been a very popular topic of discussion in the field of education [8]. It involves tracking and analyzing each student's strengths, weaknesses, learning style and more to provide recommendations for relevant educational content and suggest relevant changes to the curriculum. [9]. In this context, educators, students and other participants in the education process can make use of artificial intelligence systems to tailor educational content and strategies so that it better meets the student's needs. [9] Personalized learning systems that make use of AI may be able to provide rich, adaptive learning experiences, making use of targeted content, resources and exercises to optimize content comprehension, retention and drive engagement. Some systems may adapt the content in real time, while others may analyze the student's overall profile and improve its output over time, but in all cases, AI-driven personalization has the potential to lead to better academic results, higher motivation and incentivize them to apply self-directed learning.

3 Online-Based Teaching Platforms in Bulgaria as Sources of Educational Content

Personalized education systems are powerful tools that require enough data to function correctly. An overview of the concept for using artificial intelligence for personalized education is presented on Fig. 1.

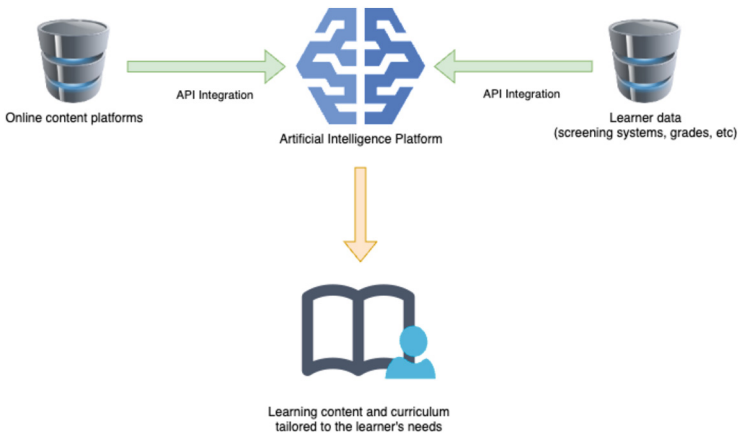


Fig. 1. Overview of the concept for using artificial intelligence for personalized learning

To achieve this, the artificial intelligence system needs three main components:

- A source of data about the learner – this can be data from a screening system, an electronic school journal or grade book, or any other data that can be used to determine the learner's needs.

- A source of educational content – existing content can be assessed, categorized and used in personalized learning, and new educational content can be created by educators based on recommendations, if needed.
- An artificial intelligence platform to bring both data sources together and provide relevant recommendations for suitable content and/or curriculum items.

Since e-learning and online-based teaching have become especially popular since the global pandemic, e-learning platforms have turned into large electronic librariess and knowledge bases that can be a great source of learning content. 'E-learning has become of increasing importance for various reasons, such as the rise of information and the global economy and the emergence of consumer culture. Students of the 21st century demand a flexible structure that allows them to study, work and participate in family life at the same time. This flexibility is reflected in alternative delivery methods, including Internet use' [10].

At the beginning of 2020, immediately after the declaration of a state of emergency in the Republic of Bulgaria, a research team consisting of Mariana Bakracheva and Yanka Takeva conducted a study, the main goal of which was: 'exploring the dynamics in pedagogical communication in the conditions of emergency in a remote situation' [3]. The survey was conducted online and includes 1345 teachers from all over the country. The most popular platforms are presented here, as they have the potential to be strong sources of educational content for building a personalized learning system powered by artificial intelligence.

One of the most used online platforms, according to the research of Bakracheva, and Totseva, 2020, supporting the process of digitalization of education in the Republic of Bulgaria, indicated by 745 teachers is 'Ucha.se' (Fig. 2).

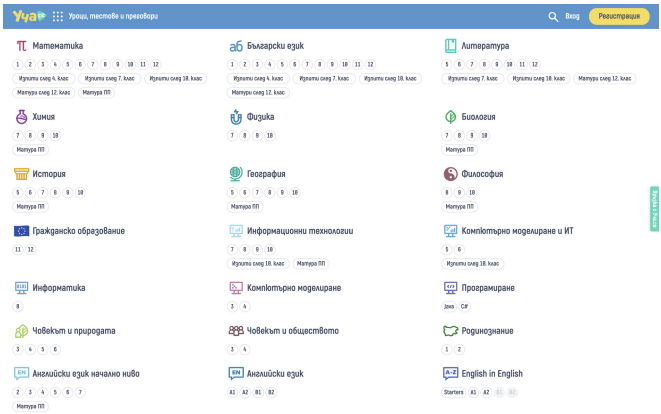


Fig. 2. A screenshot of one of the main screens of Ucha.se

Ucha.se is an online platform containing lessons and covering 97% of the learning material for students from 1st to 12th grade in Bulgarian schools. The resources are in accordance with the State Educational Standards for general education and the official school programs of the Ministry of Education and Science of the Republic of Bulgaria.

Created in 2011 by Darin Madzharov, by 2023 the platform contains more than 25,000 video lessons, tests, and discussions with mind maps, which cover 97% of the educational content studied in Bulgarian schools. The Ucha.se app is also available on Google Play, App Store, and App Gallery.

The website, as well as the mobile applications, are built on the basis of modern web and mobile application development technologies - widespread technological solutions are used, namely the PHP programming language, in addition to JavaScript, jQuery, and an architecture based on 'model-view-controller' template [11].

The other major platform mentioned by 434 of the surveyed teachers is shkolo.bg (Fig. 3).

The platform was built in 2016, and its founders are Lubomir Vanov, Miroslav Jokanov, Simeon Predov, and Alexander Stoyanov. It fulfills the requirements of the National Program 'ICT' in the system of preschool and school education', 2018.

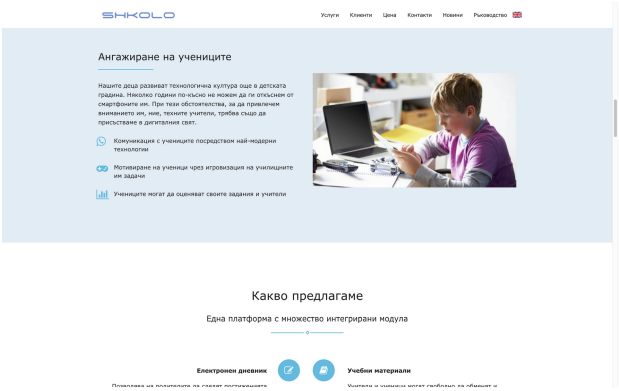


Fig. 3. A screenshot of the main screen of Shkolo.bg

Shkolo.bg unites three separate applications collected in one platform - Academy, Lessons, and Diary. The Academy application provides an opportunity for trainers to participate in face-to-face and online training to enhance their competencies. The Lessons app is e-learning-oriented and has similar functionality to some of the other platforms reviewed in the report such as ucha.se and iZZI. The 'Journal' application is aimed at the administration of the learning process and contains various administrative modules. Shkolo.bg also provides a range of educational resources: electronic lessons, maps, tests, teaching aids, publications, etc.

As of 2023, the Shkolo.bg software platform is used in over 1,800 Bulgarian schools and nearly 2,000,000 users [12]. All modules of the system are built on the basis of modern web technologies.

The following four resources presented in this post are among those identified by 262 teachers who participated in the study by Bakracheva and Totseva, 2020.

Academico.bg is a YouTube channel that hosts educational video lessons designed for students from grades 1 to 7. They are divided into topics in accordance with the state

educational standards for the relevant subject and class. The lessons were developed with the participation of experienced teachers and pedagogues (Fig. 4) [13].

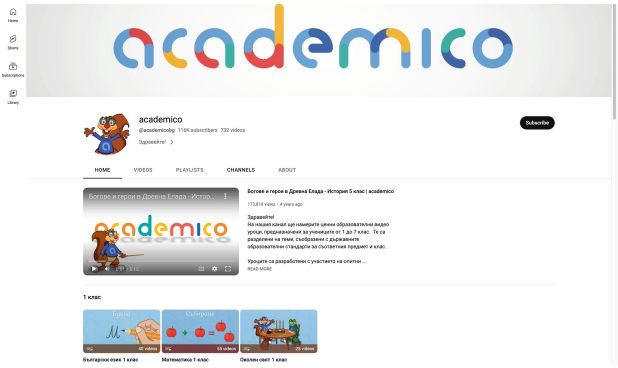


Fig. 4. A screenshot of the main page of the Academico channel

iZZI is an interactive educational environment for learning and self-training (Fig. 5).

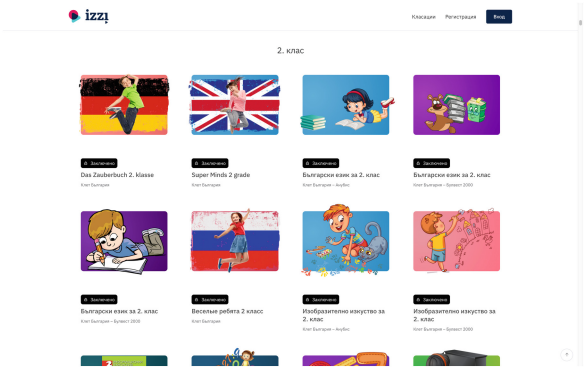


Fig. 5. The main page of iZZi

iZZI includes digital learning content from 3 major publishers gathered in one place. The digital content complements the printed textbooks approved by the Ministry of Education and Science of the Republic of Bulgaria. It also includes multimedia content such as videos, 3D visualizations and audio recordings. It is suitable for face-to-face and remote learning, as well as self-training, and is well suited to working with students with learning difficulties [14].

iZZI is a modern web-based application and integrates with the other platforms and applications of the publisher, using a single user profile to log in to all of them. Mobile apps are also available for Android, iOS, and Huawei devices. Like other platforms already discussed, iZZI is used and supported on a software-as-a-service basis, with various subscriptions available.

The next platform, founded by Alexander Popov and Stefan Kolev, is for creating and solving tests online - SmarTest (Fig. 6).

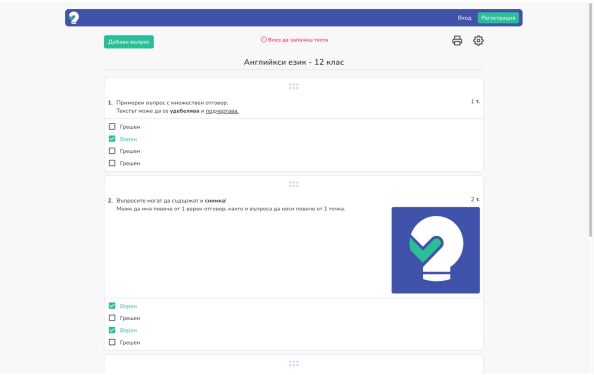


Fig. 6. A screenshot of the test creation screen of SmarTest

The platform is accessed through a web browser and is oriented towards creating tests to assess students' knowledge. Both practice modes are supported, with which students can self-test their knowledge in the process of self-training, and the 'exam' mode, which allows teachers to objectively assess the knowledge of students. It also supports the so-called 'strict mode' which automatically terminates the test when attempts to cheat are detected [15].

Like the other platforms reviewed, SmarTest is available as a web-based application built on modern web technologies, on a software-as-a-service basis, with subscription plans aimed at both individual teachers and schools. A limited free version is also available, with which anyone interested could test the application.

The last platform we will present is programiram.com. (Fig. 7).

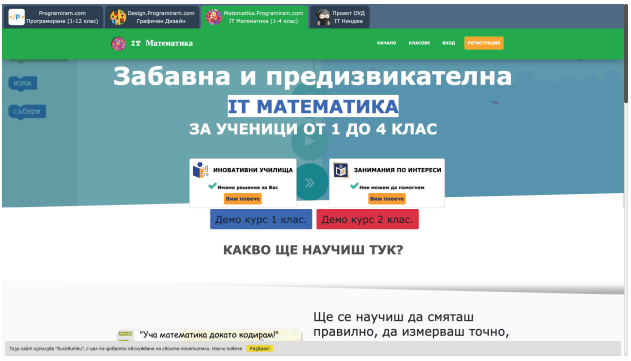


Fig. 7. The main page of Programiram.com

The platform offers online courses, training, and lessons in programming, graphic design, and mathematics for children and students. Animated clips and an automatic grading system are included, and in mathematics, educational games have been developed in accordance with the state educational standard for general education [16].

The platform is accessed via a web browser; applications have also been developed for the most common mobile operating systems, but at the time of accessing the site in June 2023, the applications were not yet available for use by the mass consumer. Like other platforms reviewed, access is done on a software-as-a-service basis, so there is no need to install any additional components, and support is provided by the development team.

4 Results from a Survey on Using Artificial Intelligence for Personalized Education

Especially for the purposes of the research, a non-standardized questionnaire was prepared, containing 10 questions - seven multiple choice and three open. The questionnaire includes questions about using online educational platforms, artificial intelligence and personalized learning. The survey was conducted through the Google Forms application and was active for 10 days. 89 respondents took part in it, and their participation was anonymous and voluntary. Only the researcher has access to the research data.

The survey was completed by 24.7% or 22 parents of students from 1 to 4 grades, 12.4% or 11 parents of students from 5 to 12 grades, 41.6% or 37 undergraduate students. About 21.3% are school students, teachers, doctoral students and university professors (Fig. 8).

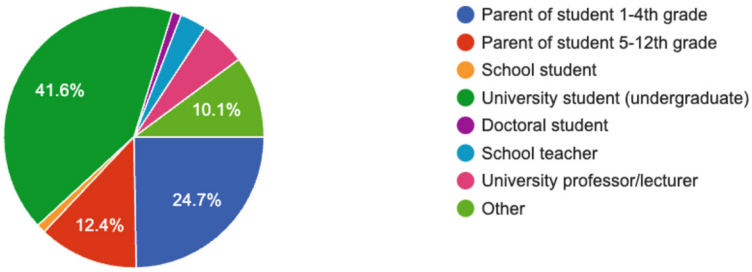


Fig. 8. Pie chart of percentages of different types of responders

The most popular among the respondents is the online educational platform Ucha.se, which was indicated by 98.9% of the respondents. The second place is taken by Shkolo.bg, which is indicated by 88.8%. E-Prosveta.bg is the platform known to 80.9% of the survey participants. After the top three, iZZi takes fourth place with 33.7%. SmarTest and Programiram.com indicate 15% of respondents (Fig. 9).

To the question Have you (or your child, if you are a parent of a student) used any of the following online educational platforms? the results are almost identical to those of the previous question because 86.5% of respondents indicate Ucha.se, 71.9% Shkolo.bg,

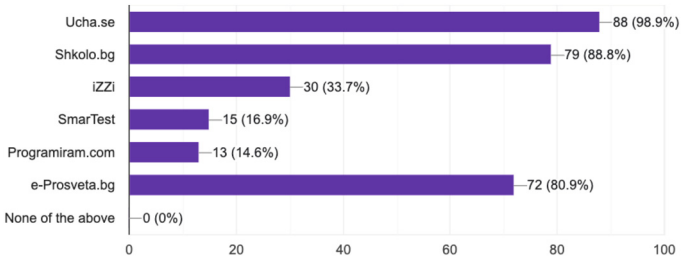


Fig. 9. Familiarity of respondents with the most popular online learning platforms

and 46.1% E-Prosveta.bg, followed by iZZi with 22.5%. Notably, 3.4%, or three total respondents, have never used any of the indicated platforms (Fig. 10).

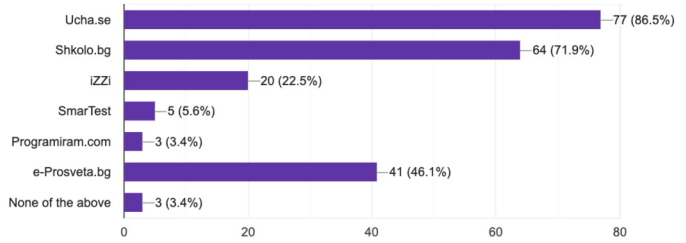


Fig. 10. Usage of the most popular online learning platforms by respondents

The combined results from the previous two questions indicate a strong familiarity of the community with online-based teaching platforms, further supporting that they can be a good starting point for content personalization.

When asked if these online platforms contribute to a higher quality of education, 80.9% of respondents answered positively, 6.7% disagreed and 12.4% could not decide (Fig. 11).

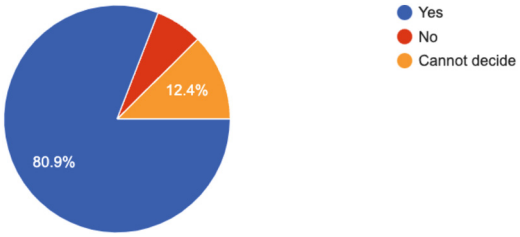


Fig. 11. Distribution of respondents based on their opinion if online-based systems have the potential of improving the education process.

Respondents who believe that platforms contribute to improving the quality of education justified their opinions by stating that electronic educational resources involve

more senses, lead to faster perception and better mastery of knowledge. Easy and timely access to information is a strong point, while information conciseness is also considered an advantage. Multimedia, 3D visualizations and virtual reality were mentioned as valuable resources available to students with online tools. Availability of tools for easy communication between teachers, parents, and students, as well as the timely disclosure of grades are also mentioned as benefits of online learning platforms.

By contrast, some of the respondents who do not agree that online educational platforms contribute to improving the quality of education state that digitization of learning materials in education should be kept to a minimum, and that traditional paper media and handwriting can sometimes be a better form of instruction. In addition, while online platforms can help with children's attention in relation to learning content, one-click accessibility of information isn't always approved, especially by parents.

When asked if they are familiar with the term 'Artificial intelligence', 48 of the respondents answered positively and showed some understanding of the topic.

Some respondents identify artificial intelligence as a "valuable aid", while others as an "enemy of the future generation" and an "obstacle". Some also go on to state that "It wouldn't replace the human mind and human skills", while some think that "progress should not be stopped".

As far as knowledge of AI-based systems, respondents indicated as follows: ChatGPT is known by 77.5%, followed by Bing Chat with 24.7%, Bard – 15.7% and DALL-E – 10.1%. 11.2% are the respondents who are not familiar with any of the mentioned artificial intelligence systems (Fig. 12).

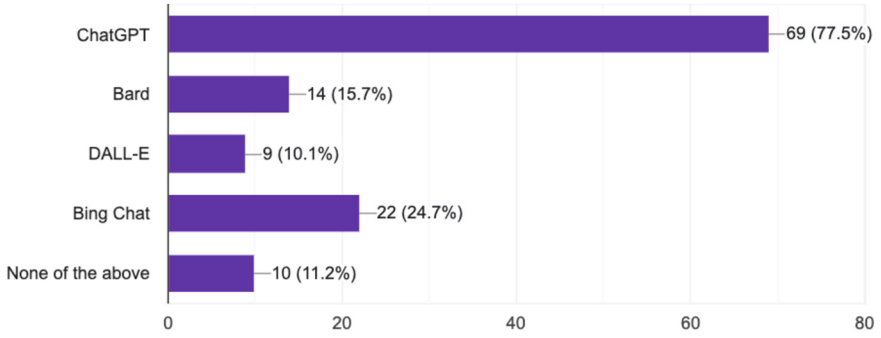


Fig. 12. Familiarity of respondents with AI-based systems

68.5% of respondents indicated that they (or their children) have not used any AI-based system in the educational process so far, and only 31.5% indicated that they have (Fig. 13).

52.3% of survey participants indicated that AI systems can contribute to improving the educational process by being used to offer appropriate learning content tailored to learners' abilities. To the same question, 23.9% answered negatively, and 23.9% could not decide (Fig. 14).

In support of their statement, those who answered positively to the previous question stated that AI has the potential to improve the educational process if used within

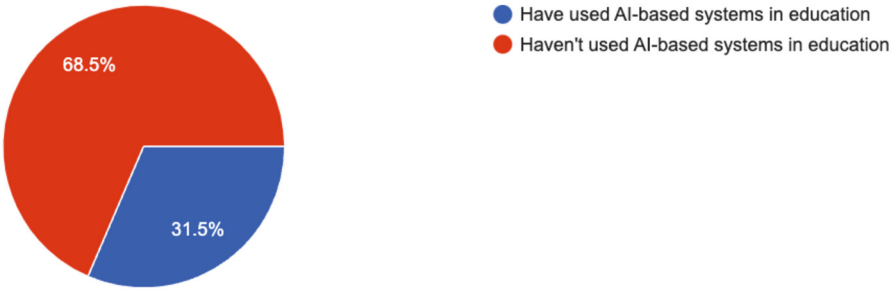


Fig. 13. Usage of AI-based systems in education by respondents

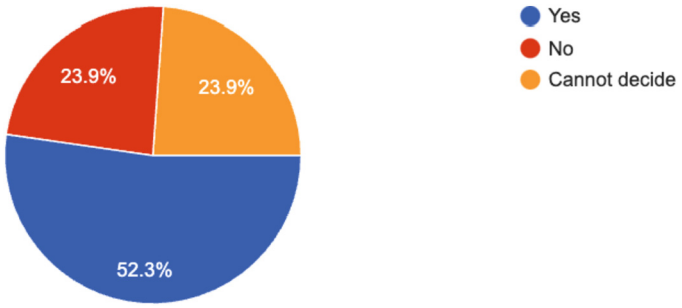


Fig. 14. Survey participants indicating whether they think AI systems can contribute to improving the educational process

reasonable limits. They go on to state that it may help get rid of outdated and unnecessary topics, and that if the system tailors content to the individual student, it can help with information perception and retention. Some respondents emphasize that artificial intelligence can greatly accelerate personalized learning, being especially helpful to children with special educational needs. Furthermore, certain respondents think that AI can help students express themselves better and more correctly and help them with their presentation skills, algorithmic and critical thinking.

Opponents of artificial intelligence also share interesting opinions, namely that too much digitization and the lack of social and human contact with teachers and students can have a massively negative impact and can lead to depression and disinterest. Some even go on to say that artificial intelligence has the potential to wipe out the teaching profession in the near future.

5 Conclusion

Rapid technological advancements open possibilities for innovation in the educational process, among which is personalized education (enabled by availability of vast amounts of learning content, as well as by analysis of learners' needs). Personalization can be further enhanced by artificial intelligence, which can be used to recommend relevant learning content quickly and efficiently. Generally, such innovations are well received

by the public, as evident from the survey, although there is a significant percentage of individuals who are either unsure or disapprove of artificial intelligence being used to personalize educational content.

The following conclusions can be drawn from the survey results:

- The educational platforms Ucha.se, Scholo.bg, e-Prosveta, iZZi, SmarTest, and Programiram.com are known and used by over 90% of respondents, indicating very high familiarity with online learning tools.
- Most respondents indicate that the above-mentioned platforms contribute to improving the quality of education, highlighting their positive aspects such as accuracy and clarity of the information presented, additional visualization, quick and easy access to resources, and support in perceiving and reproducing the educational material.
- Regarding artificial intelligence, nearly 50% of the survey participants stated that they were familiar with the concept, with certain systems, but a large percentage of respondents have never used artificial intelligence systems in an educational context. Half of the respondents believe that AI systems can contribute to improving the educational process by being used to offer relevant learning content tailored to learners' abilities, as well as to accelerate personalized learning, helping children with SEN.
- Opponents of artificial intelligence systems point out that the presentation of ready-made information leads to an inability to think logically, read, and communicate, and the lack of social contact leads to depression, disinterest and asociality.

Acknowledgements. This research is supported by the Bulgarian Ministry of Education and Science under the National Program "Young Scientists and Postdoctoral Students – 2".

References

1. Ivanova, E.: Implementation of the whole-day organization of training in pandemic conditions In: Studying the Impact of Training in an Electronic Environment in Education and the socio-Pedagogical Sphere - Alternatives and Challenges, p. 21(2022)
2. Ministry of education and science. https://www.mon.bg. Accessed July 2023
3. Bacracheva, M., Totseva, Y.: Pedagogical communication in the conditions of emergency, pp. 12–16 (2020). https://fnoi.uni-sofia.bg/wp-content/uploads/2020/04/DOKLAD.pdf
4. Ivanova G., Zlatarov, P., Baeva, D., Antonova, D.: New approaches in doctoral education at the University of Ruse – a response to the challenges of a new era. In: 59-th Annual Scientific Conference of University of Ruse "New Industries, Digital Economy, Society - Projections of the Future III", Book 9.1 Quality of Higher Education, vol. 59, pp. 58–63 (2020)
5. Brown, T., et al.: Language models are few-shot learners. Adv. Neural. Inf. Process. Syst. **33**, 1877–1901 (2020)
6. Ding, M., et al.: Cogview: Mastering text-to-image generation via transformers. Adv. Neural. Inf. Process. Syst. **34**, 19822–19835 (2021)
7. Jumper, J., et al.: Highly accurate protein structure prediction with AlphaFold. Nature **596**(7873), 583–589 (2021)
8. Zlatarov, P., Ivanova, E., Ivanova, G., Doncheva, J.: Design and development of a web-based student screening module as part of a personalized learning system. TEM J. **10**(3) (2021)

9. Zlatarov, P., Ivanova, G., Baeva, D.: A web-based system for personalized learning path tracking of doctoral students. In: 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 773–778. IEEE (2019
10. Tujarov, H.: E-learning and trends in higher education (2009). http://tuj.asenevtsi.com/EL09/EL14.htm
11. Online platform 'ucha.se'. https://www.ucha.se. Accessed July 2023
12. Online platform 'shkolo.bg'. https://www.shkolo.bg. Accessed July 2023
13. Online educational channel 'Academico'. https://www.youtube.com/channel/UCqTGIC4Lli0GZoxpAlvsNaw. Accessed July 2023
14. Educational environment Izzi Digital. https://bg.izzi.digital. Accessed July 2023
15. Online platform SmarTest. https://www.smartest.bg/. Accessed July 2023
16. Online platform Programiram. https://www.programiram.com/. Accessed July 2023



Correlation Analysis and Predictive Factors for Building a Mathematical Model

V. A. Nelyub[1,2(✉)], V. S. Tynchenko[1,3,4], A. P. Gantimurov[1], K. V. Degtyareva[5], and O. I. Kukartseva[6]

[1] Artificial Intelligence Technology Scientific and Education Center, Bauman Moscow State Technical University, 105005 Moscow, Russia
sebastian.medina@postgrado.uv.cl
[2] Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia
[3] Information-Control Systems Department, Institute of Computer Science and Telecommunications, Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia
[4] Department of Technological Machines and Equipment of Oil and Gas Complex, School of Petroleum and Natural Gas Engineering, Siberian Federal University, 660041 Krasnoyarsk, Russia
[5] Department of Information Economic System, Institute of Engineering and Economics, Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia
[6] Department of Systems Analysis and Operations Research, Institute of Informatics and Telecommunications, Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia

Abstract. The study, published in the journal Nature Medicine, looked at data on 1,000 people from China who were tracked over an average period of six years. The participants were divided into two groups: those who lived in areas with high levels of air pollution and those who lived in areas with low levels of air pollution. The study analyzed data on patients with lung cancer, including their age, gender, exposure to air pollution, alcohol consumption, dust allergy, occupational hazards, genetic risk, chronic lung disease, balanced diet, obesity, smoking, passive smoking, chest pain, cough, hemoptysis, fatigue, weight loss, shortness of breath, wheezing, difficulty swallowing, nail thickening and snoring.

Keywords: Data set analysis · Correlation analysis · Neural network prediction · Decision tree algorithm

1 Introduction

A The study, published in Nature Medicine, looked at data from 1,000 people in China who were followed for an average of six years. The participants were divided into two groups: those who lived in areas with high levels of air pollution and those who lived in areas with low levels of air pollution.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 14–25, 2024.
https://doi.org/10.1007/978-3-031-53549-9_2

The dataset contains information about the lung cancer patients, including their age, gender, exposure to air pollution, alcohol consumption, dust allergy, occupational hazards, genetic risk, chronic lung disease, balanced diet, obesity, smoking, passive smoking, chest pain, coughing. Blood, fatigue, weight loss, shortness of breath, wheezing, difficulty swallowing, nail thickening, and snoring.

The initial data collected were scored using a ten-point scale, where 0 is no symptom and 9 is the symptom is maximally expressed. Data Science.

Data science focuses on extracting value from complex data sets. The complexity of different data lies in the variety of different sources from which it has been extracted. Data, given such a challenge, must in turn be organized in such a way that it can be properly interpreted and associated. Different approaches and methods establish causality, with proper investigation establishing meaningful connections between complex interactions in given systems. Thus, the more complex the data, the more complex will be the mathematical approaches including computation, machine learning and system-based methods [1, 6–9].

2 Data Research

From the text file, the data set was loaded into Deductor for analysis. After that, correlation analysis of the data was performed. The results of the analysis are presented in Fig. 1.

Input fields		Correlation with output fields
№	Field	Degree
1	Index	0,058
2	Age	0,060
3	Gender	-0,165
4	Air pollution	0,636
5	Alcohol consumption	0,719
6	Allergy to dust	0,714
7	Harmful working conditions	0,673
8	Genetic risk	0,701
9	Chronic lung disease	0,610
10	Balanced nutrition	0,706
11	Obesity	0,827
12	Smoking rate	0,520
13	Passive smoker	0,704
14	Chest pain	0,645
15	Cough with blood	0,782
16	Patient fatigue	0,625
17	Weight loss	0,353
18	Shortness of breath	0,497
19	Wheezing breath	0,243
20	Difficulty swallowing	0,249
21	Thickening of fingernails	0,280
22	Runny nose	0,444
23	Dry cough	0,374
24	Snoring	0,289

Fig. 1. Significant factors according to the results of correlation analysis

Based on the results of correlation analysis, Kohonen maps were constructed in Fig. 2 and 3.

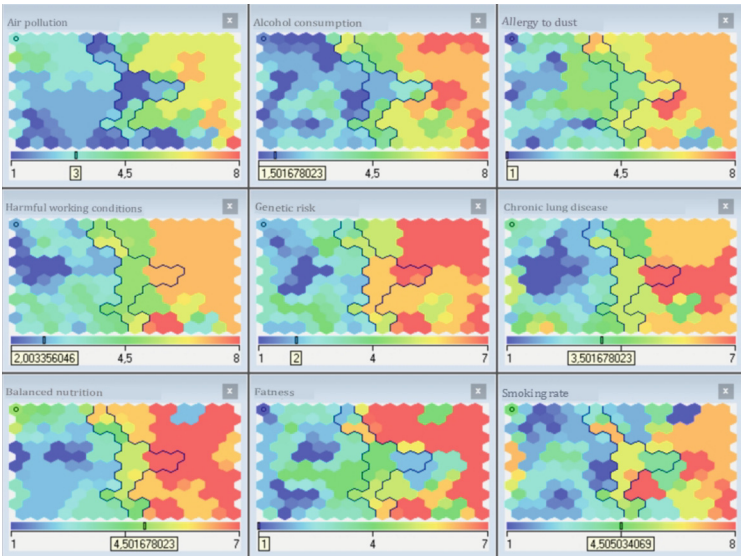


Fig. 2. Kohonen maps

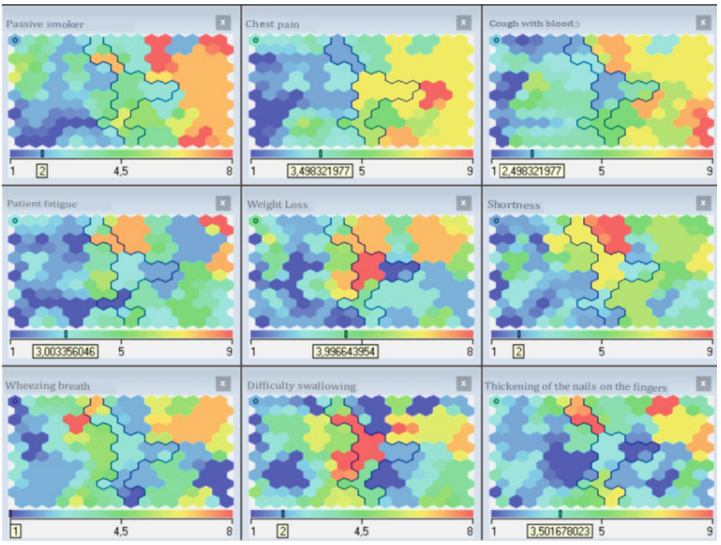


Fig. 3. Kohonen maps

The study's error rate was 0% in Fig. 4.

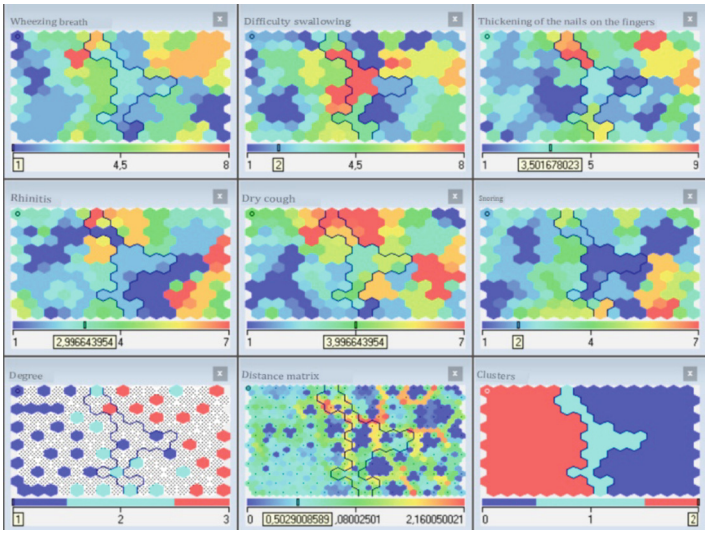


Fig. 4. Kohonen maps

The dataset was then explored using the decision tree method.

All available factors were investigated, with passive smoking (including active smokers) identified as the most significant factor. [10–15].

The error of the study was 0% in Fig. 5 (Fig. 6).

	Classified			
Actually	1	2	3	Total
1	303			303
2		332		332
3			365	365
Total	303	332	365	1000

Fig. 5. Research Error

Target attribute: Degree

№	Number	Attribute	Significance, %
1	12	Passive smoker	38,713
2	15	Patient fatigue	21,164
3	18	Wheezing breath	17,976
4	23	Snoring	11,424
5	10	Fatness	10,723
6	16	Weight Loss	0,000
7	14	Cough with blood	0,000
8	17	Shortness	0,000
9	20	Thickening of the nails on the fingers	0,000
10	21	Rhinitis	0,000
11	22	Dry cough	0,000
12	19	Difficulty swallowing	0,000
13	13	Chest pain	0,000
14	5	Allergy to dust	0,000
15	6	Harmful working conditions	0,000
16	4	Alcohol consumption	0,000
17	2	Paul	0,000
18	3	Air pollution	0,000
19	11	Smoking rate	0,000
20	1	Age	0,000
21	9	Balanced nutrition	0,000
22	7	Genetic risk	0,000
23	8	Chronic lung disease	0,000

Fig. 6. Significance of factors

All factors were then analyzed, excluding passive smoker, patient fatigue, wheezing, snoring, obesity, and weight loss in Fig. 7.

Target attribute: Degree

№	Number	Attribute	Significance, %
1	10	Cough with blood	42,683
2	14	Thickening of the nails on the fingers	14,675
3	13	Difficulty swallowing	14,302
4	1	Air pollution	8,064
5	2	Alcohol consumption	7,037
6	3	Allergy to dust	6,549
7	8	Smoking rate	3,349
8	5	Genetic risk	3,341
9	4	Harmful working conditions	0,000
10	16	Dry cough	0,000
11	15	Rhinitis	0,000
12	12	Shortness	0,000
13	9	Chest pain	0,000
14	7	Balanced nutrition	0,000
15	11	Weight Loss	0,000
16	6	Chronic lung disease	0,000

Fig. 7. Significance of factors

The study had an error rate of 0% in Fig. 8.

	Classified			
Actually	1	2	3	Total
1	303			303
2		332		332
3			365	365
Total	303	332	365	1000

Fig. 8. Research Error

The dataset was then examined for passive and active smoking factors in Fig. 9.

Target attribute: Degree				
No.	Number	Attribute	Significance, %	
1	2	Passive smoker		78,231
2	1	Smoking rate		21,769

Fig. 9. Significance of factors

The study had an error rate of 12.3% in Fig. 10.

	Classified			
Actually	1	2	3	Total
1	241	62		303
2	51	271	10	332
3			365	365
Total	292	333	375	1000

Fig. 10. Research Error

The data were then examined considering only factors independent of health status with the inclusion of smoking in Fig. 11.

Target attribute: Degree

№	Number	Attribute	Significance, %
1	8	Passive smoker	42,718
2	4	Alcohol consumption	28,191
3	7	Smoking rate	7,730
4	1	Age	7,596
5	6	Balanced nutrition	6,323
6	3	Air pollution	3,146
7	2	Paul	3,112
8	5	Harmful working conditions	1,184

Fig. 11. Significance of factors

The study had an error rate of 0.2% in Fig. 12.

	Classified			
Actually	1	2	3	Total
1	302	1		303
2	1	331		332
3			365	365
Total	303	332	365	1000

Fig. 12. Research Error

The dataset was then examined considering only factors directly related to the health of the study patients without taking smoking into account in Fig. 13 [16–21].

Target attribute: Degree

№	Number	Attribute	Significance, %
1	6	Fatness	45,271
2	12	Wheezing breath	20,564
3	17	Snoring	15,538
4	7	Chest pain	14,940
5	9	Patient fatigue	3,688
6	14	Thickening of the nails on the fingers	0,000
7	13	Difficulty swallowing	0,000
8	2	Paul	0,000
9	1	Age	0,000
10	16	Dry cough	0,000
11	15	Rhinitis	0,000
12	8	Cough with blood	0,000
13	4	Genetic risk	0,000
14	5	Chronic lung disease	0,000
15	11	Shortness	0,000
16	10	Weight Loss	0,000
17	3	Allergy to dust	0,000

Fig. 13. Significance of factors

The study's margin of error was 0% in Fig. 14.

	Classified			
Actually	1	2	3	Total
1	303			303
2		332		332
3			365	365
Total	303	332	365	1000

Fig. 14. Research Error

The dataset was then examined considering only factors directly related to the health of the study patients, active and passive smoking [22–28] (Fig. 15).

	Target attribute: Degree		
N	Number	Attribute	Significance, %
1	9	Passive smoker	39,460
2	12	Patient fatigue	20,069
3	15	Wheezing breath	18,328
4	20	Snoring	11,724
5	7	Fatness	10,418
6	19	Dry cough	0,000
7	13	Weight Loss	0,000
8	17	Thickening of the nails on the fingers	0,000
9	18	Rhinitis	0,000
10	16	Difficulty swallowing	0,000
11	14	Shortness	0,000
12	3	Allergy to dust	0,000
13	4	Genetic risk	0,000
14	1	Age	0,000
15	2	Paul	0,000
16	5	Chronic lung disease	0,000
17	10	Chest pain	0,000
18	11	Cough with blood	0,000
19	6	Balanced nutrition	0,000
20	8	Smoking rate	0,000

Fig. 15. Significance of factors

The study's margin of error was 0% in Fig. 16.

	Classified			
Actually	1	2	3	Total
1	303			303
2		332		332
3			365	365
Total	303	332	365	1000

Fig. 16. Research Error

The data were then analyzed taking into account factors directly related to the health of the study patients, active and passive smoking, without taking into account patient fatigue and wheezing in Fig. 17 [30–34].

Целевой атрибут: Степень			
№	Номер	Атрибут	Значимость, %
1	10	Пассивный курильщик	39,218
2	14	Отдышка	27,440
3	7	Хроническое заболевание легких	10,685
4	4	Аллергия на пыль	9,179
5	11	Боль в груди	6,468
6	12	Кашель с кровью	4,155
7	13	Потеря веса	1,524
8	3	Употребление алкоголя	1,331
9	17	Сухой кашель	0,000
10	18	Храп	0,000
11	15	Сложность глотания	0,000
12	16	Насморк	0,000
13	5	Вредность условий труда	0,000
14	2	Пол	0,000
15	1	Возраст	0,000
16	9	Уровень курения	0,000
17	8	Сбалансированное питание	0,000
18	6	Генетический риск	0,000

Fig. 17. Significance of factors

The study's margin of error was 0% in Fig. 18.

	Classified			
Actually	1	2	3	Total
1	303			303
2		332		332
3			365	365
Total	303	332	365	1000

Fig. 18. Research Error

The last method of researching the extent of lung cancer is research using factors that are bad habits - smoking and the level of alcohol consumption in Fig. 19 [36–41].

Target attribute: Degree			
N	Number	Attribute	Significance, %
1	5	Passive smoker	46,984
2	3	Alcohol consumption	28,897
3	4	Smoking rate	12,769
4	1	Age	7,322
5	2	Paul	4,028

Fig. 19. Significance of factors

The study had an error rate of 1.2% in Fig. 19 (Fig. 20).

	Classified			
Actually	1	2	3	Total
1	292	11		303
2	2	330		332
3			365	365
Total	294	341	365	1000

Fig. 20. Research Error

3 Conclusion

As a result of the analysis of several mathematical models, it can be excluded that passive smoking, fatigue and wheezing are possible influencing factors. At the same time, the degree of patient fatigue is a subjective assessment and can be assessed by an already existing disease, regardless of its severity. Thus, both passive and active smoking, as well as wheezing, are important for reliable prediction.

References

1. Gary, W.: Miller Chapter eight - Data science and the exposome, pp. 181–209 (2020)
2. Nagar, D., Pannerselvam, K., Ramu, P.: A novel data-driven visualization of *n*-dimensional feasible region using interpretable self-organizing maps (iSOM). Neural Netw. **155**, 398–412 (2022)
3. Tang, W., Li, W.: Frictional pressure drop during flow boiling in micro-fin tubes: a new general correlation. Int. J. Heat Mass Transfer **159**, 120049 (2020)
4. Liu, Y., Jiang, Y., Hou, T., Liu, F.: A new robust fuzzy clustering validity index for imbalanced data sets. Inf. Sci. **547**, 579–591 (2021)
5. Li, F., Zhang, X., Zhang, X., Du, C., Xu, Y., Tian, Y.C.: Cost-sensitive and hybrid-attribute measure multi-decision tree over imbalanced data sets. Inf. Sci. **422**, 242–256 (2018)
6. Menzies, T., Kocagüneli, E., Minku, L., Peters, F., Turhan, B.: Chapter 6 - Rule #4: data science is cyclic, pp. 35–38 (2015)
7. Comparison of Data Science Algorithms, pp. 523–529 (2019)
8. Zhu, C., Mei, C., Zhou, R.: Weight-based label-unknown multi-view data set generation approach. Inf. Process. Lett. **146**, 1–12 (2019)
9. Griffiths, G.W., Schiesser, W.E.: Analysis of cornea curvature using radial basis functions–Part II: fitting to data-set. Comput. Biol. Med. **77**, 285–296 (2016)
10. Mariño, L.M., de Carvalho, F.D.A.: Vector batch SOM algorithms for multi-view dissimilarity data. Knowl.-Based Syst. **258**, 109994 (2022)
11. Mariño, L.M., de Carvalho, F.D.A.: Two weighted c-medoids batch SOM algorithms for dissimilarity data. Inf. Sci. **607**, 603–619 (2022)
12. He, S.F., Zhou, Q., Wang, F.: Local wavelet packet decomposition of soil hyperspectral for SOM estimation. Infrared Phys. Technol. **125**, 104285 (2022)
13. Qiang, Z., et al.: Multi-stage design space reduction technology based on SOM and rough sets, and its application to hull form optimization. Expert Syst. Appl. **213**, 11929 (2023)

14. Kang, H., Lee, K.S., Lee, H.Y., Chung, M.J., Chin, A.Y., Kim, T.S.: CT findings of influenza A (H1N1) Pneumonia in adults: pattern analysis and prognostic correlatio. Chest **140**(4), 758A (2011)
15. Rubio-Rivas, M., Corbella, X.: Clinical phenotypes and prediction of chronicity in sarcoidosis using cluster analysis in a prospective cohort of 694 patients. Eur. J. Internal Med. **77**, 59–65 (2020)
16. Barchitta, M., et al.: Cluster analysis identifies patients at risk of catheter-associated urinary tract infections in intensive care units: findings from the SPIN-UTI Network. J. Hosp. Infect. **107**, 57–63 (2021)
17. Wang, R., Fung, B.C., Zhu, Y.: Heterogeneous data release for cluster analysis with differential privacy. Knowl.-Based Syst. **201**, 106047 (2020)
18. Carollo, A., Capizzi, P., Martorana, R.: Joint interpretation of seismic refraction tomography and electrical resistivity tomography by cluster analysis to detect buried cavities. J. Appl. Geophys. **178**, 104069 (2020)
19. Bosikov, I.I., et al.: Modeling and complex analysis of the topology parameters of ventilation networks when ensuring fire safety while developing coal and gas deposits. Fire **6**(3), 95 (2023)
20. Mikhalev, A.S., et al.: The orb-weaving spider algorithm for training of recurrent neural networks. Symmetry **14**(10), 2036 (2022)
21. Moiseeva, K., et al.: The impact of coal generation on the ecology of city areas. In: 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH), pp. 1–6. IEEE (2023)
22. Kukartsev, V., et al.: Analysis of data in solving the problem of reducing the accident rate through the use of special means on public roads. In: 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), pp. 1–4. IEEE (2022)
23. Kireev, T., et al.: Analysis of the influence of factors on flight delays in the united states using the construction of a mathematical model and regression analysis. In: 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), pp. 1–5. IEEE (2022)
24. Kukartsev, V., et al.: Prototype technology decision support system for the EBW process. In: Proceedings of the Computational Methods in Systems and Software, pp. 456–466. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-21435-6_39
25. Kukartsev, V., et al.: Methods and tools for developing an organization development strategy. In: 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), pp. 1–8. IEEE (2022)
26. Malozyomov, B.V.: Improvement of hybrid electrode material synthesis for energy accumulators based on carbon nanotubes and porous structures. Micromachines **14**(7), 1288 (2023)
27. Gutarevich, V.O., et al.: Reducing oscillations in suspension of mine monorail track. Appl. Sci. **13**(8), 4671 (2023)
28. Malozyomov, B.V., et al.: Overview of methods for enhanced oil recovery from conventional and unconventional reservoirs. Energies **16**(13), 4907 (2023)
29. Strateichuk, D.M., et al.: Morphological features of polycrystalline CdS1− xSex films obtained by screen-printing method. Crystals **13**(5), 825 (2023)
30. Bosikov, I.I., et al.: Modeling and complex analysis of the topology parameters of ventilation networks when ensuring fire safety while developing coal and gas deposits. Fire **6**(3), 95 (2023)
31. Malozyomov, B.V., et al.: Study of supercapacitors built in the start-up system of the main diesel locomotive. Energies **16**(9), 3909 (2023)
32. Malozyomov, B.V., et al.: Substantiation of drilling parameters for undermined drainage boreholes for increasing methane production from unconventional coal-gas collectors. Energies **16**(11), 4276 (2023)

33. Masich, I.S., Tyncheko, V.S., Nelyub, V.A., Bukhtoyarov, V.V., Kurashkin, S.O., Borodulin, A.S.: Paired patterns in logical analysis of data for decision support in recognition. Computation **10**(10), 185 (2022)
34. Masich, I.S., et al.: Prediction of critical filling of a storage area network by machine learning methods. Electronics **11**(24), 4150 (2022)
35. Barantsov, I.A., et al.: Classification of acoustic influences registered with phase-sensitive OTDR using pattern recognition methods. Sensors **23**(2), 582 (2023)
36. Bukhtoyarov, V.V., et al.: A study on a probabilistic method for designing artificial neural networks for the formation of intelligent technology assemblies with high variability. Electronics **12**(1), 215 (2023)
37. Rassokhin, A., Ponomarev, A., Karlina, A.: Nanostructured high-performance concretes based on low-strength aggregates. Maga. Civil Eng. **110**(2), 11015 (2022)
38. Rassokhin, A., Ponomarev, A., Shambina, S., Karlina, A.: Different types of basalt fibers for disperse reinforcing of fine-grained concrete. Maga. Civil Eng. **109**(1), 10913 (2022)
39. Shutaleva, A., et al.: Migration potential of students and development of human capital. Educ. Sci. **12**(5), 324 (2022)
40. Efremenkov, E.A., Martyushev, N.V., Skeeba, V.Y., Grechneva, M.V., Olisov, A.V., Ens, A.D.: Research on the possibility of lowering the manufacturing accuracy of cycloid transmission wheels with intermediate rolling elements and a free cage. Appl. Sci. **12**(1), 5 (2021)
41. Shutaleva, A., et al.: Environmental behavior of youth and sustainable development. Sustainability **14**(1), 250 (2021)
42. Repinskiy, O.D., et al.: Improving the competitiveness of Russian industry in the production of measuring and analytical equipment. In: Journal of Physics: Conference Series, vol. 1728, no. 1, p. 012032. IOP Publishing (2021)
43. Balanovskiy, A.E., Shtaiger, M.G., Kondratyev, V.V., Karlina, A.I.: Determination of rail steel structural elements via the method of atomic force microscopy. CIS Iron Steel Rev. **23**, 86–91 (2022)
44. Kondrat'ev, V.V., et al.: Description of the complex of technical means of an automated control system for the technological process of thermal vortex enrichment. In: Journal of Physics: Conference Series, vol. 1661, no. 1, p. 012101. IOP Publishing (2020)
45. Malozyomov, B.V., Kukartsev, V.V., Martyushev, N.V., Kondratiev, V.V., Klyuev, R.V., Karlina, A.I.: Improvement of hybrid electrode material synthesis for energy accumulators based on carbon nanotubes and porous structures. Micromachines **14**(7), 1288 (2023)
46. Potapenko, I., Kukartsev, V., Tynchenko, V., Mikhalev, A., Ershova, E.: Analysis of the structure of germany's energy sector with self-organizing kohonen maps. In: Abramowicz, W., Auer, S., Stróżyna, M. (eds.) BIS 2021. LNBIP, vol. 444, pp. 5–13. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-04216-4_1
47. Borodulin, A.S., et al.: Using machine learning algorithms to solve data classification problems using multi-attribute dataset. In: E3S Web of Conferences. EDP Sciences (2023)
48. Nelyub, V.A., et al.: Machine learning to identify key success indicators. In: E3S Web of Conferences. EDP Sciences (2023)
49. Kukartsev, V.V., et al.: Using digital twins to create an inventory management system. In: E3S Web of Conferences. EDP Sciences (2023)
50. Gladkov, A.A., et al.: Development of an automation system for personnel monitoring and control of ordered products. In: E3S Web of Conferences. EDP Sciences (2023)
51. Kukartsev, V.V., et al.: Control system for personnel, fuel and boilers in the boiler house. In: E3S Web of Conferences. EDP Sciences (2023)
52. Kozlova, A.V., et al.: Finding dependencies in the corporate environment using data mining. In: E3S Web of Conferences. EDP Sciences (2023)



Orca Predator Algorithm for Feature Selection

Camilo Ravelo(✉), Sebastian Medina, and Rodrigo Olivares

Escuela de Ingeniería Informática, Universidad de Valparaíso, Valparaíso, Chile
camilo.ravelo@postgrado.uv.cl

Abstract. In the era of data explosion, the volume and dimensionality of information pose significant challenges to the accuracy and effectiveness of machine learning systems. An efficient alternative to address this challenge is the feature selection problem, which aims to find a subset of components that provide similar results with less computational effort. Feature selection is known as an NP-hard combinatorial problem due to the exponential growth in the number of possible feature subsets as the total number of features increases. In this study, we propose a mechanism for the feature selection using the Orca Predation Algorithm, a novel metaheuristic inspired by the hunting behavior of orcas. This approach has been underexplored for solving combinatorial problems and has shown excellent results in recent applications. We evaluate the performance of the metaheuristic on an electrocardiogram dataset obtained from Kaggle using five machine-learning classification algorithms. Our results indicate that applying the bio-inspired algorithm for feature selection can enhance the performance of these algorithms. For evaluation, we employ three key metrics: F1 score, accuracy, and density. The obtained results show that 4 of 5 hybridizations exhibit improvements. This study opens up new possibilities for the use of this metaheuristic in other problems of similar complexity.

Keywords: Feature selection · orca predator algorithm · multi-objective optimization problem

1 Introduction

Nowadays, the exponential increase in the volume of data from various sources has raised several significant challenges. Among these challenges, the presence of irrelevant, noisy, and high-dimensional data stands out, leading to significantly high processing costs and negatively impacting the effectiveness and precision of machine learning systems. In the last decade, many studies have focused on improving the training phase performance. An efficient alternative is Feature Selection (FS), which aims to find a subset of features that provide similar results and require less computational effort [14]. When redundant or irrelevant features are not considered, an adequate representation of the data is achieved, dimensionality is reduced, and the model's learning process is accelerated, improving its predictive performance [2]. FS provides benefits, such as minimizing overfitting and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 26–34, 2024.
https://doi.org/10.1007/978-3-031-53549-9_3

improving model accuracy. It also allows an increase in the velocity of the training process and provides a better understanding of the complexities inherent in the data [5]. However, FS is also known to be an NP-hard combinatorial problem because it is characterized by exponential growth in the number of possible feature subsets as the total number of features increases, making an exhaustive search impractical [16].

In this work, we propose a mechanism for FS using Orca Predator Algorithm (OPA). OPA is a novel meta-heuristic inspired by the hunting behavior of orcas. This procedure is divided into three steps: (a) the establishment of an orca colony, (b) the chasing phase composed by the driving of prey, encircling of prey, and the adjustment of positions, and finally, (c) the attacking phase guided by attacking preys and the adjustment of positions. We employ this algorithm for two reasons: (a) this approach has been little explored for solving combinatorial problems, and (b) it presents outstanding results in recent applications [9].

In computational experiments, we test five classification machine learning algorithms: Support Vector Machine (SVM) [13], K-Nearest Neighbors (KNN) [3], Random Forest (RF) [7], Decision Trees (DT) [15], and Extreme Gradient Boosting (XGB) [4]. The employed dataset contains electrocardiogram signals, was taken from the Kaggle repository, and contains 187 data points.

The rest of the manuscript is divided as follows: Sect. 2 details different mechanisms for FS. Section 3 exposes the OPA applies to FS. Section 4 exhibits the computational experiments and discussion of results. Finally, in Sect. 5, we describe the main conclusions and propose interesting future line works.

2 Feature Selections

Feature selection is defined as the procedure dedicated to obtaining an optimal subset of features from a larger dataset [11]. The main aim is to find the most reduced yet sufficient set of features that can accurately describe the class label.

Feature selection can be formulated as follows: Let D a set of features defined as $\{f_1, f_2, f_3, \ldots, f_n\}$. The goal is to identify the ideal subset of D, denoted as $D' \subset D$, where $|D'| < |D|$. Here, D' represents the chosen subset of features that will be used for generating the learning models.

FS methods can be categorized into three general types: the Filter method, the Wrapper method, and the Embedded method [8]. Our work focuses on the use of the Wrapper method, which is used to generate an ideal subset of features that yield similar results but with lower computational costs. This method uses a specific learning model to evaluate different feature subsets, aiming to select the one that provides the highest classification accuracy with the least computational effort [10]. In this context, bio-inspired algorithms have already been applied to the feature selection problem [1]. Here, the representation of a solution is generally a binary encoding corresponding to a selected set of features, where 'one' represents that the feature is considered, and 'zero' represents the opposite.

We apply a recent metaheuristic to solve the FS problem. Thus, classifiers are treated as black-box evaluators, which, through their accuracy and F1 score metrics, provide information about the utility of each subset of chosen features.

Accuracy measures the ability of our model to classify examples correctly, while the F1 score provides a balanced metric of the model's performance by considering both precision, recall, and sensitivity. This approach incorporates the principles of a multi-objective strategy that evaluates the metrics provided by both the classification method and the metaheuristic. Thus, the objective is to guide the generation of efficient solutions in terms of classification performance and reduction of the selected features.

3 Orca Predation Algorithm

OPA is a bio-inspired algorithm that simulates the hunting behavior of the orca. Orcas are highly social marine animals that share prey and temporarily leave their group. They use sonar to explore the aquatic environment and communicate with their pod to plan hunting tactics and feed [9]. An orca is n-dimensional vector that stores a potential solution. The population is represented by $X = [x_1, x_2, \ldots, x_n]$.

Algorithm presents tow phases: chase phase and attack phase. The first one is divided into two types of behaviors: drive the prey (to the surface) and encircle the prey. Here, OPA uses a parameter p_1 to adjust the probability of the orca performing these two processes separately. The driving phase will be performed if $r > p_1$. Otherwise, the encircling phase will be performed. Here, r is randomly generated value in $[0, 1]$. The second phase is called attack and simulates the hunting behavior of orcas, employing tactics to close the distance, subdue the prey, and consume it. The algorithm's success depends on sensory input, decision-making, and potential collaboration with other agents.

Chase Phase. This mechanism is based on specific conditions. The first condition occurs when the fish shoal is small, resulting in a reduced swimming space dimension for the orca. The second condition is the opposite: the larger the fish shoal, the larger the hunting environment. Taking these conditions into account, two techniques are abstracted for pursuing the prey.

$$v^t_{chase,1,i} = a \times (d \times x^t_{best} - F \times (b \times M^t + c \times x^t_i)) \tag{1a}$$

$$v^t_{chase,2,i} = e \times x^t_{best} - x^t_i \tag{1b}$$

$$c = 1 - b \tag{1d}$$

$$M = \frac{\sum_{i=1}^{N} x^t_i}{N} \tag{1c}$$

$$\begin{cases} x^t_{chase,1,i} = x^t_i + v^t_{chase,1,i} \; if \; rand > q \\ x^t_{chase,2,i} = x^t_i + v^t_{chase,2,i} \; if \; rand \leq q \end{cases} \tag{1e}$$

The velocity and position are defined in Eqs. (1). Here, $v^t_{chase,1,i}$ indicates the first pursuit method, and $x^t_{chase,1,i}$ represents the position, both of the i-th orca at time t. Then, $v^t_{chase,2,i}$ and $x^t_{chase,2,i}$ represent the same for the second pursuit method. M represents the average position of the group of orcas, a, b, and d

are random numbers between $[0, 1]$, respectively, e is a random number between $[0, 2]$, the value F is equal to 2, and q is a number between $[0, 1]$ that represents the probability of choosing a particular prey pursuit method. In [9], it was determined that OPA performed better with a value of $q = 0.9$.

Now that the fish school is on the surface, the second act is to encircle the prey. Orcas use sonar to communicate with each other and determine their next position based on nearby orcas. Here, we assume that orcas are positioned based on the positions of three randomly selected orcas, and then their position after moving can be calculated, as shown in Eqs. (2).

$$x^t_{chase,3,i,k} = x^t_{j_1,k} + u \times (x^t_{j_2,k} - x^t_{j_3,k}) \tag{2a}$$

$$u = 2 \times (rand - \frac{1}{2}) \times \frac{MaxIter - t}{MaxIter} \tag{2b}$$

where $MaxIter$ represents the maximum number of iterations, j_1, j_2, and j_3 indicate the three randomly selected orcas from the population, with $j_1 \neq j_2 \neq j_3$. $x^t_{chase,3,i,k}$ is the position of the i-th orca after choosing the third pursuit method in iteration t.

During the pursuit, orcas detect the location of prey through sound and adjust their positions accordingly. Orcas will continue pursuing the fish if they perceive it approaching; otherwise, they will remain in their original position. The adjustment is made using Eqs. (3).

$$\begin{cases} x^t_{chase,i} = x^t_{chase,i} \; if \; f(x^t_{chase,i}) < f(x^t_i) \\ x^t_{chase,i} = x^t_i \; if \; f(x^t_{chase,i}) \geq f(x^t_i) \end{cases} \tag{3}$$

Attack Phase. Orcas surround their prey and then take turns entering the circuit formed to attack the prey, thereby feeding. Subsequently, they return to their original position within the enclosure to replace another orca. Let's suppose there are four orcas, corresponding to the four ideal attack positions in the circle. Based on the positions of these four orcas, other orcas may enter the circuit if they wish to do so. The direction in which the orcas wish to return to the enclosure circle after feeding, to replace other orcas, can be determined according to the positions of nearby orcas selected at random. This behavior is mathematically modeled in Eqs. (4).

$$v^t_{attack,1,i} = (x^t_{1st} + x^t_{2nd} + x^t_{3rd} + x^t_{4rd})/4 - x^t_{chase,i} \tag{4a}$$

$$v^t_{attack,2,i} = (x^t_{chase,j_1} + x^t_{chase,j_2} + x^t_{chase,j_3})/3 - x^t_i \tag{4b}$$

$$v^t_{attack,i} = x^t_{chase,i} + g_1 \times v^t_{attack,1,i} + g_2 \times v^t_{attack,2,i} \tag{4c}$$

where $v^t_{attack,1,i}$ represents the velocity vector of the i-th orca hunting the prey at iteration t, $v^t_{attack,2,i}$ indicates the velocity of the vector of the i-th orca returning to the rodeo circuit at iteration t, x^t_{1st}, x^t_{2nd}, x^t_{3rd}, and x^t_{4th} represent the four orcas in the best position. j_1, j_2, j_3 are the three orcas randomly selected from the population seen in the pursuit and where $j_1 \neq j_2 \neq j_3$. $v^t_{attack,i}$ represents

the position of the i-th orca at iteration t after the attack phase, g_1 is a random value between $[0, 2]$, and g_2 is a random number between $[-2.5, 2.5]$.

After this, the orcas use sonar to locate the prey and adjust their positions, similar to the pursuit process. The minimum limit value (lb) of the potential range of the problem is used to determine the orca's position according to the conditions of the algorithm in [9].

3.1 Adapatación de OPA

The adapted OPA pseudocode for the feature selection problem is shown in Algorithm 1.

Algorithm 1: Orca Predation algorithm

Input: $popSize$: population size; T: maximum iteration; p: probability to select driving prey or encircle prey method; q: probability to select a method when the orca group is small or large; F: constant value; n: dimensionality.
objective functions $f_k(\boldsymbol{x})$, $\boldsymbol{x} = \langle x_1, \ldots, x_n \rangle$ $(\forall\ k = \{1, \ldots, K\})$
// produce the first generation of $popSize$ orcas, randomly.
foreach *orca* $o, (\forall\ i = \{1, \ldots, popSize\})$ **do**
  foreach *variable* $j, (\forall\ j = \{1, \ldots, n\})$ **do**
    position $x_{ij}^{t=0} \leftarrow binary(Random())$ and velocity $v_{ij}^{t=0} \leftarrow Random()$;
  end
  call training phase to orca o and compute its $f(x_i^{t=0})$, via Eq. (5);
end
foreach *orca* $o, (\forall\ i = \{1, \ldots, popSize\})$ **do**
  if $f(x_i^{t=0})$ *is better than* $f(x_g)$ **then**
    $x_g \leftarrow x_i^{t=0}$;
  end
end
// produce T-generations of $popSize$ orcas.
$t \leftarrow 1$;
while $t < T$ **do**
  foreach *orca* $o, (\forall\ i = \{1, \ldots, popSize\})$ **do**
    calculate the average and select randomly three orcas;
    if $p > Random()$ **then**
      if $q > Random()$ **then**
        foreach *variable* $j, (\forall\ j = \{1, \ldots, n\})$ **do**
          chase phase for driving to the prey via Eqs. (1)
        end
      else
        foreach *variable* $j, (\forall\ j = \{1, \ldots, n\})$ **do**
          chase phase for encircling the prey via Eqs. (2) and (3);
        end
      end
    else
      foreach *variable* $j, (\forall\ j = \{1, \ldots, n\})$ **do**
        attack phase via Eqs. (4);
      end
    end
    call training phase to orca o and compute its $f(x_i^{t+1})$, via Eq. (5);
  end
  foreach *orca* $o, (\forall\ i = \{1, \ldots, popSize\})$ **do**
    if $f(x_i^{t+1})$ *is better than* $f(x_g)$ **then**
      $x_g \leftarrow x_i^{t+1}$;
    end
  end
  $t \leftarrow t + 1$;
end
return post-process results and visualization;

We highlight lines 8 and 35, where the procedure invokes the *training phase*. The generated solution vector x_{ij}^{t+1} describes *D'*, which the machine learning algorithm will use. Following this, a model is generated from which predictions are made with test sets to obtain the performance metrics F_1 and *Accuracy*. These resulting metrics are used by the orca o to contrast the efficiency of the solution. In this case, efficiency is calculated through a linear scalarization given in Eq. 5:

$$\sum_{(i,j)_{i\neq j} \in K} \overbrace{\frac{\widehat{c} - f_j(\boldsymbol{x})}{\widehat{c} - e_j(\boldsymbol{x}^{best})}\omega_j}^{min} + \overbrace{\frac{f_i(\boldsymbol{x})}{e_i(\boldsymbol{x}^{best})}\omega_i}^{max}, \quad \omega_{(i,j)} \geqslant 0 \tag{5}$$

where $\omega_{(i,j)}$ represents weight of objective functions and $\sum \omega_{(i,j)} = 1$ must be satisfied. Values of $\omega_{(i,j)}$ are defined by analogous estimating. $f_{(i,j)}(\boldsymbol{x})$ is the single-objective function and $e_{(i,j)}(\boldsymbol{x}^{best})$ stores the best value met independently. Finally, $\widehat{c}$ is an upper bound of minimization single-objective functions.

4 Computational Experiments

We conducted empirical assessments of the OPA algorithm. These trials comprised independent executions of five distinct Machine Learning methodologies: SVM, KNN, RF, DT, and XGB. In subsequent stages, we assessed the hybridization of the OPA algorithm with each of the previously mentioned classification techniques.

Experiments were conducted utilizing the "ECG Heartbeat Categorization Dataset" available on Kaggle [6]. Each record corresponds to ECG signals of heartbeats, which could be either normal or afflicted by various arrhythmias and myocardial infarctions. The type of heartbeat for each sample is stored in the last column of each row, represented by the following numbers: Normal (N) = 0, Supraventricular (S) = 1, Ventricular (V) = 2, Fusion (F) = 3, Unclassified (Q) = 4. The dataset encompasses 188 features, reflecting the dimensionality of our problem, denoted as n. We have been provided with two distinct files: a training file with 87,554 records and a test file with 21,892 records. During the training stage, for classes 1, 2, and 4, we selected a random sample of 1000 records. For class 3, we included all rows corresponding to 641 records. This approach aimed to balance the number of records for each class, ensuring accurate classification without disadvantaging classes with fewer records. Subsequently, in the testing (or prediction) stage, all available records were utilized. In both stages, all dataset features were applied in the machine learning techniques.

Finally, for experiments, we employ a computer equipped with an Intel Core i7 processor, 16 GB of RAM, and running the Windows 11 Home Edition operating system. The algorithms were developed using Python 3.9.12 and can be downloaded in [12].

To assess the performance of the algorithms, we employed three fundamental metrics: F_1 score, accuracy (Acc), and the density (Dsd) of the selected features.

With the aim of obtaining more robust and representative results, 30 executions of each algorithm were conducted. Subsequently, a range of statistical measures were applied to these metrics for a detailed analysis. These included the best value, the mean, the median, the standard deviation, and the interquartile range (see Table 1). We use abbreviations OS, OK, OR, OD, and OX, to represent OPA on SVM, KNN, RF, DT, and XGB, respectively.

Table 1. Experimental results

	SVM	OS	KNN	OK	RF	OR	DT	OD	XGB	OX
F_1 score	0.88	0.89	0.85	0.86	0.92	0.93	0.81	0.83	0.92	0.91
$\bar{F_1}$ score	0.87	0.88	0.84	0.85	0.91	0.91	0.80	0.81	0.91	0.91
$F_1(\sigma)$ score	0.87	0.88	0.85	0.85	0.91	0.91	0.80	0.81	0.91	0.91
$\breve{F_1}$ score	0.01	0.01	0	0.01	0	0.01	0.01	0.01	0	0
$F_1(\phi)$ score	0.01	0.01	0.01	0.01	0	0.01	0.01	0.01	0.01	0
Acc	0.86	0.87	0.81	0.82	0.90	0.91	0.76	0.78	0.9	0.90
$\bar{Acc}$	0.84	0.85	0.80	0.81	0.89	0.89	0.74	0.76	0.89	0.89
$Acc(\sigma)$	0.84	0.85	0.80	0.81	0.89	0.89	0.75	0.76	0.89	0.89
$\breve{Acc}$	0.01	0.01	0.01	0.01	0	0.01	0.01	0.01	0.01	0
$Acc(\phi)$	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.1	0.01
Dsd	1	0.39	1	0.40	1	0.36	1	0.40	1	0.40
$\bar{Dsd}$	1	0.44	1	0.44	1	0.43	1	0.46	1	0.43
$Dsd(\sigma)$	1	0.44	1	0.44	1	0.44	1	0.47	1	0.42
$\breve{Dsd}$	0	0.02	0	0.02	0	0.02	0	0.03	0	0.02
$Dsd(\phi)$	0	0.03	0	0.03	0	0.02	0	0.04	0	0.02

According to the results, it is possible to achieve better performances in the feature selection by reducing the density of characteristics. This optimization leads to an increase in the F_1 and Acc metrics. Therefore, the OPA algorithm successfully solves the feature selection problem for the five machine learning models.

5 Conclusion

This study addresses the problem of FS, which is a critical process in machine learning, as it enhances the efficiency of ML models by reducing data dimensionality and eliminating irrelevant or redundant features. We have proposed the use of OPA, a metaheuristic inspired by the hunting behavior of orcas, to solve the FS problem in five ML techniques. The performance of OPA was evaluated on a Kaggle ECG dataset, using various machine learning algorithms for classification. Our results show that OPA, in conjunction with FS, can help improve the

performance of these algorithms, indicating its potential utility in the field of supervised learning.

For the evaluation of the algorithms, three key metrics were used: F1 score, accuracy, and density of the selected features. An analysis of the execution results was conducted, which included determining the best value and calculating averages, medians, standard deviations, and interquartile ranges.

As future work, we plan to use other metaheuristics and comparing with the presented proposal. Further, we propose to apply this approach to cybersecurity databases, education, or similar.

Acknowledgement. R. Olivares is supported by grant ANID/FONDECYT/INICIACIÓN/11231016. C. Ravelo and S. Medina are supported by Programa de Magíster en Informática Aplicada - Universidad de Valparaíso (ReEx No 100.581/2022).

References

1. Abd Elminaam, D.S., Nabil, A., Ibraheem, S.A., Houssein, E.H.: An efficient marine predators algorithm for feature selection. IEEE Access **9**, 60136–60153 (2021)
2. Abiodun, E.O., Alabdulatif, A., Abiodun, O.I., Alawida, M., Alabdulatif, A., Alkhawaldeh, R.S.: A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities. Neural Comput. Appl. **33**(22), 15091–15118 (2021)
3. Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. Am. Stat. **46**(3), 175–185 (1992)
4. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016)
5. Dhal, P., Azad, C.: A comprehensive survey on feature selection in the various fields of machine learning. Appl. Intell. 1–39 (2022)
6. Fazeli, S.: ECG heartbeat categorization dataset (2018)
7. Fratello, M., Tagliaferri, R.: Decision trees and random forests. In: Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics, vol. 374 (2018)
8. Hoque, N., Bhattacharyya, D.K., Kalita, J.K.: MIFS-ND: a mutual information-based feature selection method. Expert Syst. Appl. **41**(14), 6371–6385 (2014)
9. Jiang, Y., Wu, Q., Zhu, S., Zhang, L.: Orca predation algorithm: a novel bio-inspired algorithm for global optimization problems. Expert Syst. Appl. **188**, 116026 (2022)
10. Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R.P., Tang, J., Liu, H.: Feature selection: a data perspective. ACM Comput. Surv. (CSUR) **50**(6), 1–45 (2017)
11. Nguyen, B.H., Xue, B., Zhang, M.: A survey on swarm intelligence approaches to feature selection in data mining. Swarm Evol. Comput. **54**, 100663 (2020)
12. Olivares, R., Medina, S., Ravelo, C.: Orca predator algorithm for feature selection (2023). https://doi.org/10.6084/m9.figshare.23624613
13. Platt, J.: Sequential minimal optimization: a fast algorithm for training support vector machines. In: Advances in Kernel Methods-Support Vector Learning, vol. 208 (1998)

14. Remeseiro, B., Bolon-Canedo, V.: A review of feature selection methods in medical applications. Comput. Biol. Med. **112**, 103375 (2019)
15. Trabelsi, A., Elouedi, Z., Lefevre, E.: Decision tree classifiers for evidential attribute values and class labels. Fuzzy Sets Syst. **366**, 46–62 (2019)
16. Xue, B., Zhang, M., Browne, W.N., Yao, X.: A survey on evolutionary computation approaches to feature selection. IEEE Trans. Evol. Comput. **20**(4), 606–626 (2015)

Skeleton Detection Using MediaPipe as a Tool for Musculoskeletal Disorders Analysis

Josef Böhm[1(✉)], Taotao Chen[2], Karel Štícha[2], Jan Kohout[2], and Jan Mareš[1,2]

[1] Faculty of Electrical Engineering and Informatics, University of Pardubice, Studentská 95, 532 10 Pardubice 2, Czech Republic
josef.bohm@upce.cz

[2] Department of Mathematics, Informatics and Cybernetics, University of Chemistry and Technology in Prague, Technická 1905/5, 166 28 Praha 6, Czech Republic

Abstract. Skeleton detection, also known as human pose estimation (HPE), is becoming more and more popular as it can be applied in a range of applications such as game entertainment, human-machine interaction, VR-based projects, medical rehabilitation, etc. Thanks to the booming development of deep learning, HPE solutions can be implemented using deep learning methods which require standard 2D RGB images or video sequences as input. That is, technology nowadays is making HPE solutions more and more lightweight and fast which is possible to run on mobile devices for the daily use of skeleton detection. This article covers a brief survey of current deep learning-based human pose estimation approaches in the first place. Then, a lightweight deep learning model – MediaPipe – will be illustrated from all the perspectives of its structure, working flow, strengths & weaknesses and the more concerned compatibility in platforms and programming languages. As a result, a multi-platform application for collecting movement data from patients suffering from musculoskeletal diseases relying on MediaPipe is introduced. Finally, there is a summary of achievements and obstacles of application development, which is significant as it can be a signpost for teams who are doing or about to do an application based on the MediaPipe library.

Keywords: Skeleton detection · MediaPipe · Desktop application · Mobile application · Musculoskeletal disorders · Deep learning · Windows · C# · Image processing

1 Introduction

Nowadays, skeleton detection or human pose estimation plays a more and more important role in many aspects of people's life such as healthcare, security, video game and animation industry, etc. Skeleton detection is actually a computer

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 35–50, 2024.
https://doi.org/10.1007/978-3-031-53549-9_4

vision technique that aims to capture or track the human skeleton within an image or video. The basic idea to achieve skeleton detection is to identify some key points or joints of a human body, such as the head, shoulders, elbows, wrists, hips, knees, and ankles, and connect them to form a skeletal structure. The primary purpose of skeleton detection is to understand the pose, movement, and spatial relationships of human bodies in visual data. And this information can then be used in practical applications.

The requirements of such human body detection have motivated a series of input devices such as Microsoft's Xbox 360 Kinect, PS4 eye, Intel RealSense Camera, etc. All these devices use the so-called depth sensor together with a standard RGB camera to sense the 3D features of the body regardless of the illumination. However, the working mechanism of the Kinect like cameras limits the use on mobile devices such as smartphones, tablets, etc. which apparently represents the trend of the future. Due to high manufacturing and production costs, Microsoft has decided to discontinue its Kinect camera.

In recent years, there has been a focus on developing lightweight and efficient models for real-time and mobile applications. These models aim to strike a balance between accuracy and computational efficiency to enable real-time skeleton detection on devices with limited resources, such as smartphones or edge devices. A typical example is MediaPipe, which is a framework integrated with pretrained models and tools to enable the applications of skeleton detection in a wide range of aspects, healthcare, gaming, sports analysis, and more. In this article, a MediaPipe-based solution for human skeleton detection applied to help patients with medical rehabilitation will be introduced.

1.1 Medical Background

Skeletal tracking and human pose estimation have emerged as valuable tools in the field of medical research and clinical applications. Accurate assessment of human movement and posture is essential for diagnosing and monitoring various conditions, particularly those related to gait disorders or musculoskeletal impairments. By analyzing skeletal data, clinicians and researchers can gain insights into patients' motor function, identify abnormalities, track progress, and develop targeted rehabilitation strategies.

The ability to detect and track the human skeleton using computer vision techniques has opened up new possibilities for non-invasive and objective analysis of movement patterns. Traditional methods for assessing gait and posture often rely on subjective observations or expensive and time-consuming motion capture systems. With skeletal tracking, healthcare professionals can obtain valuable quantitative data in a more efficient and cost-effective manner.

In the medical field, skeletal tracking has found applications in various areas. A prominent example is the diagnosis and management of gait disorders. Gait abnormalities can be indicative of underlying neurological, orthopaedic, or musculoskeletal conditions. By analysing movement patterns captured through skeletal tracking, clinicians can assess parameters such as step length, cadence, symmetry, and joint angles, aiding in the identification and evaluation of gait impairments.

Furthermore, skeletal tracking has been proven beneficial in rehabilitation settings. By closely monitoring a patient's movements and body posture, therapists can objectively evaluate the effectiveness of interventions and track the progress of rehabilitation programmes. This data-driven approach allows for personalised treatment plans and better outcomes for patients undergoing physical therapy or recovering from injuries.

In addition to clinical applications, skeletal tracking has the potential for preventive healthcare and telemedicine. It enables remote monitoring and assessment of patients' motor function, providing healthcare professionals with valuable information for early detection of movement abnormalities or changes in posture. This technology holds promise for home-based rehabilitation programmes, allowing patients to receive ongoing support and guidance from healthcare providers without the need for frequent in-person visits.

As the field of skeletal tracking continues to advance, there is a growing focus on developing lightweight and efficient models that can be deployed on portable devices. This shift to mobile applications opens opportunities for point-of-care diagnostics, ambulatory monitoring, and integration with wearable devices. These advancements hold the potential to revolutionise the way we evaluate and manage musculoskeletal conditions, enabling more accessible and personalised healthcare solutions.

In general, the integration of skeletal tracking into medical research and clinical practise has the potential to enhance diagnostics, improve treatment outcomes, and transform the way we understand and address gait disorders and musculoskeletal impairments. With continued advancements in technology and increasing adoption in healthcare settings, skeletal tracking is poised to play an integral role in improving patient care and promoting better overall musculoskeletal health.

2 Materials and Methods: Principles and Technologies of Human Skeleton Detection

For decades, people have been developing skeleton detection techniques. Motivated by advancements in computer vision, deep learning, and camera technologies, people have gained huge progress in skeleton detection with one after another improved method. An overview of how human pose estimation problem is being developed can be summarized with Table 1 below [7].

Usually, human pose estimation can be divided into 2 groups: 2D human pose estimation and 3D human pose estimation.

2D human pose estimation works with localization of key points of a human body in 2D space, namely obtaining the X and Y coordinates of human body within an image or a video frame [19].

3D human pose estimation, similarly, is associating with predicting the spatial coordinates of a human body in an image or frame of video. The purpose of human pose estimation is to reconstruct a human body models with some key

Table 1. Classification of HPE approaches.

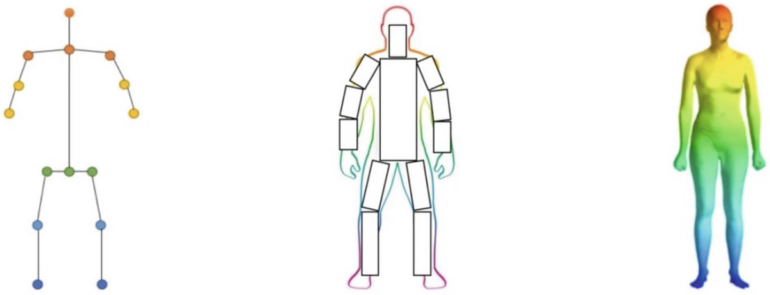
An review of classified human pose estimation approaches	
Type of input	RGB image; Depth (Time of Flight) image; Infra-red (IR) image)
Number of cameras	Single-view; Multi-view
Human body models	Kinematic; Planar; Volumetric
Type of images	Static; Frames from video sequences
Number of people	Single-person pose estimation; Multi-person pose estimation
Dimension of HPE	2D pose estimation; 3D pose estimation

points detected from the input data of human body. There are 3 types of human body models used as an aid in human pose estimation: kinematic model, planar model, and volumetric model as shown in Fig. 1.

Kinematic model, also known as a skeleton-based model, is used for 2D and 3D pose estimation. It contains less information than the other two models, as it only represents the structural information supported by human body joints. As its shortcoming, the kinematic model is not efficient for representing texture or shape information [18].

Planar model, or contour-based model, as the name tells, is used for 2D pose estimation. Unlike the kinematic model, this model is used to represent the human body's appearance and shape. As Fig. 1b shows, within this model, the body parts are illustrated by several rectangles approximating the human body contours [22].

Volumetric model, also called volume-based model, is used to represent 3D human body with geometric shapes or meshes. The common applications of this type of model are VR game, animation production, human-machine interactive



(a) Kinematic model (b) Planar model (c) Volumetric model

Fig. 1. Three types of human body models [15].

system etc. To go through the principles and technologies of human pose estimation clearly, this part will illustrate the classic approaches and deep learning based approaches human pose estimation respectively.

2.1 Traditional Approaches for Human Pose Estimation

Traditional approaches for human pose estimation rely on basic computer vision algorithms such as Histogram of Oriented Gradients (HOG), Scale Invariant Feature Transform (SIFT), canny edge detector, template matching, Pictorial Structure Model (PSM) [8] etc. HOG [20] is a feature descriptor like the Canny Edge Detector. Both are used in computer vision field to detect the edges of an object in an image. HOG features are extremely popular features for human pose estimation [9], usually HOG templates containing information of various states of human body parts are learnt in advance. Li et al. (2016) [16] propose a template matching-based method to classify detected human body poses into the existed labeled action to achieve the human action recognition. However, these approaches are involved with several shortcomings, feature extraction executions are lack of accuracy and have high complexity, as those feature extraction methods are strongly influenced by the environment such as illumination, background, etc. To overcome these shortcomings, deep learning-based approaches are proposed.

2.2 Deep Learning Approaches for Human Pose Estimation

With the rapid development of deep learning solutions for various applications, deep learning-based approaches for human pose estimation are emerging, as can be seen in Fig. 2.

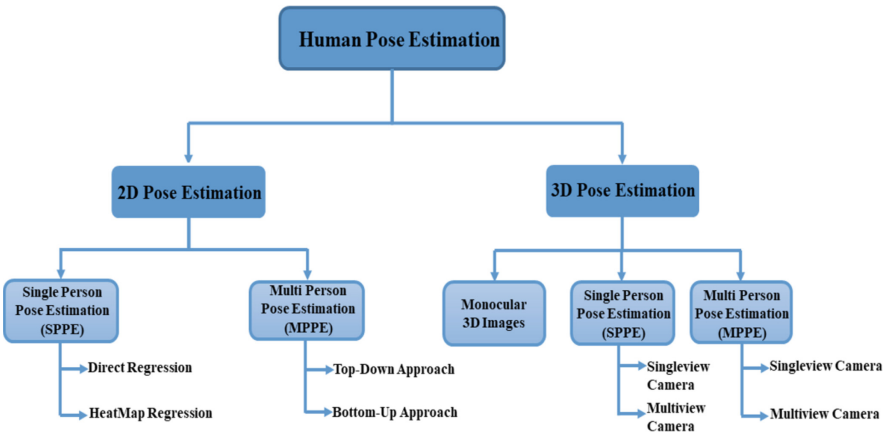
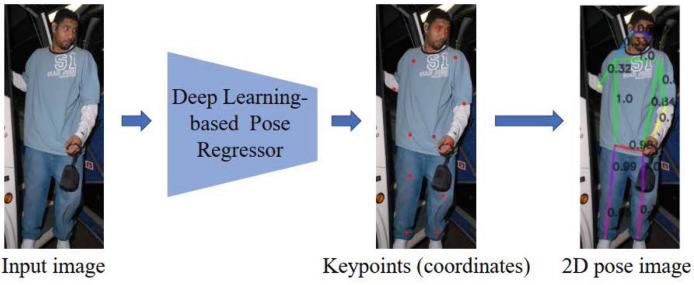
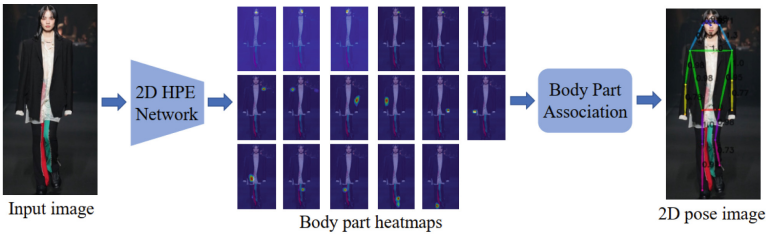


Fig. 2. An overview of deep learning-based approaches for HPE [7].

2.2.1 2D Human Pose Estimation. Situations in 2D human pose estimation are single-person pose estimation (SPPE), multi-person pose estimation (MPPE). In SPPE, only one person can be detected. If there are more than one person in an image, the image is cropped until only one person in the sub-image appears before the next step. Mainly, there are two methods which are based on deep learning for 2D SPPE: regression methods and heatmap-based methods. Regression methods use an end-to-end framework to learn the mapping from the input image to body joints. Heatmap-based methods predict approximate locations of body joints, which are supervised by heatmaps representation. The working flow of these methods can be described in Fig. 3.



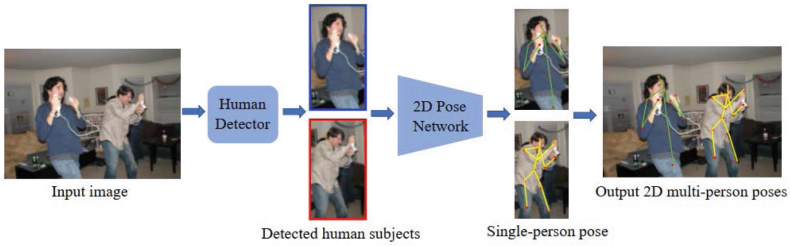
(a) Regression Methods



(b) Heatmap-based Methods

Fig. 3. 2D SPPE frameworks. (a) Regression methods build a mapping based on deep neural network between the original image and the kinematic body model. (b) Heatmap-based methods predict body joints using heatmap supervision [23].

For 2D MPPE, there are two approaches with different working principles, the top-down approach and the bottom-up approach (see Fig. 4). Top-down method runs firstly a body detector to find individual body in an image and then estimates the body joints for each body within the detected bounding boxes. In contrast to the top-down method, the bottom-up method estimates the feature points before it can form one or more bodies with these points [18].



(a) Top-Down Approaches



(b) Bottom-Up Approaches

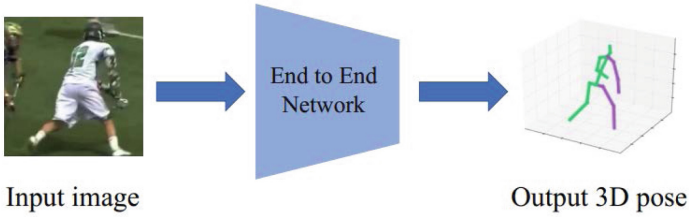
Fig. 4. Two different HPE methods in Multi-person situation [23].

2.2.2 3D Human Pose Estimation. 3D HPE can be divided into three solution groups: single-view single-person 3D HPE, single-view multi-person 3D HPE, and multi-view 3D HPE. According to the result to be obtained, the solutions for single-view single-person 3D HPE can be classified into skeleton-only and human mesh recovery categories. The main difference between these is that the human mesh recovery method employs a human model to get a 3D points estimation. For the skeleton result, there is a direct estimation method and a 2D to 3D lifting method. Figure 5 shows theoretically the two approaches for 3D HPE in situation of single-view and single-person. Similar to multi-person 2D HPE, for single-view multi-person 3D HPE, top-down approaches and bottom-up approaches are hired again. To overcome the shortcomings caused by single view strategy like: there is no detection robust when persons in images have occlusions problems, multi-view method is added to the above-mentioned approaches.

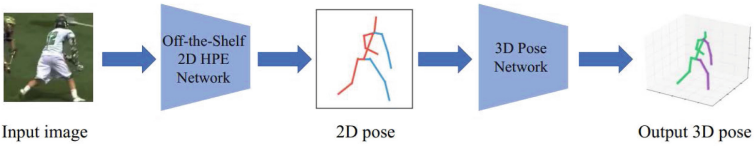
2.2.3 Major Deep Learning Models for Human Pose Estimation. Based on the approaches introduced above, many deep learning models are developed for human pose estimation in recent years. Some popular pose estimation methods based on deep learning here as examples are: OpenPose, BlazePose, DeepPose, DeepCut, PoseNet and Dense Pose.

2.3 MediaPipe: A Comprehensive Machine Learning Framework

MediaPipe is an open-source, cross-platform pipeline framework developed by Google. It was created to for original purpose of analyzing YouTube videos and



(a) Direct estimation approaches for skeleton detection



(b) 2D to 3D lifting approaches for skeleton detection

Fig. 5. Single-person 3D HPE frameworks for skeleton result. (a) Direct estimation approaches directly estimate the 3D human pose from 2D images. (b) 2D to 3D lifting approaches leverage the predicted 2D human pose (intermediate representation) for 3D pose estimation. [23]

audio in real time. Currently, this framework is still in its alpha stage and it covers multi platforms like MacOS, Windows, Android, iOS and other embedded devices like Raspberry Pi and Jetson Nano. MediaPipe offers a range of pre-built components and tools that facilitate the development of real-time applications. These components include modules for video and audio processing, 2D and 3D perception, gesture recognition, and more. Developers can leverage these components to create custom pipelines tailored to their specific application needs.

2.3.1 Architecture of MediaPipe. MediaPipe is divided into three primary parts [17]:

(a) A framework for inference from sensory input
(b) A set of tools for performance evaluation
(c) A library of reusable inference and processing components

For a better explanation of working mode, an example of an object detection application (Fig. 6) in MediaPipe is given here, also for an understandable illustration of some basic concepts. MediaPipe enables developers to prototype a pipeline incrementally. A pipeline is a directed *graph* of components, including model inference, media processing algorithms, data transformations, etc. This component is called *Calculator* in MediaPipe. As shown in Fig. 6, each transparent rectangle stands for a **Graph**, where all processing tasks take place. The components like DetectionTracking, ObjectDetection, DetectionMerging, and DetectionAnnotation in example are **Calculators**. **Streams** connect each

node in the graph to another node, which acts as carriers of **Packets**, which is the basic data unit in MediaPipe. There are so called **Side packets** which usually contain constant data for some computing situations, whereas normally streams carry data flow changing over time.

2.4 Skeleton Detection Principles in MediaPipe

In 2020, the MediaPipe developer, Google, presented a pose detection model called BlazePose, which is also known as the MediaPipe pose. According to the team's paper [5], BlazePose has been elaborately developed as a lightweight convolutional neural network architecture for HPE that can run on mobile devices for real-time detection purposes.

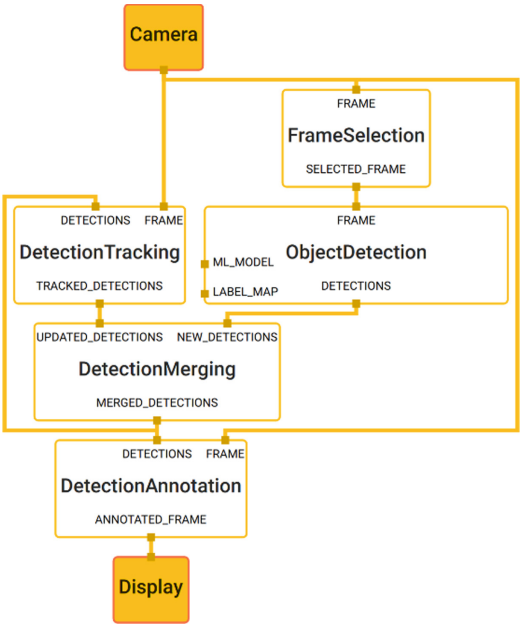


Fig. 6. Object detection using MediaPipe. The transparent boxes represent computation nodes (calculators) in a MediaPipe graph, solid boxes represent external input/output to the graph, and the lines entering the top and exiting the bottom of the nodes represent the input and output streams, respectively. The ports on the left of some nodes denote the input side packets [17].

The model consists of a lightweight body pose detector and a pose tracker network. Each still image or video frame is taken by pose detector first for human localisation in the frame, then pose tracker network will follow to predict keypoint coordinates. For the purpose of saving computing resources of devices, the body pose detector will be called only when there is no human presence in the

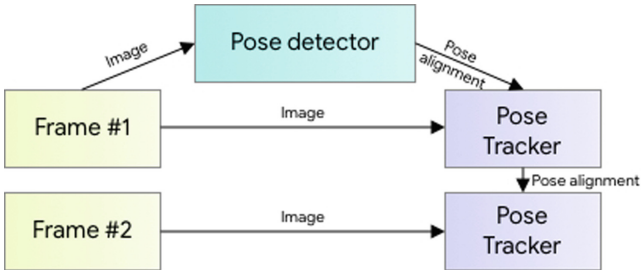


Fig. 7. Inference pipeline [5].

current frame which is reported by the face detector (see Fig. 7). From the above information, we know that BlazePose is based on a top-down HPE approach, as it detects the human body first and landmarks afterwards. The BlazePose uses a topology (see Fig. 8) that contains 33 points to reconstruct the result of skeleton detection from each frame. BlazePose works with a neural network architecture which has been pre-trained on a large dataset consisting of 60K images with a single or few people in the scene in common poses and 25K images with a single person in the scene performing fitness exercises[5].

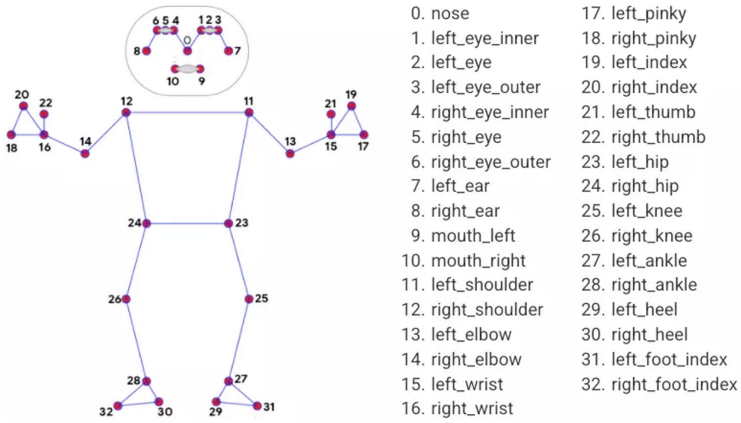


Fig. 8. Pose Landmark Model.Image source: IT Zone [1].

2.5 MediaPipe Analysis: Strengths and Weaknesses

MediaPipe is a robust machine learning platform that enables the rapid development of prototype perceptual systems with inference models and other easily reusable components. A major strength of MediaPipe is its ability to provide

developers with tools to build perceptual systems as a graph of modular components. These components can include inference models, media processing algorithms, and data transformations. The graph accepts input data, such as audio and video streams, and generates perceptual descriptions, such as object localisation and location streams.

MediaPipe is particularly useful for machine learning professionals, including researchers, students, and software developers [4]. They can work on implementing production-ready machine learning applications, publishing code related to research projects, and creating technology prototypes. The framework facilitates the deployment of perceptual technology in demonstrations and applications on a variety of hardware platforms, including mobile and edge devices such as Google Coral [11].

MediaPipe is characterised by its ability to address the challenges associated with adapting a perceptual application to include additional processing steps or inference models. This is often challenging due to the tight coupling of the steps [10]. By abstracting and linking individual perceptual models into sustainable systems, MediaPipe enables easy reuse of components in different systems and across different applications. In addition, MediaPipe is designed to work effectively across platforms, allowing developers to build applications on workstations and then deploy them on mobile devices, for example.

Although MediaPipe has its strengths, there are also some limitations. One is that while WebAssembly performance is generally faster than pure JavaScript, it is typically slower than native C++, which can affect the user experience [10]. MediaPipe tries to address this issue by using the GPU for image operations when possible and prefers to run as lightweight versions of all ML models as possible, although this can sometimes lead to a reduction in quality [14].

Another limitation of MediaPipe, especially its web version, is that it currently only supports calculators in sample charts and does not allow users to create their own charts from scratch, nor to add or change assets [10]. Additionally, the graph executor must be single-threaded, and TensorFlow Lite inference on GPUs is not supported. Despite these limitations, MediaPipe developers plan to continue to evolve the platform to give developers more control and potentially eliminate many of these limitations in the future [4].

2.6 MediaPipe Compatibility: Platforms and Languages

Primary implementation of the MediaPipe library is implemented in C++. Based on this developer decision, the library can be used among the available operating systems, which traditionally include Linux, macOS, and Windows. However, it should be mentioned that it also depends on the language because as the following text describes the various uses in specific languages, certain limitations will gradually emerge (Fig. 9).

MediaPipe also supports Python, as it is a language widely used among machine learning experts and faces a high popularity in the developer community in general [12]. There is an API that again, as in the C++ variant, offers the possibility to use pre-made ML [2]. However, we should not neglect C++

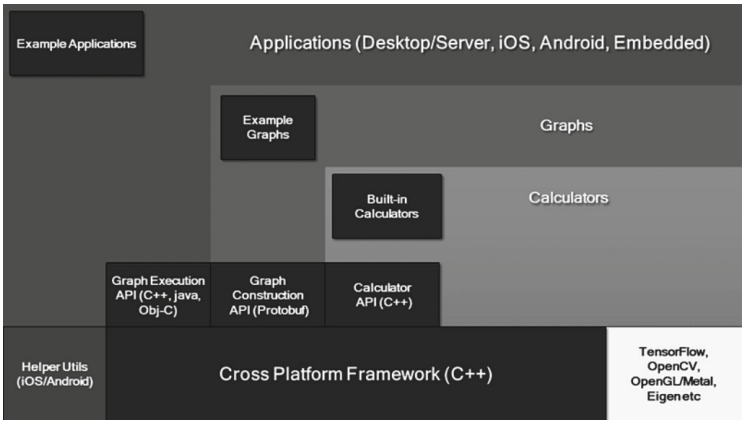


Fig. 9. MediaPipe framework architecture. Image source: Learn OpenCV [3].

itself, in which this library can also be used simply because it is based on this language [2]. Another well-known language supported by this library is Java and Kotlin. I deliberately mention Kotlin along with Java because these are languages that are compatible with each other and can use the libraries with each other. This is particularly useful if the implementation is orientated toward native mobile applications for android devices [11]. In the world of web services, JavaScript is also officially supported and is the main domain of all existing websites [13]. The use of this language is made possible by using the Emscripten tool, which compiles C++ into WebAssembly [13].

Table 2. Compatibility of MediaPipe with different languages and operating systems.

	Linux	macOS (Intel)	macOS (Apple Silicon)	Windows	Android	iOS
C++	Yes	Yes	Yes	Yes	Yes	Yes
Python	Yes	Yes	Yes	Yes	No	No
Java/Kotlin	No	No	No	No	Yes	Yes
JavaScript	No	No	No	No	No	No
C# (non-official)	Yes	No	No	Yes	No	No

The C# programming language is the last language mentioned here, and this is because there is no official support from developers. Anyway, this language can still be used thanks to the C++ Wrapper for C#, which is available on GitHub as a third-party MediaPipe.NET library based on the MediaPipeUnityPlugin library, which is again a third-party library [21]. However, the actual compatibility on different operating systems ceases to be as varied as the variants mentioned above. On macOS systems it can still be used on Intel-based hardware, however for M1 processors based on a modified ARM architecture this

variant is treated as experimental and native support is non-existent. According to official information, Linux should not be a barrier for this library, nor should the latest Windows 11. It should also be noted that MediaPipe on Windows, as with macOS, is still an experimental feature with no native developer support [11,14].
The summary of compatibility of MediaPipe is shown in Table 2.

3 Results

The aim of this application is to assess whether the movement of a patient contains distinct anomalies that are difficult to detect by simple observation, and it is expected that each musculoskeletal disease will have specific attributes that will allow predicting whether the patient is suffering from a particular disease or is yet to show symptoms and is at risk of a full-blown outbreak. To some extent, this feature can be seen as a predictive tool to help detect the disease before it becomes fully manifest, allowing the physician to react in time before permanent damage is done or at least to slow the disease to a point where the patient can continue to function for as long as possible.

In order to understand how this application can be used, it is necessary to briefly explain what its main functions are. First of all, this app can detect a person's skeleton from a bitmap and store this information in memory for future analysis. This image information is obtained through a camera attached to a device. In the next step, this image data is analysed and used for a teacherless AI model. The expected output is an output summary regarding the collected data that has been evaluated by the AI based on previous experiences (Fig. 10).

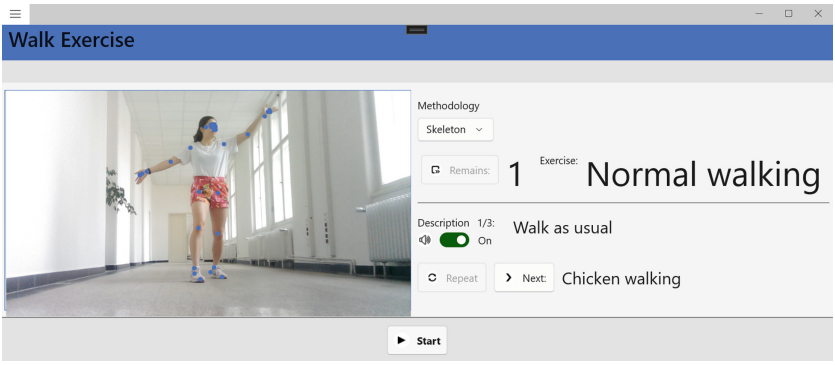


Fig. 10. Demonstration of detecting a human skeleton from an image from application Image source: Own

Simply put, a complete model of human movement needs to be collected. Each frame will contain information about the current position of the skeleton,

and this will be recorded over time. The patient being analysed will receive instructions to follow in sequence, such as walking straight from the camera to a certain distance and back. These instructions will be executed sequentially and stored by the app for later use. Once the patient has completed all the necessary steps, the analysis of these movements follows. Since this is a type of neural network learning without a teacher, it is expected to first look for various anomalies based on the previously collected information.

In a real-world setting, this application will mainly be in the hands of experts, which in this context means doctors, but there is also the option of the data evaluation taking place at the patient's home. The app will include detailed instructions on how to perform this analysis, and then all that is needed is to evaluate the collected information on a remote server, which will return the result.

4 Conclusion

Although this article was mainly orientated toward the theoretical level, it is also worth mentioning what the progress of this project has been so far. Currently, this application has taken a somewhat non-standard direction by choosing the Windows platform along with the C# programming language. Although a significant part of the development team is aware that this choice is not optimal in terms of official developer support, this choice was mainly dependent on the technology stack of the individual team members who are simply proficient in this language, and learning other technologies would have fundamentally slowed down the development cycle of the entire project to an unacceptable point.

In any case, the current implementation is in the testing phase, where the application handles some of the important aspects already mentioned, such as facial and skeletal analysis. It can efficiently collect image information and has implemented operations to work with this data and modify it for the core AI that can work with this information. Another important observation is that the application contains a relatively simple graphical design that is particularly suitable for workers no longer completely technically orientated, thus offering an easy to use and intuitive approach, although that it is still considering further improvements, this is for the time being at a sufficient level from a subjective point of view.

A summary of this phase of the project could be seen as a functioning ecosystem that offers medical professionals the opportunity to examine the patient on their personal PC with a connected webcam and obtain valuable data from which to evaluate the patient's condition.

4.1 Future Work

As MediaPipe.NET does not yet support the Android and iOS platform for pose detection (see Table 3), team will focus on implementing the necessary parts to get MediaPipe.NET to work on these platforms.

Table 3. Compatibility Matrix of MediaPipe.NET [6].

	Linux (x86_64)	macOS (x86_64)	macOS (ARM64)	Windows (x86_64)	Android	iOS
Linux (AMD64)	✓				?	
Intel Mac		✓			?	?
M1 Mac			✓		?	?
Windows 10 (AMD64)	✓			✓	?	

Acknowledgement. This work was supported from the grant of Specific university research - grant No A1 FCHI 2023003 (UCT Prague) and grant No SGS 2023 016 (University of Pardubice).

References

1. A bit about the pose classification (2023). https://itzone.com.vn/en/article/a-bit-about-the-pose-classification/
2. How to use mediapipe in c++? (2023). https://stackoverflow.com/questions/72014807/how-to-use-mediapipe-in-c. Accessed 07 June 2023]
3. Introduction to mediapipe (2023). https://learnopencv.com/introduction-to-mediapipe/
4. ARXIV LABS: Ar5iv: Pose tracking with 3d and rgb-d cameras (2023). https://ar5iv.labs.arxiv.org/html/1906.08172. Accessed 07 June 2023
5. Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., Grundmann, M.: Blazepose: on-device real-time body pose tracking. (2020) CoRR abs/2006.10204. https://arxiv.org/abs/2006.10204
6. cosyneco: cosyneco/mediapipe.net: Pure.net bindings for google's mediapipe. (2023). https://github.com/cosyneco/MediaPipe.NET. přístup dne 24. června 2023
7. Dubey, S., Dixit, M.: A comprehensive survey on human pose estimation approaches. Multimedia Syst. **29** (2022)
8. Felzenszwalb, P.F., Huttenlocher, D.P.: Pictorial structures for object recognition. Int. J. Comput. Vision **61**, 55–79 (2005)
9. Gong, W., et al.: Human pose estimation from monocular images: a comprehensive survey. Sensors **16**(12) (2016). https://www.mdpi.com/1424-8220/16/12/1966
10. GOOGLE DEVELOPERS: Mediapipe on the web (2023). https://developers.googleblog.com/2020/01/mediapipe-on-web.html. Accessed 07 June 2023
11. GOOGLE DEVELOPERS: Mediapipe: Setup android (2023). https://developers.google.com/mediapipe/solutions/setup_android. Accessed 07 June 2023
12. GOOGLE DEVELOPERS: Mediapipe: Setup python (2023). https://developers.google.com/mediapipe/solutions/setup_python. Accessed 07 June 2023
13. GOOGLE DEVELOPERS: Mediapipe: Setup web (2023). https://developers.google.com/mediapipe/solutions/setup_web. Accessed 07 June 2023
14. GOOGLE DEVELOPERS: Pose landmarker (2023). https://developers.google.com/mediapipe/solutions/vision/pose_landmarker. Accessed 07 June 2023
15. Kanjee, R.: Top 9 pose estimation models of 2022 (2022). https://medium.com/augmented-startups/top-9-pose-estimation-models-of-2022-70d00b11db43

16. Li, C., Hua, T.: Human action recognition based on template matching. Procedia Eng. **15**, 2824–2830 (2011). https://www.sciencedirect.com/science/article/pii/S1877705811020339
17. Lugaresi, C., et al.: Mediapipe: a framework for building perception pipelines (2019)
18. Odemakinde, E.: Human pose estimation with deep learning - ultimate overview in 2023 (2023). https://viso.ai/deep-learning/pose-estimation-ultimate-overview/. Accessed 18 June 2023
19. Rastogi, K.: Know all about 2d and 3d pose estimation (2022). https://www.analyticsvidhya.com/blog/2022/04/comprehensive-guide-for-pose-estimation/. Accessed 18 June 2023
20. Tyagi, M.: Hog (histogram of oriented gradients): an overview (2021). https://towardsdatascience.com/hog-histogram-of-oriented-gradients-67ecd887675f. Accessed 18 June 2023
21. VIGNETTEAPP: Mediapipe.net (2023). https://github.com/vignetteapp/MediaPipe.NET. Accessed 07 June 2023
22. Zatolokina, L.: Human pose estimation technology capabilities and use cases in 2023 (2022). https://mobidev.biz/blog/human-pose-estimation-technology-guide. Accessed 18 June 2023
23. Zheng, C., et al.: Deep learning-based human pose estimation: a survey (2020). CoRR abs/2012.13392. https://arxiv.org/abs/2012.13392

Model Design Research Plan for Warehouse Barcode Image Recognition in Smart Systems

Jan Tyrychtr[2(✉)], Shady Aly[1], Adéla Hamplová[2], and Tomáš Benda[2]

[1] Faculty of Engineering, Helwan University, Helwan, Egypt
[2] Department of Information Engineering, Faculty of Economics and Management, Czech University of Life Sciences Prague, Prague, Czech Republic
tyrychtr@pef.czu.cz

Abstract. This article presents a comprehensive plan for creating a research model of a neural network for barcode recognition in a warehouse environment. The introduced new methodology enables the development of an efficient model with high accuracy for barcode recognition and decoding. In the proposed methodology, we consider current approaches and present a research plan focusing on convolutional neural networks and the Roboflow tool. The new methodology has been divided into several steps, including data collection, dataset compilation, selection of a suitable neural network model, network training, validation, and optimization, and finally, testing the model on a test dataset.

Keywords: Convolutional Neural Networks · Barcode Image Recognition · CNN · Artificial Intelligence · YOLO · AI research plan

1 Introduction

This article presents the research plan for designing a model to recognize barcode images in a warehouse (e.g. [4]), including a description of the research process steps, objectives, methods, data collection plan, test scenario design, and expected results upon research completion.

The research goal is to create an application for barcode recognition in a warehouse from photographs taken with a mobile device. The purpose is to design and develop a neural network model for barcode recognition (e.g., as [5]). A prerequisite for the model is that it must be efficient and reliable in recognizing barcodes and should enable easy and fast data processing from photographs.

Currently, there are no methodological approaches that allow researchers to create a barcode recognition model quickly and efficiently.

2 Methodology

In this article, a series of methodological steps and procedures were undertaken, which are detailed here. Our methodology is based on the analysis of available literature and data and employs modern tools for model and dataset creation.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 51–57, 2024.
https://doi.org/10.1007/978-3-031-53549-9_5

The first step in the process of creating this scientific article was a thorough literature review. For this, we utilized reputable scientific databases, including Scopus and Web of Science, to identify previous research in the field of barcode recognition. This literature review provided us with a comprehensive overview of the current state of this issue.

To design a neural network model for barcode recognition, we drew inspiration from the latest developments in machine learning. Specifically, we focused on Convolutional Neural Networks (CNNs), which are known for their ability to process visual data. A convolutional neural network (CNN) is one of the most significant networks in the deep learning field [1–3]. For dataset creation, we utilized the Roboflow tool [6], which facilitates the easy creation, annotation, and management of datasets.

3 Results

In this section, we present our innovative methodological approach for creating a research neural network model for barcode recognition (see Fig. 1).

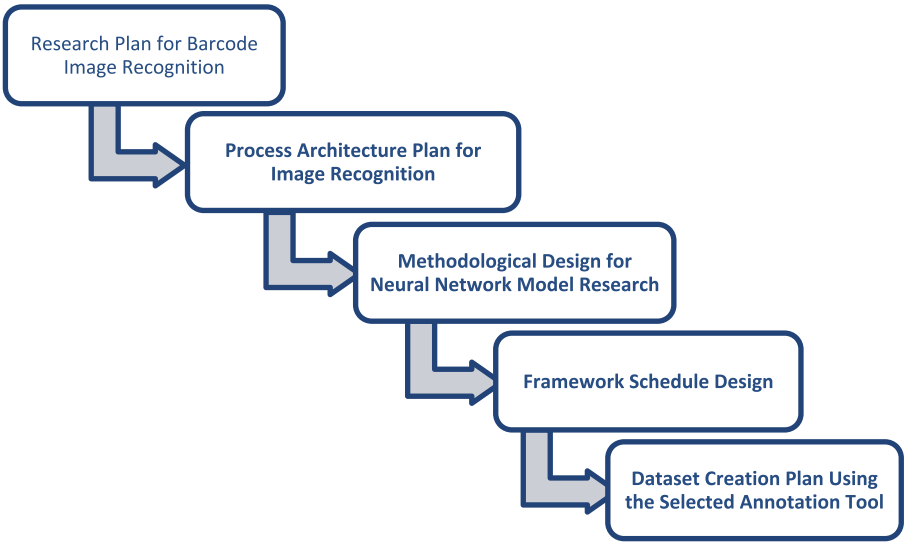


Fig. 1. Methodological process of designing a neural network model for barcode recognition.

3.1 Research Plan for Barcode Image Recognition

Formulate Objectives, e.g.:

- Develop an efficient model for recognizing and decoding barcode images in the warehouse.
- Improve the accuracy of barcode recognition.

Identify the Methods:

- Research and analyse available algorithms for image recognition and machine learning.
- Select a tool for creating datasets.
- Create training and testing datasets.
- Train, cross-validate, and test the model using selected standard AI methods.

Create a Data Collection Plan:

- Capture many barcodes in the warehouse under various lighting conditions, distances, and angles.
- Collect data using a mobile phone with a high-resolution camera.
- Ensure an adequate amount of data for training, validation, and testing of the model.

Design a Test Scenario Proposal:

- Test the optimal rotation step of the code sent to the decoder (which accepts images only when they are flat).
- Compare the model's performance with existing barcode recognition algorithms (this comparison should be based on a literature review of scientific articles, where compare the parameters of different algorithms).
- Test the impact of various factors on barcode recognition accuracy, such as distance, angle, lighting, and image quality.

Formulate the Expected Results:

- Development of an efficient model for barcode image recognition with high accuracy.
- Comparison of the performance of the new model with existing barcode recognition algorithms (this will be done theoretically before selecting an algorithm, e.g., YOLO).
- Identification of factors influencing barcode recognition accuracy and their impact on model performance.

3.2 Process Architecture Plan for Image Recognition

The process architecture for the barcode recognition application involves several critical steps necessary for successful barcode recognition and data processing. Here we present a sample process architecture:

1. Data Collection: Users capture photos of products with barcodes using a mobile device with a barcode recognition application.
2. Data Preprocessing: Images are pre-processed using image processing algorithms to enhance image quality and contrast, remove noise, and address other anomalies.
3. Barcode Recognition: A neural network model is employed to recognize the barcode in the image. The neural network analyses the processed image and attempts to locate the barcode.
4. Data Interpretation: After successful barcode recognition, data is processed and interpreted. This may involve rotating the code to a readable orientation for the decoder, decoding the code, and associating the information contained in the code with the warehouse product.

5. Displaying Results: As per technical specifications, the result should be an image highlighting the detected code, aiding in locating the item. Barcode recognition results are displayed to users in the mobile application, including the product name, description, price, stock availability, and other information.
6. Data Storage: Data about recognized codes and products are stored in a database for future use, such as inventory tracking and product demand analysis.
7. This process architecture includes steps crucial for successful barcode recognition and data processing. Each step should be thoroughly tested and optimized to ensure the best accuracy and reliability of the application.

3.3 Methodological Design for Neural Network Model Research

Step 1: Data Collection

The first step involves capturing images of products (boxes) with barcodes (labels), totalling approximately 2000 + photos.

Step 2: Dataset Assembly

The next step is assembling a dataset containing images of barcodes. The dataset should be sufficiently large and diverse to enable the neural network to recognize various types of codes and deformations.

Step 3: Selection of Neural Network Model

The next step is deciding on a suitable neural network model for this task. The recommended model should be a convolutional neural network (CNN) designed specifically for processing image data.

Step 4: Neural Network Training

The next step is training the neural network on the created dataset. This process involves dividing the dataset into training, validation, and testing subsets. As this is supervised learning, the network is provided with expected outputs (bounding boxes around individual barcodes) in addition to input images. The actual training of the convolutional network is done, for example, using the backpropagation algorithm.

Step 5: Validation and Optimization

Upon completing training, the network is validated on a testing dataset and optimized to achieve the best results. This may involve adjusting network hyperparameters such as the number of layers, convolutional window size, number of filters, learning rate, epochs, and batch size.

Step 6: Neural Network Testing

Finally, the neural network should be tested on a separate testing dataset to evaluate its real-world performance. The testing dataset should be distinct from the training and validation datasets to assess how well the network generalizes to new data.

3.4 Framework Schedule

The proposed schedule, including estimated workloads for each step in the project's second phase:

1. **Data Collection**

   Duration: 3 months.

Workload: Involves 2–5 team members.
Workload Level: Light, primarily involves capturing warehouse photos.

2. **Dataset Assembly**

   Duration: 2 months.
   Workload: Involves 2–5 team members.
   Workload Level: Moderate, requires image processing and coding. Builds upon Phase I where warehouse photos were taken.

3. **Selection of Neural Network Model**

   Duration: 1 month.
   Workload: Involves 2–5 team members.
   Workload Level: Moderate, recommends conducting a literature review and comparing various types of neural network models for image recognition.

4. **Neural Network Training**

   Duration: 4 months.
   Workload: Involves 4–7 team members.
   Workload Level: High, requires deep knowledge of neural networks, supervised learning, and optimization.

5. **Validation and Optimization**

   Duration: 1 month.
   Workload: Involves 2–4 team members.
   Workload Level: Moderate, requires expertise in model validation and optimization.

6. **Neural Network Testing**

   Duration: 2 months.
   Workload: Involves 5–7 team members.
   Workload Level: Low, involves simple testing of the neural network on a testing dataset.

7. **Setting up Server Environment for API**

   Duration: 1 month.
   Workload: Involves 1–3 team members.
   Workload Level: Low, configuring the server environment for barcode analysis on an internal web server.

8. **API Development**

   Duration: 3 months.
   Workload: Involves 3–5 team members.
   Workload Level: Moderate, preparing the decoding script for real-time image recognition.

9. **Integration with Warehouse Management System**

   Duration: 2 months.

Workload: Involves 2–4 team members.
Workload Level: Moderate, involves integrating the API with the existing warehouse management system.

10. **Testing and Debugging**

    Duration: 2 months.
    Workload: Involves 3–6 team members.
    Workload Level: Moderate, involves extensive testing, identifying, and resolving issues.

11. **Documentation and Training**

    Duration: Ongoing throughout the project.
    Workload: Involves 1–3 team members.
    Workload Level: Light, involves documenting the project's progress, preparing user guides, and providing training to warehouse staff.
    The above schedule is an estimate and may be subject to change based on unforeseen challenges and developments during the AI project. Regular progress reviews and adjustments to the schedule are advisable.

3.5 Dataset Creation Plan Using the Selected Annotation Tool

There are several tools available for data annotation, such as labelimg, V7labs, keymakr, Label Studio, Roboflow, and others. We have chosen Roboflow as our annotation tool because it allows team collaboration during annotation, provides a simple and intuitive interface, and offers the option to export datasets using a key, which can be used for training externally after selecting the dataset.

Roboflow is an online platform for managing, preparing, and training data for computer vision, including barcode recognition. This tool can be utilized to complete the dataset for training the neural network. Roboflow can be used for barcode recognition in the warehouse as follows:

1. Initially, we will upload photos from the warehouse containing barcodes to Roboflow. When creating a project, we will select the project type, which, in this case, is "object detection." Roboflow also offers classification, segmentation, and other project types.
2. Next, we need to select or create a dataset. Roboflow offers various tools for dataset creation and editing. The dataset should include photos with barcodes and annotations that specify the location and identifier of the barcode in each image.
3. Each photo will then be annotated. This step can be time-consuming, especially considering the use of 2000 + photos from the warehouse.
4. Subsequently, we will prepare the dataset for training the neural network.
5. After training, it will be necessary to test the model on a testing dataset and evaluate the results.
6. Finally, the created model can be used for barcode recognition in real-world environments, such as the warehouse or other locations.

4 Conclusion

The paper presented a novel methodological approach for creating an efficient neural network model for high-precision barcode recognition. The methodology was tested on real-world data, where it demonstrated its ability to create a barcode recognition model for products in a real warehouse of an e-commerce company in the Czech Republic. This article introduced a comprehensive methodology for building and training a neural network for image recognition and has the potential to be applied in various applications within the field of computer vision and image recognition.

Acknowledgement. This work was conducted within the project "Smart environments - modelling and simulation of complex decision-making problems in intelligent systems" (2022B0010) funded through the IGA foundation of the Faculty of Economics and Management, Czech University of Life Sciences in Prague, and the project "Precision agriculture and digitization in the Czech Republic" (QK23020058) funded through the NAZV Program ZEMĚ 2022 of the Ministry of Agriculture of the Czech Republic.

References

1. Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J.: A survey of convolutional neural networks: analysis, applications, and prospects. IEEE Trans. Neural Netw. Learn. Syst. **33**, 6999–7019 (2022)
2. Wu, J.: Introduction to convolutional neural networks. Natl. Key Lab Novel Softw. Technol. **5**(23), 495 (2017)
3. Chauhan, R., Ghanshala, K.K., Joshi, R.C.: Convolutional neural network (CNN) for image detection and recognition. In: 2018 first International Conference on Secure Cyber Computing and Communication (ICSCCC), pp. 278–282. IEEE (2018)
4. Chowdhury, A.I., Rahman, M.S., Sakib, N.: A study of multiple barcode detection from an image in business system. Int. J. Comput. Appl. **181**(37), 30–37 (2019)
5. Ventsov, N.N., Podkolzina, L.A.: Localization of barcodes using artificial neural network. In: 2018 IEEE East-West Design & Test Symposium (EWDTS), pp. 1–6. IEEE (2018)
6. Lin, Q., Ye, G., Wang, J., Liu, H.: RoboFlow: a data-centric workflow management system for developing AI-enhanced robots. In: Conference on Robot Learning, pp. 1789–1794. PMLR (2022)



Robust H_∞ Controller Design for Satellite Systems with Uncertain Inertia Matrix: A Linear Matrix Inequality Approach

Ibrahim Shaikh(✉), Samuel Emebu, and Radek Matušů

Department of Automation and Control Engineering, Faculty of Applied Informatics, Tomas Bata University in Zlín, nám. T. G. Masaryka 5555, 760 01 Zlín, Czech Republic

shaikh@utb.cz

Abstract. This paper discusses the design of a robust H_∞ controller for satellite systems that exhibit changes in its inertia matrix within a range of ±5%. Using MATLAB Simulink, the proposed approach is a Linear Matrix Inequality (LMI) by LMILAB Semidefinite programming solver in YALMIP. Simulation results demonstrate the controller's effectiveness in stabilizing the system against disturbance and maintaining performance despite variations in the inertia matrix.

Keywords: Robust Control · H_∞ Controller · Satellite Systems · Dynamic Inertia Matrix · Linear Matrix Inequality (LMI) · MATLAB Simulink Model · YALMIP Solver · Disturbance Rejection · Robustness Analysis

1 Introduction

In the last decades, there has been a significant focus on addressing the challenges associated with controlling the orientation of satellites in space [1–4]. Investigated a subset of sampled-data systems related to satellite control, where variations in the continuous-time systems they originate from introduce uncertainty. Solving a set of linear matrix inequalities (LMIs) or a dynamic optimization problem subject to LMI constraints led to the development of both a robust H_∞ control approach and an optimally robust H_∞ control approach [5]. Novel delay-dependent stabilization criteria are introduced by employing the parallel distributed compensation method, which effectively mitigates conservatism without imposing an additional computational load [6]. Investigation into the domain of robust stabilization and H_∞ control, with a specific focus on the challenges presented by uncertain stochastic time-delay systems and the development of a memoryless state feedback control approach that ensures the closed-loop system's robust and asymptotic stability [7]. Use of a Proportional Plus Derivative (PD) state feedback controller. The objective is to achieve quadratic normality and quadratic stability while adhering to a specified H∞ norm bound for

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 58–66, 2024.
https://doi.org/10.1007/978-3-031-53549-9_6

the closed-loop descriptor system [8]. This study appears to be directed towards tackling the difficulties associated with ensuring the stability and accuracy of spacecraft orientations when confronted with nonlinear dynamics and external disruptions [9]. A robust H_∞ controller assures specific H_∞ performance even in bounded parameter uncertainty and unmodeled dynamics within the closed-loop system [10,11]. Diverse noise inputs often lead to adopting distinct performance criteria, such as the H_2 or the H_∞ norm. The H_2 norm is commonly employed when noise input signals possess identifiable spectral density. Conversely, in scenarios where statistical information regarding the noise input is unavailable, the H_∞ norm is considered more suitable [12]. The incorporation of performance criteria has reinforced the significance of the subject, offering optimal filtering designs applicable to areas like satellite attitude systems, radar systems, fault detection, and signal processing [13,14]. It introduces a self-adaptive control parameter multiobjective differential evolution algorithm, avoiding LMIs. For systems with polytopic uncertainties, it computes worst-case norms through implicit optimization, employing the self-adaptive differential evolution method [15,19,20]. The mixed H_2/H_∞ control synthesis issue represents a crucial multiobjective control design challenge. In tackling this problem, the LMI approach is frequently employed [16–18].

The inertia matrix of a satellite is a critical parameter that affects its attitude dynamics. The inertia matrix is the product of the mass matrix and the moment of inertia tensor, and it determines the way in which the satellite's moments of inertia are distributed. The inertia matrix can change due to various factors, such as the loss of fuel, the deployment of solar panels, or the impact of micrometeorites. Changes in the inertia matrix can have a significant impact on the attitude dynamics of a satellite. The satellite's natural frequencies and mode shapes will change if the inertia matrix changes. This can lead to instability and make controlling the satellite's attitude more difficult. Robust control is a type of control design designed to be insensitive to changes in the system parameters. Robust controllers can be used to ensure that a satellite's attitude remains stable even in the presence of inertia matrix changes.

2 Formulation for Satellite Systems and Design

H_∞ control has been a very active field for the last two decades, and there have been tremendous results by now. As an introduction to H_∞ control, contemplate the continuous linear time-invariant system characterized by the subsequent state space representation:

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}_1\boldsymbol{u}(t) + \boldsymbol{B}_2\boldsymbol{d}(t) \\ \boldsymbol{z}_\infty(t) = \boldsymbol{C}_1\boldsymbol{x}(t) + \boldsymbol{D}_1\boldsymbol{u}(t) + \boldsymbol{D}_2\boldsymbol{d}(t) \end{cases} \tag{1}$$

where, $\boldsymbol{x}$ is the vector of state variables, $\boldsymbol{u}$ is the vector of control inputs, $\boldsymbol{d}$ is the vector of disturbance inputs, z_∞ is the output vectors, others are the system coefficient matrices of appropriate dimensions.

Design a state feedback control law

$$\boldsymbol{u} = \boldsymbol{K}\boldsymbol{x} \,. \tag{2}$$

Such that with some possibly small γ_∞ positive scalar the H_∞ performance condition is satisfied.

$$\|G_{z_\infty d}\|_\infty = \left\| \left(\boldsymbol{C}_1 + \boldsymbol{D}_1\boldsymbol{K}\right)\left(s\boldsymbol{I} - \left(\boldsymbol{A} + \boldsymbol{B}_1\boldsymbol{K}\right)\right)^{-1}\boldsymbol{B}_2 + \boldsymbol{D}_2 \right\|_\infty \leq \gamma_\infty \,. \tag{3}$$

The definitions of H_∞ norm for transfer function $G_{z_\infty d}$ is

$$\|G_{z_\infty d}(s)\|_\infty = \sup_\omega \sigma_{\max}\left(G_{z_\infty d}(j\omega)\right) \,. \tag{4}$$

The main objective of this study is to formulate a robust H_∞ controller. This involves determining a state feedback gain matrix $\boldsymbol{K}$ for a given positive scalar γ_∞ in such a way that the stability of the linear time-invariant system is assured while adhering to the constraint $\|G_{z_\infty d}\|_\infty \leq \gamma_\infty$ is an emphasis on minimize.

$$\|G_{z_\infty d}(s)\|_\infty = \sup_\omega \sigma_{\max}\left(G_{z_\infty d}(j\omega)\right) \,. \tag{5}$$

2.1 Satellite Modelling

Dynamic equations of satellite modelling are fundamental equations that describe the motion of satellites in space. They are derived from Newton's laws of motion and include gravitational forces as well as non-gravitational perturbations [21].

$$\dot{\boldsymbol{H}} = \boldsymbol{T}_c + \boldsymbol{T}_g + \boldsymbol{T}_d \,. \tag{6}$$

where, $\boldsymbol{T}_\mathrm{c}, \boldsymbol{T}_\mathrm{g}$ and $\boldsymbol{T}_\mathrm{d}$ are the control input torque, the gravitational torque and the disturbance torque, respectively. $\boldsymbol{H}$ is the total momentum acting on the satellite. The derivative of the total momentum $\boldsymbol{H}$ can be given as follows

$$\dot{\boldsymbol{H}} = \boldsymbol{I}_b\boldsymbol{\omega} + \boldsymbol{\omega} \times \left(\boldsymbol{I}_b\boldsymbol{\omega}\right). \tag{7}$$

where, $\boldsymbol{I}_b$ and $\boldsymbol{\omega}$ denote the inertia matrix and the angular velocity respectively. Equation(6) and Eq.(7) yield the following dynamic equation:

$$\boldsymbol{I}_b\boldsymbol{\omega} + \boldsymbol{\omega} \times \left(\boldsymbol{I}_b\boldsymbol{\omega}\right) = \boldsymbol{T}_c + \boldsymbol{T}_g + \boldsymbol{T}_d. \tag{8}$$

Mathematical representation for inertia matrix, the control input torque, the gravitational torque and the disturbance torques are

$$I_b = \mathrm{diag}\left(I_x, I_y, I_z\right),$$

$$T_c = \begin{bmatrix} T_{cx} \\ T_{cy} \\ T_{cz} \end{bmatrix}, \quad T_g = \begin{bmatrix} T_{gx} \\ T_{gy} \\ T_{gz} \end{bmatrix}, \quad T_d = \begin{bmatrix} T_{dx} \\ T_{dy} \\ T_{dz} \end{bmatrix}.$$

After then, Eq.(8) can be converted into

$$\begin{cases} I_x\dot{\omega}_x + (I_z - I_y)\,\omega_y\omega_x = T_{cx} + T_{gx} + T_{dx} \\ I_y\dot{\omega}_y + (I_x - I_z)\,\omega_z\omega_x = T_{cy} + T_{gy} + T_{dy} \\ I_z\dot{\omega}_z + (I_y - I_x)\,\omega_x\omega_y = T_{cz} + T_{gz} + T_{dz}. \end{cases} \tag{9}$$

Utilizing the small angle approximation, the angular velocity of the satellite within the inertial coordinate system, as expressed in the body coordinate system, is provided by [21]

$$\omega = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \begin{bmatrix} \dot{\varphi} - \omega_0\psi \\ \dot{\theta} - \omega_0 \\ \dot{\psi} + \omega_0\varphi \end{bmatrix}, \tag{10}$$

where $\omega_0 = 7.292115 \times 10^{-5}$ rad/sec is the rotational-angular velocity of the earth. ϕ, θ and ψ are the three Euler angles. By substituting Eq.(10) into Eq.(9), then we get

$$\begin{cases} I_x\ddot{\varphi} + (I_y - I_z)\,\omega_0^2\varphi + (I_y - I_z - I_x)\,\omega_0\dot{\psi} = T_{cx} + T_{gx} + T_{dx} \\ I_y\ddot{\theta} = T_{cy} + T_{gy} + T_{dy} \\ I_z\ddot{\psi} + (I_x + I_z - I_y)\,\omega_0\dot{\varphi} + (I_y - I_x)\,\omega_0^2\dot{\psi} = T_{cz} + T_{gz} + T_{dz}, \end{cases} \tag{11}$$

where the gravitational torques are easily shown to be given by

$$\begin{cases} T_{gx} = -3\omega_0^2\,(I_y - I_z)\,\varphi \\ T_{gy} = -3\omega_0^2\,(I_x - I_z)\,\theta \\ T_{gz} = 0. \end{cases} \tag{12}$$

By merging Eq. (11) and Eq. (12), we can derive the final attitude dynamics equation obtained

$$\begin{cases} I_x\ddot{\varphi} + 4\,(I_y - I_z)\,\omega_0^2\varphi + (I_y - I_z - I_x)\,\omega_0\dot{\psi} = T_{cx} + T_{dx} \\ I_y\ddot{\theta} + 3\omega_0^2\,(I_x - I_z)\,\theta = T_{cy} + T_{dy} \\ I_z\ddot{\psi} + (I_x + I_z - I_y)\,\omega_0\dot{\varphi} + (I_y - I_x)\,\omega_0^2\dot{\psi} = T_{cz} + T_{dz}. \end{cases} \tag{13}$$

Representation of some vectors are

$$q = \begin{bmatrix} \varphi \\ \theta \\ \psi \end{bmatrix}, \quad u = \begin{bmatrix} T_{cx} \\ T_{cy} \\ T_{cz} \end{bmatrix}, \quad d = \begin{bmatrix} T_{dx} \\ T_{dy} \\ T_{dz}. \end{bmatrix}$$

Then the Eq.(13) can be written compactly in the following second-order matrix form:

$$M\ddot{q} + H\dot{q} + Gq = L_1u + L_2d \tag{14}$$

where for the matrix M, we have:

$$M = \mathrm{diag}\,(I_x, I_y, I_z), \quad L_1 = L_2 = I_{3\times3}$$

The matrix H is defined as:

$$H = \omega_0 \left(I_y - I_x - I_z\right) \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

Lastly, the matrix G can be expressed as:

$$G = \operatorname{diag}\left(4\omega_0^2\left(I_y - I_z\right), 3\omega_0^2\left(I_x - I_z\right), \omega_0^2\left(I_y - I_x\right)\right).$$

proposed state variable

$$x = \begin{bmatrix} q & \dot{q} \end{bmatrix}^{\mathrm{T}}$$

and output variable

$$z_\infty = 10^{-3} M\ddot{q}.$$

Compare Eq.(1) and Eq.(13), and we get into the following state-space matrixes as $A, B_1, B_2, C_1, C_2, D_1$, and D_2:

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -\frac{4\omega_0^2 I_{yz}}{I_x} & 0 & 0 & 0 & 0 & -\omega_0 \frac{I_{yzx}}{I_x} \\ 0 & -\frac{3\omega_0^2 I_{xz}}{I_y} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{\omega_0^2 I_{yx}}{I_z} & \omega_0 \frac{I_{yxz}}{I_z} & 0 & 0 \end{bmatrix},$$

$$B_1 = B_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{I_x} & 0 & 0 \\ 0 & \frac{1}{I_y} & 0 \\ 0 & 0 & \frac{1}{I_z} \end{bmatrix},$$

$$C_1 = 10^{-3} \times \begin{bmatrix} -4\omega_0^2 I_{yz} & 0 & 0 & 0 & 0 & -\omega_0 I_{yxz} \\ 0 & -3\omega_0^2 I_{xz} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\omega_0^2 I_{yx} & \omega_0 I_{y\times z} & 0 & 0 \end{bmatrix},$$

$$C_2 = \begin{bmatrix} I_{3\times 3} & 0_{3\times 3} \end{bmatrix},$$

$$D_1 = 10^{-3} \times L_1, \quad D_2 = 10^{-3} \times L_2.$$

2.2 H_∞ Control Law Based on LMI

It has a solution if and only if there exists a matrix W, and a symmetric positive definite matrix X, such that

$$\begin{bmatrix} (AX + B_1 W)^{\mathrm{T}} + AX + B_1 W & B_2 & (CX + D_1 W)^{\mathrm{T}} \\ B_2^{\mathrm{T}} & -\gamma I & D_2^{\mathrm{T}} \\ CX + D_1 W & D_2 & -\gamma I \end{bmatrix} < 0$$

When such a pair of matrices W and X are found, a solution can be given as

$$K = WX^{-1}$$

This the second condition in the problem is satisfied if and only if there exists a symmetric positive definite matrix X, such that,

$$\begin{bmatrix} X\left(A+B_1K\right)^{\mathrm{T}}+\left(A+B_1K\right)X & B_2 & X\left(C+D_1K\right)^{\mathrm{T}} \\ B_2^{\mathrm{T}} & -\gamma I & D_2^{\mathrm{T}} \\ \left(C+D_1K\right)X & D_2 & -\gamma I \end{bmatrix} < 0$$

that is,

$$\begin{bmatrix} \left(AX+B_1KX\right)^{\mathrm{T}}+\left(AX+B_1KX\right) & B_2 & \left(CX+D_1KX\right)^{\mathrm{T}} \\ B_2^{\mathrm{T}} & -\gamma I & D_2^{\mathrm{T}} \\ CX+D_1KX & D_2 & -\gamma I \end{bmatrix} < 0.$$

By defining

$$W = KX.$$

The aforementioned inequality is turned into the LMI form. With a prescribed attenuation level, the problem is turned into an LMI feasibility problem. The problem with a minimal attenuation level γ_∞ can be sought via the following optimization problem:

The following corollary reveals that the condition $\|G_{zw}(s)\|_\infty < \gamma_\infty$ implies the closed-loop stability.

This system with a minimal attenuation level γ_∞ can be sought via the following optimization problem:

$$\begin{cases} \min \gamma_\infty \\ \text{s.t. } X > 0 \\ \quad \begin{bmatrix} \left(AX+B_1W\right)^{\mathrm{T}}+AX+B_1W & B_2 & \left(CX+D_1W\right)^{\mathrm{T}} \\ B_2^{\mathrm{T}} & -\gamma I & D_2^{\mathrm{T}} \\ CX+D_1W & D_2 & -\gamma I \end{bmatrix} < 0 \end{cases}. \quad (15)$$

By minimizing γ_∞, the state feedback gain matrix can be constructed as $K = WX^{-1}$

3 Simulation Result

In this section, the efficacy of the proposed H_∞ controller is demonstrated through simulations performed on a satellite's attitude system. The given values for moment of inertia are utilized: $Ix = 1030.70$ kg·m^2, $Iy = 3015.65$ kg·m^2, and $Iz = 3030.43$ kg·m^2. Additionally, the satellite is assumed to orbit at an altitude of 300 km. The desired state is set to zero, while the initial state is chosen as $x(0)$ = [1000 rad, 2000 rad, 3000 rad, 4000 rad/s, 5000 rad/s, 6000 rad/s]. To account for disturbances $d = [sin(t/200); cos(t/200); sin(t/250)]$. By applying the above formula and utilizing the MATLAB LMI toolbox, we can derive two values for

γ_∞ and H_∞ as 0.001 its same value due to using Semidefinite programming solver in YALMIP. This leads to the determination of the controller gain

$$\boldsymbol{K} = \begin{bmatrix} -0.0007 & 0.0000 & 0.0000 & -1.7233 & 0.0000 & -0.0325 \\ 0.0000 & -0.0006 & 0.0000 & 0.0000 & -2.6366 & 0.0000 \\ 0.0000 & 0.0000 & -0.0005 & 0.0370 & 0.0000 & -2.4680 \end{bmatrix}.$$

The matrices W and X in their respective relationship

$$\boldsymbol{W} = \begin{bmatrix} 3.6556 & 0.0000 & 51.4433 & -0.9725 & 0.0000 & -0.0206 \\ 0.0000 & -0.8013 & 0.0000 & 0.0000 & -0.3314 & 0.0000 \\ 0.6316 & -0.0000 & 35.4444 & 0.0000 & 0.0000 & -0.3371 \end{bmatrix}$$

$$\boldsymbol{X} = 1 \times 10^6 \times \begin{bmatrix} 3.3370 & -0.0000 & 0.0598 & -0.0014 & 0.0000 & -0.0001 \\ -0.0000 & 3.3987 & -0.0000 & 0.0000 & -0.0007 & -0.0000 \\ 0.0598 & -0.0000 & 3.3572 & 0.0000 & -0.0000 & -0.0007 \\ -0.0014 & 0.0000 & 0.0000 & 0.0000 & -0.0000 & 0.0000 \\ 0.0000 & -0.0007 & -0.0000 & -0.0000 & 0.0000 & 0.0000 \\ -0.0001 & -0.0000 & -0.0007 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}.$$

The attenuation level represented by $\gamma_\infty = 0.0010$ is exceedingly small, suggesting that the disturbance effect d is effectively suppressed. However, with this particular solution, the locations of the closed-loop system poles are -0.000927, -0.000745, -0.000454, -0.000360, -0.000542 -0.003314 which all are stable, but are all very close to the imaginary axis. Therefore, the system response must be too slow and must have a long settling time. It is seen in Fig. 1, and Fig. 2, that within 2000 s the system responses are still far from settling down.

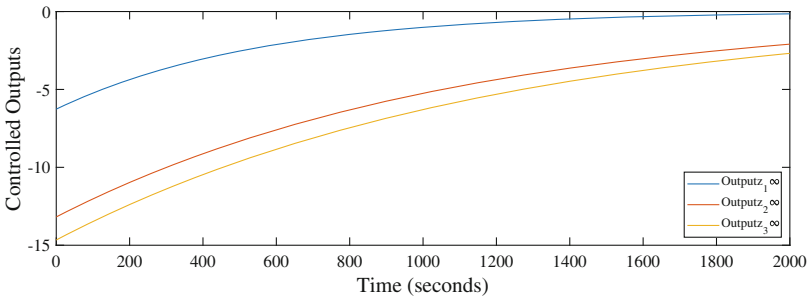


Fig. 1. Closed-loop outputs$(z_{1\infty}, z_{2\infty}, z_{3\infty})$ due to nominal inertia matrix (Ix, Iy, Iz).

Subsequently, the formulated controller is implemented within the closed-loop attitude control system of the satellite. The controller gain is maintained at a consistent value, and the inertia of the matrix has changed within a range of $\pm 5\%$ over 1331 possibilities using Matlab code. In Fig. 2, despite these modifications, the system consistently converged to stability.

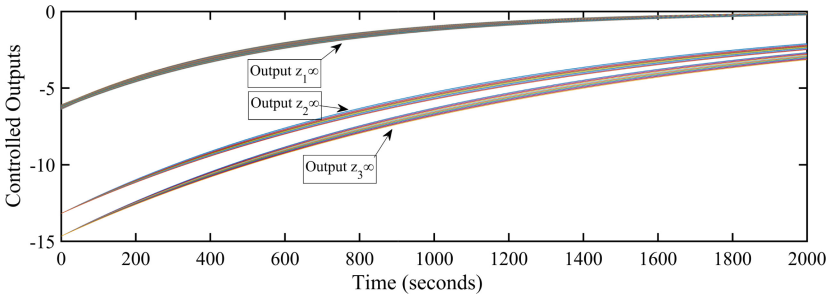


Fig. 2. Closed-loop outputs($z_{1\infty}, z_{2\infty}, z_{3\infty}$) with variation inertia matrix (Ix, Iy, Iz) as 1331 possibilities.

4 Conclusion

In conclusion, this paper has presented a comprehensive study on designing a robust H_∞ controller favourable for satellite systems facing uncertainty in their inertia matrix within a $\pm 5\%$ range. MATLAB Simulink is the proposed methodology that employs an LMI framework facilitated by the LMILAB positive semidefinite programming solver within YALMIP. The simulation outcomes sharply illustrate the controller's remarkable efficacy in ensuring the satellite system's stability against disturbances and sustaining optimal performance even when fluctuations within the inertia matrixes exist. The significance of a robust H_∞ control strategy in enhancing the reliability and resilience of satellite systems, thereby contributing to the advancement of aerospace technology and mission success. Future research directions could explore the applicability of this approach to other dynamic systems and further refine the controller design for broader practical implementation.

Acknowledgments. This work was supported by the Internal Grant Agency of the Tomas Bata University in Zlín, under the project number IGAICebiaTech/2023/004.

References

1. Govindaraj, T., Humaloja, J. P., Paunonen, L.: Robust controllers for a flexible satellite model. arXiv preprint. arXiv:2204.03292 (2022)
2. Liu, C., Sun, Z.W., Shi, K.K., Wang, F.: Mixed H2/H∞ control approach and its application in satellite attitude control system. Int. J. Eng. Res. Afr. **25**, 89–97 (2016)
3. Liu, C., Shi, K., Wang, F.:. Mass and mass center identification of target satellite after rendezvous and docking. In Proceeding of the 11th World Congress on Intelligent Control and Automation, pp. 5802–5807. IEEE (2014)
4. Shi, K., Liu, C., Sun, Z., Wang, F.: Mixed H 2/H∞ approach of full order state observer design for satellite attitude control system. In: The 27th Chinese Control and Decision Conference (2015 CCDC), pp. 4513–4517. IEEE (2015)

5. Liu, S., Zhang, R., Zhang, J.: Robust H (infinity) control of a class sample-data system for satellite attitude control with structured uncertainty. Aerosp. Control **30**(6), 50–53 (2012)
6. Gassara, H., Hajjaji, A.E., Chaabane, M.: Robust H_∞ control for T-S fuzzy systems with time-varying delay. Int. J. Syst. Sci. **41**(12), 1481–1491 (2010)
7. Xia, Y., Qiu, J., Zhang, J., Gao, Z., Wang, J.: Delay-dependent robust $\mathcal{H}\infty$ control for uncertain stochastic time-delay system. Int. J. Syst. Sci. **39**(12), 1139–1152 (2008)
8. Ren, J., Zhang, Q.: Robust H_∞ control for uncertain descriptor systems by proportional-derivative state feedback. Int. J. Control **83**(1), 89–96 (2010)
9. Capua, A., Shapiro, A., Choukroun, D.: Spacecraft Attitude Control using Nonlinear H-infinity Output-Feedback. In AIAA Guidance, Navigation, and Control (GNC) Conference, p. 4793 (2013)
10. Mahmoud, M., Zribi, M.: H∞-controllers for time-delay systems using linear matrix inequalities. J. Optim. Theory Appl. **100**, 89–122 (1999)
11. Isidori, A., Kang, W.: H/sub/spl infin control via measurement feedback for general nonlinear systems. IEEE Trans. Autom. Control **40**(3), 466–472 (1995)
12. Palhares, R.M., Peres, P.L.D.: LMI approach to the mixed H/sub 2//H/sub/spl infin//filtering design for discrete-time uncertain systems. IEEE Trans. Aerosp. Electron. Syst. **37**(1), 292–296 (2001)
13. Grigoriadis, K.M., Watson, J.T.: Reduced-order H/sub/spl infin//and L/sub 2/-L/sub/spl infin//filtering via linear matrix inequalities. IEEE Trans. Aerosp. Electron. Syst. **33**(4), 1326–1338 (1997)
14. Li, H., Fu, M.: A linear matrix inequality approach to robust h/sub/spl infin//filtering. IEEE Trans. Signal Process. **45**(9), 2338–2350 (1997)
15. Wu, L., Wang, Y., Zhou, S., Yuan, X.: Design of mixed H-two/H-infinity optimal control systems using multiobjective differential evolution algorithm. J. Control Theory Appl. **11**, 521–528 (2013)
16. Molina-Cristóbal, A., Griffin, I.A., Fleming, P.J., Owens, D.H.: Linear matrix inequalities and evolutionary optimization in multiobjective control. Int. J. Syst. Sci. **37**(8), 513–522 (2006)
17. Hwang, C.L., Han, S.Y.: Mixed H/sub 2//H/sub/spl infin//design for a decentralized discrete variable structure control with application to mobile robots. IEEE Trans. Syst. Man Cybernet. Part B (Cybernet.) **35**(4), 736–750 (2005)
18. Moghadam, M.G., Jamshidi, F., Beheshti, M.T.: The multiobjective H_2/H_∞/impulse-to-peak output-feedback control via LMI optimization. In: 2009 International Conference on Advanced Computer Control, pp. 281–286. IEEE (2009)
19. Sun, C.C., Chung, H.Y., Chang, W.J.: H_2/ H_∞ Robust Static Output Feedback Control Design via Mixed Genetic Algorithm and Linear Matrix Inequalities (2005)
20. Chung, H.Y., Chan, S.C.: GA-based H_2/H_∞ static output feedback design with average performance concept and techniques of family of polynomials. Expert Syst. Appl. **36**(3), 5859–5865 (2009)
21. Duan, G.R., Yu, H.H.: LMIs in Control Systems: Analysis, Design and Applications. CRC Press (2013)

Pipeline Leakage Detection via Extreme Seeking Entropy

Jakub Steinbach(✉), Jakub Seiner, and Jan Vrba(✉)

Department of Computing and Control Engineering, University of Chemistry and Technology in Prague, Technická 1905/5, 166 28, Praha 6, Czech Republic
{jakub.steinbach,jan.vrba}@vscht.cz

Abstract. In this article, we present a case study of pipeline leakage detection via an adaptive novelty detection algorithm. The detection is based on an affine projection filter applied to a sound recording of the pipeline outlet. Adaptive weight increments of the filter are evaluated by the Extreme Seeking Entropy algorithm to estimate the amount of novelty in data. The suitability of the proposed approach is presented in three case studies where the leakage was simulated in the degassing process model by increasing the discharge flow.

Keywords: novelty detection · adaptive filtering · pipeline leakage detection

1 Introduction

Pipeline transportation, owing to its safety and cost-effectiveness, stands as a prevalent means for conveying liquids or gases across extensive distances. However, the occurrence of leaks resulting from structural impairment in pipelines presents a substantial hazard to the vicinity, stemming from attributes of the conveyed material or the elevated pressure propelling it.

Hence, the significance of early detection of leaks is apparent. Multiple strategies exist for the detection of pipeline leaks. Hao et al. [12] and Diao et al. [7] employed acoustic wave measurements along the pipeline, employing SVM to identify potential leakages. Brunone et al. [3] introduced the transient test-based technique (TTBT) for pinpointing leakage location and severity within a pipeline. Muggleton et al. [15] delved into the use of optical fiber cables to detect lateral acoustic waves generated by leaks.

In this article, we introduce an innovative method for detecting leaks by employing an algorithm rooted in novelty detection principles. Novelty detection (ND) has gained recent prominence as a research area, with various methods emerging to address the challenge. In a recent ND review, authors categorized approaches into deep and shallow methods [18]. Pimentel's work [17] introduced categories like probabilistic, distance-based, domain-based, reconstruction-based, and information theoretic. Another perspective, presented

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 67–74, 2024.
https://doi.org/10.1007/978-3-031-53549-9_7

in [13,14], classifies ND methods into statistical and neural-network based approaches.

ND methods find applications across diverse domains. In healthcare, they identify disease outbreaks [23] and aid medical diagnoses [11]. Industries use ND to detect faults and prevent breakdowns [1,21], as well as foresee natural disasters [24]. In economics, ND helps uncover potential market manipulations [8].

A group of significant algorithms, such as Extreme Seeking Entropy [22], Learning Entropy [4], employ adaptive systems (mainly various adaptive filters) to detect novelty in data. For the purpose of showcasing its applicability in pipeline leakage detection, this study opted to focus on the Extreme Seeking Entropy algorithm.

2 Methods

As indicated, our intention is to assess the effectiveness of the affine projection filter (AP) along with the adaptive novelty detection technique for identifying pipeline leakage. This will be accomplished through an analysis of the audio signal recorded at the pipeline outlet.

2.1 Adaptive Filter

The input of AP filter at the disrete time index k is given as

$$\mathbf{X}_{AP}(k) = [\mathbf{x}(k), \mathbf{x}(k-1), \ldots, \mathbf{x}(k-L)] \tag{1}$$

where L is the projection order, $\mathbf{x}(k)$ is the input vector that contains last n samples.

$$\mathbf{x}(k) = [x(k), x(k-1), \ldots, x(k-n+1)]^T \tag{2}$$

The output of this filter at discrete time index k is expressed as follows.

$$\mathbf{y}_{AP}(k) = \mathbf{X}_{AP}^T(k)\mathbf{w}(k) \tag{3}$$

where $\mathbf{y}(k)$ denotes output at a sample time k, $w_i(k)$ denotes a weight for i-th input and $\mathbf{X}_{AP}^T(k)$ denotes an input at a sample time k. The vector of targets $\mathbf{d}_{AP}(k)$ for affine projection contains $L+1$ target values, and is given as

$$\mathbf{d}_{AP} = [d(k), d(k-1), \ldots, d(k-L)]^T. \tag{4}$$

The error of the affine projection filter is a vector, that is estimated as

$$\mathbf{e}_{AP}(k) = \mathbf{d}_{AP}(k) - \mathbf{y}_{AP}(k). \tag{5}$$

The adaptation of filter weights $\Delta\mathbf{w}(k)$ can be expressed as

$$\Delta\mathbf{w}(k) = \mu\mathbf{X}_{AP}(k))(\mathbf{X}_{AP}^T(k)\mathbf{X}_{AP}(k) + \epsilon\mathbf{I})^{-1}\mathbf{e}_{AP}(k) \tag{6}$$

where μ is learning rate, ϵ is initialization constant, $\mathbf{I}$ is identity matrix. Note, that the adaptive weights update is performed with every new sample, and is given as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \Delta\mathbf{w}(k). \tag{7}$$

2.2 Extreme Seeking Entropy Algorithm

Extreme Seeking Entropy (ESE) is an algorithm that estimates the novelty in data via the changes of adaptive weights. This algorithm utilizes Generalized Pareto distribution (GPD) [6] to estimate the probability of adaptive weight increment. The central notion is that the emergence of novelty within the data should manifest in the magnitudes of adaptive weight increments, indicating the necessity to grasp new process dynamics. ESE algorithm does not require a precise model with precisely optimized weights - as long as the adaptive algorithm can appropriately respond by adjusting model weights.

ESE algorithm quantifies the novelty in the obtained data sample as

$$ESE(|\Delta \mathbf{w}(k)|) = -\log \prod_{i=1}^{n} (1 - f_{cdf_i}(|\Delta w_i(k)|)) \tag{8}$$

where

$$f_{cdf_i}(|\Delta w_i(k)|) = \begin{cases} 0, \ |\Delta w_i(k)| < \zeta_i \\ F_{(\xi_i,\mu_i,\sigma_i)}(|\Delta w_i(k)|), \ |\Delta w_i(k)| \geq \zeta_i. \end{cases} \tag{9}$$

Consistent with GPD $\mu_i \in (-\infty, +\infty)$ is the location parameter, $\sigma_i \in (0, \infty)$ is the scale, and $\xi_i \in (-\infty, \infty)$ is the shape parameter of the underlying GPD. The cumulative distribution function of GPD is given as

$$F_{(\xi_i,\mu_i,\sigma_i)}(|\Delta w_i(k)|) = \begin{cases} 1 - \left(1 + \frac{\xi_i(|\Delta w_i(k)| - \mu_i)}{\sigma_i}\right)^{-\frac{1}{\xi_i}} & \text{for } \xi_i \neq 0, \\ 1 - \exp\left(-\frac{|\Delta w_i(k)| - \mu_i}{\sigma_i}\right) & \text{for } \xi_i = 0. \end{cases} \tag{10}$$

Note, that parameter ζ_i is a threshold for Peak-Over-Threshold (POT) method that enables the adaptive weight increments to be approximated by GPD (for more details see Pickands-Balkema-de Haan theorem [2,16]) as only l largest increments of adaptive weight w_i that met the condition $w_i \geq \zeta_i$ are used for GPD parameters estimation. Those l largest increments are obtained from a floating window with lenght n_s

The ESE algorithm is summarized as follows.

Algorithm 1. Extreme Seeking Entropy Algorithm

1: set n_s and choose POT method
2: initial setting of the GPDs parameters ξ_i, μ_i, σ_i for each adaptable parameter
3: **for** each obtained $d(k)$ **do**
4: update adaptive filter to get $\Delta \mathbf{w}(k)$
5: apply POT method
6: **if** $|\Delta w_i|(k) > \zeta_i$ **then**
7: update parameters of GPDs ξ_i, μ_i, σ_i
8: **end if**
9: compute ESE according to (8)
10: **end for**

2.3 Experiment Description

The length of the AP filter was set to $n = 5$ with order $L = 3$. The learning rate was $\mu = 0.5$ and the initialization constant $\epsilon = 0.005$. Those values were chosen experimentaly. Regarding the ESE algorithm, we employed a floating window with length $n_s = 1000$. The parameter l for the POT method was chosen so that 10% of the largest adaptive weight increments in the floating window are used to estimate the GPDs parameters. The estimation of GPD parameters is obtained via the maximum likelihood method, which is computationally demanding, but does not need any assumptions and limitations, compared to other methods [19].

The experimental setup employed is depicted in Fig. 1. To replicate a pipe leak scenario, a valve was opened, inducing a change in sound attributed to water flow. The audio emanating from the pipe was captured at the outlet of the water tank, resulting in three brief recordings obtained using this method. During the analysis, a challenge arose from the presence of noise caused by the nearby pump's sound, which interfered with the recordings. To mitigate this issue, a contoured polystyrene barrier was introduced between the recording device and the pump. For the recording itself, a Trust TYRO Full HD webcam microphone with a sampling frequency of 48 kHz was utilized.



Fig. 1. Experimental setup used in the experiments (the green dot represents the location where the measuring device was placed).

3 Results

All experiments were implemented in Python 3.9 [20], using NumPy [9] library for fast data manipulation and Padasip library [5] providing various adaptive filters and novelty detection methods. The graphs were plotted by Matplotlib library [10]. The values of ESE were thresholded with value 40, respecting the logarithmic character of the ESE values. The results from 3 different experiments are depicted in the following Figs. 2, 3 and 4, where the peak corresponds with the beginning of pipeline leakage. The key finding of this research is the direct connection between the maximum value of ESE and the onset of fluid leakage from a pipeline. This pivotal discovery highlights the potential role of ESE in early pipeline leakage detection systems based on sound analysis.

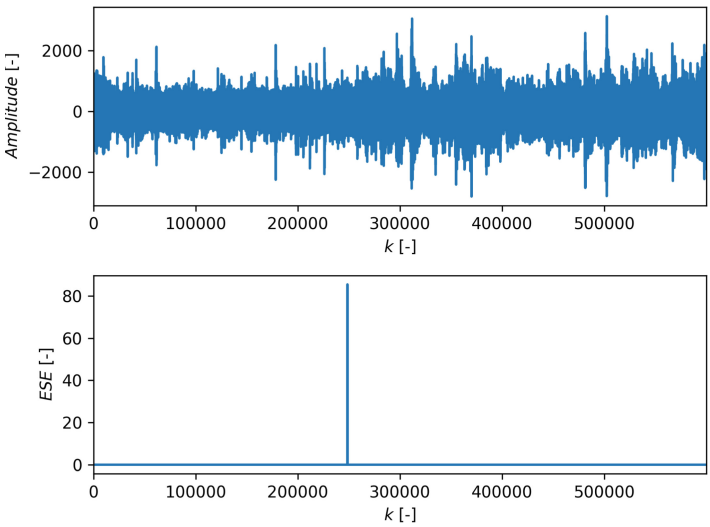


Fig. 2. The first experiment. The upper plot depicts recorded sound data. The lower plot is the result of ESE algorithm. The peak in ESE corresponds with the beginning of the leakage.

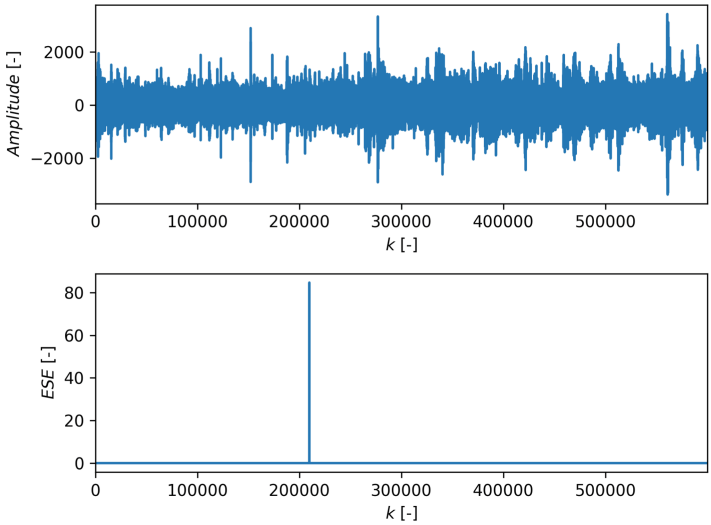


Fig. 3. The second experiment. The upper plot depicts recorded sound data. The lower plot is the result of ESE algorithm. The peak in ESE corresponds with the beginning of the leakage.

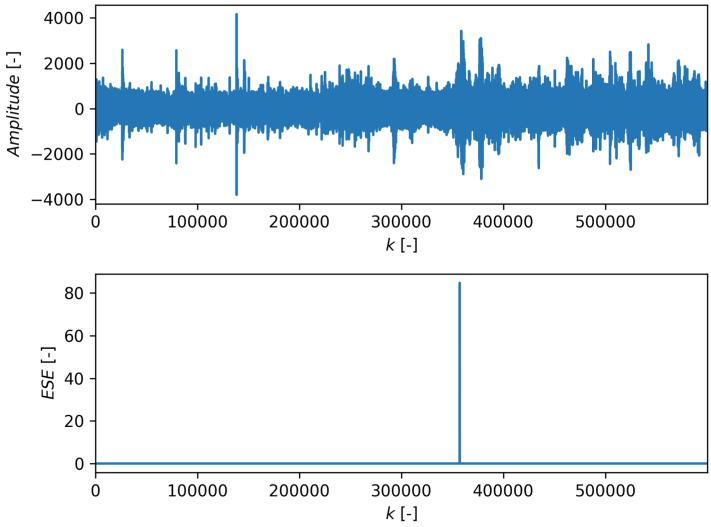


Fig. 4. The third experiment. The upper plot depicts recorded sound data. The lower plot is the result of ESE algorithm. The peak in ESE corresponds with the beginning of the leakage.

4 Conclusion

In summary, we have introduced an innovative approach to address pipeline leakage detection through our laboratory-based study, employing the novelty detection algorithm Extreme Seeking Entropy, which is based on evaluating adaptive filter parameters simulating the monitored signal's changes. Our investigations have yielded promising outcomes, as exemplified by the successful detection of simulated pipeline faults across three distinct experiments.

This method signifies a notable advancement in the field of pipeline monitoring. Its adaptability and non-invasive characteristics render it particularly fitting for implementation in scenarios where direct sensor attachment to the target pipeline presents challenges. The capacity to detect leaks and anomalies without the need for physical sensor installation on the pipeline itself offers a substantial advantage, particularly in situations where conventional monitoring equipment installation proves to be logistically complex or unfeasible.

By identifying potential leakages at their nascent stages, our inventive approach presents the potential to curtail costs by addressing issues prior to their escalation, potentially leading to more cost-effective maintenance strategies. Moreover, this method may enhance personnel safety by enabling prompt leak detection and intervention, thereby mitigating the risks associated with pipeline failures and providing a safer operational environment.

Acknowledgements. Jakub Steinbach acknowledges his specific university grant (IGA) A1_FCHI_2023_003.

References

1. Atha, D.J., Jahanshahi, M.R.: Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. Struct. Health Monit. **17**(5), 1110–1128 (2017). https://doi.org/10.1177/1475921717737051
2. Balkema, A.A., De Haan, L.: Residual life time at great age. Ann. Probabil. 792–804 (1974)
3. Brunone, B., Meniconi, S., Capponi, C.: Numerical analysis of the transient pressure damping in a single polymeric pipe with a leak. Urban Water J. **15**(8), 760–768 (2018). https://www.tandfonline.com/doi/full/10.1080/1573062X.2018.1547772
4. Cejnek, M., Bukovsky, I.: Concept drift robust adaptive novelty detection for data streams. Neurocomputing **309**, 46–53 (2018)
5. Cejnek, M., Vrba, J.: Padasip: an open-source python toolbox for adaptive filtering. J. Comput. Sci. **65**, 101887 (2022)
6. Coles, S., Bawa, J., Trenner, L., Dorazio, P.: An Introduction to Statistical Modeling of Extreme Values, vol. 208. Springer, Heidelberg (2001). https://doi.org/10.1007/978-1-4471-3675-0
7. Diao, X., et al.: An improved variational mode decomposition method based on particle swarm optimization for leak detection of liquid pipelines. Mech. Syst. Signal Process. **143** (2020). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85081914559&doi=10.1016%2fj.ymssp.2020.106787&partnerID=40&md5=e3f9e1ffe0d0f4437b3760d89a194177

8. Golmohammadi, K., Zaiane, O.R.: Time series contextual anomaly detection for detecting market manipulation in stock market. In: 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE (2015). https://doi.org/10.1109/dsaa.2015.7344856
9. Harris, C.R., et al.: Array programming with NumPy. Nature **585**, 357–362 (2020)
10. Hunter, J.D.: Matplotlib: a 2d graphics environment. Comput. Sci. Eng. **9**(3), 90–95 (2007)
11. Iakovidis, D.K., Georgakopoulos, S.V., Vasilakakis, M., Koulaouzidis, A., Plagianakos, V.P.: Detecting and locating gastrointestinal anomalies using deep learning and iterative cluster unification. IEEE Trans. Med. Imaging **37**(10), 2196–2210 (2018). https://doi.org/10.1109/tmi.2018.2837002
12. Jin, H., Zhang, L., Liang, W., Ding, Q.: Integrated leakage detection and localization model for gas pipelines based on the acoustic wave method. J. Loss Prevent. Process Ind. **27**, 74–88 (2014). https://www.sciencedirect.com/science/article/pii/S0950423013002349
13. Markou, M., Singh, S.: Novelty detection: a review-part 1: statistical approaches. Signal Process. **83**(12), 2481–2497 (2003). http://www.sciencedirect.com/science/article/pii/S0165168403002020
14. Markou, M., Singh, S.: Novelty detection: a review-part 2: neural network based approaches. Signal Process. **83**(12), 2499–2521 (2003). http://www.sciencedirect.com/science/article/pii/S0165168403002032
15. Muggleton, J., Hunt, R., Rustighi, E., Lees, G., Pearce, A.: Gas pipeline leak noise measurements using optical fibre distributed acoustic sensing. J. Natural Gas Sci. Eng. **78**, 103293 (2020). https://www.sciencedirect.com/science/article/pii/S1875510020301475
16. Pickands III, J., et al.: Statistical inference using extreme order statistics. Ann. Stat. **3**(1), 119–131 (1975)
17. Pimentel, M.A., Clifton, D.A., Clifton, L., Tarassenko, L.: A review of novelty detection. Signal Process. **99**, 215–249 (2014). https://doi.org/10.1016/j.sigpro.2013.12.026
18. Ruff, L., et al.: A unifying review of deep and shallow anomaly detection. Proc. IEEE **109**, 756–795 (2021)
19. Steinbach, J., Vrba, J.: Computational time evaluation of the various methods of parameter estimation for pareto distribution. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) CoMeSySo 2021. LNNS, vol. 231, pp. 818–827. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-90321-3_67
20. Van Rossum, G., Drake Jr, F.L.: Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam (1995)
21. Vrba, J., Cejnek, M., Steinbach, J., Krbcova, Z.: A machine learning approach for gearbox system fault diagnosis. Entropy **23**(9), 1130 (2021)
22. Vrba, J., Mareš, J.: Introduction to extreme seeking entropy. Entropy **22**(1), 93 (2020). https://www.mdpi.com/1099-4300/22/1/93
23. Wong, W. K., Moore, A., Cooper, G., Wagner, M.: What's strange about recent events (wsare): An algorithm for the early detection of disease outbreaks. J. Mach. Learn. Res. **6**, 1961–1998 (2005). https://www.scopus.com/inward/record.uri?eid=2-s2.0-29144530575&partnerID=40&md5=39780b7a3e54eba0ee3f95d54def644d
24. Wu, Y., Lin, Y., Zhou, Z., Bolton, D.C., Liu, J., Johnson, P.: DeepDetect: a cascaded region-based densely connected network for seismic event detection. IEEE Trans. Geosci. Remote Sens. **57**(1), 62–75 (2019). https://doi.org/10.1109/tgrs.2018.2852302

Holographic Discs-Volumetric Media and Quality Data Storage Media

Salahaddin Yusifov[3], Elnare Firdus[1], Durdana Rustamova[1], Veyis Aliyev[2], Sabina Sharifli[2], Rauf Mayilov[3], and Almaz Mehdiyeva[3(✉)]

[1] Mechanics and Mathematics Department of Western Caspian University, Baku, Azerbaijan
[2] Azerbaijan Technical University, Baku, Azerbaijan
[3] Electronics and Automation‖ Department of Azerbaijan State Oil and Industry University, Baku, Azerbaijan
almaz.mehdiyeva@asoiu.edu.az

Abstract. 3D hologram technology is one of the most advanced technologies today. Holograms are our real future, and the study of this method is one of the most important areas of our time. Three-dimensional holograms have a wide range of applications. Spatial holograms are already being used more and more at presentations, conferences, exhibitions, even at meetings. Nowadays the industry demands a new type of carriers with volume. That is why holographic methods of information processing attract a lot of attention. Based on the interference system of initial data recording they are used for creation of high capacity holographic storage devices, fast, also, reliable and providing protection from unauthorized access to data. Therefore, recording/reading from discs can be obtained by changing the wavelength or phase shift of the reference beam as discussed in the article but also considering the ability of the volume hologram to multiplex, that is the ability to store several different impressions of data in the same recording volume, the following methods have been developed as well. This article tells about these, Recording/Reading methods and also discusses a mathematical model for writing to holographic disks, using the example of the "holographic library" project. #COMESYSO1120.

Keywords: Hologram · Data · Holographic Library

1 Introduction

New modern technologies are forcing people to reach a new stage of development. They make life much easier and more popular and allow us to do things much faster and better. Current technologies are our past innovative ideas [1–3].

The possibility of "volumetric preservation" of information has interested scientists for a long time [4]. Back in 1963, Peter van Heerden proposed this method. But the proposed theory of high-speed recording/reading and storing large amounts of information for half a century failed to be realized. In the first British company "New Medium Enterprises" at CeBIT in 2006 introduced the format of HD VMD media, which aroused

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 75–81, 2024.
https://doi.org/10.1007/978-3-031-53549-9_8

interest in Hollywood. Michael J. Solomon of Warner Bros. Company expressed his intention to promote HD VMD in the movie industry.

The data density on the medium is currently reaching 390 bits/mkm [4], considering the parameter for DVDs not exceeding 5 bits/mkm, this is much higher.

Today, the market for enterprise storage media is growing due to the increasing volumes of data used in business operations, which need to be backed up and stored in case of unforeseen losses. Previously hard drives or tapes were used for this purpose, which with the progress made in science and technology are passing the baton to new drives that have a greater capacity, are more reliable and provide protection against unauthorized access to data. Holographic disks can most likely become such media [5–8].

The holographic processor is implemented on a helium-neon laser (LG-52-1)/at a wavelength of 0.63 μm according to the standard hologram reconstruction scheme (Fig. 1).

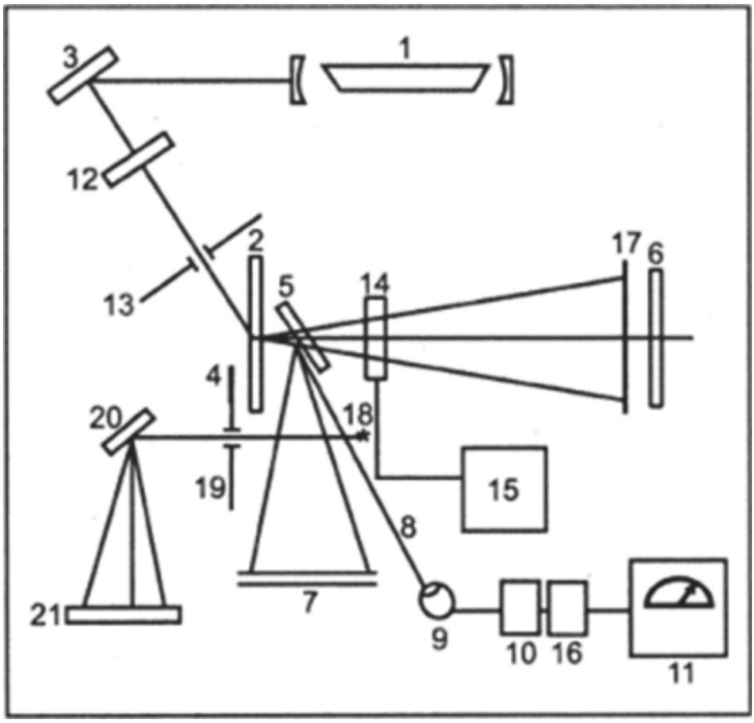


Fig. 1. Optical scheme of the holographic processor: 1 - laser; 2 - block of holograms; 3 - quotation mirror; 4 - hologram with a coordinate mechanism; 5 - flow divider; 6 - photo cassette; 7 - visual control screen; 8 - light guide; 9 - photomultiplier: 10 - amplifier; 11 - indicator; 12 - polarizer; 13 - diaphragm; 14 - photogate with control unit (75); 16 - synchronous detector; 17 - light filter; 18 - light source; 19 - the second diaphragm; 20 - mirror; 21 – screen.

With the help of an adjusting mirror, laser radiation is directed to the block of holograms at an angle of −45°. At the same time, the beam hits the hologram, restoring the image of the search matrix corresponding to one of the request descriptors [9]. The light beam from the hologram is divided by a mirror into two parts. The image of the search matrix is projected simultaneously onto the film and the visual control screen. The name of the descriptor and its numerical code are read from the screen and written into the search matrix on the reconstructed hologram. Precise setting of the hologram is carried out by a coordinate mechanism for moving the block of holograms and is aligned with the image of the search matrix with control marks on the screen.

The brightness of the image on each hologram is controlled by a light guide and a photomultiplier, the signal from which is amplified by the amplifier and indicated on the indicator. The maximum brightness of the reconstructed image is set according to the maximum value of the indicator, provided that the center of the hologram coincides with the axis of the laser beam. The brightness of the reconstructed image can be corrected using a rotating polarizing filter. Aperture 13 reduces noise interference. Exposure of the image of the search matrix is carried out using a photogate. The exposure mode is selected so that the total exposure in the output plane of the processor does not depend on the number of descriptors in the request. The correcting light filter 17 is designed to equalize the brightness of the elements of the reconstructed image in front of the plane of the photographic material. The exposure was made on a highly sensitive photographic film (of the Mikrat-300 type).

The position of the block of holograms was controlled using a point light source IS, a mirror 20, and a screen with a coordinate grid 2/. A cell with the code of the reconstructed hologram is displayed along this optical path. The light spot on the screen moves simultaneously with the block of holograms.

The elements of the optical system of the holographic processor were developed on the basis of micrometric screws of the original design on rigid metal bases to fix them anywhere on the optical table (size 1000 × 540 mm).

The request is entered using linear scales from the processor control unit. The exposed frame of a microfilm (or a separate frame in a cassette) with the results of a search for the semantic (or formal) part of the query is processed in the original photo process and decoded. During numerical decoding, the coordinates of the signals are read from the microfilm frame, which serve as the addresses of the required documents [10, 11]. The search for document addresses is performed using the tables of correspondence of coordinates and document addresses.

An interactive mode or multiple search is possible with sequential adjustment of the query based on the results of the primary search. Such an adjustment was made according to the composition of the request descriptors, their weight coefficients and the relevance threshold.

In the case of searching for secondary documents on one holographic block, you can record up to 5 × 105 abstracts with a volume of more than I03 printed characters each.

The found documents were read on a specially designed reading holographic apparatus with a liquid crystal screen. Such an optical processor had high performance, efficiency in searching for relevant information, and the ability to quickly obtain a copy

of a document. However, the weak point of the holographic processor was the built-in liquid photochemical process for processing photographic films.

2 Recording/Reading Discs

Holographic media is recorded using a laser beam onto a three-dimensional substrate. Different information can be written to the same area of the hologram surface, which provides a high density of data (Fig. 2). Data can be read from holographic disks at very high speeds. It creates compact, portable and high-quality storage devices. The recording types can be very different - alphanumeric, numeric, graphic, object, etc.

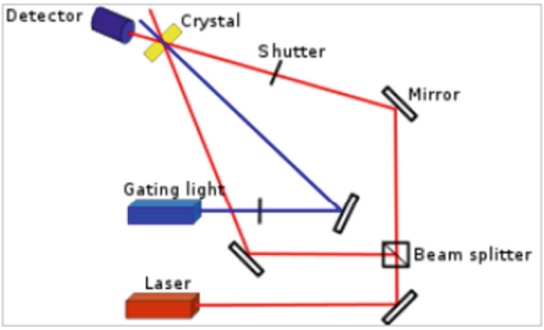


Fig. 2. Set up for holographic recording.

Holographic discs are recorded by means of two parallel beams of red and green lasers where a green laser reads data close to the disc surface and a red laser reads auxiliary signals from the depth of the disc (Fig. 3). The auxiliary information is used to track the reading position.

All information is recorded in one photosensitive layer of the disk. The disk records flat, non-volumetric light information focused by a lens. Recording takes place by interference, which gives the right to call the recorded information a volume hologram. A volume hologram has multiplexing capability.

Multiplexing is the ability to store many different casts of data in the same recording volume.

Recording multiple data to the same volume is achieved by varying the angle of the planes burned across the volumetric photo layer, resulting in extremely high densities.

The recording/reading from discs can also be achieved by changing the wavelength or phase shift of the reference beam. The algorithm for recording/reading information is disclosed by the example of the “holographic library” project in the article.

This recording method makes their commercial application much more difficult. Because this data recording/reading method requires complex optical systems and thick media.

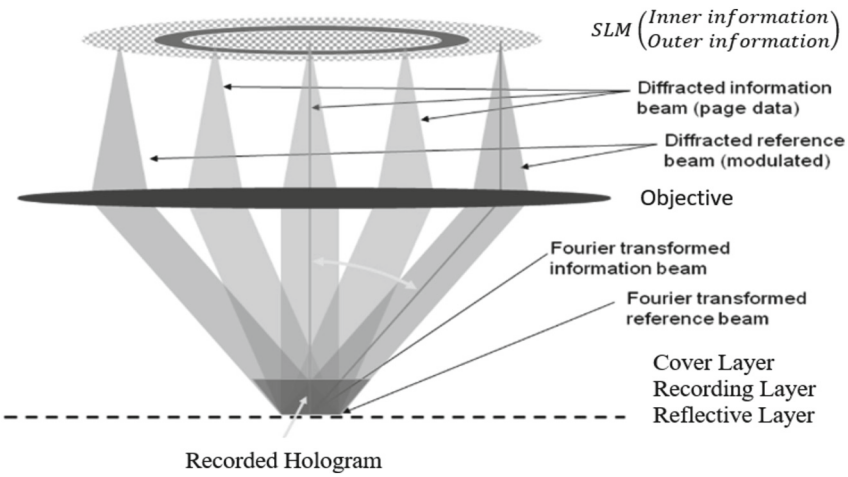


Fig. 3. Holographic technology.

That is why the following multiplexing methods have been developed:

1. Shear;
2. Aperture;
3. Correlation multiplexing.

In all three multiplexing methods, in order to record/read information from the disks, the position of the carrier with respect to the light beams is changed. The light beams used for shear and aperture multiplexing are spherical - reference beams. In correlation multiplexing these beams have a more complex shape.

For a higher compaction of the data recording is also used overlapping books. The essence of this method is to record multiplexed arrays of pages on top of each other. But using this method it is necessary to take into account the ability of the equipment to distinguish information on each separate page, because as the number of recorded pages increases the total transparency of the hologram decreases. Therefore, the degree of density is limited.

In addition, after the formation of information pages with the help of a lens by optical method they can be reduced and when restoring in the same way to increase to the size of the reader.

This method also allows a significant increase in the data access speed. Since to read/write from holographic disks the entire page of data, which contains up to a million or more bits, is accessed simultaneously.

3 Mathematical Model for Holographic Disk Recording

Let us consider this on the example of the "holographic library" project. Let's show it as a tensor T^k. Let's divide the books in our national library into two categories: Azerbaijani and world literature. Let's place information about each of the authors in tensors T^{α}

and T^{d} where, α and d is a set of authors' names, and each of them occupies a certain position in the tensor:

$\alpha \epsilon A$ where, A = {Nizami, Vurqun, Mushviq, Vidadi and others};

$d \epsilon D$ where, D = {Mann, Markes, Bulqakov, Remark and others}.

$$T^{\alpha} = \begin{vmatrix} Nizami\ Vurqun \\ Mushviq\ Vidadi \end{vmatrix} \tag{1}$$

$$T^{d} = \begin{vmatrix} Mann\ Markes \\ Bulqakov\ Remark \end{vmatrix} \tag{2}$$

Now let us assemble a collection of each author's works in the sub-tensors $\mathbf{T}^{\boldsymbol{\alpha}}_{\mathbf{ij}}$ and $\mathbf{T}^{\mathbf{d}}_{\mathbf{ij}}$, but on the condition that they should be located in the elements corresponding to the coordinates of their placement in the tensors $\mathbf{T}^{\boldsymbol{\alpha}}$ and $\mathbf{T}^{\mathbf{d}}$ and the information should be written, respectively, in the following way:

- Azerbaijani literature in the sub-tensor T^{α}_{ij} with shift multiplexing method;
- World literature in the sub-tensor T^{d}_{ij} aperture multiplexing method

$$T^{\alpha}_{ij} = \begin{vmatrix} T^{Nizami}_{11} & T^{Vurqun}_{21} \\ T^{Mushviq}_{12} & T^{Vidadi}_{22} \end{vmatrix} \tag{3}$$

$$T^{d}_{iy} = \begin{vmatrix} T^{Mann}_{11} & T^{Markes}_{21} \\ T^{\mathbf{Bulqakov}}_{12} & T^{Remark}_{22} \end{vmatrix} \tag{4}$$

As we can see, the elements of sub-tensor (3) T_{11}consist of the set of works by Nizami, T_{11} Vidadi are collected in accordance with tensor (1). The same can be said about the sub-tensor (4), where all the elements are collected from the set of works corresponding to the authors specified in the tensor (2).

$$T^{k} = \begin{vmatrix} \begin{vmatrix} T^{Nizami}_{11} & T^{Vurqun}_{21} \\ T^{Mushviq}_{12} & T^{Vidadi}_{22} \end{vmatrix} & \begin{vmatrix} Nizami & Vurqun \\ Mushviq & Vidadi \end{vmatrix} \\ \begin{vmatrix} T^{Mann}_{11} & T^{Markes}_{21} \\ T^{\mathbf{Bulqakov}}_{12} & T^{Remark}_{22} \end{vmatrix} & \begin{vmatrix} Mann & Markes \\ \mathbf{Bulqakov} & Remark \end{vmatrix} \end{vmatrix} \tag{5}$$

For a higher compaction of each author's data record, the principle of overlapping books is applied. In this way a volumetric recording of information on a holographic disk of a "holographic library" is obtained.

Volumetric reading differs from reading CDs or magnetic media in that it is impossible to read the information directly. Each disk has to have a built-in chip to store the data map. In our case it is a tensor T^{k}.

As you can see in this digital library all data about the format and position of the data is stored. When reading data from the disk, the first thing to do is to read the information from this chip. If this data is lost, it becomes almost impossible to read the information [12–15]. They are encrypted using cryptographic methods and only the owner of the information has access to them. As you can see, the recording produced by this method also becomes a means of protecting the information from intruders.

4 Conclusion

New modern technologies are forcing people to reach a new stage of development. They make life much easier and more popular, allowing us to perform familiar activities much faster and more efficiently. Today's technologies are our past innovative ideas.

By creating a "holographic library" we are creating a new future, without crossing out or forgetting our past. A man without a past is a blind and limited man - without a future.

Holographic disks are the speed, quality and ability to store data for a long time. They will allow us to transmit our history to future generations in a more colorful form. All valuable books, journals, information, documents will be preserved in quality and for a long time for future readers.

References

1. Holographic discs (hvd). studfile.net. Accessed 05 May 2023
2. Rustamova, D.F., Elnare, F., Aliyev, V.K., Talibov, Z.A.: Virtual reality in training impact on students of educational programs in augmented reality. J. Pharmaceut. Negat. Results **14**(2), 1 (2023)
3. Holographic information recording systems, studfile.net. Accessed 05 May 2023
4. https://www.pnrjournal.com/index.php/home/article/view/6536. Accessed 05 May 2023
5. Holographic data storage, findout.su. Accessed 05 May 2023
6. Baranov, Y., Gindilis, A.L., Kiselev, B.S., Koblov, A.N., Mikaelyan, A.L., Salakhutdinov, V.K.: Optical memory on the biochemical structure. Quant. Electron. **9**, 1 (1989)
7. Zheenbekov, A.A., Sarybaeva, A.A.: Optimal conditions for the convergence of the inverse functioning algorithm. Sci. New Technol. **1**, 8–11 (2000)
8. Akhmanov, S.A., Nikitin, S.Yu.: Physical optics. Publishing House of Moscow State University, Nauka (2004)
9. Bespalova, V.G., Andreeva, O.V.: Optoinformatics, parts 1, 2. ITMO, St. Petersburg, Russia (2003)
10. Gurova, I.P., Kozlova, S.A.: Problems of Coherent and Nonlinear Optics. ITMO, St. Petersburg, Russia (2006)
11. Ermakov.: Applied optoelectronics. Technosfera, Moscow, Russia (2004)
12. Holographic discs. Information storage technology. Writing and reading an optical disc hologram, studfile.net. Accessed 05 May 2023
13. Neimark, Y.: H 45 Mathematical Modeling as a Science and Art. Publishing House of the Nizhny Novgorod State University, Nizhny Novgorod (2010)
14. Holographic discs store information even after a breakdown. Blog Imena, UA (2023)
15. Rozensher, E., Winter, B.: Optoelectronics. Technosfera Moscow, Russia (2004)
16. Gochuyeva, A.F.: Thermophysical and structural properties of manganese ferrite nanoparticles. Mod. Phys. Lett. B **36**, 2150542 (2022). https://doi.org/10.1142/S0217984921505424
17. Gochuyeva, A.F.: Structural and magnetic properties of manganese ferrite nanoparticles. Mod. Phys. Lett. B **37**, 2350005 (2023). https://doi.org/10.1142/S0217984923500057



Ensuring Operational Continuity and Safety in Refineries Through a Robust Emergency Shutdown System Design and Implementation

Salahaddin Yusifov, Rauf Mayilov, Abdulaga Gurbanov, Ijabika Sardarova, Kerim Bagirzade(✉), Mehriban Mammadova, Sevil Ahmadova, and Zarifa Mahmudova

Azerbaijan State Oil and Industry University, Baku, Azerbaijan
kerim.bagirzada@asoiu.edu.az

Abstract. In summary, the guidelines presented in this research paper offer critical insight into the design and implementation of an effective ESD system. Adherence to these guidelines will ensure that the ESD system is capable of protecting refinery processes in emergency situations while maintaining a high level of availability and low spurious trip rates. The self-diagnostic and modular design features will further enhance the reliability and efficiency of the ESD system, contributing to a safer and more secure refinery operation. Therefore, this paper serves as a valuable resource for refinery engineers and designers in developing an effective ESD system that adheres to industry standards and regulations. The modular design and self-diagnostic features of the ESD system will provide valuable insights into system performance and ensure a robust and reliable ESD system that can help safeguard refinery processes.

Keywords: ESD System · Safety · refinery process · Programmable Electronic Systems · refinery operation

1 Introduction

The aim of this thesis is to establish the fundamental necessities for Emergency Shut Down (ESD) systems utilizing microprocessor equipment (specifically, Programmable Electronic Systems) to serve as the primary process protection system for the refinery.

This specification includes any safety system that is PES (Programmable Electronic Systems) based and designed according to IEC 61508, including all its components.

Consistency is required throughout the entire refinery for the ESD design, configuration, and hardware, including using standardized manufacturers, model numbers, software, and operating philosophy.

The ESD system provides:

- The plant, equipment, or machinery will shut down automatically in the event of a potentially hazardous abnormal condition;

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 82–92, 2024.
https://doi.org/10.1007/978-3-031-53549-9_9

- An interlock system will safeguard valuable equipment throughout all operational phases;
- The system will have functions to reduce false shut downs;
- The system must have testing facilities that are automatic, manual, and online;
- The system will have self-diagnostic capabilities and monitoring tools for system performance such as line monitoring strategies;
- High-priority alarm and ESD management systems, as well as control room alarm interfaces, will be provided.

According to the aforementioned engineering standard, there are specific technical requirements related to the following areas:

- ESD system design and assembling;
- Programming tasks;
- Installation and testing procedures.

2 Design Requirements

The ESD system will utilize a PLC that features a Fault Tolerant (MMR - multiple modular redundant), Fail-Safe (De-Energize-To-Trip) architecture.

The ESD must be designed, supplied, installed, and tested in accordance with the requirements of IEC 61508 standards. The configuration of the logic solver will be based on two requirements:

- The target for availability (or SIL);
- The frequency of false shutdowns (spurious trip rate).

If a high level of availability is necessary, or the impact of false shutdowns is significant, 1oo2D redundant, 2oo3 Triple Modular Redundant (TMR), Quad (2oo4), or other appropriate architectures must be employed [1–3].

The ESD system will be a Quadruple Modular Redundant (QMR) system, which is more dependable than TMR systems because they are usually implemented using two-fault tolerant architectures.

In essence, even if two processors experience a simultaneous fault, the system can still execute the safety function at the highest SIL level.

The ESD components must be separate from the basic control system components (DCS, process PLCs, etc.) so that any failure of these system components will not negatively affect the ESD's integrity or functionality, except as outlined below. These components will consist of sensors, logic, final elements, manual initiators, power supplies, and communication networks [4, 5].

In some particular instances, it is acceptable to use the same components for both the ESD and control system. If shared components are used, they must be deemed as part of the ESD and be powered by it. All instances of shared components must be explicitly identified in instrument records and through unique tag IDs in the control system to ensure that proper operating, maintenance, and control application procedures are followed throughout the system's lifecycle. The acceptable situations where component sharing is permitted are specified below. For situations not listed, formal review and approval is necessary [6–9]:

- If prior engineering practice has shown that shared components result in considerable cost savings without compromising safety;
- In cases where specific equipment control requires a dedicated stand-alone control and protective system package that is separate from the main Control System, the control system equipment shall be designed using hardware that meets the same standards as a conventional PES Logic Solver. In such situations, the controlling logic or associated inputs, or failures within the controlling logic, shall not compromise the system's ability to achieve the required SIL or protective function. The integrated protective and control system package shall be engineered and designed in accordance with ESD system requirements, including change management and access control. Changes to the control part of the logic should be made at the same low frequency as changes to the protective system part of the logic and follow the same change management procedures.

This refers to situations where a separate protective system package is needed for specific equipment control, and it is not feasible or necessary to incorporate process control functionality into the main Process Control System.

The design of ESD should enable testing of its components at regular intervals to reduce the possibility of failure when required. Unless instructed otherwise, the design should allow for maintenance and testing to be performed online.

External connections to the ESD system should not compromise its safety integrity or protective function, which necessitates signal isolation and read-only communications [1, 20]. In exceptional cases where written communications are required between the DCS and ESD, the FTE network and Safety Network should be utilized for peer-to-peer communication from the HMI/OWS to the ESD controller [4]. The Fault Tolerant Ethernet (FTE) network is integrated into the Experion PKS, unifying all systems into one architecture, which can enhance safety, process availability, and efficiency. This solution offers an optimal level of safety and process integration while still maintaining functional safety separation as required by IEC 61511 and IEC 61508.

The ESD system must be designed to prevent unauthorized access to the bypass or override functions and unauthorized modification of the protective function, such as program or set point changes. The ESD logic design should read all field device status at 'power-on' or after a processor replacement, and it should not incorporate logic states of field devices within the program that depend on transitions that may not be maintained during power-up or processor change. ESD components must be appropriate for the specified electrical area classification in which they are installed, and they should not be installed in locations where vibration is present. If a portion of a package unit is evaluated as a protective system, then this portion must be implemented according to this specification. ESD trip initiators should be independent and specific to a single piece of equipment or have provisions for redundant voting of the sensor if used to protect multiple pieces of equipment [7, 10, 11]. If there is a need to trip one system as a result of another system tripping, this should be done directly using a signal from the first ESD to the second ESD. The trip of the second system should not be inferred from cascading effects from the results of the trip of the first system (e.g., waiting for trip initiators to react).

3 Logic Architecture

The design should consider the following factors:

- Typically, the logic for a single system should be located in a single cabinet.
- The design of the logic equipment should include all necessary test facilities, such as bypass switches.

In cases where non-independent process units are functionally part of other units, they can share a common architecture if the following conditions are met:

- One cabinet may contain multiple systems, but it should be evident that each system is separate and distinct.
- Wiring, fusing, and terminations for each system must be separate and identified clearly by labeling the equipment. Terminal dividing plates can be used in terminal strips to differentiate signals from each system, as long as the separate systems are clearly identified.

A fault-tolerant logic must raise an alarm in case of any single fault, but it should not initiate any protective action unless required by the implementation conditions [9].

To prevent electrical noise from causing nuisance trips, all sensor inputs must be electrically isolated.

Time delays should be properly used to avoid nuisance trips and the time delay on sensor inputs should be set to required seconds unless specified otherwise.

ESD logic and the relevant final elements should remain in their protective state even after a trip until manually reset, including a trip caused by any reason, such as power loss, even if power source(s) and/or trip initiators return to their normal operating positions.

To perform a manual reset, an electrical switch must be installed for this specific purpose in the field or in dedicated spaces on operator control desks in control rooms. It should send a momentary 24 VDC discrete input signal to the safety system.

After a successful manual reset, any involved final element should stay in the safe position to which ESD logic is forced to go [1, 12–15].

The design criteria should be shared with any other control systems that have access to the same final elements to prevent commands from changing the status of the final elements.

After the control system logics have been aligned by the successful activation of a manual reset, operators shall have full control of the final elements [7, 16, 18].

Dedicated Auxiliary Consoles (Matrix design) will be provided for each Control Room, equipped with:

- Emergency Pushbuttons to trigger main equipment trip function;
- Critical Equipment reset, override, or direct commands;

Lamp indications to provide feedback on the above-listed functions.

4 ESD Hardware Requirements

4.1 ESD Structure

The Emergency Shutdown (ESD) system shall possess the subsequent features:

- The design shall be modular and include separate I/O circuits, processor modules, and communication modules, allowing I/O cards to be installed remotely via redundant fiber optic cables
- The system shall be capable of replacing any module online without affecting the process or requiring any system downtime
- The system shall feature an "auto spare" function that automatically switches to a backup I/O module in the event of a module failure
- Self-testing and system diagnostics shall be an inherent part of the standard system and require no additional application logic. The self-testing and diagnostics shall be transparent to the user during system implementation
- Self-diagnostic tests must include at least the following:

  - Serial and parallel communications must be checked for errors;
  - There must be an internal watchdog timer that can detect when the processor execution has halted or is looping;
  - Every functional cycle must include a set of instructions that can exercise the active system components, including the processor;
  - The memory must be checked periodically;
  - Prior to a "read" or "write" operation to an input or output component, each signal line of a parallel bus must be checked;
  - The system must be able to detect the removal or defect of any logic unit, communication module, processor, I/O module, or power supply;
  - The system's logic-solving ability must be checked, including a test that can exercise the active system components, such as the processor(s). This test must be run before each logic cycle;
  - A set of power-up initialization and communications checks must be performed.

4.2 System Cabinet

Each ESD's system cabinet, which contains the main chassis, shall consist of:

- Redundant power supply units;
- Power distribution panels with circuit breakers;
- Redundant communication modules;
- Redundant input/output modules (IOM);
- If needed, the first ESD system cabinet may include two additional expansion chassis, each containing redundant power supply units, redundant system bus communication modules, and input/output modules.

4.3 Marshalling Cabinets

To ensure proper functionality, the I/O Terminal Assemblies shall be situated in separate Marshalling Cabinets unless there is more than one System Cabinet, in which case, it may be permissible to position them at the back of the additional System Cabinet(s) [5, 17, 19, 20]. The Marshalling cabinets shall be divided into separate units based on the type of

- AI/AO signals;
- DI/DO signals;
- Intrinsically safe signals.

Additionally, I/O with varying voltage levels must be kept separate.

Each marshaling cabinet shall have dedicated grounding bars for each of the following types:

- A non-insulated AC/frame safety earth bar (PE);
- An insulated instrument master reference earth bar (IE) for field cable shields, with a resistance to earth of less than 1 Ω;
- An insulated instrument IS reference earth bar for field IS cable shields, with a resistance to earth of less than 1 Ω.

4.4 Network Cabinets

The cabinets for network equipment will provide space for fiber optic converters, switches, patch cords, routers, power supply modules, and other necessary equipment. These network cabinets should follow the same design criteria as the system/marshalling cabinets, and ideally, any patch panels for field fiber optic circuits should be installed in separate dedicated cabinets or panels.

4.5 CPU

The CPU must possess the following functionalities and accessories:

- The CPU must have the capabilities to process I/O, perform online diagnostics at every scan cycle, and notify of any abnormal CPU operation;
- The CPU design should include a "watchdog" timer to detect any halt or loop in CPU execution or exceeding CPU loading beyond 100%;
- The CPU program logic and data registers must be saved during power outages, and if battery backup is used, the system must have a low battery alarm and allow for online battery replacement;
- The CPU must provide all program logic and computational functions necessary to fit the application effectively and efficiently.
- The selection of the quantity and nature of CPU communication ports shall be determined by the following criteria:

  - Interfacing with remote I/O channels, with potential growth requirements taken into account;
  - Interfacing with programming devices such as portable engineering workstations;
  - Interfacing with peer-to-peer networks is restricted only to safety systems;
  - Interfacing with DCS equipment, requiring support for applicable protocols.

5 ESD Logic Requirements - Software

5.1 Design Requirements

To meet the minimum requirements, the following items shall be produced:

- A program listing;
- A system configuration;
- Application logic diagrams;
- A cross-referenced list linking equipment tag numbers to their respective program use locations;
- The application logic may be developed using one or a combination of the following programming languages:

  - Cause-Effect Matrix;
  - Function Block Diagram;
  - Sequential Function Chart;
  - Ladder Diagram.

- A library of standard logic blocks and configurations shall also be included.

5.2 Program Structure

The following deliverables must be produced to meet the minimum requirements:

A function block diagram format for the overall application program, which should be divided into blocks in the following order:

- Diagnostics and Grouping of Alarms;
- Standard Function Blocks;
- Custom Function Blocks (such as 1oo1 AI, 2oo3 AI, etc.);
- Miscellaneous Global Functions (including Timer Values, Lamp Test, Blink Coils, Dead-band Constants, etc.);
- Main Program;
- Input Processing;
- Input Alarm Detection;
- Voting (if required);
- Logic Execution;
- Output Processing;
- Output Alarm Detection;
- First-Out Alarming;
- Communications to DCS;
- Other (if required).

All logic should be written in the "negative" or "de-energized" direction, and the use of positive logic is strictly forbidden.

5.3 Voting and Deviation Alarms

If the use of voting redundancy is required for sensors, 2oo3 voting is the recommended method to achieve high availability and low spurious trip rate [9]. When using 2oo3 voting for analog values, "median" selection shall be used, and an alarm shall be generated if any value falls outside of a predetermined deadband above or below the median value. For ESD 2oo3 voted discrete devices, a logic function shall be implemented to differentiate between one device and the other two (Fig. 1, Fig. 2).

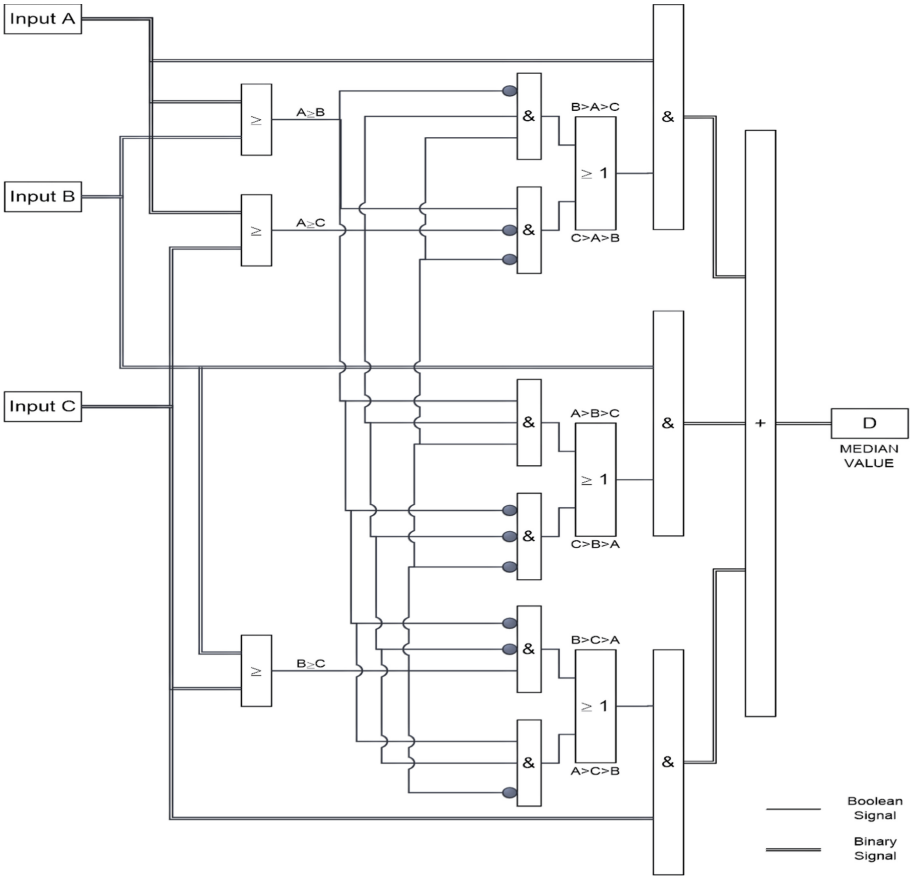


Fig. 1. 2003 median select voting algorithm.

To ensure diversity, each voted input shall be allocated to separate I/O modules, and separate racks shall be used unless the backplane design allows input to adjacent cards without a common mode failure scenario. All process variables (PV) values for ESD analogs shall be transmitted via the communication link to the DCS system, including the median value [10]. A 5-s delay will be applied to deviation and discrepancy alarms. The voting of groups will be compromised if any failures are detected, as shown in Table 1.

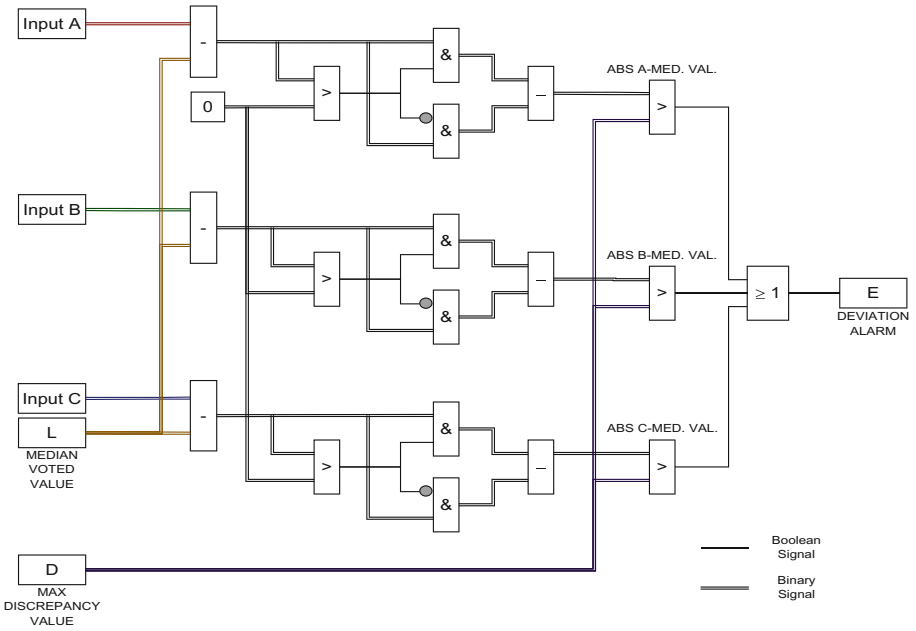


Fig. 2. Deviation alarm of 3 AI from median value.

Table 1. Detected failures.

Original Voting	Voting (1 failure)	Voting (2 failures)	Voting (3 failures)
1oo1	Trip	n/a	n/a
1oo2	1oo1	Trip	n/a
2oo2	1oo1	Trip	n/a
2oo3	1oo2	1oo1	Trip

6 Conclusion

In conclusion, this research paper has successfully established the fundamental criteria for designing and implementing an Emergency Shutdown (ESD) system for refinery processes. The use of microprocessor equipment, specifically Programmable Electronic Systems, has been identified as the primary process protection system for the refinery. The adoption of Fault Tolerant (MMR), Fail-Safe (De-Energize-To-Trip) architecture, and specific redundancies will ensure high availability and low spurious trip rates of the ESD system. Furthermore, adherence to IEC 61508 standards is imperative for safe and reliable ESD system operation. The logic architecture design, with consideration for the separation and distinctiveness of each system, has been identified as critical for effective ESD operation. The modular design and self-diagnostic features of the ESD system will

provide valuable insights into system performance and ensure a robust and reliable ESD system that can help safeguard refinery processes.

In summary, the guidelines presented in this research paper offer critical insight into the design and implementation of an effective ESD system. Adherence to these guidelines will ensure that the ESD system is capable of protecting refinery processes in emergency situations while maintaining a high level of availability and low spurious trip rates. The self-diagnostic and modular design features will further enhance the reliability and efficiency of the ESD system, contributing to a safer and more secure refinery operation. Therefore, this paper serves as a valuable resource for refinery engineers and designers in developing an effective ESD system that adheres to industry standards and regulations.

References

1. Crowl, D.A.: Introduction to Process Safety for Undergraduates and Engineers, 416 p. Wiley-AIChE (2016). ISBN-13: 978-1119453951
2. Goble, W.M., Cheddie, H.: Safety Instrumented Systems Verification: Practical Probabilistic Calculations, 223 p. CRC Press, Boca Raton (2010). ISBN-13: 978-1439837035
3. Gruhn, P.: Safety Instrumented Systems: Design, Analysis, and Justification, 608 p. ISA (2016). ISBN-13: 978-1937560783
4. Holloway, M.D., Nwaoha, C.: Process Plant Equipment: Operation, Control, and Reliability, 452 p. CRC Press (2019). ISBN-13: 978-1498763782
5. Goble, W.M., Cheddie, H.: Safety Instrumented Systems: Verification - Practical Probabilistic Calculations, 2nd edn, 246 p. CRC Press, Boca Raton (2020). ISBN-13: 978-0367272767
6. Manuele, F.A.: Advanced Safety Management: Focusing on Z10 and Serious Injury Prevention, 568 p. Wiley, Hoboken (2010). ISBN-13: 978-0470935746
7. Bahadori, A.: Hazardous Area Classification in Petroleum and Chemical Plants: A Guide to Mitigating Risk, 252 p. Elsevier, Amsterdam (2013). ISBN-13: 978-0127999675
8. Sutton, I.: Process Risk and Reliability Management: Operational Integrity Management, 370 p. Gulf Publishing Company (2014). ISBN-13: 978-0128001810
9. Rumane, A.R.: Emergency Shutdown System: A Systematic Approach to Shutdown Process Safety, 302 p. CRC Press, Boca Raton (2018). ISBN-13: 978-1138574299
10. Mordvinov, A.A., Miklina, O.A.: Gas lift operation of oil and gas wells. Ukhta. 2013; 39. 2. Edgar Camargo and Others. Production Improving in Gas Lift Wells using Nodal Analysis, Signal Processing, Robotics and Automation. 2008; 99–102. 3. Java Native Interface Specification, Version 1.1. Sun Microsystems, Inc. (1997)
11. Mirzajanzadeh, A.X., et al.: Processing and exploitation of oil and gas resources
12. Torres-Echeverria, A.C., Martorell, S., Thompson, H.A.: RESS, p. 106 (2012)
13. Mehdiyeva, A.M., Quliyeva, S.V.: Mathematical model for estimation the characteristics of the noise immunity. J. Phys. Conf. Ser. Cybern. IT. **2094**, Ser. 2094 032060 (2022)
14. Hu, B.: Characterizing gas-lift instabilities. Department of Petroleum Engineering and Applied Geophysics Norwegian University of Science and Technology Trondheim, Norway, pp. 1–178 (2004)
15. Camponogara, E., Nakashima, P.H.: Solving a gas-lift optimization problem by dynamic programming. Eur. J. Oper. Res. **174**, 1220–1246 (2006)
16. Mehdiyeva, A.M., Baxtiyarov, I.N., Bakhshaliyeva, S.V.: Increasing the immunity of information transmission and fault tolerance of the path. In: Shakya, S., Papakostas, G., Kamel, K.A. (eds.) Mobile Computing and Sustainable Informatics. LNDECT, vol. 166, pp. 775–784. Springer, Singapore (2023). https://doi.org/10.1007/978-981-99-0835-6_55

17. Forero, G., McFadyen, K., Turner, R., Waring, B., Steenken, E.: Gas Lift Design Guide Management of Artificial Lift Systems, pp. 1–155 (1993)
18. Rausand, M.: Reliability of Safety-Critical Systems: Theory and Applications, Wiley, New York (2014)
19. Metso Automation, ESD valve selection guide general ESD valve definition (Metso Automation) (2005)
20. Mehdiyeva, A.M., Bagirzadeh, K.E.: Refinery emergency shutdown system based on high safety analysis. J. Eng. Res. Rep. **23**(7), 37–41 (2022). Article no. JERR.92158



Cyber-Physical Fire Detection and Recognition System with Smart Glasses

Nikolay Gospodinov(✉) and Georgi Krastev

University of Ruse, Ruse, Bulgaria
{ngospodinov,geork}@uni-ruse.bg

Abstract. The presence of fire is one of the main prerequisites for the occurrence of a fiery calamity. Early warning of the occurrence of even a small to larger fire is a good prevention to reduce fire casualties. The paper aims to present a prototype module for real-time fire detection and recognition. The prototype uses concepts of artificial intelligence, neural networks, and computer vision. The use of these technologies is not only useful, but also shows the extent of their development and how useful they are in solving problems of life around us. The module works with augmented reality smart glasses. The modular prototype quickly and in real time detects a fire and sends a message to the user, alerting him of the fiery situation that has occurred. Detecting, recognizing, and signaling when a fire occurs makes a cyber-physical system applicable to the work of firefighters by helping them locate fires from a distance and reducing the likelihood of accidental casualties. The research is based on the study of a lot of data related to the recognition and analysis of fire. Its composition and the reasons for its occurrence are also the subject of research. All accumulated knowledge leads to the successful creation of the prototype.

Keywords: Fire · Recognition · Computer vision · Artificial intelligence · Smart Glasses · Cyber-physical system

1 Introduction

Combustion is a rapid self-accelerating chemical transformation of substances, which is accompanied by an intense release of heat and light. For the occurrence and development of combustion, the interaction of a combustible substance, an oxidizer and an ignition source is necessary. The most common oxidizing agent is oxygen. Its concentration should be less than 14–18% by volume. Only substances such as acetylene, ethylene, hydrogen, carbon disulfide can burn at a concentration of up to 10% by volume. For the combustion process to occur, the chemical reaction must be exothermic and accelerate with increasing temperature. Therefore, combustion occurs at one point and spreads, covering the entire mass of the combustible substance.

Fires can be caused intentionally or through carelessness. Every year, thousands of people around the world fall victim to fires. Fires can by of any nature: accidental, burning car, burning apartment, burning three, etc. Locating, recognizing, and preventing

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 93–102, 2024.
https://doi.org/10.1007/978-3-031-53549-9_10

fires in their initial stages is extremely important and is a prerequisite for reducing fire casualties.

The proposed prototype implements the capabilities of artificial intelligence and works with smart glasses with augmented reality.

The research is conducted, using a third generation of augmented reality glasses Moverio BT-300 (Fig. 1). Their benefits include energy economy, portability, great luminance, and visual contrast. The inactive portion of the display appears fully transparent, and the projected picture is seamlessly transferred to physical objects.

The development of the prototype includes a module for localization, detection and signalling of fire.



Fig. 1. Moverio BT-300 smart glasses

2 Methods

2.1 Causes for Fire Occurrence

There are different sources of ignition. Some of them are presented in Table 1. In a number of cases, chemical interactions lead to self-heating.

A similar phenomenon is found in some biological processes. In them, combustible substances are a nutrient medium for microorganisms. During biological processes, heat is released, which can cause spontaneous combustion [1].

Table 1. Sources of ignition

Sources of ignition
Open flame
Heated surfaces
Electrical discharges
Light, ionising and laser
Radiation
Electromagnetic fields
Compression of gas mixtures
Shock wave Mechanical and electrical sparks
Chemical and biological process

There are several substances prone to ignition. Knowing them helps in locating and identifying them more quickly. Table 2 presents some of them.

Table 2. Substances prone to ignition

Name	Information
Substances of plant origin	For example, hay, sawdust, etc. In them at a temperature of 60–70 °C, biological processes take place, which pass into oxidative chemical processes
Brown and hard coal	They contain organic substances that oxidize. Coal is heated and when the heat generated is not dissipated, it ignites
Fats and oils	For example, cooking oils it low smoking point
Substances that ignite on contact with air	White phosphorus, aluminium and zinc powder, iron sulphides
Substances that ignite on contact with water	Alkali metals, calcium carbide, quicklime, magnesium powder, etc.
Substances that ignite on contact with other substances (mainly oxidizing agents)	For example, organic substances in contact with concentrated nitric acid

2.2 Smart Glasses Moverio BT-300

With Moverio BT-300 smart glasses running on the Moverio operating system, a prototype intelligent system for fire localization, detection and signalling is developed. In addition to supporting WAV, MP3, and AAC music formats, they have a light sensor, camera, and 9-axis motion sensor (accelerometer, gyroscope, and others) (Table 3).

Table 3. Moverio BT-300 Specification

Specification name	Information
OS Version	Moverio OS
Gyroscope	Yes, 3-axis in both Headset and Controller
RAM	2 GB
Firld of View	23° (diagonal)
Screen Size (Projected Distance)	40 in. at 2.5 m–320 in. at 20 m

The Moverio BT-300 glasses can show 3D material side-by-side on the left and right sides of the screen and project images for both eyes. For instance, when the

receiver is a direct participant in a live performance, utilizing the built-in video camera and microphone, he may not only send real-time sound information on the event's ambiance (through social networks), but also receive one in which emotional empathy is simultaneous [2].

3 Implementation

The prototype software is built from a single module: Module for localization, detection and signalling of fire.

In Fig. 2, a diagram of the software implementation is shown. The Implementation is carried out in six stages: Creating/Modifying a dataset of fire images samples for training and testing the fire localization and detection model based on artificial intelligence; Develope a model for localization and detection of fire; Training the model with the dataset; Implementing the trained model in Android-based software; Fire localization and detection; Signalling for fire detection.

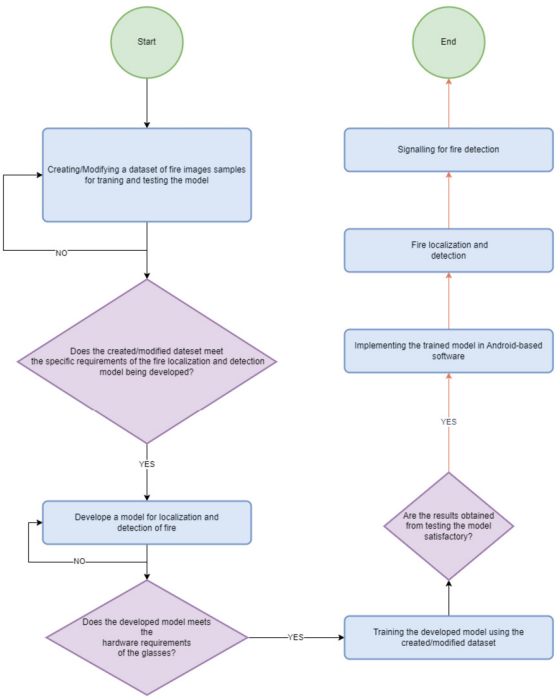


Fig. 2. Software implementation architecture

3.1 Creating/Modifying a Dataset of Fire Images Samples for Training and Testing the Fire Localization and Detection Model Based on Artificial Intelligence

This step is crucial for the successful operation of the developed model. The dataset of image samples must be rigorous and meet operable criteria.

A good selection of images is crucial for the training and testing phase of the programmed model. Close monitoring of the images used is required because the likelihood of the neural network being spoofed is quite high at this stage. Some of the criteria are:

- The largest possible set of image samples for training and testing.
- Image samples must be of high quality.
- The image samples should portray the searched object from all possible positions and angles.

The dataset contains two categories of images: images of fire indoors and outdoors, images resembling fire (flashing lights, headlights, etc.). The second category is added to prevent false-positive outcomes.

3.2 Develop a Model for Localization and Detection of Fire

At this stage, an artificial intelligence-based neural network has been developed. The neural network must meet the following criteria:

- The neural network model should work on devices with undemanding hardware requirements (mobile devices, smart glasses).
- The model must be real-time executable.
- The neural network recognizes medium fires of about 60 × 60 pixels.
- The model needs to distinguish other objects with color and similar shape to the fire: flashing lights, headlights, glare in the image.
- Figure 3 shows the principle scheme of the fire localization and detection model.

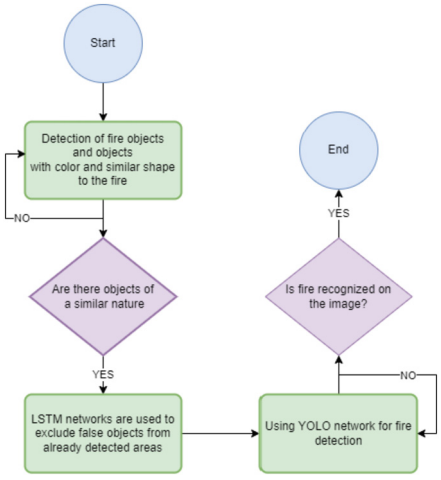


Fig. 3. Scheme of the fire localization and detection model

YOLO V2 and Residual Neural Network model for Detection and Recognition of Fire Objects and Objects with Color and Similar Shape to the Fire

At this stage, the user, using the camera built into the smart glasses, observes the environment that surrounds him.

With the help of the camera, a series of frames are captured. In some of the frames there are objects of different types and in other frames there are no objects present. At this stage YOLO convolutional neural network model is used. The classifier layer is remains and the feature extraction layer is replaced with ResNet-18 neural network. (Fig. 4).

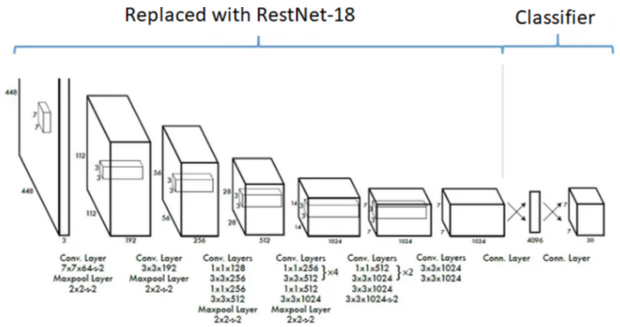


Fig. 4. YOLO Architecture with ResNet-18 as feature extraction

The working process of the model is presented in Fig. 5. The work of the model is divided into several steps:

- Predicting bounding boxes.
- Using IoU to calculate the distance of similarity between the bounding box of target and predicted output.
- Using ResNet-18 for feature extraction
- Using YOLO v2 for classification

The model divides the image into a V × V grid.

Once an item's center falls inside a grid cell, that cell is responsible for detecting the object. Each grid cell forecasts Z bounding boxes and associated confidence ratings. These confidence ratings indicate how certain the model is that the box contains an object, as well as how accurate the model believes the box to be. The confidence is defined with formula 1:

$$Pr(Objects) * ION \tag{1}$$

If there is no item in a given cell, the confidence ratings should be set to zero.

If not, the confidence score has to be equal to the intersection over union (IOU) of the projected box and the actual box. Each bounding box consists of the following five predictions: x, y, w, h, and confidence. The (x, y) coordinates reflect the box's centroid with relation to the boundaries of the grid cell. With relation to the whole picture,

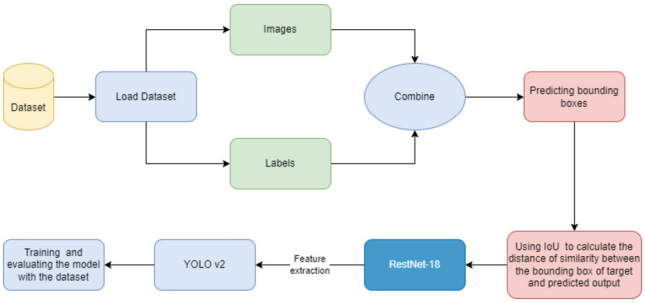


Fig. 5. Working process of the model

width and height are anticipated. Lastly, the confidence prediction shows the debt owed between the anticipated box and any actual box. Moreover, each grid cell forecasts the C conditional class probabilities (Formula 2):

$$Pr(Classi|Object) \tag{2}$$

These probabilities are dependent on whether or not a given grid cell contains an item. Regardless of the number of boxes Z, we only anticipate a single set of class probabilities per grid cell.

After this step a ResNet-18 is used for feature extraction. ResNet-18 is a 72-layer architecture with 18 deep layers [3]. Each layer has convolution size 1 × 1 and 3 × 3 with feature maps (64, 128, 256, 512) as shown in Fig. 6. The detection network (YOLO v2) consists of several convolutional layers, a transform layer, and an output layer. The transform layer pulls activations from the convolutional layer and increases the network's stability [4].

LSTM Network for Excluding False Objects from Already Detected Areas

After completing the previous stage, the model recognizes objects from an image.

In this image there are objects of different nature (fire, luminosity, glare, etc.). In order for the model to distinguish fire from other objects resembling its characteristics, a Long short-term memory (LSTM) is used (Fig. 7). LSTM is a type of Recurrent neural network (RNN).

The cell remembers values across arbitrary time periods, and the three gates control the flow of information into and out of the cell.

By comparing a prior state to the current input and assigning it a value between 0 and 1, forget gates determine what information to delete from a previous state. A value of 1 indicates to retain the information, whereas a value of 0 indicates to discard it. Using the same principle as forget gates, input gates determine which new bits of data to store in the existing state. Output gates determine whether information in the current state should be output by assigning a value between 0 and 1 to the information based on its previous and current states. In this case, fire object is 1 and non-fire object is 0 [5]. By selectively outputting pertinent information from the present state, the LSTM network

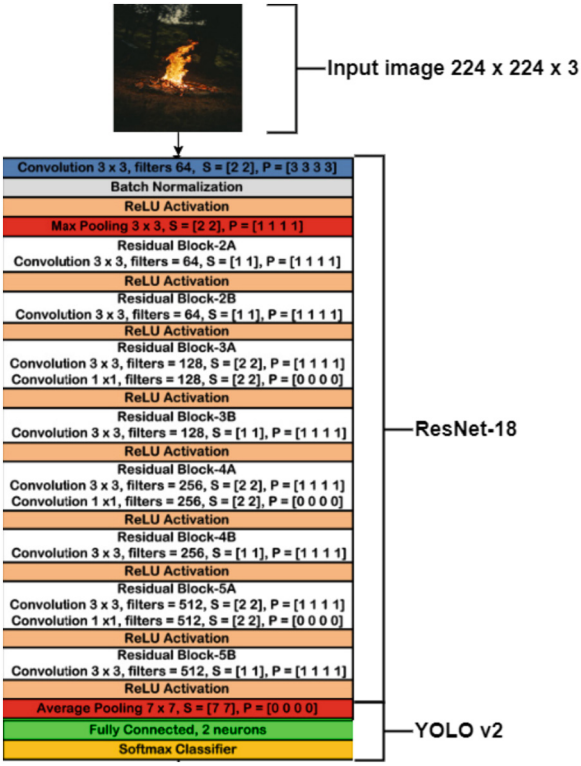


Fig. 6. YOLO v2 with ResNet-18

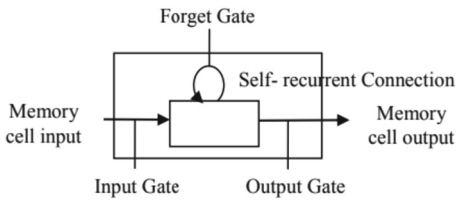


Fig. 7. LSTM cell

is able to preserve important, long-term dependencies for making predictions in both current and future time steps.

3.3 Training and Evaluating the Model with the Dataset

Once the model is created, the trained process follows. This process is extremely important, and the results show to what extent the model works.

A prepared set of images is used to train the model to be tested. It contains images with fire, images with objects resembling fire and images without fire. Once the model

returns a satisfying result, it is now ready for implementation in the Android-based smart glasses application.

3.4 Implementing the Trained Model in Android-Based Software

The creation of a prototype program for augmented reality smart glasses necessitates the usage of several software tools and platforms.

Software platforms for generating and training a model for localization and recognition of dynamic life-threatening events and Software platforms for implementing the trained model and creating the user interface of the prototype software are the two types of software platforms employed. The first software platform utilized is JupyterLab, followed by Android Studio as the second software platform.

3.5 Fire Localization and Detection

Once the model is implemented in the application, the user adds their email to which a message is sent when a fire is detected.

3.6 Signalling for Fire Detection

When a fire is detected, the user receives an email notification that a fire has been detected. This is useful in case a third party needs to be notified. In addition, the user receives an audible alert to their ear through the headset of the glasses.

4 Conclusions

The creating of prototype software for localization and recognition of fire for smart glasses with augmented reality is an attempt to prevent fire events that can lead to a fatal end using innovative computer technologies.

The use of neural networks and artificial intelligence makes the cyber-physical system flexible in decision-making and easily modified. The creation of such a cyber-physical system is an attempt to significantly improve the fight against fire disasters.

In the future, this cyber physical system is planned to be improved by increasing the accuracy of the model and sending the coordinates of the detected fire to the email address provided by the user.

Acknowledgements. This publication is developed with the support of Project BG05M2OP001-1.001-0004 UNITe, funded by the Operational Programme "Science and Education for Smart Growth", co-funded by the European Union trough the European Structural and Investment Funds.

References

1. Verstockt, S.: Methods and Techniques for Fire Detection. Academic Press, London (2016)
2. Epson Moverio. https://moverio.epson.com
3. Khalifa, N.E.M., Taha, M.H.N., Ezzat, Ali D., Slowik, A., Hassanien, A.E.: Artificial intelligence technique for gene expression by Tumor RNA-Seq data: a novel optimized deep learning approach. IEEE Access (2020). https://doi.org/10.1109/access.2020.2970210
4. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. IN: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788 (2016). https://doi.org/10.1109/CVPR.2016.91
5. Hochreiter, S., Schmidhuber, J.: LSTM can solve hard long time lag problems. In: Advances in Neural Information Processing Systems (1996)
6. Hochreiter, S.: Untersuchungen zu dynamischen neuronalen Netzen (PDF) (diploma thesis). Technical University Munich, Institute of Computer Science, advisor: J. Schmidhuber (1991)

Trends and Challenges in Surveillance - A Systematic Review of Camera Systems Implementing Artificial Intelligence

Boyana Ivanova, Kamelia Shoilekova(✉), and Rumen Rusev

Angel Kanchev University of Ruse, 8 Studentska Street, 7000 Ruse, Bulgaria
{bivanova,kshoylekova,rir}@uni-ruse.bg

Abstract. Among the many areas where Artificial Intelligence (AI) is bringing new opportunities for innovation and efficiency, one of the most exciting is the impact it is having on the video surveillance industry. Indeed, the rapid growth in the video surveillance camera market is being driven by smart systems and analytical software applications. The areas of AI that are used in surveillance are Machine learning, Deep learning, Neural networks, Natural language, and Expert systems.

The focus of this paper is put on the newest applications in the field of camera surveillance that are implementing AI. The purpose of this paper is to introduce briefly AI technology and how it is used in surveillance - the trends and challenges are identified, what are the advantages of the usage of AI, and how these kinds of applications increase the quality of surveillance and security.

Keywords: Surveillance · AI · challenges · technology · trends · camera

The Global Video Surveillance Systems Market is expected to grow by 9.3% over the forecast period. Introducing new IP-based digital technologies to detect and prevent undesirable behaviors such as shoplifting, theft, vandalism, and terror attacks are expected to fuel the growth of the video surveillance market. Video surveillance is used in various industries, including manufacturing, banking, financial services, transportation, and retail. Many applications, such as crime prevention, industrial process monitoring, and traffic management, are increasingly utilizing video surveillance systems.

Integrating advanced AI-powered solutions into video surveillance systems allows security staff to keep watch over the entire premises 24/7 and have better situational awareness without having to be glued to monitors all the time.

Deep learning is an AI-backed discipline by which computers learn through exposure to data and execute tasks such as identifying objects or recognizing an object throughout a video. In order for deep learning to be successful, you need vast amounts of data, which must be processed and annotated. This data is then used to train the network until it is able to repeat what it has been trained to do. The field of Deep learning is used in surveillance by the implementation of various algorithms for face recognition and person and object detection in a crowd. A National Institute of Standards and Technology (NIST) Interagency Report said that the improvements over the last 20 years of face recognition

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 103–112, 2024.
https://doi.org/10.1007/978-3-031-53549-9_11

error rates have decreased by three orders of magnitude. Most of today's top-performing commercial face recognition products are based on deep learning. The accuracy has reached 99.9% for controlled environments like airport immigration face recognition applications [1].

The use of Machine learning (ML) and Artificial Intelligence has improved the ability to search the hundreds or thousands of hours of video recorded by those systems. ML video content analysis software with deep learning can extract, classify, and quickly index targeted objects - such as humans or vehicles - making video feeds significantly more searchable, actionable, and quantifiable [19, 22].

Improvements in machine learning (ML) technology can both improve the efficiency of gleaning data from surveillance camera feeds, while also going a long way toward protecting the privacy of people who appear in those feeds. A smart camera can have the intelligence to know the difference between what it should be capturing and what it should ignore (Fig. 1).

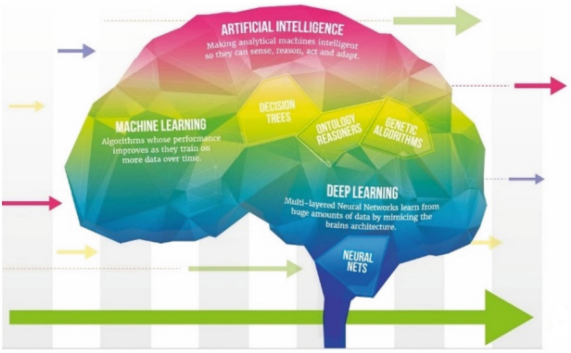


Fig. 1. Machine learning &Deep learning capability [21]

Another subfield of AI used in camera surveillance is Neural Networks (Fig. 2). Neural networks can deliver great advantages for video surveillance applications: from retail shops to "Safe City" solutions. For security systems market transition from assumptions based on mathematical analysis of the geometry and colour characteristics of a set of pixels to pattern recognition can be considered as a serious step in the situational analysis development. Analysis results clearly show people crowding, people on the rails, or trucks near the gate. Neural network technology can be effectively used for searching and analysing data during accident investigations. Also, you can sort objects by classes to save time and obtain more clear results.

Last but not least in surveillance used Expert systems (ES) technology enables the development of intelligent tools which perform real-time events analysis and correlation, detect and highlight unusual situations and suggest recovery actions. The power of expert systems stems primarily from the specific knowledge about a narrow domain stored in the expert system's knowledge base.

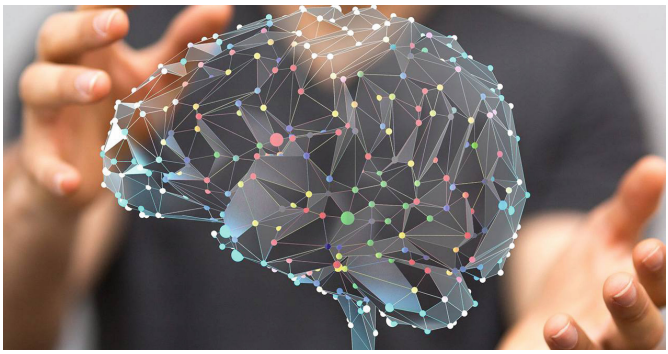


Fig. 2. Using neural networks for video surveillance [20]

1 The Coexistence Between Surveillance and Artificial Intelligence

Figure 3 shows how AI is implemented in camera surveillance and this is going to be the classification for the research presented in this paper. The next sections present applications, system prototypes, and research made from 2019 to 2023. All the subsets of AI work together when we are talking about surveillance. Machine learning uses neural networks and both of them are parts of Expert systems.

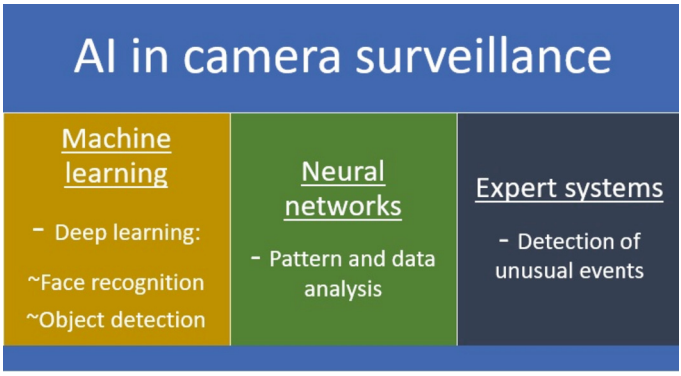


Fig. 3. Classification of camera surveillance using Artificial Intelligence

2 Camera Surveillance Using Machine Learning

Services through a reduction in resource consumption can be delivered in smart cities by using machine learning and artificial information. Machine learning models which are cloud-based enable resource-restricted devices to interconnect and optimize efficiency. This increased production of device designs that are targeted at reducing energy savings [2].

Research made in 2019, proposes a Distributed Intelligent Video Surveillance (DIVS) system using Deep Learning (DL) algorithms and deploying it in an edge computing environment. A multi-layer edge computing architecture is established and a distributed DL training model for the DIVS system is created. An approach for dynamic data migration is proposed to address the imbalance of workload and computational power of edge nodes. The experimental results had shown that the architecture can provide elastic and scalable computing power, and the proposed DIVS system can efficiently handle video surveillance and analysis tasks [3].

In 2020 are proposed two object detection algorithms such as You Only Look Once (YOLOv3 and YOLOv4) which are implemented for traffic and surveillance applications. YOLO model variants such as YOLOv3 are implemented for images and YOLOv4 for video datasets. The algorithm effectively detects the objects approximately with an accuracy of 98% for the image dataset and 99% for the video dataset according to the results of the experiments [4].

Some of the object recognition efforts are put on health, specially Covid-19 social distancing. In a paper published in 2021, a new multicore programming model developed and tested at the Al Zaytoonah University of Jordan is proposed. The multicore programming model implements data of real-time moving objects to provide an effective social distance monitoring solution. The YOLO algorithm, which employs a convolutional neural network is used for the detection process. This algorithm is already trained on the Common Objects in Context dataset which consists of 80 labels where the "PERSON" class is used to recognize humans from other objects in the scene. Tracking the mapped bounding boxes and determining the distance between moving/unmoving people are the stages that ensure no violation of the assumed physical distance. The proposed parallel social distancing system combined three techniques: object detection, object tracking, and distance measurement on a real-time base for multiple cameras. The results showed the high efficiency of this system in tracking people and determining the distances between them in accordance with public safety instructions [5].

A paper, presented in 2020, put focus on intelligent video retrieval technology. It integrates video processing, computer vision, and artificial intelligence, which greatly improves the efficiency of monitoring and the accuracy and linkage of the monitoring system. Based on deep learning theory and face detection neural network, this paper proposes a video-oriented cascaded intelligent face detection algorithm, which builds a deep learning network by cascading multiple features, from edge features, contour features, and local features to semantic features, and advances layer by layer. According to the last semantic feature, the information of the input data is obtained to accurately realize the face detection under the non-ideal condition. The algorithm resulted in good detection performance for single-face and multi-face images and it has strong robustness for rotating faces. Another advantage of the proposed algorithm is the speed - it is fast and can basically meet the requirements of real-time face detection [6].

A study regarding Transport infrastructure is implementing computer vision, neural networks, and object recognition in video surveillance cameras and laser scanners. The paper investigates a version of the monitoring system, which includes 2 main types of information received. The first implies constant monitoring of the object to determine its maximum permissible parameters using video cameras, information from which is

processed in real time, using computer vision technology. A striking example of the application of this technology is the use of cameras for fixing traffic violations. This kind of monitoring opens up great prospects for obtaining comprehensive, highly informative data on engineering structures, as well as their reliable automatic transformation into objective information without direct human participation [7].

Another application of machine learning is described in a multi-camera video surveillance system that automates video monitoring and minimizes human involvement in monitoring which can make significant impacts on the security industry. This system detects and recognizes a human target from videos taken from cameras mounted on the wall. The methodology applied is the basis for the functionality of the systems and reduces the difficulty for the end user of the system to operate the equipment. The the architectural design of the system is accurate enough to provide the user detailed information in different levels of the operation including motion detection capabilities, tracking systems, analysis system and decision-making processing system [8].

A paper presented in 2021, showed a smart video surveillance system executing AI algorithms in low-power consumption embedded devices. This application requires a distributed smart camera system. The proposed AI application allows the detection of people in the surveillance area using a MobileNet-SSD architecture. The algorithm can keep track of people in the video and also provide people with counting information by the usage of a robust Kalman filter bank. The selected architecture for the edge node is based on a UpSquared2 device that includes a vision processor unit (VPU) capable of accelerating the AI CNN inference. The results show the usefulness of deploying this smart camera node throughout a distributed surveillance system [9].

Machine learning and its basic conceptual of deep learning has a huge impact on surveillance. Some of the researchers put efforts to investigate the privacy and security aspects of implementing deep learning in camera surveillance.

A paper from 2021, highlights the many issues confronting the use of ML in the context of national security and mass surveillance, with a focus on the ethical challenges posed by ML. The first set of issues stems from the data used to train ML algorithms and the second set of issues concerns the efficacy of ML, without which it would be unethical to use. Due to the opacity of the features causing a particular ML decision, there are social issues surrounding human accountability and responsibility [10].

Because of the wide usage of facial recognition technology, the general public's worry about privacy increased significantly. These concerns have prompted numerous federal authorities to propose new legislation to protect personal privacy. In the described paper, it is proposed an accurate person localization scheme to enable law enforcement agencies to identify the locations visited by wanted and suspected people using surveillance cameras placed in various public places. Unlike the existing techniques that measure the Euclidean distance between two images to determine if they are for the same person, the proposed technique is based on a machine learning model to determine whether the features of two images belong to the same person. The performance evaluations showed excellent results and demonstrated that the model outperforms the Euclidean distance-based schemes when operating in a multi-source camera environment with changing recording quality [11].

3 Camera Surveillance Using Neural Networks

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of artificial neural networks (ANN), most commonly applied to analyze visual imagery. CNN are the most common neural networks used in camera surveillance.

One of the applications of CNN is street crime snatching and theft detection. The main target is to select features/objects which usually occurs at the time of snatching. The number of moving targets imitates the performance, speed, and amount of motion in the anomalous video. The dataset used in the paper, published in 2020, is Snatch 101; the videos in the dataset are further divided into frames. The frames are labeled and segmented for training. The VGG19 Convolutional Neural Network architecture algorithm is applied and it is extracted the features of objects and compared them with original video features and objects. The main contribution of the research is to create frames from the videos and then label the objects. The proposed system outperformed with 81% accuracy as compared to state-of-the-art systems such as VGG16 and AlexNet. A series of tests were carried out with the snatching theft videos that were not used in the training set. Another important aspect of the system is the processing time, which is better compared to the other systems with the same experimental setup. The experiment was carried out on 300 video frames of the same data set [12].

One of the goals of a proposed study in 2020, is a demonstration of the potential of C3D - a 3D CNN pre-trained with a large dataset of sport activities, as a feature extractor for violent scenes classification. The results have shown that the approach reaches high accuracy on both the Hockey Fight and the Crowd Violence datasets, scoring better than person-to-person fights, and in line with the best approach on crowd fights. The proposed method has a good generalization capability, being versatile and usable in different cases [13].

A novel computer vision framework for airport-airside surveillance is proposed, using cameras to monitor ground movement objects for safety enhancement and operational efficiency improvement. This is the first surveillance system that monitors runway to apron using CV models and the prototype is presented in 2022. The framework adopts Convolutional Neural Networks and camera calibration techniques for aircraft detection and tracking, push-back prediction, and manoeuvring monitoring. The proposed framework is applied on video camera feeds from Houston Airport, USA and Obihiro Airport, Japan. The object detection models of the proposed framework achieve up to 73.36% average precision on Houston airport and 87.3% on Obihiro airport. The framework estimates aircraft speed and distance with low error (up to 6 m), and aircraft push-back is predicted with an average error of 3 min from the time an aircraft arrives with the error-rate reducing until the aircraft's actual push-back event [14].

In 2021, it is presented a real-time video-based fire and smoke detection using YOLOv2 Convolutional Neural Network (CNN) in antifire surveillance systems. The training stage is processed off-line with indoor and outdoor fire and smoke image sets in different indoor and outdoor scenarios. Deep Neural Designer tool in MATLAB is used to build YOLOv2 neural network layers. The constructed CNN has 21 layers to establish a light-weight deep learning model to fit the embedded system. The proposed approach includes the input layer, middle layers, and subnetwork of YOLOv2 layers. The proposed solution achieved promising results [15].

A paper, published in 2020, investigates the performance of deep Convolutional Neural Network (CNN) for recognizing highway traffic congestion state in surveillance camera images. The purpose of the research is focus on the reliability of convolutional networks and their ability to classify such images, without any special previous processing such as segmentation of objective roads. In the paper are used AlexNet and GoogLeNet convolutional networks. It is built a highway imagery dataset using real-life traffic videos to evaluate the CNNs recognition performance. The results indicate that under the current strategy of feeding images directly into networks, both AlexNet and GoogLeNet can achieve an excellent recognition accuracy of 98% on held-out test samples. Many of the misclassified images turn out to be borderline cases. It can be concluded that scale and perspective in photography could affect the recognition result [16].

A holistic solution for Smart Home Security is implemented in 2021 which helps in improving privacy and security using two independent and emerging technologies of facial authentication and speech recognition. The developed application allows the user to monitor home through his mobile phone, tablet or PC.

The facial recognition is implemented by taking a real-time feed of the person at the door and analysis of the live feed is carried out where the face recognized is authenticated with the data of owners in the database, which matches the face to a name. Speech recognition is used as a two-factor authentication after the facial authentication. These recognitions are developed with neural networks. The overall accuracy of the proposed model is 82.71% with an accuracy of 87.5% for Facial Authentication and 84.62% for Speaker Authentication.

The innovation in the project is to identify faces through masks which will help to properly verify the identity of the person. The training dataset for the facial authentication model consists of 50 images each of 8 people. The testing dataset includes both masked and unmasked 8 images of 3 people.

The facial features are extracted from the live input and fed into the FaceNet model. In the case of masked facial recognition, the features from the upper region of the face like eyes and eyebrows are extracted and taken as input for the FaceNet model for authentication. The testing results showed that for small datasets, the proposed models are more efficient than the state-of-the-art models which require larger datasets for training. The proposed model reports a final accuracy of 82.71% for the entire Home Security system [17].

When we talk about smart cities, visual sensors are deployed almost everywhere, generating a massive amount of surveillance video data in smart cities that can be inspected intelligently to recognize anomalous events. In a proposed work in 2022, it is presented an efficient and robust framework for recognition of anomalies from surveillance Big Video Data (BVD) using Artificial Intelligence of Things (AIoT). It is used a two-stream neural network. The first stream comprises instant anomaly detection that is functional over resource-constrained IoT devices, whereas second phase is a two-stream deep neural network allowing for detailed anomaly analysis, suited to be deployed as a cloud computing service. Firstly, a self-pruned fine-tuned lightweight convolutional neural network (CNN) classifies the ongoing events as normal or anomalous in an AIoT environment. Upon anomaly detection, the edge device alerts the concerned departments and

transmits the anomalous frames to cloud analysis center for their detailed evaluation in the second phase. The cloud analysis center resorts to the proposed two-stream network, modelled from the integration of spatiotemporal and optical flow features through the sequential frames. Fused features flow through a bi-directional long short-term memory (BD-LSTM) layer, which classifies them into their respective anomaly classes, e.g., assault and abuse. It is reported a 9.88% and 4.01% increase in accuracy when compared to state-of-the-art methods evaluated over the aforementioned datasets [18].

4 Trends and Challenges in AI Surveillance

This research identified some of the trends in camera surveillance. Most methods and prototypes presented from 2019 to 2023 in camera surveillance put focus on developing real-time systems and terms such as Internet of Things (IoT) and Big Data are used more frequently. These are two of the most innovative technologies used in world. By using AI, it is revealed a new term "Intelligent Video Surveillance" - it sums up the usage of edge computing, deep learning, neural networks, big data infrastructure, IoT devices, cloud services.

As challenges could be identified: the price of the equipment/infrastructure, the automation of the process, and the adoption of AI camera surveillance - end users are not familiar with this type of algorithm. Also, all the challenges that are faced with all the above mention technologies are valid for camera surveillance, too. For example, Big data processing and analysis. The successful analysis of Big Data provides the executives with new information from external and internal sources that improves and speeds up the process of decision-making.

5 Advantages of AI Surveillance

AI-based tools such as facial recognition, crowd management, and object detection offer public and private security the following advantages: Prediction with evidence-based information; Boosts productivity; Discernment through patterns and similarities; Optimized workflows.

People are the priority of video surveillance, so they are put at the center of the usage of these technologies that consider the inclusion of multiple actors by providing safer and more pleasant environments.

AI video surveillance could deploy special attention in marginalized areas where crime is more prevalent.

6 Conclusion

As a result of the made review, it can be concluded that camera surveillance with implementing AI improved accuracy - video analytics-enabled surveillance systems analyse video footage in real-time and discover unusual activities; reduced expenses - with the high automation of processes; improvement of video search capabilities - the power of AI is making video footage searches as simple as using internet search engines.

Artificial Intelligence (AI) is the next evolution in security and surveillance. Technology has advanced in recent years and is expected to continue to improve in terms of reliability, accuracy, and power. Transit from surveillance into an intelligent security system is now capable of ensuring improved levels of safety, precision, and cost-effectiveness.

Acknowledgements. This publication is developed with the support of Project BG05M20P001-1.001-0004 UNITe, funded by the Operational Program "Science and Education for Smart Growth" co-funded by the European Union trough the European Structural and Investment Funds.

References

1. Sun, P.: Deep learning technology applications for video surveillance. Sourcesecurity. https://www.sourcesecurity.com/insights/deep-learning-technology-applications-video-surveillance-co-14319-ga.21460.html. Accessed 31 Jan 2023
2. Ahmed, S., Hossain, M.F., Kaiser, M., Noor, M.B.T., Mahmud, M., Chakraborty, C.: Artificial intelligence and machine learning for ensuring security in smart cities. In: Chakraborty, C., Lin, J.C.-W., Alazab, M. (eds.) Data-Driven Mining, Learning and Analytics for Secured Smart Cities. ASTSA, pp. 23–47. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72139-8_2
3. Chen, J., Li, K., Deng, Q., Li, K., Yu, P.S.: Distributed deep learning model for intelligent video surveillance systems with edge computing. IEEE Trans. Industr. Inf. (2020). https://doi.org/10.1109/TII.2019.2909473
4. Kumar, B.C., Punitha, R., Mohana: YOLOv3 and YOLOv4: multiple object detection for surveillance applications. In: 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, pp. 1316–1321 (2020). https://doi.org/10.1109/ICSSIT48917.2020.9214094
5. Alqatawneh, S., Jaber, K.M., Salah, M., Dalal, B.Y., Alqatawneh, O., Abulahoum, A.: Employing of object tracking system in public surveillance cameras to enforce quarantine and social distancing using parallel machine learning techniques. Int. J. Adv. Soft Comput. Appl. **13**(3), 170–180 (2021)
6. Dong, Z., Wei, J., Chen, X., Zheng, P.: Face detection in security monitoring based on artificial intelligence video retrieval technology. IEEE Access **8**, 63421–63433 (2020). https://doi.org/10.1109/ACCESS.2020.2982779
7. Gura, D., Markovskii, I., Khusht, N., Rak, I., Pshidatok, S.: A complex for monitoring transport infrastructure facilities based on video surveillance cameras and laser scanners. Transp. Res. Procedia **54**, 775–782 (2021). https://doi.org/10.1016/j.trpro.2021.02.130
8. Alshammari, A., Rawat, D.B.: Intelligent Multi-camera video surveillance system for smart city applications. In: IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), pp. 0317–0323 (2019). https://doi.org/10.1109/CCWC.2019.8666579
9. Cob-Parro, A.C., Losada-Gutiérrez, C., Marrón-Romera, M., Gardel-Vicente, A., Bravo-Muñoz, I.: Smart videosurveillance system based on edge computing. Sensors **21**(9), 2958 (2021). https://doi.org/10.3390/s21092958
10. Robbins, S.: Machine learning, mass surveillance, and national security: data, efficacy, and meaningful human control. In: The Palgrave Handbook of National Security. Palgrave Macmillan, Cham (2021). https://doi.org/10.1007/978-3-030-53494-3_16
11. Nabil, M., Sherif, A., Mahmoud, M., Alsmary, W., Alsabaan, M.: person localization using machine learning in multi-source camera surveillance system. In: SoutheastCon, pp. 110–116 (2022). https://doi.org/10.1109/SoutheastCon48659.2022.9763985

12. Butt, U., Letchmunan, S., Hafinaz F., Zia, S., Baqir A.: Detecting video surveillance using VGG19 convolutional neural networks. Int. J. Adv. Comput. Sci. Appl. **11**(2) (2020). https://doi.org/10.14569/IJACSA.2020.0110285
13. Accattoli, S., Sernani, P., Falcionelli, N., Mekuria, D.N., Dragoni, A.F.: Violence detection in videos by combining 3D convolutional networks and support vector machines. Appl. Artif. Intell. **34**(4), 329–344 (2020). https://doi.org/10.1080/08839514.2020.1723876
14. Thai, P., Alam, S., Lilith, N., Nguyen, B.T.: A computer vision framework using Convolutional Neural Networks for airport-airside surveillance. Transp. Res. Part C: Emerg. Technol. **137** (2022). https://doi.org/10.1016/j.trc.2022.103590
15. Saponara, S., Elhanashi, A., Gagliardi, A.: Real-time video fire/smoke detection based on CNN in antifire surveillance systems. J. Real-Time Image Proc. **18**, 889–900 (2021). https://doi.org/10.1007/s11554-020-01044-0
16. Cui, H., Yuan, G., Liu, N., Xu, M., Song, H.: Convolutional neural network for recognizing highway traffic congestion. J. Intell. Transp. Syst. **24**(3), 279–289 (2020). https://doi.org/10.1080/15472450.2020.1742121
17. Saxena, N., Varshney, D.: Smart home security solutions using facial authentication and speaker recognition through artificial neural networks. Int. J. Cogn. Comput. Eng. **2**, 154–164 (2021). https://doi.org/10.1016/j.ijcce.2021.10.001
18. Ullah, W., et al.: Artificial Intelligence of Things-assisted two-stream neural network for anomaly detection in surveillance Big Video Data. Futur. Gener. Comput. Syst. **129**, 286–297 (2022). https://doi.org/10.1016/j.future.2021.10.033
19. Krastev, G., Voinohovska, V.: Application of Nonhierarchical Cluster Analysis in the Machine Learning 4th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, HORA 2022 (2022). https://doi.org/10.1109/HORA55278.2022.9799959
20. https://www.videonet9.com/using-neural-networks-for-video-surveillance.html
21. https://memoori.com/portfolio/global-market-video-analytics-2018/
22. Petrova, M., Atanasova, D.: Training through the implementation of a web-based laboratory for Artificial Intelligence exercises 29th Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEEIE) (Scopus), pp. 1–4 (2019). https://doi.org/10.1109/EAEEIE46886.2019.9000417

Usage of the Summary Model DELIS-CH for Starting the Design Process of an Educational Video Game for Cultural Heritage

Yavor Dankov[1](✉) and Andjela Dankova[2]

[1] Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohdridski", Sofia, Bulgaria
yavor.dankov@fmi.uni-sofia.bg

[2] Faculty of Architecture, University of Architecture, Civil Engineering and Geodesy, Sofia, Bulgaria
andjela.dankova@gmail.com

Abstract. Increasing the popularity of using educational video games in the educational process, the desire of more people to design and create educational video games in the Cultural Heritage (CH) domain for their students is rising. Often, the designers are educators who are not specialists in IT or CH and are not familiar with the design and creation process of educational video games in detail. This paper is based on the use of the **Summary Model DELIS-CH**. The purpose of the model is to support designers who are not specialists in IT and CH to design educational video games in the domain of tangible CH. Using the **Summary Model DELIS-CH** and following the steps of designing educational video games for CH presented in the model, this paper describes the initial stages of starting the design process of an example educational video dedicated to an object of tangible CH. As a result, the paper presents **a conceptual design model of the content** of the educational video game for the Boyana Church located in Sofia, Bulgaria, representing a tangible CH object of world importance. The Boyana Church is on the UNESCO list of CH sites of world significance. The application of the model provides the opportunity for designers to use and follow the specialized *Design Recommendations* (included in the model), which support designers in choosing the appropriate approach for searching and selecting relevant educational and gaming content according to the importance of the CH object.

Keywords: Serious Games · Game-Based Learning · Summary Model DELIS-CH · Cultural Heritage · Architecture · UNESCO

1 Introduction

The design and creation of educational video games are gaining in popularity precisely because of the benefits of using video games in the educational process of learners of different ages and fields of learning [1–3]. Educational video games apply game-based learning, which provides an interactive way of presenting the necessary knowledge and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 113–120, 2024.
https://doi.org/10.1007/978-3-031-53549-9_12

skills for learning by the learners [3]. Educational video games are, in their essence, serious video games that have been applied to the field of education. There are also various examples of the application of serious video games in diverse areas such as security, industry, education, architecture, Cultural Heritage, geoscience, etc. [4–11]. Therefore, the paper is focused on educational video games.

In essence, educational video games integrate the application of modern interactive methods of interacting with learners based on the application of game-based learning [1, 2, 12]. The benefits of game-based learning for learners are generally associated with an increased degree of adoption of the learning material, increased engagement of learners to the educational video game, as well as the increasing success of learners through playing the educational video game, and many others [12–14]. In this regard, the design and creation of educational video games in different fields of education require a variety of specialists familiar with these processes. Professionals with expertise in the educational field and knowledge of the overall game design process are necessary to create educational video games. Software development and creating an educational video game often require information technology (IT) specialists to develop (encode) the designed educational video game and generate an executable file that learners can use (play). One person can possess these skills, but often the educational video game design and creation teams consist of many members, each of whom has its professional position and skill and participates in the overall process of the design and creation of educational video games. In this paper, the focus is established on educational video games in the field of tangible Cultural Heritage (CH).

Increasing the popularity of using educational video games in the educational process, the desire of more people to design and create educational video games for their students is rising. Often, the designers are educators who are not specialists in IT or CH and are not familiar with the design and creation process of educational video games in detail. In this case, it is possible to use specialized platforms to automatically generate educational video games and specialized instruments that support these processes, such as the APOGEE software platform [15–19]. In the domain of tangible CH, a possible solution for dealing with the challenges in designing educational video games by educators provides the "Summary Model of the **D**esign proc**E**ss for educationa**L** v**I**deo game**S** for **CH**" [20], named the **Summary Model DELIS-CH**. The model aims to support designers (in particular, educators and individuals who are not specialists in IT and CH) to design educational video games in the tangible CH domain. The tangible CH refers to objects such as monuments, buildings, artifacts, etc., with significant value and importance for the preservation of future generations [21, 22]. The *Summary Model DELIS-CH* is presented in detail in a prior study by the authors of this paper [20]. The model focuses on the CH objects' importance and describes the essential and important steps in the design process of educational video games [20].

This paper is based on the use of the *Summary Model DELIS-CH* and the *Classification of Educational Video Games for CH* [22], introduced in earlier works by the authors of this paper [20, 22]. The classification consists of three different categories of educational video games for CH, depending on the defined significance of the CH object. The classification and the *Summary Model DELIS-CH* are closely interconnected, and it is necessary to utilize them together to design and create educational video games for tangible CH. Using the **Summary Model DELIS-CH** and following the steps of

designing educational video games for CH presented in the model, this paper presents the description of the initial stages of starting the process of designing an example educational video game dedicated to an object of tangible CH from designers who are not specialists in IT and CH. As a result, the paper presents a **conceptual design model of the content** of the educational video game for the object of tangible CH - the Boyana Church. The Boyana Church is located in Sofia, Bulgaria, representing a tangible CH object of world importance. The Boyana Church is on the UNESCO list of CH sites of world significance [23].

Using the *Summary Model DELIS-CH* supports educators and individuals who are not specialists in IT and CH to start designing educational video games dedicated to objects of tangible CH. The application of the model provides the opportunity for designers to use and follow the specialized *Design Recommendations* [22] (included in the model), which support designers in choosing the appropriate approach for searching and selecting relevant educational and gaming content according to the importance of the CH object.

The paper continues with the following sections: Section 2 presents the application of *Summary Model DELIS-CH* for starting the initial stages of designing an example educational video game dedicated to the tangible CH object of global importance - the Boyana Church, located in Sofia, Bulgaria. The section presents the model and the stages for starting the initial design process of the game by designers who are not specialists in IT and CH. Section 3 describes the conceptual design model of the content of the educational video game for the Boyana Church. The paper ends with a conclusion.

2 Application of the Summary Model DELIS-CH

The current section presents the usage of the *Summary Model DELIS-CH* to design an exemplary educational video game dedicated to a tangible CH object. Figure 1 illustrates a view of the *Summary Model DELIS-CH* based on [20]. The presented view of the model is focused on the initial and foundational essential steps required to start the design process of an educational video game of a CH object, as well as the appropriate and correct definition of the CH object, according to the Classification of Educational Video Games for CH [20].

This paper presents the initiation of an educational video game design for CH by designers who are not specialists in IT and CH. Following the model, the designer's first step is to start the process through the clear desire to make an educational video game about a tangible CH. Following the *Summary Model DELIS-CH*, the designer must determine the educational discipline for which the game will be intended. The educational domain (discipline) is fundamental to starting the process and moving through the following stages described in the model. Upon successful discipline definition, the process can continue with the next stages presented in the model.

This paper presents the initial stage of the design process of an exemplary educational video game for the discipline of History intended for teaching children in primary education. Once the educational discipline (domain) is defined, the next critical stage of the model is the determination of the objectives, users, and related requirements [20]. The game's primary goal will be to familiarize learners with the tangible CH object. Users of the game will be primary education children between 8 and 11 years old.

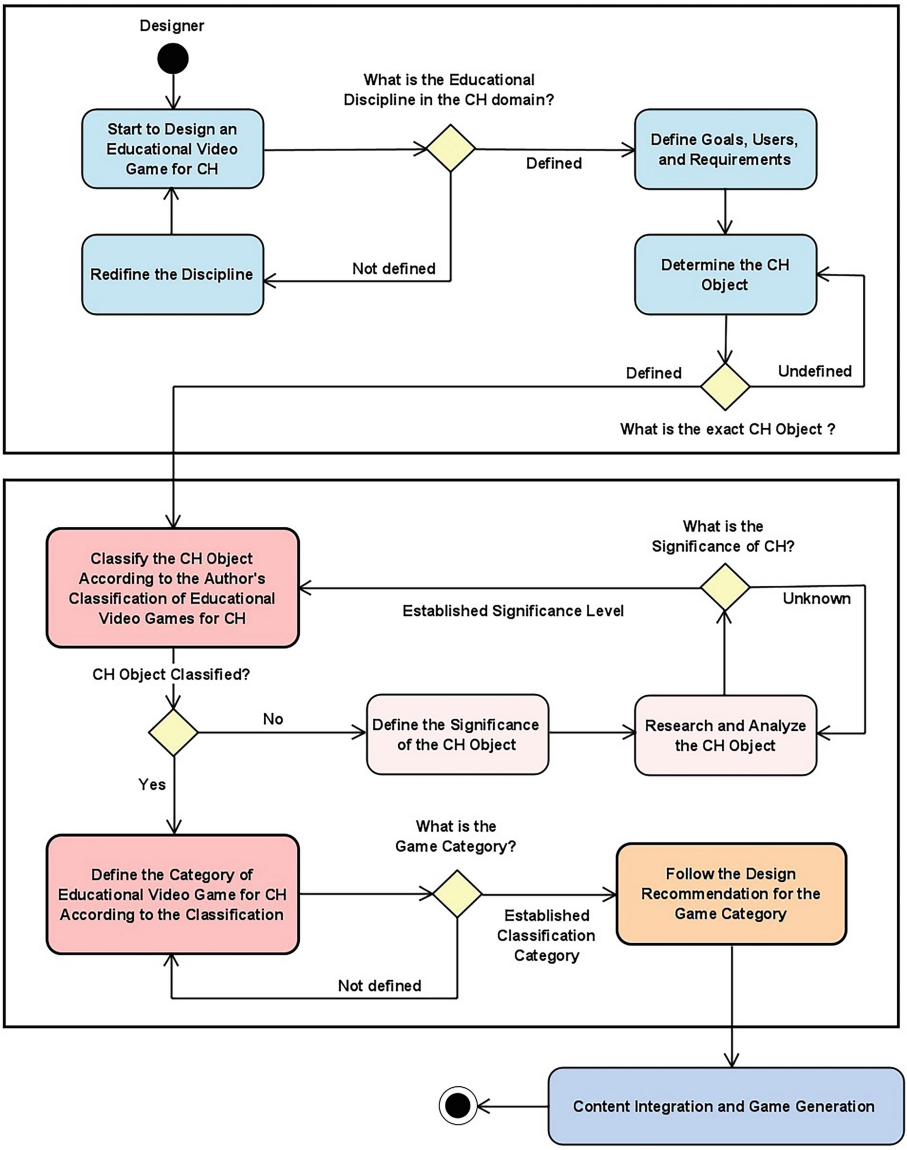


Fig. 1. View of The Summary Model DELIS-CH, based on [20]

The designer must determine specific game constraints depending on the learning class or the learner's characteristics. For the current study, the educational video game will not have strictly defined requirements for individual users or school classes. This enables future testing of a developed game prototype to analyze and evaluate the game design and, if necessary, to go through the game design stages again and improve individual components or implement personalization of the learning and gaming content

integrated into the game. The restrictions set in the example meet the age requirements of the learners (between 8 and 11 years of age), and these users are primary education learners in the respective schools.

After the educational discipline, users, and goals of the game have been defined, designers - educators who are not specialists in IT or CH, must make a specific choice of a given object (site) of tangible CH that is suitable for the requirements set in the previous stages presented in the *Summary Model DELIS-CH*. Designers can use the *Classification of Educational Video Games for CH* [22] to achieve this goal. Depending on the category and significance of the CH object (site), designers can use specialized *Design Recommendations* (included in the model) to integrate the appropriate learning and gaming content into the game. A designer must first classify the tangible CH object (site) for which an educational video game will be designed and developed. In this regard, the designer must consistently investigate the significance of the selected CH object (site) and classify it according to the *Classification of Educational Video Games for CH* [22]. Based on the consistent observance of the stages presented in the *Summary Model DELIS-CH* and the defined category of educational video games, according to the *Classification of Educational Video Games for CH* [22], the designer can use the specialized *Design Recommendations* [20, 22] and efficiently and concretely to determine the essential learning and gaming content to be integrated into the educational video game for the realization of the set goals.

This paper presents the initial design of an educational video game for the Boyana Church, located in Sofia, Bulgaria, representing a tangible CH site of national importance, according to the national legislation of Bulgaria. In addition, the significance of the Boyana Church is recognized globally, as the site is included in the UNESCO list [23] of CH sites of world importance. This testifies to the global significance of the site and its value for the history of the world. Since the Boyana church represents a tangible CH object of international importance, this object is classified in the first category of the *Classification of Educational Video Games for CH* [22]. Therefore, the educational video game that will be designed corresponds to the characteristics of *Category 1: Educational Video Games for CH of Global Significance* [22]. The designer can use specialized *Design Recommendations* [20, 22] to determine and select the appropriate learning and gaming content in the following stages of the *Summary Model DELIS-CH*.

3 Conceptual Design Model of the Content of the Educational Video Game for Boyana Church

Following the successful completion of the definition of the importance of the CH object and the definition of the category of educational video game for CH, according to the *Classification of Educational Video Games for CH* [22], **an example conceptual design model of the content of the educational video game for Boyana Church** is presented in this section. The model is illustrated in Fig. 2.

As the presented example falls into *Category 1: Educational Video Games for CH of Global Significance* of the *Classification of Educational Video Games for CH* [22], a variety of information is available from various literature sources [24–26], including internet sources such as the UNESCO website and the list of UNESCO Cultural Heritage

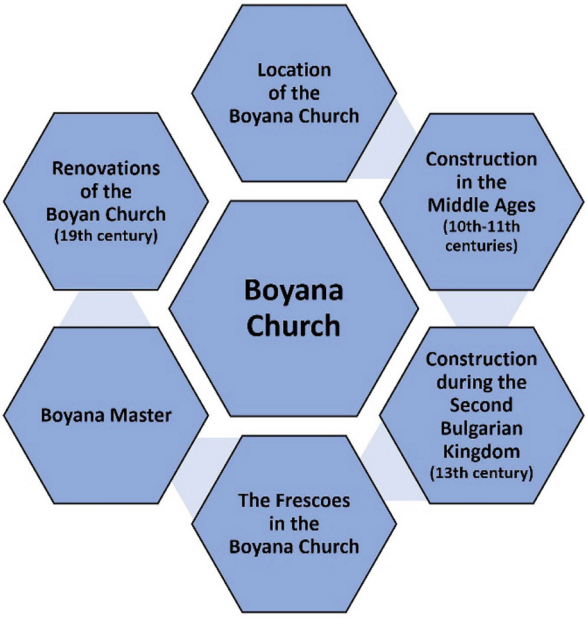


Fig. 2. Conceptual Design Model of the Content of the educational video game for Boyana Church

sites [23]. After an in-depth study of various sources of information, the present paper defines the essential and required topics of the educational video game's learning content for the Boyana Church, which, according to the authors, would be suitable for the defined target group and the set goals.

The model presents six main topics of learning content that designers can integrate into the educational video game dedicated to the Boyana Church for the discipline of History studied in schools by primary education learners between 8 and 11 years old. The conceptual design model of the educational content of the educational video game dedicated to the Boyana church includes the following topics:

- Location of the Boyana Church
- Construction of the Boyana Church in the Middle Ages (10th-11th centuries)
- Construction of the Boyana Church during the Second Bulgarian Kingdom (13th century)
- The frescoes in the Boyana Church
- Boyana Master
- Renovations of the Boyana Church (19th century)

Depending on the goals set, the game users defined, and the constraints imposed, designers can reach varying degrees of detail on each topic in the model.

4 Conclusion

This paper has presented the use of the *Summary Model DELIS-CH* to start the initial steps of the design process of an example educational video game dedicated to a tangible CH object by designers who are not specialists in IT and CH. As a result, the paper presented a conceptual design model of the learning content of the educational video game for the CH object of global importance - the Boyana Church, located in Sofia, Bulgaria. Thanks to the *Summary Model DELIS-CH* application and the presentation of this particular example, designers can better understand the model and its application and benefits for future work on creating educational video games in the CH domain.

Acknowledgements. This research is supported by Project BG05M2P001-1.001-0004 "Universities for Science, Informatics and Technologies in the e-Society (UNITe)" financed by Operational Program "Science and Education for Smart Growth", co-financed by the European Regional Development Fund.

References

1. Pan, L., et al.: How to implement game-based learning in a smart classroom? A model based on a systematic literature review and Delphi method. Front. Psychol. **12**(749837) (2021). https://doi.org/10.3389/fpsyg.2021.749837
2. Alshar'E, M., Albadi, A., Jawarneh, M., Tahir, N., al Amri, M.: Usability evaluation of educational games: an analysis of culture as a factor affecting children's educational attainment. Adv. Hum.-Comput. Interact. (2022). https://doi.org/10.1155/2022/9427405
3. Pozo, J.I., Cabellos, B., Sánchez, D.L.: Do teachers believe that video games can improve learning? Heliyon **8**(6), e09798 (2022). https://doi.org/10.1016/J.HELIYON.2022.E09798
4. Duin, H., Hauge, J.B., Hunecker, F., Thoben, K.: Application of serious games in industrial contexts. In: Cruz-Cunha, M., Varvalho, V., Tavares, P. (eds.) Business, Technological, and Social Dimensions of Computer Games: Multidisciplinary Developments, pp. 331–347. Information Science Reference, Hershey, PA (2011). https://doi.org/10.4018/978-1-60960-567-4.ch020
5. Pourabdollahian, B., Taisch, M., Kerga, E.: Serious games in manufacturing education: evaluation of learners' engagement. Procedia Comput. Sci. **15**, 256–265 (2012). https://doi.org/10.1016/J.PROCS.2012.10.077
6. Messaadia, M., Bufardi, A., Le Duigou, J., Szigeti, H., Eynard, B., Kiritsis, D.: Applying serious games in lean manufacturing training. In: Emmanouilidis, C., Taisch, M., Kiritsis, D. (eds.) APMS 2012. IAICT, vol. 397, pp. 558–565. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40352-1_70
7. Bontchev, B.: Serious games for and as cultural heritage. Digit. Present. Preserv. Cult. Sci. Herit. **5**, 43–58 (2015). https://doi.org/10.55630/dipp.2015.5.3
8. Riedel, J., Feng, Y., Hauge, J., Hansen, P., Tasuya, S.: The adoption and application of serious games in corporate training - the case of manufacturing. In: 2015 IEEE International Conference on Engineering, Technology and Innovation/International Technology Management Conference (ICE/ITMC) (2016). https://doi.org/10.1109/ICE.2015.7438684
9. Fonseca, D., et al.: Mixed assessment of virtual serious games applied in architectural and urban design education. Sensors **2021**(21), 3102 (2021). https://doi.org/10.3390/s21093102

10. Capecchi, I., et al.: The combination of serious gaming and immersive virtual reality through the constructivist approach: an application to teaching architecture. Educ. Sci. **2022**(12), 536 (2022). https://doi.org/10.3390/educsci12080536
11. McGowan, E.G., Alcott, L.J.: The potential for using video games to teach geoscience: learning about the geology and geomorphology of Hokkaido (Japan) from playing Pokémon Legends: Arceus. Geosci. Commun. **5**, 325–337 (2022). https://doi.org/10.5194/gc-5-325-2022
12. Camacho-Sánchez, R., Rillo-Albert, A., Lavega-Burgués, P.: Gamified digital game-based learning as a pedagogical strategy: student academic performance and motivation. Appl. Sci. **12**, 11214 (2022). https://doi.org/10.3390/app122111214
13. Tavares, N.: The use and impact of game-based learning on the learning experience and knowledge retention of nursing undergraduate students: a systematic literature review. Nurse Educ. Today J. **117**, 105484 (2022). https://doi.org/10.1016/j.nedt.2022.105484
14. Behnamnia, N., Kamsin, A., Ismail, M., Hayati, S.: A review of using digital game-based learning for preschoolers. J. Comput. Educ. 1–34 (2022). https://doi.org/10.1007/S40692-022-00240-0/TABLES/15
15. Bontchev, B., Vassileva, D., Dankov, Y.: The APOGEE software platform for construction of rich maze video games for education. In: Proceedings of the 14th International Conference on Software Technologies – ICSOFT 2019, pp. 491–498. SciTePress (2019). https://doi.org/10.5220/0007930404910498
16. Dankov, Y., Bontchev, B.: Software instruments for management of the design of educational video games. In: Ahram, T., Taiar, R., Groff, F. (eds.) IHIET-AI 2021. AISC, vol. 1378, pp. 414–421. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-74009-2_53
17. Dankov, Y., Bontchev, B.: Designing software instruments for analysis and visualization of data relevant to playing educational video games. In: Ahram, T., Taiar, R., Groff, F. (eds.) IHIET-AI 2021. AISC, vol. 1378, pp. 422–429. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-74009-2_54
18. Dankov, Y., Bontchev, B., Terzieva, V.: Design and Creation of Educational Video Games Using Assistive Software Instruments. In: Ahram, T.Z., Karwowski, W., Kalra, J. (eds.) AHFE 2021. LNNS, vol. 271, pp. 341–349. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-80624-8_42
19. APOGEE Project site. https://apogee.online/. Accessed 15 June 2023
20. Dankov, Y.: The design process of educational video games in cultural heritage. Digit. Present. Preserv. Cult. Sci. Herit. **13**, 229–238 (2023). https://doi.org/10.55630/dipp.2023.13.22.
21. Paolis, L., Chiarello, S., Gatto, C., Liaci, S., Luca, V.: Virtual reality for the enhancement of cultural tangible and intangible heritage: the case study of the Castle of Corsano. Digit. Appl. Archaeol. Cult. Herit **27**, e00238 (2022). https://doi.org/10.1016/j.daach.2022.e00238
22. Dankov, Y., Dankova, A.: Educational video games as tools for raising awareness of the protection and preservation of cultural heritage. Digit. Present. Preserv. Cult. Sci. Herit. **13**, 219–228 (2023). https://doi.org/10.55630/dipp.2023.13.21.
23. UNESCO: World Heritage List (2023). https://whc.unesco.org/en/list/. Accessed 15 March 2023
24. Dimitrov, D.P., et al. (eds.): KratkaIstoriya na Balgarskata Arhitektura [A Brief History of Bulgarian Architecture]. Bulgarian Academy of Sciences (BAS), Sofia, Bulgaria (1965)
25. Mavrodinov, N.: Starobalgarskoto Izkustvo: XI-XIII v. [Old Bulgarian Art: XI-XIII c.]. Balgarski Hudozhnik, Sofia, Bulgaria (1966)
26. Stancheva, M.M. (ed.): World Heritage of Bulgaria, vol. 978-954-322-372-5. Prof. Marin Drinov Academic Publishing House (2010)

User-Oriented Dashboard Design Process for the DIZU-EVG Instrument for Visualizing Results from Educational Video Games

Yavor Dankov(✉)

Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohdridski", Sofia, Bulgaria
yavor.dankov@fmi.uni-sofia.bg

Abstract. In the modern society of technological advancement, educational video games as a means of interacting with the younger generation are gaining a larger scale of application. Educational video games aim to educate users of specific knowledge in an entertaining way. This paper presents the **initial phase of designing** a **Dashboard for learners to visualize learning results** from educational video games and a **Dashboard for players to visualize gaming results**. Visualizing Dashboards represents one of the main functionalities of the **DIZU-EVG** (**D**ata visual**IZ**ation instr**U**ment for **E**ducational **V**ideo **G**ames) instrument for visualizing learning and gaming results from educational video games. This paper presents the **primary requirements for the user-oriented dashboard design,** goals and restrictions**,** and **the core functionalities** that each Dashboard provides to the relevant users – players, and learners of educational maze video games generated through the APOGEE software platform. The paper presents **formal models of using the core functionalities** of the individual Dashboards of the DIZU-EVG instrument. The user-oriented (players and trainees) dashboard design, through which the gaming and learning results of the users are visualized within the DIZU-EVG instrument, will contribute to the development of the instrument and its improvement. These functionalities will increase the DIZU-EVG instrument's efficiency and its users' benefits. The ability to visualize the appropriate data of the respective users (players or learners) will provide opportunities for timely analysis of the results achieved by users and for analyzing and evaluating the game's design.

Keywords: Educational video games · Serious games · Game-based learning · DIZU-EVG instrument · Data Visualization · User-oriented design

1 Introduction

In the modern society of technological advancement, educational video games as a means of interacting with the younger generation are gaining a larger scale of application [1–5]. In essence, educational video games aim to educate users of specific knowledge in an entertainment way [6]. These games are not only intended for fun but also for learning [7].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 121–129, 2024.
https://doi.org/10.1007/978-3-031-53549-9_13

Games that entertain consumers are a permanent part of the daily routine of young people [8]. The individual character of the games provokes players to achieve higher results and go further in the game's history, pass the upper play levels, solve more complex strategic puzzles, etc. Achieving higher individual results in terms of games that have gameplay with levels, for example, and specific requirements for passing a level (such as minimum time to pass to the next level, for example), make players replay the level to reach the achievement of shorter time to pass the given level. Different games and their gameplay correlate with the variety of different genres of games, and for this reason, they are not subject to research in this paper. The individual character of the games is undoubtedly interconnected with the element of competition (collective nature) of the users of the relevant game, who have also played it and have achieved certain specific results. The ability to compete to achieve better results in a game between different users is one of the essential factors for achieving high financial profits in the game industry and increased indicators of gaming user experience and a growing number of users [9–11].

In this regard, the visualization of gaming and learning results to users who have played an educational video game is essential for realizing satisfactory and appropriate designs of such games for the respective target group of learners and players. Designing educational video games that provide users with suitable learning and gaming content in a virtual environment is necessary [12]. In addition, these games must be visually attractive and offer a high degree of gaming and learning user experience [6, 13, 14]. Therefore, it is essential to use the appropriate tools to provide the opportunity to visualize the results achieved by users after the game sessions. In this paper, the focus is placed on **DIZU-EVG (D**ata visual**IZ**ation instr**U**ment for **E**ducational **V**ideo **G**ames) [15], which provides the functionality for the visualization of learning and gaming results from educational maze video games generated within the APOGEE software platform [15]. The application of the DIZU-EVG tool develops within the APOGEE platform as part of various software tools that support designers in designing, analyzing, and creating educational video games generated on the platform and played by the respective users [16, 17]. The DIZU-EVG instrument studies have been presented in the previous publications of the author of this paper [15, 18]. Generally, the DIZU-EVG tool provides a wide range of functionalities that can be used by players and learners, including designers and creators of educational video games [15]. Among the essential functionalities of the DIZU-EVG instrument is the functionality that provides customized, according to the user, dashboards for visualization of gaming and learning results.

Therefore, this paper presents the **initial phase of designing** a **Dashboard for learners to visualize learning results** from educational video games and a **Dashboard for players to visualize gaming results**. Visualizing Dashboards represents one of the main functionalities of the DIZU-EVG instrument. This paper presents the **primary requirements for the user-oriented dashboard design,** goals and restrictions, and **the core functionalities** that each Dashboard provides to the relevant users – players, and learners of educational maze video games generated through the APOGEE software platform. The paper presents **formal models of using the core functionalities** of the individual Dashboards of the DIZU-EVG instrument.

The paper continues with Sect. 2, presenting the general requirements for the user-oriented design of the dashboards for players and learners of the DIZU-EVG instrument.

Section 3 introduces the designed core functionalities of the Dashboards and the formal models of using these functionalities. The paper ends with the conclusion and future work.

2 User-Oriented Requirements for the Design of the DIZU-EVG Instrument Dashboards

Designing and creating educational video games frequently requires a specialized software platform that provides diverse software tools supporting designers in these processes. The DIZU-EVG tool aims to give users interactive methods for visualizing gaming and learning results, using dashboards that can be used by learners, players, and designers of educational video games [18].

The main objective of the Dashboards of the DIZU-EVG tool is to visualize the data from the game sessions of players and learners in a way that is appropriate for the relevant users and understandable to them [18]. The visual presentation of the results of each player or learner is among the factors that contribute to the achievement of good educational results (by the learners in the game) and the increase in the learners' success. The dualistic role of users is that a user of an educational video game can only play the game for pleasure without perceiving and exercising (for example, through mini-games) the didactic content integrated into the game. On the other hand, the player can also be perceived as a learner who, when playing the game, has fun and purposefully reads, learns, and exercises the integrated didactic content into the entire educational video game.

With the help of appropriate tools that enable the visualization of the results achieved by the users, the analysis and evaluation of the game's design by the designers of educational video games are possible. This also contributes to the timely perception of what players or learners have achieved during and after game sessions. Visualizing the results of users will give the appropriate feedback to the respective learner or player for the positive or negative results he has achieved. Therefore, this will demonstrate whether the learner has learned the educational content in the game and whether he has learned this knowledge through playing an educational video game. In addition, this will give a prerequisite for analyzing whether the learner, while playing the game, was intrigued and engaged by the game (user gaming experience).

Generally, the primary requirements for the user-oriented design of the Dashboards of the DIZU-EVG instrument can be divided into two directions:

- **Primary user-oriented design requirements** for **Dashboards for visualizing learning results** from the **learners'** game sessions.
- **Primary user-oriented design requirements** for **Dashboards for visualizing playing results** from the **players'** game sessions.

To design a dashboard for learners, this dashboard needs to visualize the data related to the learning results achieved in the game. These are the data from the playing sessions that are tied to the integrated learning content in the game and show the degree of the learning material perceived. For example, these data can be generated when the learners play certain mini-games incorporated in an educational maze video game developed

within the APOGEE platform [17]. Mini-games are different types of small games integrated into the educational video game the learner can play. Among the principal goals of the mini-games is to test or review the degree of perception of the learning material by the learner after playing the mini-games. An example of such educational video games (with mini-games included) that have been generated within the APOGEE software platform is "Valchan Voivoda" [19] and "Let's Save Venice" [20]. With the accomplishment of the tasks in the game, the learner comprehends and learns the didactic content of the game and plays mini-games for which a certain number of points are obtained and measured time for the achieved result. The goal of the DIZU-EVG instrument and the Dashboard for visualizing the learning results is to show the learners, in an interactive and accessible manner, the learning results achieved for a specific user who has played the relevant educational video game [18]. Depending on the educational video game and its various mini-games, the Dashboard for learners should visualize information about success, playing time, accurate results, and the possibility of comparability with other learners who have played the same video game. This should allow the comparability of the achieved learning results of a user with the achieved learning results of all other learners who have played the same educational video game and mini-games.

Similarly, the primary user-oriented requirements for the dashboard design for the visualization of players' game results are directly related to the generated data from the player's behavior during the game, but without considering the results achieved related to didactic content in the game. In the example of educational video games generated within the APOGEE platform, the games developed contain mini-games intended to exercise and verify didactic content throughout the game. The mini-games can be optional to play by the players and can be missed and not played. The player can play (repeatedly) the educational video game for the sole purpose of entertaining by deciding to skip the learning content and to play the whole game and the mini-games included in it that are not intended for the exercise and verification of didactic content. The DIZU-EVG instrument and the Dashboard for the players must visualize the results from the game sessions in a way that enables players to visually perceive their success or failure from their educational video game session and take action to improve their results. In addition, the dashboard must provide the opportunity to visualize summary data from the game sessions of the other players who played it. Therefore, the dashboard should visualize the ranking of the gameplay results of all the players who have played a particular educational video game.

It is important to note that the DIZU-EVG instrument can also be integrated into other software platforms that are applied to create educational video games that use additional tools that increase the platform's capacity with diverse functionalities. The DIZU-EVG tool is used within the APOGEE platform [15]. The software platform is intended for designing and creating educational maze video games with mini-games included. The APOGEE software platform is also enriched with numerous assistive and analytics tools, each supporting the complete process of designing and creating educational video games [17]. In this regard, depending on the target group of learners and players for which the game will be developed, the Dashboards of the DIZU-EVG instrument can be further customized, depending on the defined personalization characteristics such as a learning style or gaming style, age, preferences, educational discipline or class, etc. [21]. This

provides a future opportunity for Dashboard personalization, which is the subject of future research. This can also be defined as one of the restrictions placed in the design of the Dashboard of the DIZU-EVG instrument.

3 Core Functionalities and Models

Based on the primary user-oriented dashboard design requirements and the previous studies of the DIZU-EVG instrument [17, 18], the **core functionalities** and **formal models of using the core functionalities of the individual Dashboards** of the DIZU-EVG instrument are presented in the current section.

The core functionality of the *Learner Dashboard* for visualization of learning results from the learners' game sessions and the process of their use is illustrated in the model of Fig. 1. The following process model outlines how to utilize the Dashboard's essential features for learners based on [17, 18]. An example of the application of the DIZU-EVG instrument is the APOGEE software platform and its profile system. In short, the process starts with the learner's entry in his profile with the designated role of a learner in the system. For this purpose, the DIZU-EVG tool must check that the user has an account created in the APOGEE platform [18]. If the user is not registered, the process proceeds with creating an account for a learner in the system. Once the learner has entered his learner profile, this user can use the DIZU-EVG tool's functionality to visualize data concerning the learning results from an educational video game [18]. Therefore, this is done by the functionality of the DIZU-EVG instrument for visualizing the *Learner Dashboard* [15]. The user can choose to use three core *Learner Dashboard* functionalities. These functionalities can visualize the learner's learning results after playing an educational video game. The data that the *Learner Dashboard* can show to the learners is related to the individually achieved results of the user (logged in to their profile with a role as a learner) who played educational video games, including the integrated mini-games intended for exercise and verification of didactic content in the whole game. The DIZU-EVG *Learner Dashboard* also provides a feature for visualizing the results of all other learners who have achieved some learning results after playing a specific educational video game [17, 18]. This enables the learner to visually analyze their success/failure compared to the other learners. Depending on the requirements of the particular target group of learners, the *Learner Dashboard* provides functionality for adjusting its interface.

The presented process model of Fig. 2 illustrates the core functionalities of the *Player Dashboard* for visualizing gaming results from players' game sessions and the process of using them. The process model outlines how to utilize the *Players Dashboard's* essential features for players based on[17, 18]. Again, an example of the application of the DIZU-EVG instrument is the APOGEE software platform and its profile system.

In the illustrated process model of Fig. 2, the user of the DIZU-EVG instrument is the player who necessarily begins with entering a player's profile in the APOGEE platform. The player enters his unique profile with a player's role and can use the DIZU-EVG tool's functionality to visualize data about gaming results from playing sessions [18]. At this stage, it is essential to note that the player can only use the functionality of the DIZU-EVG tool to visualize the *Player Dashboard*.

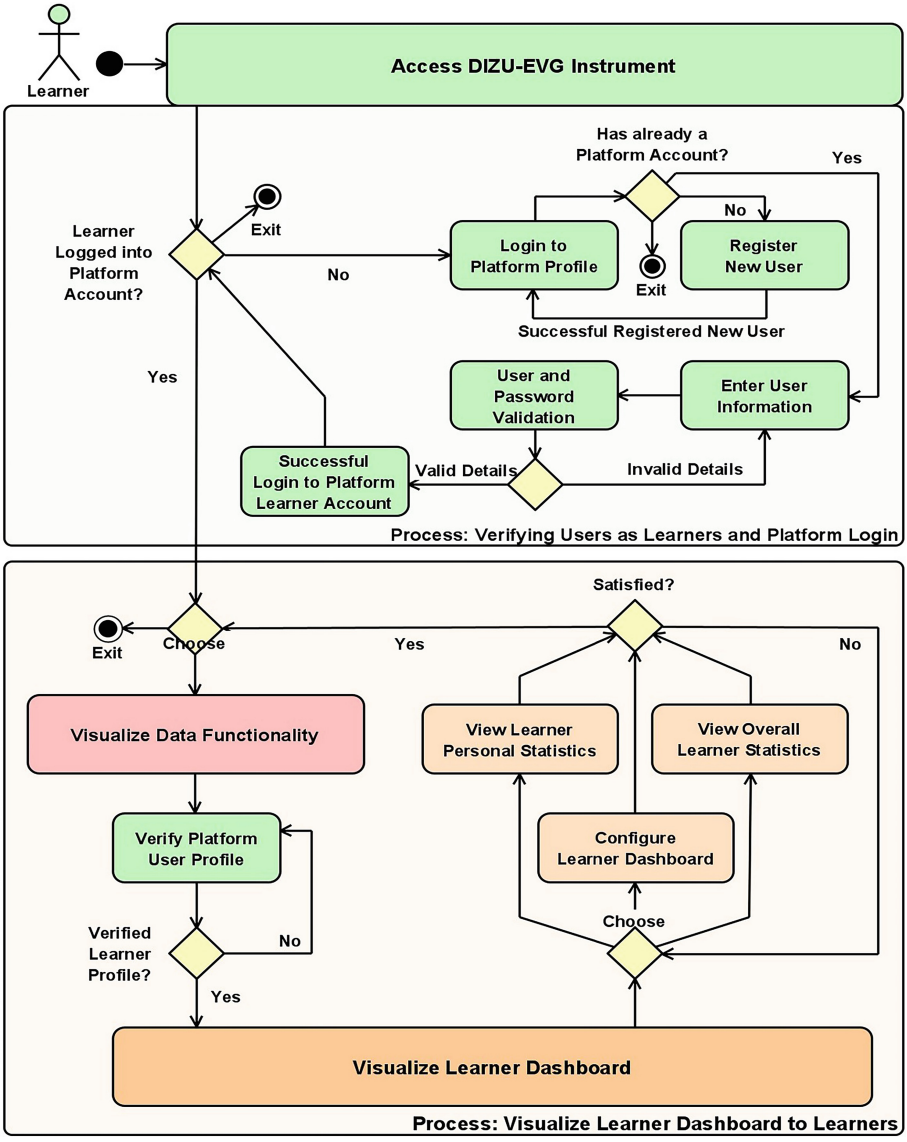


Fig. 1. Process Model of Using the Core Functionalities of the Learner Dashboard of the DIZU-EVG Instrument

Likewise, to the process model illustrated in Fig. 1, the learner has access only to the *Learner Dashboard* [18]. The user can utilize the three core functionalities of the *Player Dashboard*, Illustrated in the process model of Fig. 2. Depending on whether the player has played a particular educational video game in the system, generated data from game sessions are available, which can be shown to the player. If no video game is played, the player must access and play the appropriate educational video game

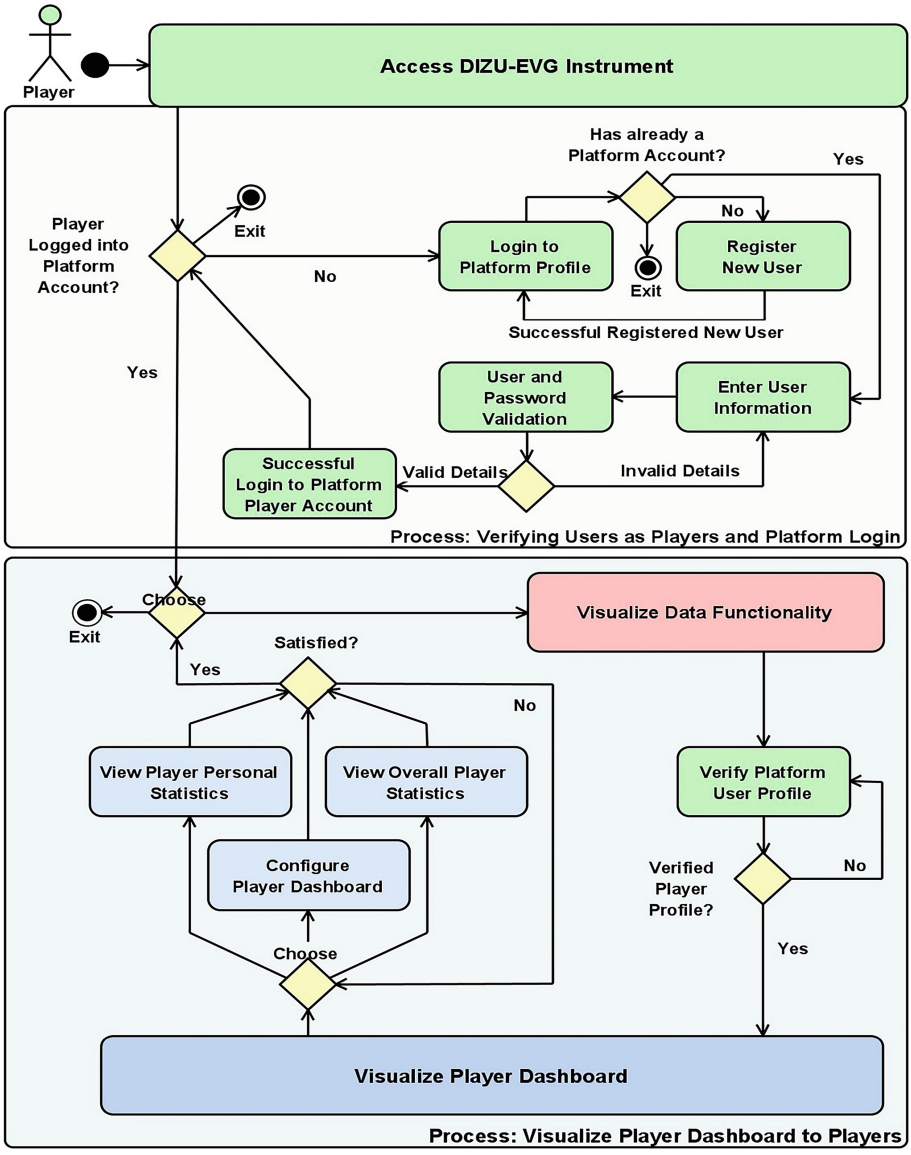


Fig. 2. Process Model of Using the Core Functionalities of the Player Dashboard of the DIZU-EVG Instrument

generated on the APOGEE platform, as this will provide the relevant game data from ended game sessions. After playing the game, the player can use the three core features of the *Player Dashboard*, which are directly related to the visualization of the individually achieved game results and the collective game results of all players who played the same educational video game. This provides a prerequisite for players to visually compare their success or failure according to data on the results of other players. Based on visually

analyzed results, players can take the necessary action to improve their game results and play the educational video game again.

4 Conclusion and Future Work

This paper presented the initial phase of designing the Dashboards of the DIZU-EVG instrument: one for learners to visualize their learning results from educational video games and one for players to visualize their gaming results. This paper also presented the primary requirements for the user-oriented dashboard design, goals and restrictions, and the core functionalities that each Dashboard provides to players and learners of educational maze video games generated through the APOGEE software platform. The paper presented two formal models of using the core functionalities of the individual Dashboards of the DIZU-EVG instrument.

The user-oriented (players and trainees) dashboard design, through which the gaming and learning results of the users are visualized within the DIZU-EVG instrument, will contribute to the development of the instrument and its improvement. These functionalities will increase the DIZU-EVG instrument's efficiency and its users' benefits. The ability to visualize the appropriate data of the respective users (players or learners) will provide opportunities for timely analysis of the results achieved by users and for analyzing and evaluating the game's design.

Developing additional Dashboard functionalities and improving the presented models have been planned as future work. This will provide prerequisites for future research on the DIZU-EVG instrument, expanding the benefits of the tool for software platforms for designing and creating educational video games and further improving and developing the instrument.

Acknowledgements. This research is supported by the Bulgarian Ministry of Education and Science under the National Program "Young Scientists and Postdoctoral Students – 2".

References

1. Cheng, M., Lin, Y., She, H.: Learning through playing Virtual Age: exploring the interactions among student concept learning, gaming performance, in-game behaviors, and the use of in-game characters. Comput. Educ. **86**, 18–29 (2015). https://doi.org/10.3390/informatics6030030
2. Mawas, El N., Truchly, P., Podhradský, P., Muntean, C.: The effect of educational game on children learning experience in a Slovakian School. In: CSEDU - 7th International Conference on Computer Supported Education, May 2019, Heraklion, Greece (2019). (hal-02249542)
3. Cheung, S., Ng, K.: Application of the educational game to enhance student learning. Front. Educ. **6**, 623793 (2021). https://doi.org/10.3389/FEDUC.2021.623793/BIBTEX
4. Martinez, L., Gimenes, M., Lambert, E.: Entertainment video games for academic learning: a systematic review. J. Educ. Comput. Res. **60**(5), 1083–1109 (2022). https://doi.org/10.1177/07356331211053848
5. Xinogalos, S., Satratzemi, M.: The use of educational games in programming assignments: SQL island as a case study. Appl. Sci. **12**(13), 6563 (2022). https://doi.org/10.3390/app12136563

6. Vlachopoulos, D., Makri, A.: The effect of games and simulations on higher education: a systematic literature review. Int. J. Educ. Technol. High. Educ. **14**, 22 (2017). https://doi.org/10.1186/s41239-017-0062-1
7. Pan, Y., Ke, F., Xu, X.: A systematic review of the role of learning games in fostering mathematics education in K-12 settings. Educ. Res. Rev. **36**, 100448 (2022). https://doi.org/10.1016/J.EDUREV.2022.100448
8. Rüth, M., Kaspar, K.: Commercial video games in school teaching: two mixed methods case studies on students' reflection processes. Front. Psychol. **11**, 594013 (2021). https://doi.org/10.3389/fpsyg.2020.594013
9. Nazirah, S., Othman, T., Tengku, N., Adura, I., Roselan, B.: Male students and digital game: reason, motivation and feeling. Int. J. Inf. Educ. Technol. **4**(1), 6–11 (2014). https://doi.org/10.7763/IJIET.2014.V4.359
10. Nagalingam, V., Ibrahim, R.: User experience of educational games: a review of the elements. Procedia Comput. Sci. **72**, 423–433 (2015). https://doi.org/10.1016/J.PROCS.2015.12.123
11. Emihovich, B., Roque, N., Mason, J.: Can video gameplay improve undergraduates' problem-solving skills? Int. J. Game-Based Learn. **10**(2), 21–38 (2020). https://doi.org/10.4018/ijgbl.2020040102
12. Rahimi, S., et al.: Timing of learning supports in educational games can impact students' outcomes. Comput. Educ. **190**, 104600 (2022). https://doi.org/10.1016/J.COMPEDU.2022.104600
13. Tan, J.W., Zary, N.: Diagnostic markers of user experience, play, and learning for digital serious games: a conceptual framework study. JMIR Serious Games **7**(3), e14620 (2019). https://doi.org/10.2196/14620
14. Jackson, L., O'Mara, J., Moss, J., Jackson, A.: Analysing digital educational games with the games as action, games as text framework. Comput. Educ. **183**, 104500 (2022). https://doi.org/10.1016/J.COMPEDU.2022.104500
15. Dankov, Y.: DIZU-EVG – an instrument for visualization of data from educational video games. In: Silhavy, R., Silhavy, P. (eds.) CSOC 2023. LNNS, vol. 722, pp. 769–778. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35311-6_73
16. Bontchev, B., Vassileva, D., Dankov, Y.: The APOGEE software platform for construction of rich maze video games for education. In: Proceedings of the 14th International Conference on Software Technologies – ICSOFT 2019, pp. 491–498. SciTePress (2019). https://doi.org/10.5220/0007930404910498
17. Dankov, Y., Bontchev, B.: Towards a taxonomy of instruments for facilitated design and evaluation of video games for education. In: Proceedings of the 21st International Conference on Computer Systems and Technologies (CompSys-Tech'20), pp. 285–292. ACM, NY, USA (2020). https://doi.org/10.1145/3407982.3408010
18. Dankov, Y.: User-oriented process analysis of using the DIZU-EVG instrument for educational video games. In: Silhavy, R., Silhavy, P. (eds.) CSOC 2023. LNNS, vol. 723, pp. 684–693. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_61
19. Bontchev, B., Terzieva, V., Dankov, Y.: Educational video game for Valchan Voivoda. Bulletin 'Heritage BG' - Research Announcements, 1, Association Centre for Excellence 'Heritage BG' (2021). https://cloud.nasledstvo.bg/s/62J7E9wLQBD8Fmx
20. Bontchev, B., Antonova, A., Terzieva, V., Dankov, Y.: "Let Us Save Venice"—an educational online maze game for climate resilience. Sustainability **14**(1), 1–23 (2022). https://doi.org/10.3390/su14010007
21. Bontchev, B., Vassileva, D., Aleksieva-Petrova, A., Petrov, M.: Playing styles based on experiential learning theory. Comput. Hum. Behav. **85**, 319–328 (2018). https://doi.org/10.1016/J.CHB.2018.04.009



Heuristic Approaches to Delivering Cloud Resources at Minimal Cost in IT Infrastructure Management

Georgi Shipkovenski(✉) and Oleg Asenov

University of Veliko Tarnovo, Veliko Tarnovo, Bulgaria
g.shipkovenski@gmail.com

Abstract. This paper presents a survey of various heuristic algorithms for load balancing in cloud resource provisioning. The main goal is to reduce the task execution time of available virtual machines and thereby achieve lower cloud resources cost. The specific load-balancing problem is formulated in the context presented. Using the Proxmox virtual infrastructure, the efficiency of four selected heuristic algorithms was investigated under different circumstances in a production lab environment consisting of five virtual machines. The experimental results show that both the greedy strategy approach and the combined greedy and probabilistic strategy approach are potentially effective solutions for achieving load balancing in cloud resource provisioning at minimum cost.

Keywords: Heuristic Algorithms · Load Balancing · Cloud Resources · Minimal Cost · Proxmox Infrastructure

1 Introduction

In the world of the modern computer technology, cloud-computing services are extremely important. Leading companies in the field of information technology such as Amazon with AWS, HP, Oracle, Microsoft and Google have large data centers equipped with a large-scale hardware network in order to effectively provide cloud services to their customers. In order for customers to use cloud services over the Internet, cloud service providers must ensure proper resource management and provisioning.

Cloud computing is an Internet-based computing model in which resources (such as networks, servers, storage, applications and services), software and information are made available on-demand to users' devices, similar to basic utilities such as water, gas, electricity and telephony. Cloud computing technology offers three main types of service models (called tiers) - Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS), as well as four deployment models - private, public, hybrid and community.

In recent years, cloud service providers have focused on dynamic resource management in order to ensure the sharing of cloud computing resources among different users.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 130–140, 2024.
https://doi.org/10.1007/978-3-031-53549-9_14

The dynamic cloud computing technique provides the ability to allocate resources to different clients based on their current need. This turns the cloud into a computing platform with unlimited storage capacity, which improves the performance of cloud services.

In order to achieve the best allocation of resources in dynamic hosting frameworks, cloud service providers must ensure the efficient provisioning of resources to all their customers. This efficient allocation of resources is called workload balancing in cloud service frameworks. Cloud service providers use a variety of resource provisioning strategies to improve quality of service.

The load-balancing problem can be defined in two different ways. The first way is by distributing a finite number of tasks to different physical machines (PMs), which are in turn distributed to different virtual machines (VMs) of the corresponding PM. The second way is related to managing the migration of virtual machines or tasks. In cloud computing, virtual machine migration is the process of moving a virtual machine from one PM to another PM to improve data center resource utilization, especially when a physical server is overloaded. In addition to virtual machine migration, cloud computing also performs task migration, which involves migrating the current state of a task from one VM to another VM or from a VM on one host to a VM on another host.

One of the ways to achieve load balancing is to use an efficient Load Balancer. The following model describes the workflow of a cloud load balancer. The user request is analyzed and routed to the selected data center depending on the available resources. Both overloading and underloading of the servers (virtual machines) should be avoided in order to achieve an even load distribution among them. This is where the concept of load balancing comes into play.

Uneven load distribution can be caused by a variety of factors, one of which is task scheduling. If task scheduling is not done properly, data center resources will not be used efficiently. To address this problem, a load balancer uses algorithms to distribute tasks and requests among resources, such as virtual machines or physical machines, depending on the current load and available capacity. Figure 1 shows a model that describes the workflow of a cloud load balancer [1].

There are two types of load balancing algorithms: static and dynamic. Static balancing algorithms are best suited for stable environments with homogeneous systems, and load balancing does not depend on the state of the system at the time the task is executed. Dynamic algorithms, on the other hand, are more adaptive and efficient in both homogeneous and heterogeneous environments and depend on the state of the system at the time of task execution.

To determine whether a load-balancing algorithm is optimal in terms of resource allocation and task makespan improvement, it is necessary to use different metrics. These metrics typically measure the response time for each request and the overall completion time of tasks. Appropriate metrics need to be chosen so that load-balancing algorithms can provide flexibility, low overhead, and optimal and efficient resource allocation.

Many authors [2–4] consider basic qualitative and quantitative metrics. Listed below are several metrics that are considered when analyzing load balancing algorithms.

- **Throughput:** Throughput refers to the number of user requests (tasks) served by the virtual machine per unit of time. The higher the throughput, the better the performance of the cloud computing system.

- **Resource Utilization:** In order to achieve efficiency, the load-balancing algorithm must take into account the resources in the system (such as memory, processor, etc.) and ensure increased and even utilization of these resources.
- **Scalability:** This is the ability of the cloud infrastructure to dynamically expand or contract in response to changing demands for computing resources, such as storage, processing power, or bandwidth. Any effective load-balancing algorithm must ensure optimal resource utilization to achieve high scalability.
- **Response time:** This is the time it takes for the cloud framework to execute and respond to user requests.
- **Makespan:** This is the total time it takes to complete all tasks that are submitted to the system.
- **Fault Tolerance:** Load balancing must be able to function effectively even if one or more system elements fail. This means that if one virtual machine becomes overloaded, there must be an opportunity for another available virtual machine to take over the tasks.
- **Migration Time:** This is the total time required to migrate a task from one virtual machine to another or the total time it takes to migrate a virtual machine from one host to another within a data center or across data centers.

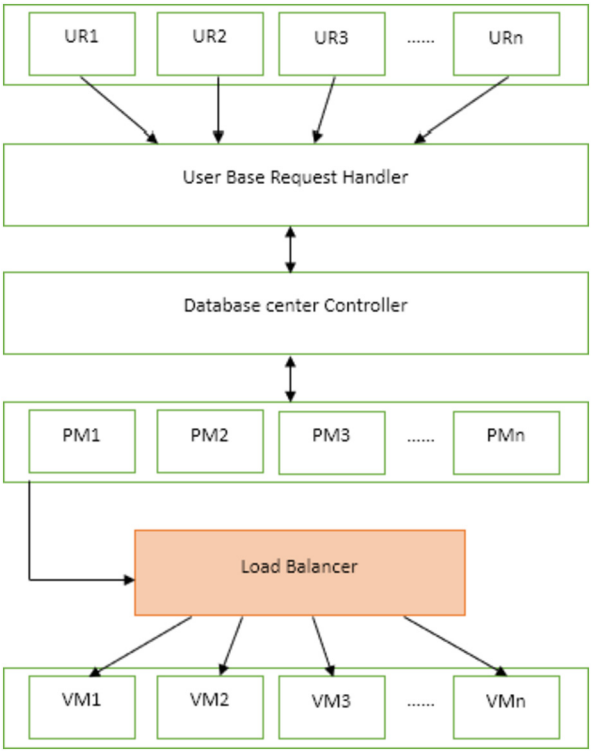


Fig. 1. Load Balancing Model

One of the most common and simplest load balancing techniques is the Round Robin algorithm. In this technique, tasks are allocated to different resources based on units of time [5]. It selects a random node in the cloud system and assigns it the first task. After this allocation, the next task will be allocated to the next free node. This load balancing technique ensures that all free cloud resources are used in a Round Robin manner. If all nodes are busy during task allocation, then the task is queued to the node with the fewest tasks. The task allocation cycle continues in the above manner until all tasks have been allocated to a node. Opportunistic Load Balancing (OLB) is an algorithm similar to Round-Robin, but it randomly distributes tasks among available nodes [6]. These algorithms have some major drawbacks, such as not taking into account the length of the tasks, the computing power of the virtual machine, or the priority of the tasks when they are allocated.

Common static algorithms used in the cloud environment, such as Round-Robin and OLB, are not suitable and efficient due to many drawbacks, such as uneven load distribution across nodes, where some machines may be overloaded while others remain underloaded. In contrast, dynamic algorithms consider the workload and response time of nodes in real time through a dynamic feedback mode.

DLB (Dynamic Load Balancing) algorithm is used to distribute the tasks in the system, taking into account the load of the virtual machines and redirecting the tasks to virtual machines with lower load. However, the algorithm does not take into account the length of the tasks, which can lead to an increase in the completion time and the duration of the system operation [7].

The algorithm that takes into account the load of VMs and the length of tasks is Min-Min. Min-Min starts with a set of all unallocated tasks and consists of two phases. In the first phase, for each task in the set the minimum expected completion time is calculated (this is the time in which the task is expected to complete on the corresponding machine). In the second phase, the task with the smallest total expected completion time is selected from the set and allocated to the corresponding resource. This task is then removed from the set and the process is repeated until all tasks have been allocated [8]. The Min-Min algorithm is only suitable for distributed systems with limited dimensions, and one of the improved variants is proposed by Chen et al. (2013) [9]. The improved Min-Min algorithm aims to balance the load in the system and optimize the total execution time of tasks, while improving the utilization of available resources.

The Max-Min algorithm is similar to the Min-Min heuristic algorithm, but with a difference in the second phase of execution. Instead of selecting the task with the minimum expected completion time, the Max-Min algorithm selects the task with the maximum expected completion time and assigns it to the corresponding resource. In this way, the task with a larger size (heavier process) is allocated to a cloud resource (virtual machine) that provides a minimum execution time. Studies [10, 11] present an extended version of the Max-Min algorithm that uses a task state table. This table is used to measure the real-time workload of virtual machines, as well as the expected time to complete tasks.

In some studies [12, 13] a Genetic Algorithm (GA) is used. It is based on a population of individual chromosomes representing potential distributions. Each distribution can be evaluated by a fitness value, which can represent various optimization parameters,

such as power consumption, makespan, throughput, and others. Through an iterative process of evolution and selection of the best solutions, the GA progressively improves the quality of the distributions and strives to find an optimal solution. This allows the GA to be used to optimize resource allocation, taking into account various factors and fitness values. Two other metaheuristic algorithms are used to solve the load balancing problem - Simulated Annealing [14] and Tabu Search [15].

This paper examines various load balancing algorithms for providing cloud resources with minimum cost. Section 2 formulates the load balancing problem and presents greedy algorithms and their applicability to solve it. Section 3 evaluates the performance of the max-min algorithm, the sorted greedy algorithm, the improved greedy probabilistic algorithm, and the simulated annealing algorithm through experimental studies on a Proxmox virtual infrastructure with five virtual machines assigned to ten tasks.

2 Heuristic Algorithms for Providing Cloud Resources with Minimum Cost

The main factor in reducing the cost of cloud resources is the optimization of the execution time of the tasks. When tasks are optimally distributed and executed quickly on available virtual machines, the potential of those machines is used more efficiently, resulting in cost reductions as no additional machines or resources need to be leased. Faster task execution also frees up cloud resources for other tasks or clients, optimizing resource utilization and reducing energy and maintenance costs of virtual machines. Another important consideration is also that some cloud service providers base their pricing on the time it takes to complete the tasks. If the execution time is reduced, this can lead to lower prices for the resources used. In other words, customers can benefit from lower prices for services if their tasks are completed faster and more efficiently in the virtual environment of the cloud service provider.

The load balancing problem can be formulated as follows [16]. Given a set of m machines $M_1,\ldots., M_m$ and a set of n tasks $J_1,\ldots., J_n$. Each task J_i has a processing time t_j. The goal is to allocate each task to one of the machines so that the load on all machines is as balanced as possible.

Specifically, for each task allocation among machines, we can define the set A(i) of tasks that are allocated to machine M_i. With this distribution, the M_i machine must run for a total time:

$$T_i = \sum_{j \in A(i)} T_j, \tag{1}$$

This total time is declared as the machine load M_i. The goal is to minimize the makespan magnitude. This is the maximum load on each machine, $T = \max_i T_i$.

In most implementations of heuristics, the main factors considered to obtain an accurate estimate of the execution time of a task are the size of the task, measured in million instructions (MI), and the execution speed of the virtual machine, measured in million instructions per second (MIPS). The execution time of task j on virtual machine i is determined by dividing the size of the task by the speed of the ith virtual machine.

This information is represented in a matrix E_{ij} with the execution times of the tasks from each machine.

For a more realistic estimation of the execution time, a more efficient allocation of tasks and in order to achieve task completion time optimization, a new matrix is compiled with the expected completion time of each task from each virtual machine C_{ij}, which includes the time for execution E_{ij} and the ready time r_i. The ready time is determined based on data volume and connection bandwidth, as Ready Time (seconds) = Data Volume (MB)/Bandwidth (Mbps). The expected completion time C_{ij} of task j on virtual machine i is defined as $C_{ij} = E_{ij} + r_i$.

The sorted greedy algorithm starts by assigning the longest and most complex task to the most efficient virtual machine in order to speed up the solution of complex tasks and to reduce the total execution time of all tasks. This is followed by the first step of the algorithm, in which the tasks are sorted from the largest to the smallest taking into account the size of the task (MI), and the virtual machines are sorted from the smallest to the largest according to their execution speed (MIPS). Thus, a sorted matrix E_{ij} is constructed, which contains the execution times of the tasks by each machine. The data volume (MB) for each task and the bandwidth (Mbps) for each machine are used to calculate the ready times (r_i). They are added to the matrix of task execution times from each machine E_{ij} to obtain the matrix of expected completion times (C_{ij}) of the tasks. In a second step of the algorithm, the greedy strategy is implemented by selecting the minimum from each column of the matrix C_{ij} and assigning it to the machine with the least current processing time.

This approach optimizes the expected completion time of tasks by prioritizing virtual machines with higher execution speed and selecting an optimal combination of tasks that can be completed as quickly as possible. However, it has its drawbacks and sometimes does not achieve full optimality or efficiency and cannot handle complex tasks that involve large changes in demand.

The improved greedy probabilistic algorithm starts by implementing greedy behavior as it selects minimum values from the columns in the sorted matrix C_{ij}. By adding the probabilistic aspect (by using the probability p) it gets some degree of randomness and variation in the choice of values. The purpose of introducing probability is to allow the algorithm an opportunity to try different solutions, which may lead to finding better results. The probabilistic aspect allows the algorithm to "look" at several alternatives, which can help avoid local minima and converge to the global minimum (or best solution).

The main steps of the improved greedy probabilistic algorithm are:

1. Creation of a set of virtual machines and tasks and define the sorted matrix C_{ij}.
2. Initialization of the sorted matrix C_{ij} containing the expected completion time of the tasks of the corresponding virtual machines.
3. Distribution of tasks. The algorithm selects the first task to be executed by choosing the smallest completion time from the first column of the time matrix C_{ij}. This is done at the beginning, when no tasks have been assigned yet.
4. Selection of a virtual machine for task distribution. The selected task is allocated to a virtual machine with minimum completion time. The corresponding completion times are updated to reflect the completion of the machine's task.

5. Using the probabilistic aspect of the algorithm. The algorithm uses a probability p to determine whether to select the next task from the current column, applying a greedy task selection approach, or to move to the next task selection VM.
6. Repeat steps 4 and 5. The task selection and allocation process is repeated until all tasks are allocated to the virtual machines.
7. Displaying the results. Finally, the completion time of each task on each virtual machine and the total completion time of all tasks (makespan) are displayed.

Figure 2 shows the block diagram of the improved greedy probabilistic algorithm.

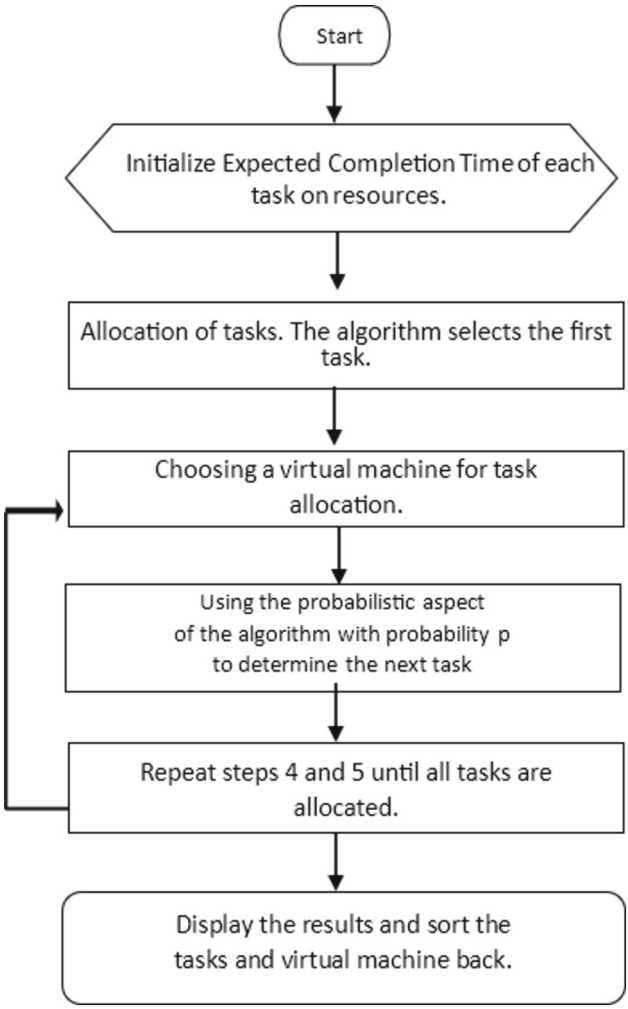


Fig. 2. Improved greedy probabilistic algorithm

3 Experimental Results

A server on which Proxmox virtual infrastructure is configured is used and five virtual machines are created. The processing speed of the 5 virtual machines is VM1-2700, VM2-2900, VM3-1300, VM4-2000 and VM5-2800. To get the exact speeds for each virtual machine in the Proxmox web interface, hardware - processor options section, limits are set on each core. Limits were set on the network interfaces of the virtual machines, with the bandwidth for each being VM1-800 Mbps, VM2-800 Mbps, VM3-400 Mbps, VM4-400 Mbps, and VM5-80 Mbps. The tasks used in the experiment have the following sizes: Task 1 – 203650, Task 2 – 508090, Task 3 – 312180, Task 4 – 451570, Task 5 – 177540, Task 6 – 193360, Task 7 – 210450, Task 8 – 324930, Task 9 – 317270 and Task 10 - 320170. The volume in megabytes for each task is Task 1 - 795.51, Task 2 - 1984.73, Task 3 - 1219.45, Task 4 - 1763.95, Task 5 - 693.52, Task 6 - 755.31, Problem 7 - 822.07, Problem 8 - 1269.26, Problem 9 - 1239.34 and Problem 10 - 1250.66.

The results of the allocation of tasks to the virtual machines when using the sorted greedy algorithm are shown in Table 1. They coincide with the results obtained after running the Max-Min algorithm. The minimum total time to complete all tasks on the specific virtual machines is 372.17.

Table 1. The allocation of tasks with sorted greedy algorithm

T_4		T_3			
1	0.00	1	184,89		
	184,89		312,7		
T_2		T_7		T_6	
2	0.00	2	195,05	2	275,84
	195,05		275,84		350,07
T_9					
3	0.00				
	268,84				
T_8		T_1			
4	0.00	4	187,85		
	187,85		305,59		
T_{10}		T_5			
5	0.00	5	239,41		
	239,41		372,17		

When using the improved greedy probabilistic algorithm, good results are achieved in the allocation of tasks. The minimum total time to complete all tasks is 370.00. The obtained results are shown in Table 2.

The optimal score achieved for completing all tasks using the simulated annealing algorithm was 337.67. The distribution of tasks with the simulated annealing algorithm is presented in Table 3.

Table 2. Task distribution with the improved greedy probabilistic algorithm

T_4		T_3			
1	0.00	1	184,89		
	184,89		312,7		
T_2		T_1		T_6	
2	0.00	2	195,05	2	273,23
	195,05		273,23		347,46
T_{10}					
3	0.00				
	271,30				
T_8		T_7			
4	0.00	4	187,85		
	187,85		309,52		
T_9		T_5			
5	0.00	5	237,24		
	237,24		370,00		

Table 3. Allocation of tasks with the simulated annealing algorithm

T_3		T_5		T_9	
1	0.00	1	127,817	1	200,507
	127,82		200,51		330,41
T_2		T_8			
2	0.00	2	195,051		
	195,05		319,79		
T_1		T_6			
3	0.00	3	172,56		
	172,56		336,41		
T_7		T_{10}			
4	0.00	4	121,666		
	121,67		306,76		
T_4					
5	0.00				
	337,67				

The experiment shows that the best results are obtained by using the simulated annealing algorithm. The results obtained when applying an improved greedy probabilistic algorithm are also comparable. The worst results are obtained by using sorted greedy algorithms.

When using the simulated annealing algorithm in the experiment, optimal solutions are not always achieved, since this technique is based on probabilistic principles and may depend on the initial parameters and settings of the algorithm. On the other hand, the greedy probabilistic algorithm always achieves a near-optimal solution within the current parameters and conditions. It can be concluded that the improved greedy probabilistic

algorithm is chosen as the preferred algorithm for task distribution to virtual machines in the Proxmox virtual infrastructure.

4 Conclusion

This paper presents an investigation of various heuristic algorithms with the aim of reducing the execution time of the available virtual machines, which further reduces the cost of the provided cloud resources. Emphasis is placed on the efficient way to distribute the load by using a sorted greedy algorithm and an improved greedy probabilistic algorithm. To provide cloud resources with minimum cost in Proxmox virtual infrastructure, the improved greedy probabilistic algorithm is preferred.

This algorithm shows good results and approaches the optimal load balancing solutions, which leads to minimizing the execution time of the tasks and reducing the cost of the provided cloud resources.

References

1. Shafiq, D.A., Jhanjhi, N., Abdullah, A.: Load balancing techniques in cloud computing environment: a review. J. King Saud Univ.-Comput. Inf. Sci., 1–24 (2021)
2. Afzal, S., Ganesh, K.: A Taxonomic Classification of Load Balancing Metrics: A Systematic Review. 33rd Indian Eng. Congr., pp. 85–90 (2019)
3. Shahid, M.A., Islam, N., Alam, M.M., Su'Ud, M.M., Musa, S.: A comprehensive study of load balancing approaches, in the cloud computing environment and a novel fault tolerance approach. IEEE Access **8**(c), 130500–130526 (2020)
4. Roy, S., Hossain, A., Sen, S., Hossain, N., Asif, R.: Measuring the performance on load balancing algorithms. Global J. Comp. Sci. Technol. **19**, 41–49 (2019)
5. Pasha, N., Agarwal, A., Rastogi, R.: Round robin approach for VM load balancing algorithm in cloud computing environment. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **4**(5), 34–39 (2014)
6. Joshi, V.: Load Balancing Algorithms in Cloud Computing. International Journal of Research in Engineering and Innovation (2019)
7. Kumar, S., Rana, D.: Various dynamic load balancing algorithms in cloud environment: a survey. Int. J. Comput. Appl.Comput. Appl. **129**(6), 16 (2015)
8. Etminani, K., Naghibzadeh, M.: A min-min max-min selective algorithm for grid task scheduling. In: The Third IEEE/IFIP International Conference on Internet, pp. 1–7 (2007)
9. Chen, H., Wang, F., Helian, N., Akanmu, G.: User-priority guided min-min scheduling algorithm for load balancing in cloud computing. In: Proceedings of the National Conference on Parallel Computing Technologies, pp. 1–8 (2013)
10. Li, X., Mao, Y., Xiao, X., Zhuang, Y.: An Improved Max-Min Task-Scheduling Algorithm for Elastic Cloud. Computer, Consumer and Control (IS3C), IEEE Computer Society, pp. 340–343 (2014)
11. Elzeki, O., Reshad, M., Elsoud, M.: Improved max-min algorithm in cloud computing. Int. J. Comput. Appl. (0975 – 8887) **50**(12), 22–27 (2012)
12. Dasgupta, K, Mandal, B., Dutta, P., Mondal, J., Dam, S.: A Genetic Algorithm (GA) based load balancing strategy for cloud computing. In: Proceedings of Computational Intelligence: Modeling, Techniques and Applications, pp. 340–347 (2013)
13. Tripathi, A., Kumer, B., Kumar, N., Vidyarthi, D.: A GA based multiple task allocation considering load. Int. J. High Speed Comput. **11**(04), 203–214 (2000)

14. Mondal, B., Choudhury, A.: Simulated Annealing (SA) based load balancing strategy for cloud computing. (IJCSIT) Int. J. Comput. Sci. Inf. Technol. **6**(4), 3307–3312 (2015)
15. Alam, M., Pandey, M., Rautaray, S.: A proposal of resource allocation management for cloud computing. Int. J. Cloud Comput. Serv. Sci. (IJ-CLOSER) **3**(2), 79–86 (2014)
16. Phi, N., Hieu, L., Hung, T.: Load balancing algorithm on cloud computing for optimize response time. Int. J. Cloud Comput. Serv. Architecture (IJCCSA) **10**(3) (2020)

Vocal Folds Image Segmentation Based on YOLO Network

Jakub Steinbach[1(✉)], Zuzana Urbániová[2(✉)], and Jan Vrba[1(✉)]

[1] University of Chemistry and Technology in Prague, Department of Mathematics, Informatics and Cybernetics, Technická 1905/5, 166 28 Praha 6, Czech Republic
{jakub.steinbach,jan.vrba}@vscht.cz
[2] Charles University, 3rd Faculty of Medicine, Department of Otorhinolaryngology, University Hospital Královské Vinohrady, Prague, Czech Republic
zuzana.urbaniova@lf3.cuni.cz

Abstract. The focus of this article is on utilizing YOLOv8 segmentation models for the detection of vocal fold openness in laryngoscopic videos, eliminating the need for extra image enhancement. The evaluation and comparison of different models are carried out based on accuracy metrics such as box mean average precision and mask mean average precision. The outcomes indicate the potential applicability of YOLOv8 segmentation models in objectively quantifying vocal fold openness, offering a potential avenue for integration into clinical practice.

Keywords: vocal cords · vocal folds · YOLO · neural networks · machine learning · image segmentation

1 Introduction

In accordance with [18], the vocal cords play a crucial role in processes like speaking, breathing, and swallowing. When their mobility is compromised, individuals face significant clinical challenges related to essential physiological functions. Disorders affecting vocal cord movement can stem from various factors, including injury, issues with the arytenoid joint (such as arthritis, ankylosis, or dislocation), muscular diseases (like myasthenia gravis), or nerve-related problems. Neurological dysfunctions primarily involve the recurrent laryngeal nerve or superior laryngeal nerve, along with frequent vagus nerve abnormalities. Conditions like strokes, Parkinson's disease, and multiple sclerosis affecting the central nervous system can lead to central paresis of the vocal cords and neurogenic voice problems like spasmodic dysphonia should also be considered within this group [3,11].

Individuals experiencing vocal cord dysfunction frequently encounter dysphonia. The crucial component of their physical assessment and subsequent therapeutic monitoring involves the visualization of the larynx and vocal cords [16].

Within the realm of daily otorhinolaryngological practice, clinicians employ a range of visualization techniques, spanning from indirect mirror laryngoscopy to high-speed kymography. Among the most prevalent and pertinent methods for functionally assessing the vocal tract are stroboscopic visualization of the vocal

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 141–149, 2024.
https://doi.org/10.1007/978-3-031-53549-9_15

mucosal wave and the dynamic appraisal of vocal performance through flexible laryngoscopy. The integration of these methodologies is recommended for a comprehensive understanding of both the vibrational dynamics of the vocal cord mucosa and the overall functioning of the vocal tract. Rigid laryngostroboscopy offers heightened magnification of the vocal cords and impeccable optical clarity. Nevertheless, it comes with the drawback of requiring the patient to assume a somewhat constrained posture with tongue protrusion. Moreover, rigid laryngoscopy presents limitations in its inability to provide a comprehensive functional assessment of the entire vocal tract [14].

The described visualization techniques inherently involve a level of subjectivity. The ultimate observation is impacted by the otorhinolaryngologist's expertise. Currently, no practical and objective method exists to consistently aid in the assessment of patient treatment outcomes. Within clinical settings, the evaluation of vocal cord mobility impairment remains qualitative, and a standardized vocabulary is absent to precisely characterize movement abnormalities. Consequently, clinicians may encounter challenges when attempting to effectively communicate essential clinical details.

Numerous research studies have introduced algorithms aimed at detecting vocal fold disorders [6]. The conventional approach relies on the analysis of sound or speech and involves the utilization of diverse voice attributes like fundamental frequency, jitter, shimmer, Mel-frequency cepstral coefficients, and others [17]. Additionally, some investigations have explored the use of mobile phones as voice recording devices [22]. Nevertheless, a pivotal aspect of these methodologies revolves around accurately capturing patients' vocal recordings [19]. Another approach, that has not attracted as much attention as the conventional one, is the analysis of laryngoscopical images. The study [15] is devoted to the analysis of high-speed videoendoscopy recordings and the glottal area segmentation is approached via Glottis Analysis Tools. In [20], the authors propose a method for laryngeal images, that is based on blood vessel centerline extraction and evaluation. The application of deep learning applied to videolaryngoscopical images for real-time laryngeal cancer detection is introduced [2]. In [1], the authors propose automated glottic action tracking based on open-source computer vision toolbox DeepLabCut.

In our previous work [18] we aimed to detect and localize the position of vocal folds. We utilized YOLOv5 object detector [7], which was trained on the dataset of images, obtained during laryngoscopical examination. We achieved the 99.5% mean average precision at 0.5 intersection over union threshold with all models. In this article, we build on the previous one and extend the vocal fold detection with the vocal fold openness detection based on the YOLOv8 segmentation models, which we evaluated on images of 3 different patients with different diagnoses.

1.1 YOLOv8-Based Segmentation

Incorporating machine learning methods to objectively asses vocal cords, requires to locate their position in an image. While object detection allows finding an

approximate location defined by a bounding box, image segmentation allows to find the exact location of the cords within the bounding box.

The YOLO algorithm, which was originally introduced by Redmon in 2016 [13], has emerged as a prominent state-of-the-art object detection methodology. Operating as a one-stage detector, it formulates object detection as a regression task. Notably, both the predictive bounding box parameters and their corresponding class labels are simultaneously determined in a single computation step, leading to its nomenclature 'You Only Look Once'. YOLO, including its iteration YOLOv5 [7], holds significant prominence within diverse range of research topics. Its application is particularly well-suited for scenarios involving static images or stable video sequences. The most recent version, YOLOv8 [8], extends the application of the YOLO architecture by incorporating segmentation models.

While the research paper on this novel model is yet to be published, Glenn Jocher, the founder of Ultralytics, provided a brief introduction to the architecture (in his commentary of the issue #189). The backbone of the YOLOv8-Seg model is a CSPDarknet53 feature extractor. The traditional YOLO neck architecture has been replaced by a novel C2f module. The segmentation heads then learn to predict the semantic segmentation masks for the input image. The whole architecture puts emphasis on being fast and efficient while achieving state-of-the-art results [8].

2 Materials and Methods

The dataset under analysis in this study comprises video recordings obtained by the Department of Otorhinolaryngology, 3rd Faculty of Medicine, situated within the University Hospital Královské Vinohrady in Prague. The acquisition process involved the utilization of two flexible Olympus rhino-laryngoscopes for capturing a pair of videos, while the third video was captured using a rigid laryngendoscope. All recording procedures incorporated stroboscopic examination via the Highlight Plus Invisia diagnostic video chain. Subjects were instructed to produce the phonemic sound /e:/ followed by a deep inhalation.

In terms of the patients' medical conditions, the initial recording depicted a patient diagnosed with left-sided vocal cord paralysis. The second video documented a routine follow-up examination on the second day post-total thyroidectomy, revealing a mild partial paresis of the right vocal cord. The third recording showcased a healthy patient for comparative purposes.

Each video recording was conducted at a sampling rate of 25 Hz, with dimensions measuring 960 × 720 pixels. The duration of the longest video was 49 s, while the shortest spanned 19 s. All recordings were saved in the mp4 format. Detailed information concerning frame utilization from each video is outlined in Table 1.

Table 1. Data Source Information

Diagnosis	left-sided vocal cord paralysis	partial right vocal cord paresis	healthy patient
Framerate	25 fps		
Frame Size	960 × 720		
Total Length	49 s	38 s	19 s
Frames Used	181	89	36

The entirety of the video footage was partitioned into distinct frames. Following this segmentation, an algorithm based on the YOLO network, as expounded in [18], was applied for the purpose of image cropping.

Subsequently, the cropped images underwent manual annotation through utilization of the Roboflow Annotate application [4]. Throughout the annotation phase, images characterized by closed vocal folds or low image quality were eliminated, yielding a dataset containing 306 jpg-format images.

The finalized dataset was subsequently retrieved from the Roboflow Annotate application, complete with labels formatted suitably for the training of a YOLO network.

3 Results

3.1 Model Validation

To estimate the performance of various YOLOv8-Seg architectures, we evaluated the mean average precision (threshold 0.5) [5] and mean average precision over multiple thresholds (from 0.5 to 0.95 with step 0.05) [10] for both the bounding boxes and the masks. We also evaluated the training time (TT, see Table 2). The comparison of architectures are in Table 2. Examples from all three videos comparing the manual label and label inferred by a trained YOLOv8n-Seg model are in Fig. 2, Fig. 1, and Fig. 3.

3.2 Implementation Specification

In this experimental study, we used Pytorch [12] based YOLOv8 release implementation that is publicly available at https://github.com/ultralytics. We conducted experiments with different YOLO models, namely YOLOv8l-Seg, YOLOv8m-Seg, YOLOv8s-Seg, and YOLOv8n-Seg. The number of weights (Parameters) for each model is in Table 2. All models utilize input images with a size of 640 × 640 (pixels). To train the models, the Adam optimizer [9] with a learning rate of 0.001 and decay of 10^{-5} was used. All experiments were carried out with a batch size of 16. The computational time was measured for every single YOLOv8 segmentation model. The experiments were performed on a PC using an AMD Ryzen 7700X 8-Core CPU running at 4.5 GHz with 32 GB RAM. The YOLOv8 was trained using NVIDIA GeForce RTX 4070 Ti 16 GB with CUDA 11.7 driver. The operating system was Windows 10 Enterprise 64-bit

version 10.0.19044 and the code was written in Python 3.9.13 [21] with Pytorch-2.0.1. The results obtained are summarized in Table 2. Examples of successful detection are shown in Fig. 2 and Fig. 1.

Table 2. Results of Model Validation for Selected Models

Model	Parameters (M)	Time (min)	Box		Mask	
			mAP (0.5)	mAP (0.5 - 0.95)	mAP (0.5)	mAP (0.5 - 0.95)
YOLOv8n-Seg	3.3	7.02	0.995	0.826	0.995	0.771
YOLOv8s-Seg	11.8	11.34	0.995	0.833	0.995	0.778
YOLOv8m-Seg	27.2	26.82	0.994	0.815	0.994	0.767
YOLOv8l-Seg	45.9	40.02	0.995	0.828	0.995	0.774

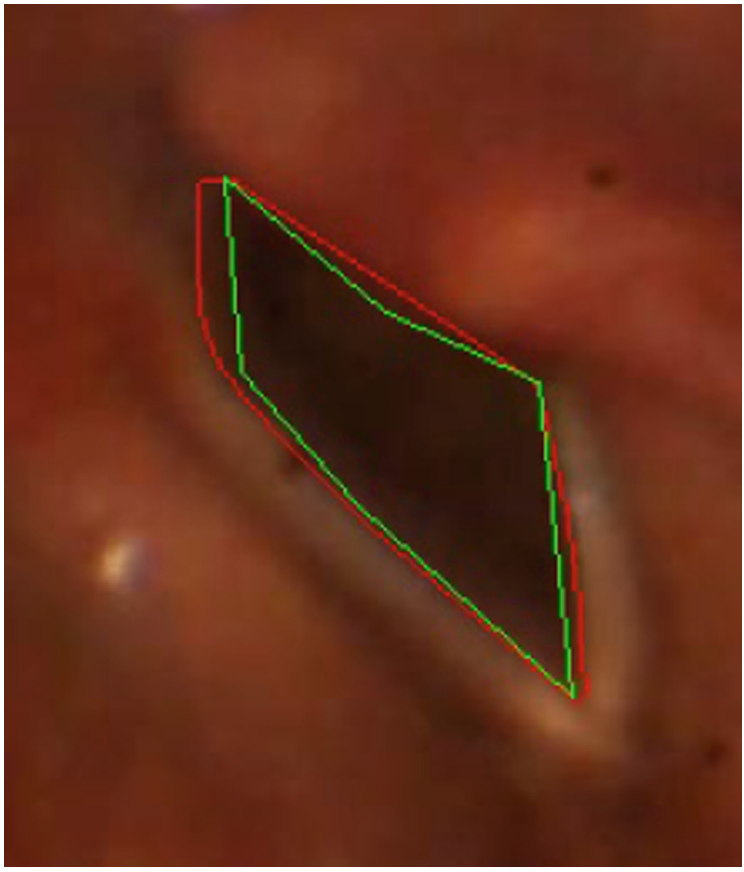


Fig. 1. Comparison of the manual label (green) and the inferred (red) label - left-sided vocal cord paresis

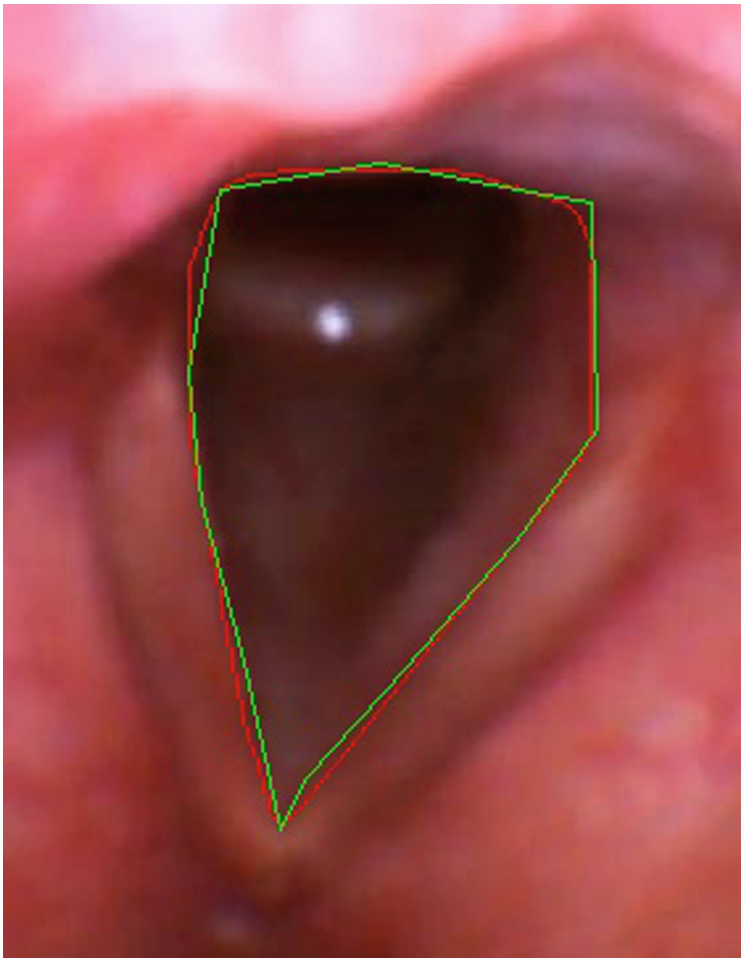


Fig. 2. Comparison of the manual label (green) and the inferred (red) label - partial right vocal paresis after a total thyroidectomy

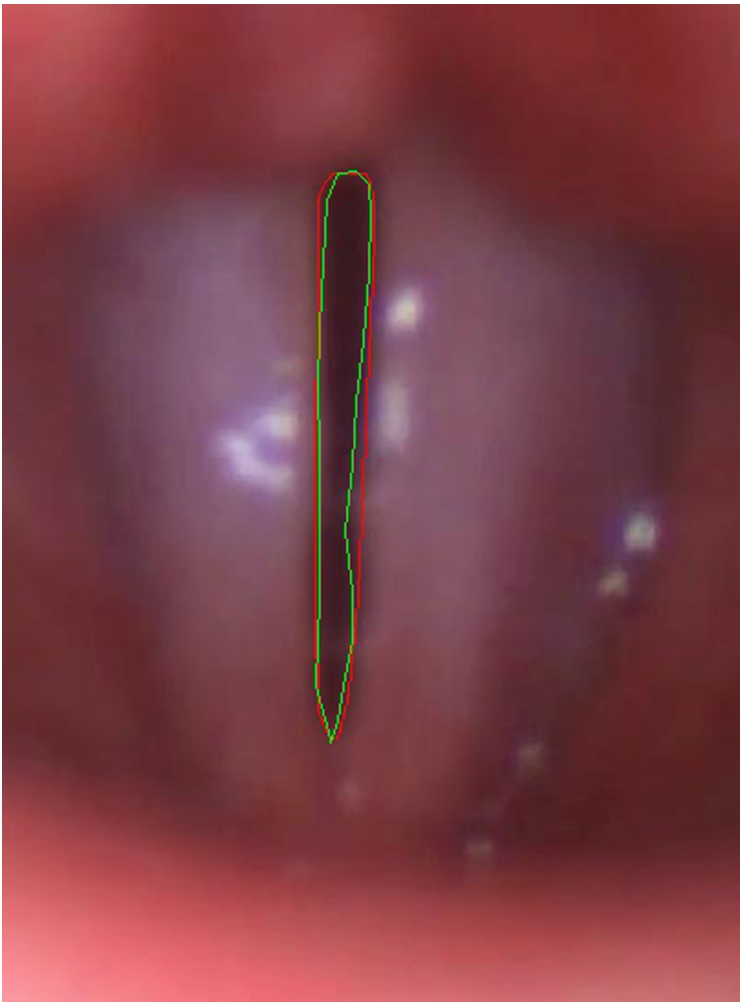


Fig. 3. Comparison of the manual label (green) and the inferred (red) label - healthy patient

4 Conclusion

In this article, we introduced the use of the YOLOv8 segmentation models to estimate the area corresponding to the opening of the vocal cords in images obtained from the rhino-laryngoscopic and rigid laryngoendoscopic examination.

Before training, the YOLOv5 based detection model was used, to select images containing vocal folds. Those images were manually labeled using Roboflow Annotate application. We evaluated four different YOLOv8 segmentation models and compared them while measuring the train time, box mean

average precision and mask mean average precision with multiple thresholds from 0.5 to 0.95.

According to our results, the performance of various YOLOv8 segmentations models is very similar in both precision and recall as same as in box mean average precision and mask mean average precision. Our results show, that those models, in combination with YOLOv5 model for vocal folds detection, are potentially usable for measuring vocal fold openness. This can lead to a fully automatic objective evaluation of the patient's vocal folds condition.

Acknowledgements. Jakub Steinbach acknowledges his specific university grant (IGA) A1_FCHI_ _2023_003. Zuzana Urbániová acknowledges her specific research project of Charles University COOPERATIO 43 - Surgical disciplines.

References

1. Adamian, N., Naunheim, M.R., Jowett, N.: An open-source computer vision tool for automated vocal fold tracking from videoendoscopy. Laryngoscope **131**(1), E219–E225 (2021)
2. Azam, M.A., et al.: Deep learning applied to white light and narrow band imaging videolaryngoscopy: toward real-time laryngeal cancer detection. Laryngoscope **132**(9), 1798–1806 (2022). https://onlinelibrary.wiley.com/doi/abs/10.1002/lary.29960. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/lary.29960
3. Dršata, J.: Foniatrie - hlas. Medicína hlavy a krku, Tobiáš, Havlíčkův Brod, 1 edn. (2011). http://arl.uhk.cz/arl-hk/cs/detail-hk_us_cat-0014865-Foniatrie-hlas/
4. Dwyer, B., Nelson, J., Solawetz, J., et. al.: Roboflow. https://roboflow.com
5. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. Int. J. Comput. Vision **88**(2), 303–338 (2010)
6. Hegde, S., Shetty, S., Rai, S., Dodderi, T.: A survey on machine learning approaches for automatic detection of voice disorders. J. Voice **33**(6), 947-e11 (2019)
7. Jocher, G., et al.: ultralytics/yolov5: v7.0 - yolov5 sota realtime instance segmentation (2022). https://zenodo.org/record/7347926
8. Jocher, G., Chaurasia, A., Qiu, J.: YOLO by Ultralytics, January 2023. https://github.com/ultralytics/ultralytics
9. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization (2017)
10. Lin, T., et al.: Microsoft COCO: common objects in context. CoRR abs/1405.0312 (2014). http://arxiv.org/abs/1405.0312
11. Merati, A.L., Heman-Ackah, Y.D., Abaza, M., Altman, K.W., Sulica, L., Belamowicz, S.: Common movement disorders affecting the larynx: a report from the neurolaryngology committee of the AAO-HNS. Otolaryngology-Head Neck Surgery **133**(5), 654–665 (2005). https://onlinelibrary.wiley.com/doi/10.1016/j.otohns.2005.05.003
12. Paszke, A., et al.: Pytorch: An imperative style, high-performance deep learning library. In: Advances in Neural Information Processing Systems 32, pp. 8024–8035. Curran Associates, Inc. (2019). http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf
13. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788. IEEE, Las Vegas, NV, USA, June 2016. http://ieeexplore.ieee.org/document/7780460/

14. Rosen, C.A., et al.: Nomenclature proposal to describe vocal fold motion impairment. European Archives of Oto-Rhino-Laryngology **273**(8), 1995–1999 (2016). http://link.springer.com/10.1007/s00405-015-3663-0
15. Schlegel, P., Kniesburges, S., Dürr, S., Schützenberger, A., Döllinger, M.: Machine learning based identification of relevant parameters for functional voice disorders derived from endoscopic high-speed recordings. Sci. Rep. **10**(1), 10517 (2020)
16. Stachler, R.J., et al.: Clinical Practice Guideline: Hoarseness (Dysphonia) (Update). Otolaryngology-Head Neck Surg. **158**(S1) (2018). https://onlinelibrary.wiley.com/doi/10.1177/0194599817751030
17. Steinbach, J., Mazúr, R., Vrba, J.: Trends in voice recording classification-comparison of conventional features and image analysis approach. In: Proceedings of the Computational Methods in Systems and Software, pp. 627–635. Springer (2022)
18. Steinbach, J., Urbániová, Z., Vrba, J.: Detection of vocal cords in endoscopic images based on yolo network. In: Computer Science On-line Conference, pp. 747–755. Springer (2023)
19. Steinbach, J., Vrba, J., Urbániová, Z.: Voice recording setup in clinical practice. In: Proceedings of the Computational Methods in Systems and Software, pp. 475–483. Springer (2022)
20. Turkmen, H.I., Karsligil, M.E., Kocak, I.: Classification of laryngeal disorders based on shape and vascular defects of vocal folds. Comput. Biol. Med. **62**, 76–85 (2015)
21. Van Rossum, G., Drake, F.L.: Python 3 Reference Manual. CreateSpace, Scotts Valley, CA (2009)
22. Verde, L., De Pietro, G., Sannino, G.: Voice disorder identification by using machine learning techniques. IEEE Access **6**, 16246–16255 (2018)

Knowledge Discovery Systems: An Overview

Serafeim A. Triantafyllou(✉)

Greek Ministry of Education and Religious Affairs, Athens, Greece
ser.triant@gmail.com

Abstract. Knowledge Discovery Systems are computer systems that form the backbone of the process of producing explicit or implicit knowledge from analysis of data or from prior knowledge synthesis. Knowledge Discovery in Databases is a promising research area which can bring positive outcomes when properly implementing basic machine learning and discovery methods concerning data storage in relational databases. This paper tries to contribute to a better understanding of the research field by conducting a useful discussion after thoroughly examining decision tree learning algorithms to find implicit and explicit knowledge in databases.

Keywords: knowledge discovery systems (KDS) · Decision Trees · Learning Systems

1 Introduction

Knowledge Discovery Systems are computer systems that form the backbone of the process of producing explicit or implicit knowledge from analysis of data or from prior knowledge synthesis [1]. KDS are used in a wide variety of industries, including finance, healthcare, marketing, and manufacturing. They are also used in scientific research to analyze large datasets in fields such as genomics, climate science, and particle physics. They are computer-based systems that assist users in discovering new knowledge and insights from large amounts of data. These systems are also known as *Knowledge Discovery in Databases (KDD)* or *Data Mining* systems. "*Knowledge Discovery in Databases – KDD*" is a promising research domain and in many cases can bring positive rewards in scientific and business context [2, 3].

Research in Knowledge Discovery Systems is in many cases concentrated on facets of manual implementation of basic machine learning and discovery methods concerning data storage in relational databases [4]. Knowledge Discovery Systems cover two subprocesses closely related to knowledge discovery: (i) *combination* that enables the discovery of new forms of explicit knowledge, (ii) *socialization* that helps to synthesize tacit knowledge and therefore acting as a vehicle to the desirable destination of discovering of new implicit knowledge [4–6]. Knowledge Discovery Systems are useful for analyzing raw data, finding the similar features and relationships among attributes that represent the gained knowledge, and presenting them in an understandable and readable format [3–6].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 150–159, 2024.
https://doi.org/10.1007/978-3-031-53549-9_16

Decision tree is a technique that helps to model discovered knowledge. Through this work specific algorithms for learning decision trees are examined with main purpose to study the searching and modeling forms of knowledge stored in databases [7].

2 Knowledge Discovery Systems

Knowledge Discovery Systems (KDS) are software systems designed to extract useful knowledge and insights from large volumes of data [14]. KDS use a variety of techniques from machine learning, data mining [18], and artificial intelligence [6, 19–22] to analyze data and uncover patterns. These techniques include clustering, classification, regression analysis, association rule mining, and anomaly detection.

The process of knowledge discovery typically involves several steps, including data selection and cleaning, data transformation and preprocessing, pattern discovery, and knowledge representation and visualization. KDS provide tools and algorithms to support each of these steps, allowing analysts to quickly and efficiently extract valuable insights from their data. KDS use a combination of algorithms and statistical techniques to analyze data and identify patterns, trends, and relationships that may not be readily apparent. The process of knowledge discovery involves several steps, including data cleaning and preparation, data mining, and the interpretation and evaluation of the results. Some common techniques used in KDS include clustering, classification, association rule mining, and anomaly detection. These techniques can be applied to a wide range of data sources, including structured data in databases, unstructured data in text documents, and semi-structured data in social media.

Knowledge discovery systems (KDS) play an important role in enabling data-driven decision making and facilitating the discovery of new knowledge and insights from complex data sets. KDS are becoming increasingly important as organizations seek to leverage the vast amounts of data, they collect to gain a competitive advantage.

However, there are also concerns about privacy and ethical issues surrounding the use of personal data in these systems. Some examples are the following [17]:

1. **Privacy**: KDS often rely on data that includes personal information, and there is a risk that this information could be misused or exploited. It's important to ensure that data is collected and used in an ethical and responsible manner, with appropriate consent and privacy safeguards in place.
2. **Bias**: KDS may generate biased results if the data used to train them is biased or if the algorithms used are biased. This can result in discriminatory outcomes or perpetuate existing inequalities. It's important to carefully evaluate the data used and the algorithms employed to identify and mitigate bias.
3. **Transparency**: KDS can be difficult to understand, particularly if they use complex algorithms. It's important to make sure that the results generated by KDS are transparent and explainable, so that users can understand how they were arrived at and make informed decisions.
4. **Accountability**: KDS can have significant impacts on individuals and society. It's important to ensure that there is accountability for the decisions made based on KDS-generated insights and that those decisions are transparent, fair, and just.

In summary, while KDS offer many potential benefits, it's important to consider the ethical implications of their use and ensure that they are designed and implemented in an ethical and responsible manner. This includes taking steps to safeguard privacy, identify and mitigate bias, ensure transparency, and establish accountability.

3 Decision Trees

Knowledge discovery systems (KDS) are computer programs that use data mining and machine learning techniques to extract knowledge and insights from large and complex datasets. One common technique used in KDS is the decision tree [7, 9]. A decision tree is a hierarchical model that represents a sequence of decisions and their possible consequences. Each decision node in the tree represents a particular attribute of the data, and each leaf node represents a decision or outcome. The goal of a decision tree algorithm is to find the best split points for each attribute, such that the resulting tree maximizes the classification or prediction accuracy.

Decision tree is a technique that helps to model discovered knowledge. It allows to present dependencies between elements' properties, and the decision representing the class within each element is included [9, 10]. Dependencies are shown as in the form of a tree graph without the existence of any cycles. The decision tree contains nodes that generate a *rooted tree*, and this is logical because it is a *directed tree*, starting from the basic node called as "*root*" that has no incoming edges. Every other node has one accurate incoming edge. A node with no leaving edges is called as *internal* node. Every other node is called a leaf or decision node. A leaf represents the decision-class according to the value of properties on the route starting from the root node to a specific leaf. Every edge in the graph, has a label indicating a value of the property of each node.

A decision tree is a way of classifying a recursive part of the instant space [9]. Each decision tree is generated in a recursive way by using a learning sample called as *training stage* [8]. In each recursive subroutine the learning sample (at first stage) or a partition of this learning sample (in steps in succession) is divided into smaller entities according to the value of the property that is considered as the most suitable. The most suitable has the meaning of seeking the property value that maximizes the "*knowledge or information gain*" over the present partition of the sample. There are tools to evaluate the "*knowledge or information gain*" and specifically, these tools are functions [11]. A lot of these functions fit best to the algorithm that is chosen to generate the tree from the input data. The maximization of "*knowledge gain*" property leads to the shorter route in the graph starting from the root to the next closer node. The decision tree that is created, can be used to exploit the gained knowledge to make predictions and classify various cases (not known in the learning process) according to specific criteria that can be applied to find a quick decision based on classification of the property values [12]. Classifying begins from the root and steps down the tree until a leaf is met. At each node that is not a leaf, the case's result is the decision that needs to be taken and continue to the process of comparison to the root node of the subtree according to the previously taken decision [12].

An important question that comes next is to answer through the process of evaluation if a tree is functional or without practical use. Decision trees are often used in KDS

because they are easy to interpret and can provide valuable insights into the relationships between different attributes in a dataset. They can also be used for both classification and regression tasks, making them versatile tools for analyzing different types of data [12]. There are methods to help in the process to rate the total classification error for not known cases that is called *percentage of misclassified cases*. Two well-known methods that are proven to provide the best outcomes are *independent testing sample* and *cross-validation*. The first method is implemented to an independent testing sample with a predetermined class property for every case and detects the count of misclassified cases after the comparison of the true class value with class prediction derived from the tree. The second method is more knowledgeable, due to the calculation of estimation error using only the learning sample. At first stage, it divides by randomness, the sample on N subsets. Next, tree is created N times using N-1 subsets, and each of N time another subset is separated. The subset that has not been included to the process of constructing the tree is an independent testing sample and partial classification error can be estimated by implementing independent testing sample. Cross-validation classification error is the average of partial classification errors. The second method is implemented when separate testing sample is out of availability.

Any route in the decision tree is an accurate hypothesis representing the subset of learning sample [9]. During the training stage, an algorithm is generating a hypothesis that corresponds to the learning sample. This results to significant errors during the process of classification of the new cases. The above-mentioned refer to the *pruning process* – lessen fitting to learning sample and not enlarging the total classification error for the not known cases simultaneously. Some popular decision tree algorithms used in KDS include ID3, C4.5, and CART [13]. These algorithms differ in their splitting criteria, pruning techniques, and other parameters, and may perform differently depending on the specific dataset.

4 Decision Tree Algorithms

ID3, C4.5, and CART are all algorithms used for building decision trees, which are a type of machine learning model that can be used for classification or regression tasks [13, 15–16].

ID3 (**Iterative Dichotomiser 3**) was developed in the 1980s by Ross Quinlan and was one of the first decision tree algorithms [15]. ID3 uses a heuristic approach to recursively split the data into subsets based on the attribute that provides the most *information or knowledge gain*. The algorithm stops when all the data is classified into homogeneous classes or when no more attributes can be selected to split the data. It works by recursively partitioning the data based on the attribute that provides the most *information or knowledge gain* (i.e., the attribute that best separates the classes) at each node. ID3 is limited in that it can only handle categorical data and does not handle missing data well.

C4.5 is an extension of ID3 that was also developed by Ross Quinlan [14]. It can handle both categorical and continuous data and can handle missing data by using surrogate splits. C4.5 works by recursively partitioning the data based on the attribute that provides the highest gain ratio (i.e., the attribute that best separates the classes while also taking into account the number of possible outcomes for each attribute) at each

node. C4.5 also has a better handling of missing data and incorporates pruning to reduce overfitting. Additionally, as mentioned above, C4.5 introduces the concept of gain ratio, which measures the information gain relative to the intrinsic information of the attribute.

CART (**Classification and Regression Trees**) is a decision tree algorithm that was developed by Leo Breiman, Jerome Friedman, Richard Olshen, and Charles Stone in the 1980s [16]. CART can be used for both classification and regression tasks and handles both continuous and categorical variables. Unlike ID3 and C4.5, CART constructs binary trees, where each internal node corresponds to a binary decision, and each leaf node corresponds to a class or a numeric value. CART can handle both categorical and continuous data and can handle missing data. CART works by recursively partitioning the data based on the attribute that provides the best split, which is defined as the split that minimizes the impurity of the resulting subsets (i.e., the split that best separates the classes for classification tasks or minimizes the variance of the targets for regression tasks).

All three algorithms are used for knowledge discovery and decision making in various fields, including engineering, medicine, and business. However, there are newer algorithms that have been developed since ID3, C4.5, and CART that can handle more complex data and are more accurate.

The following flow diagrams provide a general overview of the three algorithms.

1. **ID3 Flow Diagram:**

   START
   1. Select the best attribute (based on information or knowledge gain)
   2. Create a new branch for each possible value of the selected attribute
   3. Recursively apply steps 1-2 to each subset until all instances belong to a single class or a stopping criterion is met
   END

2. **C4.5 Flow Diagram:**

   START
   1. Select the best attribute (based on information gain ratio)
   2. Create a new branch for each possible value of the selected attribute
   3. Prune the tree to improve accuracy on a validation set
   4. Recursively apply steps 1-3 to each new branch until all instances belong to a single class or no more attributes are left
   END

3. **CART Flow Diagram:**

   START
   1. Select the best attribute to split the data
   2. Divide the data into two subsets based on the selected attribute
   3. Recursively apply steps 1-2 to each subset until all instances belong to a single class or a stopping criterion is met
   4. Prune the tree to improve accuracy on a validation set
   END

In terms of implementation, these algorithms are available in various programming languages, such as Python and R. There are also several open-source machine learning libraries that implement these algorithms, such as scikit-learn in Python and caret in R. These libraries provide pre-built functions to train and evaluate decision trees, making it easier to implement these algorithms in practice. Specifically, there are several Python libraries that can be used to implement the ID3, C4.5, and CART algorithms for knowledge discovery. Below are some of the commonly used libraries and their implementation examples:

1. **Scikit-learn**: Scikit-learn is a popular machine learning library that provides efficient tools for data analysis and modeling. It includes DecisionTreeClassifier class that can be used to implement decision tree algorithms such as ID3, C4.5, and CART. Here is an example of using DecisionTreeClassifier for ID3 algorithm in Python programming language:

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.datasets import load_iris

iris = load_iris()

X = iris.data
y = iris.target

clf = DecisionTreeClassifier(criterion="entropy")
clf.fit(X, y)
```

2. **Orange**: Orange is an open-source data analysis and visualization tool that includes several machine learning algorithms, including decision trees. Here is an example of using Orange for C4.5 algorithm in Python programming language:

```
import Orange

data = Orange.data.Table("iris.tab")
tree = Orange.classification.tree.TreeLearner(data, algorithm="c45")
```

3. **PyDatalog**: PyDatalog is a Python library that can be used for rule-based programming and logical inference. It includes an implementation of the CART algorithm for decision trees. Here is an example of using PyDatalog for CART algorithm:

```
from pyDatalog import pyDatalog
from pyDatalog import functions as fn

pyDatalog.create_terms('data, Tree')
data = [(1, 'A', 10), (2, 'B', 20), (3, 'A', 30), (4, 'B', 40), (5, 'A', 50)]

# Define the variables
Tree.decision['attribute']=fn.minimal_hamming_tree(data,
Tree.attribute, Tree.value)

# Print the decision tree
print(Tree.decision[None])
```

5 Knowledge Discovery Systems' Significant Role in a Data-Driven World

Knowledge Discovery Systems (KDS) are computer-based tools and techniques used to analyze data and extract useful knowledge from it [6, 9, 23–27]. They play a significant role in today's data-driven world and use various techniques and algorithms to discover patterns, trends, and insights in large data sets. The primary goal of KDS is to transform raw data into useful and actionable knowledge that can inform decision-making processes.

KDS are significant for various reasons, including:

1. **Improved decision-making**: KDS helps organizations to make informed decisions by identifying hidden patterns and insights that may not be apparent through manual analysis.
2. **Data analysis**: KDS can analyze large datasets quickly and accurately, enabling organizations to gain insights into their operations, customers, and markets.
3. **Prediction**: KDS can predict future trends and behavior patterns based on historical data, enabling organizations to make informed decisions and develop effective strategies.
4. **Optimization**: KDS can identify patterns and relationships in data that can be used to optimize business processes, reduce costs, and improve efficiency.
5. **Innovation**: KDS can help organizations identify new opportunities and develop innovative products and services based on customer needs and preferences.
6. **Personalization**: KDS can be used to personalize customer experiences, such as recommending products or services based on their past purchases or browsing history.
7. **Increased efficiency**: KDS automates the process of data analysis, which saves time and resources that would otherwise be spent on manual analysis.
8. **Competitive advantage**: KDS helps organizations gain a competitive advantage by identifying opportunities and trends before their competitors.
9. **Better customer service**: KDS can analyze customer data to identify their preferences and behaviors, which can inform targeted marketing campaigns and improve customer service.

10. **Scientific discoveries**: KDS can be used in scientific research to identify patterns and trends in data, leading to new discoveries and innovations.

6 Discussion

KDS has become an essential tool in various fields, including science, healthcare, finance, and marketing, as it enables organizations to make better-informed decisions and gain a competitive edge. Knowledge discovery systems are powerful tools for extracting useful knowledge and insights from large volumes of data. Here are some key conclusions about these systems:

Knowledge discovery systems can be used in a wide range of applications, including business intelligence, scientific research, healthcare, and social media analysis. These systems use a variety of techniques, such as data mining, machine learning, and natural language processing, to identify patterns and relationships in data. Knowledge discovery systems require high-quality data to produce accurate and meaningful results. Data cleansing and preprocessing are critical steps in the knowledge discovery process.

The results produced by knowledge discovery systems are only as good as the questions asked and the data used. Careful planning and analysis are required to ensure that the right questions are being asked and the appropriate data sources are being used.

The insights generated by knowledge discovery systems can be used to inform decision-making, improve operational efficiency, and drive innovation.

In addition, ethical considerations are important when using knowledge discovery systems, particularly when dealing with sensitive or personal data. It is essential to protect the privacy and security of individuals and ensure that the results generated are not biased or discriminatory.

7 Conclusion

Knowledge Discovery Systems (KDS) are mainly used to extract useful knowledge from large datasets. Specifically, these systems are mainly used in cases where it is needed to discover trends, patterns, and relationships within data that may not be obvious at first sight. To achieve the above-mentioned, Knowledge Discovery Systems exploit basic techniques of machine learning, data mining, database management and statistics.

One common technique used in KDS is the decision tree and the ID3, C4.5, and CART algorithms which are used for building decision trees were described in detail through this paper. In terms of implementation, these algorithms are available in programming languages such as Python or R programming language. There are also several open-source machine learning libraries that implement these algorithms, such as scikit-learn in Python and caret in R which were described in detail through this paper.

To conclude, Knowledge Discovery Systems play a significant role in various areas, as they give organizations' the opportunity to gain important knowledge for decision-making from complex datasets. Thus, continuous research on this subject area seems necessary and future studies and their findings will contribute to a better understanding of this continuously evolving domain.

References

1. Silberschatz, A., Tuzhilin, A.: What makes patterns interesting in knowledge discovery systems. IEEE Trans. Knowl. Data Eng. **8**(6), 970–974 (1996). https://doi.org/10.1109/69.553165
2. Matheus, C.J., Chan, P.K., Piatetsky-Shapiro, G.: Systems for knowledge discovery in databases. IEEE Trans. Knowl. Data Eng. **5**(6), 903–913 (1993). https://doi.org/10.1109/69.250073
3. Piatetsky-Shapiro, G., Frawley, W.J. (eds.): Knowledge Discovery in Databases. AAAIIMIT, Cambridge, MA (1991)
4. Piatetsky-Shapiro, G. (ed.): Special issue on: knowledge discovery in data and knowledge bases. Int. J. Intell. Syst. **7**(7) (1992)
5. Becerra-Fernandez, I., Sabherwal, R.: Knowledge Management: Systems and Processes. M.E. Sharpe, Armonk (N.Y.), London (2010)
6. Triantafyllou, S.A.: Work in progress: educational technology and knowledge tracing models. In: 2022 IEEE World Engineering Education Conference (EDUNINE) (2022). https://doi.org/10.1109/edunine53672.2022.9782335
7. Fayyad, U.M., Irani, K.B.: Multi-interval discretization of continuous-valued attributes for classification learning. Artif. Intell. **13**, 1022–1027 (1993)
8. Murthy, S.K.: Automatic construction of decision trees from data: a multi-disciplinary survey. Data Min. Knowl. Discov. **2**(4), 345–389 (1998)
9. Rokach, L., Maimon, O.: Decision Trees. In: Maimon, O., Rokach, L. (eds.) Data Mining and Knowledge Discovery Handbook. Springer, Boston, MA (2005). https://doi.org/10.1007/0-387-25465-X_9
10. Quinlan, J.R.: Decision trees and multivalued attributes. In: Richards, J. (ed.), Machine Intelligence, vol. 11, pp. 305–318, Oxford, England, Oxford Univ. Press (1988)
11. Elwakil, E., Zayed, T.: Construction knowledge discovery system using fuzzy approach. Can. J. Civ. Eng. **42**(1), 22–32 (2015)
12. Cios, K.J., Pedrycz, W., Swiniarski, R.W.: Data Mining Methods for Knowledge Discovery, vol. 458. Springer, New York (2012). https://doi.org/10.1007/0-387-25465-X_9
13. Gupta, S.K., Suresh, P., Bhatanagar, V.: Incremental information mining. In Data Mining and Knowledge Discovery: Theory, Tools, and Technology IV, vol. 4730, pp. 448–459. SPIE (2002)
14. Dunkel, B., Soparkar, N., Szaro, J., Uthurusamy, R.: Systems for KDD: From concepts to practice. Futur. Gener. Comput. Syst. **13**(2–3), 231–242 (1997)
15. Hssina, B., Merbouha, A., Ezzikouri, H., Erritali, M.: A comparative study of de-cision tree ID3 and C4. 5. Int. J. Adv. Comput. Sci. Appl. **4**(2), 13–19 (2014)
16. Breiman, L.: Classification and Regression Trees. Routledge, London (2017)
17. Brankovic, L., Estivill-Castro, V.: Privacy issues in knowledge discovery and data mining. In: Australian Institute of Computer Ethics Conference, pp. 89–99 (1999)
18. Domdouzis, K., Akhgar, B., Andrews, S., Gibson, H., Hirsch, L.: A social media and crowd-sourcing data mining system for crime prevention during and post-crisis situa-tions. J. Syst. Inf. Technol. **18**(4), 364–382 (2016)
19. Triantafyllou, S.A.: Magic squares in order 4K+2. In: 2022 30th National Conference with International Participation (TELECOM) (2022). https://doi.org/10.1109/TELECOM56127.2022.10017312
20. Triantafyllou, S., Georgiadis, C.: Gamification of Moocs and security awareness in corporate training. In: Proceedings of the 14th International Conference on Computer Supported Education (2022). https://doi.org/10.5220/0011103000003182

21. Triantafyllou, S.A.: A quantitative research about MOOCs and EdTech tools for distance learning. In: Auer, M.E., El-Seoud, S.A., Karam, O.H. (eds.) Artificial Intelligence and Online Engineering. REV 2022. LNNS, vol. 524, pp. 521–532. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-17091-1_52
22. Triantafyllou, S.A.: TPACK and Toondoo digital storytelling tool transform teaching and learning. In: Florez, H., Gomez, H. (eds.) Applied Informatics. ICAI 2022. CCIS, vol. 1643, pp. 338–350. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19647-8_24
23. Stergiou, C.L., Psannis, K.E., Gupta, B.B.: InFeMo: flexible big data manage-ment through a federated cloud system. ACM Trans. Internet Technol. **22**(2), 1–22 (2021). https://doi.org/10.1145/3426972
24. Demetriou, A., Efklides, A., Papadaki, M., Papantoniou, G., Economou, A.: Structure and development of causal experimental thought: from early adolescence to youth. Dev. Psychol. **29**(3), 480–497 (1993). https://doi.org/10.1037/0012-1649.29.3.480
25. Triantafyllou, S.A., Sapounidis, T.: Game-based learning approach and serious games to learn while you play. In: 2023 IEEE World Engineering Education Conference (EDUNINE) (2023). https://doi.org/10.1109/EDUNINE57531.2023.10102872
26. Triantafyllou, S.A., Georgiadis, C.K.: Gamification Design Patterns for user engagement. Inform. Educ. 655–674 (2022). https://doi.org/10.15388/infedu.2022.27
27. Triantafyllou, S.A.: A detailed study on the 8 Queens problem based on algorithmic approaches implemented in PASCAL programming language. In: Silhavy, R., Silhavy, P. (eds.) Software Engineering Research in System Science. CSOC 2023. LNNS, vol. 722, pp. 160–173. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35311-6_18



Novel Radio Scheduling Framework for Optimal Energy Efficiency in Wireless Sensor Network

K. Deepa Mathew(✉) and T. Anita Jones Mary Pushpa(✉)

Karunya University, Coimbatore, India
deepamathewk@gmail.com, tajmp8576@gmail.com

Abstract. Wireless Sensor Network (WSN) has been penetrating deeper into various commercial monitoring services/application which demands substantial amount of residual energy to assure superior network lifetime. In this perspective, there are various Medium Access Control (MAC) as well as Carrier Sense Multiple Access (CSMA) scheme which is frequently adopted to address this problem; however, they too have shortcomings. Hence, this paper introduces a novel computational model of radio scheduling where a network model and neighborhood exploration are carried out using graph-based method. Further, the scheme introduces a novel management of routing to effectively control timeslots and message. Further, a simplified packet prioritization scheme is implemented balancing the demands of normal and urgent services. The simulation outcome shows that proposed scheme excels better data transmission performance in contrast to existing MAC and CSMA-based schemes.

Keywords: Medium Access Control · Carrier Sense Multiple Access · Wireless Sensor Network · energy · radio scheduling

1 Introduction

Wireless Sensor Network (WSN) is increasingly adopted in commercial application that demands seamless monitoring of environment without human intervention [1]. In order to do this continuous process of monitoring, a sensor is required to possess a maximum lifetime which unfortunately is one the critical challenges in WSN apart from limited processing capability and restricted availability of resources [2]. There are multiple reasons towards energy consumption in WSN e.g., communication, data aggregation, routing protocol processing, network topology, computation and processing of data, scheduling of sleep and wake state, and various environmental factors [3–5]. In order to deal with this issue, there are various energy-based routing schemes being introduced [6–10], which is characterized by certain beneficial features as well as shortcomings too. In this perspective, Medium Access Control (MAC) based approaches are widely used to address specifically the energy problems in WSN [11–15]. Majority of the existing MAC-based scheme primarily targets towards radio scheduling with advantages of energy efficiency, scalability, collision avoidance, latency reduction, and adaptability. However, they are also characterized by shortcomings e.g., overhead, complexity, fairness issues,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 160–171, 2024.
https://doi.org/10.1007/978-3-031-53549-9_17

demands of synchronization, and security considerations. In order to develop a robust radio scheduling approach, there is a need to develop a completely revised version of existing MAC scheme suiting with the dynamic demands of nodes in WSN, which is not much witnessed to be reported in existing scheme.

Therefore, the proposed study presents a novel radio scheduling approach that emphasize on improving the performance of data transmission in energy efficient way for WSN. The organization of the manuscript is: Sect. 2 highlights existing MAC-based methodologies adopted for radio scheduling in WSN followed by highlights of identified research gap in Sect. 3. Proposed Methodology is briefed in Sect. 4 while system implementation is discussed in Sect. 4. Result obtained is present in Sect. 5 and Sect. 6 gives summary of this paper.

2 Related Work

Various existing research methodologies towards an effective radio scheduling in WSN has been reviewed in our prior study [16]. Sakib et al. [17] have presented a MAC scheme where multihop routing is considered for transmitting prioritized data in WSN for ensuring better forwarding of multiple packets and limiting idle listening time. Hai et al. [18] have used a hybrid scheme towards improving the network lifetime of WSN using MAC scheme with low latency. Further, the work of Tomovic and Radusinovic [19] have used deep learning scheme to develop an adaptive MAC scheme that is capable of learning an efficient transmission strategy without dependency of detailed environmental data. Similar set of adoption of deep learning is also witnessed in work of Fu and Kim [20], Sarang et al. [21] as well as Kherbache et al. [22]. Adoption of CSMA scheme towards collision avoidance is reported by Miskowicz [23] which is also used for formulating bandwidth allocation strategy in local operating networks. The work carried out by Su et al. [24] have presented a unique MAC scheme considering channel condition using Markov process in order to allocate optimal channel resource and minimize rate of collision. Adopting a mobility case study, Tripti and Jibukumar [25] have presented a priority-based MAC scheme with cooperative feature towards resisting forward collision. Adoption of substitutional revised version of CSMA is reported by Khun et al. [26] towards data transmission in Internet-of-Things for higher throughput. The work carried out by Sadeq et al. [27] have developed a conceptual model to leverage MAC scheme for energy efficiency in WSN applications. Zhang et al. [28] have presented discussion about MAC protocols and stated its scope towards varied sensory application using multi-channel and multi radio frequencies for better data transmission. Uthayakumar et al. [29] have presented a framework discussion which improvises the existing MAC protocol for accomplishing energy efficiency in WSN. Pradhan and Chaudhari [30] have presented a scheduling approach in autonomous mode in large sensory network where slot frames are divided into segments followed by allocating them to nodes. Subramanyam et al. [31] have presented another MAC scheme where duty cycle is used in variable form in order to acquire better network lifetime in adaptive manner. MAC-based scheme is also discussed in unique way by Thippun et al. [32] and Khalifeh et al. [33] while study towards CSMA-based scheme is reported in work of Nisha et al. [34] and Zhong et al. [35] towards energy optimization.

3 Problem Description

The identified problem associated with existing radio scheduling approaches using varied MAC-based schemes are as follows: i) existing radio scheduling offers an emphasis towards static slot management with less flexibility when exposed to dynamic network, ii) irrespective of various energy-efficient MAC scheme, the models lacks consideration of attributes required to address the dynamicity problems in WSN, iii) although throughput is found improved in existing scheme but there is no significant improvement of delay attribute over uncertain traffic circumstances, iv) there are no explicit schemes of radio scheduling towards considering the mobility attribute of sink node which makes it less applicable for large scale sensory application that uses mobility towards data aggregation. The next section discusses about the methodology adopted as a solution to above-mentioned research problems.

4 Proposed Methodology

The proposed implementation study model is an extension of our prior research model [36] where the current study objective to introduce a simplified and yet novel framework towards radio scheduling in WSN considering the optimization of data transmission over complex network condition. An analytical research methodology is developed for this purpose that is exhibited in Fig. 1.

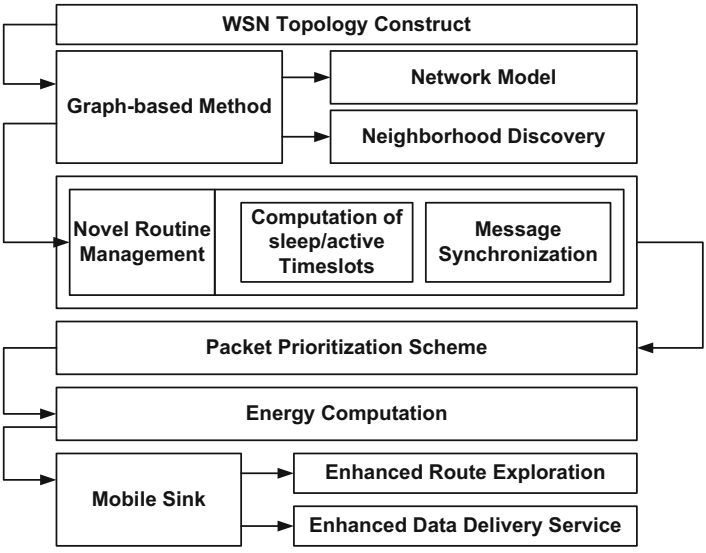


Fig. 1. Proposed Architecture of Radio Scheduling in WSN

The specification of the techniques adopted for the implementation approach is as following:

- The first operation of the proposed methodology is to construct a novel *WSN topology* where a *graph-based method* is adopted to construct the topology of WSN which is meant for representing both staticness and mobility perspective with response to node and sink respectively.

  - The graph-based method is used for developing a *network model* and *neighborhood discovery process.* The network model is meant for constructing a novel form of a cluster defined by user on the basis of application that is completely different from conventional clustering mechanism. Each node in the cluster is then followed by establishment of communication vector that leads to sink node. Hence, a single-channel based connectivity is formulated for each cluster nodes which performs aggregation and forward the aggregated data to the sink node.
  - The *neighborhood discovery model* is meant for acquiring information of all the sensors situated in the vicinity of each other's transmission range in order to formulate an optimal chain of well-connected nodes.

- In the second part of operation, a *novel routine management* is formulated by harnessing the MAC scheme where a computation towards *timeslots* and *message synchronization* is carried out. The term 'routine management' refers to organization of the timeslots meant for the node to be sleep or awake state.

  - The system performs *computation of either active or sleep state of timeslots* which are required to be allocated and finetuned during every routing management operation.
  - The task of *message synchronization* is meant for internal structuring of the control message towards scheduling the resources during hop-to-hop communication.

- The operational block of *Packet Prioritization Scheme* is another set of novelty which allocates a priority of varied ranges towards radio scheduling among sensor nodes. The user set the priority of packet by embedding 1-bit message within the packet that distinguishes it from rest other form of packets disseminated in network.
- Further, a standard *energy modelling* is carried out towards data transmission on mobile sink with an objective of enhanced exploration of route and enhanced delivery of aggregated data.
- In the last part of implementation, the nodes forward the data to the *mobile sink* adhering to the proposed radio scheduling approach with novel routine management.

  - *Enhance route exploration* process refers to operation carried out by nodes to find out the best possibilities of communication vector with reduced energy consumption
  - *Enhanced data delivery service* refers to mechanism of data transmission carried out towards balancing both computational demands and communication demands in WSN.

Further elaboration of the above-mentioned methodologies is carried out in next section towards system design involved in proposed scheme.

5 System Design

The system design adopted in proposed scheme emphasizes on a unique operation of radio scheduling in WSN to improve the overall data delivery performance. This agenda is accomplished by developing a primary and secondary radio scheduling approach where the former one focuses on baseline design construction while the latter one focuses on optimizing the baseline design. Following is the discussion towards individual system design:

5.1 Primary Radio Scheduling Approach

This is the primary scheduling model where the sensor nodes are widely distributed in the simulation area with a variable possibility of location of the sink node. The first stage of implementation of this module is associated with the *discovery of the adjacent nodes* in order to accomplish an initial communication. The second stage of implementation is associated with the *computation of the time-slots* for sleep and awake for the sensor node associated with the MAC protocol. The scheme incorporates a cut-off in duty cycle in order to sync multiple forms of frame scheduling associated with the phases of activeness or passiveness of state of sensor node. The third stage of implementation performs *synchronization of the message* followed by *establishing routes*. The process of synchronization of message is carried out by unique form of clustering different from all conventional clustering approach. In this scheme, the clusters are formulated for all sensors who have similar network and physical properties. It will mean that proposed scheme formulates a heterogeneous cluster of varied nodes where all heterogeneous nodes are randomly dispersed in simulation area unlike the existing clustering where all homogeneous nodes combined in circular form to generate a cluster. It is imperative that such form of circular formation of cluster is quite impractical in real-life random deployment of WSN and hence this form of clustering presented in proposed scheme makes more sense. The consecutive steps to formulate a direct communication vector from all the nodes indexed with specific cluster performs data aggregation towards the sink node. The fourth stage of implementation is associated with the *energy modelling* in proposed scheme where an inter-arrival time is computed by differentiating timestamp for route discovery message and timestamp of acknowledgement received by node. This inter-arrival time is used for controlling the flow of the message which has a positive influence on transmittance energy. The final step of implementation is associated with *prioritizing the data transmission* while performing radio scheduling in WSN. A multi-classification scheme of high, medium, and low priority is defined for the data transmission on the basis of priority set by user by adding 1 bit to control message. Interestingly, the proposed scheme uses multiple communication channels to perform propagation of data in such a way that high prioritized data are forwarded in one channel and less prioritized data are forwarded in different channel on the basis of size of queue in each channel. This adoption permits the data transmission to have an equal emphasis towards both high and low/medium prioritized data in entire network thereby offering a consistency in rate of data transmission performing in proposed radio scheduling.

5.2 Secondary Radio Scheduling Approach

This is the consecutive framework developed on the top of prior primary framework with an agenda of optimizing the performance of proposed radio scheduling in WSN. Considering the input of total sensors, time-to-live, and size of packet, the first process is towards performing *optimization of the route buffers* to yield an outcome of final route buffer to be used. Route buffers refers to a matrix consisting of all the routes to be used for data transmission in WSN to be carried out depending upon the available energy of sensors. This is quite useful towards route exploration during a heavy traffic condition in WSN. The second process is associated with the *optimization of energy of buffer* module considering a feed of residual energy level. This process contributes towards addressing excessive consumption of power in frequently used CSMA. Further, the process generates a matrix which store the information of energy efficient routes that is helpful especially in presence of mobile sink. The third process involved in the secondary radio scheduling is towards *organizing updates of dynamic locations* of nodes in WSN which takes the input feed of location of sensors, energy level, and route buffer obtained from first process to finally generate updated information of dynamic location. The novelty of this approach is the newly incorporated auxiliary node which can access the location information of nodes and performs delivery of the finally aggregated data thereby minimizing the computational burden associated with conventional data aggregation scheme in WSN. The final process is to perform *communication management* which takes all the same input like the prior process along with update of dynamic location in order to yield a communication vector that is used for routing in WSN. In this stage of processing, the system generates a new optimal position of node followed by computing the distance among them. Further, the proposed scheme addresses the issues of scalability problem in CSMA by generating a unique communication channel that doesn't possess any form of redundant routes leading to faster data propagation. This also leads to faster processing of routes which overcomes the problems of long waiting time associated with conventional CSMA based schemes. The scheme further facilitates evaluation of distinctive routes which are reduced in number compared to sensor node present. This conditional assessment is incorporated in order to monitor the number of generated distinctive communication channel information that will be further subjected to reevaluation process to narrow down the list of optimal channels.

Therefore, the proposed design methodology is essentially meant towards considering power and buffer for route information which offers a novel solution towards hosting a greater number of analytical operations in the process of route exploration while performing radio scheduling. Another bigger contribution of proposed scheme is towards addressing the power dissipation and channel quality issues associated with frequently used CSMA scheme in WSN. Further, incorporation of dynamic location information and its updating process make the proposed study model application to a network system using mobile sink. Hence, a closer look into the proposed study model exhibit that it offers a good balance between routing performance and computational demands required during radio scheduling in WSN in cost effective manner. The next section outlines the results accomplished in this process.

6 Results Discussion

The implementation of the proposed scheme is carried out considering 1000 sensor nodes in 1200 x 1100 m^2 simulation area with an initialized energy of 10J for each node. All the nodes are randomly distributed while the sink is positioned at the peripheral position of simulation area. Each nodes uses MAC 802.11 standard with omni-directional antenna. The scripting of the proposed implementation is carried out in MATLAB on normal windows system environment of 64-bit. The primary analysis is carried out for delay with respect to arrival time. As the proposed scheme introduces a scheduling of forwarding data that is controlled by arrival time; hence, observation is carried out towards assessing increasing arrival time and its possible effect on delay (Fig. 2). Further, the analysis is carried out with existing standard of scheduling using two frequently adopted MAC scheme i.e., S-MAC and Q-MAC. Similar observation is also carried out for energy consumption on similar test environment as noted in Fig. 3

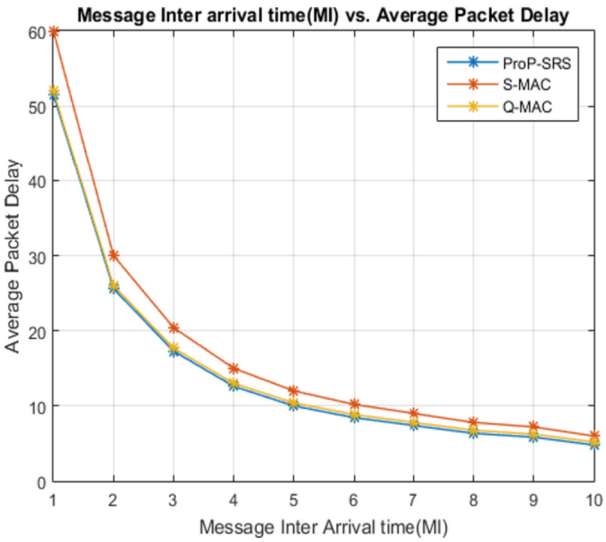


Fig. 2. Comparative Analysis of Delay

The outcome in Fig. 2 exhibits that proposed scheme ProP offers reduced delay in contrast to existing S-MAC as well as Q-MAC. The rationale behind this outcome is that although S-MAC can offer better scalability in radio scheduling performance by handling large sensors but the mechanism of sleep scheduling S-MAC deliberately introduces extra latency owing to its dependency towards synchronization with sensors. Further, the performance of proposed scheme ProP and Q-MAC is found to be nearly equal, but proposed scheme excel better than Q-MAC in closer observation of the trends. Although Q-MAC can facilitate better management of priority queues but it introduces multiple priority queues which are difficult to manage leading to higher delay then proposed scheme.

The trend of outcome exhibited in Fig. 3 is nearly same as that of Fig. 2 with respect to energy utilized per bit of data transmission. The reason behind this is S-MAC assists in extending the networking lifetime of sensors by minimizing the collision and idle listening; however, exchange of control beacons in S-MAC induces additional deployment of resources leading to extensive energy utilization. On the other hand, Q-MAC performs assignment of timeslots on the basis of priority set to address contention issue, but its performance degrades when the network starts increasing in its dimension.

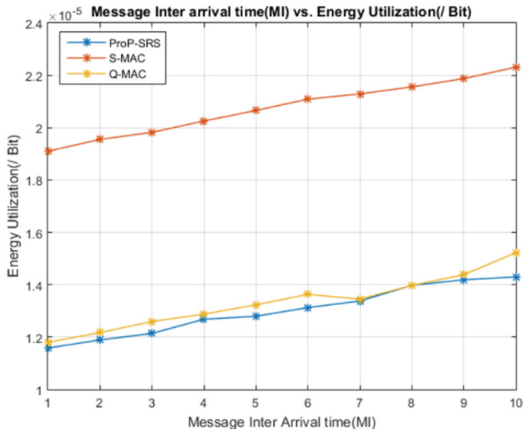


Fig. 3. Comparative Analysis of Energy w.r.t. Inter-arrival time

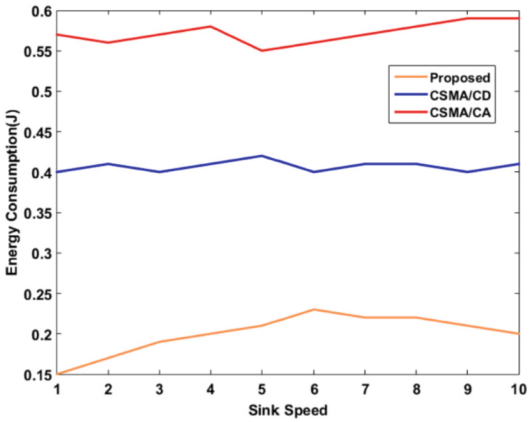


Fig. 4. Comparative Analysis of Energy with respect to sink speed

Based on this observation, it can be stated that proposed scheme performs better delay minimization and reduced energy utilization in comparison to existing Q-MAC and S-MAC protocol. However, it is equally important to assess the contribution of proposed scheme with existing CSMA scheme both with respect to collision detection

(CSMA/CD) and collision avoidance (CSMA/CA). As the proposed scheme is an improvisation over conventional CSMA scheme, hence this comparative analysis is important. Further, the analysis is carried out in presence of sink node mobility to investigate the impact of increasing sink node mobility over energy consumption.

Figure 4 showcase that proposed scheme ProP offers reduced energy consumption in comparison to both the existing CSMA scheme. The existing CSMA/CD scheme facilitate usage of same channel for data transmission in its idle mode and hence offers better consistency in energy consumption curve; however, the collision probability rises when it is exposed to larger area. On the other hand, CSMA/CA schemes offers better acknowledgement mechanism as well as better reliability in data transmission to retain better signal quality. It also offers minimization of collision, but it incorporates an extra computational dependency as well as burden while it carry out channel sensing. However, proposed system is not witnessed with any of such flaws, on the contrary, it offers a simplified computational scheme which uses simplified parameters to perform routing considering the uniqueness of routing and node information and resist encountering any condition that might affect the data transmission in peak traffic condition in WSN. Hence, the radio scheduling approach presented in proposed scheme offers better performance in contrast to existing schemes frequently reported. The discussion of the outcome accomplished is carried out as follows:

- *Network Lifetime*: This parameter is used to evaluate the duration for which the sensor network can work effectively without any need of maintenance or replacement of power source. This parameter can be justified from graphical outcome of Fig. 4 exhibiting proposed scheme to have considerably lesser extent of energy consumption (J) compared to both the variants of conventional CSMA schemes. The prime reason behind this is CSMA/CD scheme is more likely to exhibit retransmission on increasing traffic condition leading to spontaneous degradation of energy; while CSMA.DA scheme introduces additional overhead while performing detection leading to reduce energy retention among the sensors. Proposed scheme mitigates both this issues by dynamically reconfiguring the routines without any additional overhead thereby significantly reducing the need of extra resources.
- *Scalability*: This parameter is used for assessing the capability of a sensor node to efficiently handle increasing amount of traffic with consistent performance without any significant degradation. The scalability performance can be seen in Fig. 2, Fig. 3, and Fig. 4 on various test parameters of traffic. The prime reason behind this is existing schemes demands usage of sleep scheduling for energy management while it doesn't mitigate effectively the likelihood of increasing contention and collision. Existing schemes also cannot handle priority-based queue along with non-inclusion of fairness in access control over increasing traffic. However, proposed scheme addresses this by balancing a priority and regular traffic over same network topology with its novel routine management.
- *Robustness to varying Network Condition*: The complete analysis is carried out with increasing arrival time of message and sink speed to consider the practical traffic condition. The outcome shown in Fig. 3, and Fig. 4 eventually shows robustness of performance parameters over this testcases signifying its robust over varied network conditions.

7 Conclusion

The proposed study has presented a solution towards the shortcomings of existing MAC schemes towards facilitating an optimal radio scheduling performance in large scale dynamic WSN. An essential finding of this study implementation is that the most frequently used CSMA schemes suffers from critical shortcoming associated with unnecessary transmission and hidden terminal problem as well as higher dependencies towards beacon management. From this perspective, the proposed radio scheduling scheme addresses these problems by introducing following notable contributions viz. i) the presented scheme offers better route availability emphasizing towards the node location and message session to offer better transmittance of an aggregated data using route buffer, ii) the scheme facilitates allocation of demanded energy only for transmission requirement using buffer power which doesn't disrupt any progressive communication, iii) a mechanism to perform update of the dynamic link is provided in proposed scheme that cater up the demands of both prioritized and normal sensors, and iv) in contrast to existing scheme, proposed scheme shows approximate minimization of 70% of consumption of transmittance energy compared to existing scheme.

References

1. Naskar, S.: Wireless sensor networks challenges and solutions. In: Sen, P.J. (ed.), Wireless Sensor Networks - Research Issues and Effective Smart Solutions [Working Title]. IntechOpen (2023)
2. Eghonghon Ukhurebor, K., Odesanya, I., Soo Tyokighir, S., George Kerry, R., Samson Olayinka, A., Oluwafemi Bobadoye, A.: Wireless sensor networks: applications and challenges. In: Yellampalli, S.S. (ed.), Wireless Sensor Networks - Design, Deployment and Applications. IntechOpen (2021)
3. Li, J., Lv, J., Zhao, P., Sun, Y., Yuan, H., Xu, H.: Research and application of energy-efficient management approach for wireless sensor networks. Sensors (Basel, Switzerland) **23**(3), 1567 (2023). https://doi.org/10.3390/s23031567
4. Evangelakos, E.A., Kandris, D., Rountos, D., Tselikis, G., Anastasiadis, E.: Energy sustainability in wireless sensor networks: an analytical survey. J. Low Power Electron. Appl. **12**(4), 65 (2022). https://doi.org/10.3390/jlpea12040065
5. Bahadur, D.J., Lakshmanan, L.: A novel method for optimizing energy consumption in wireless sensor network using genetic algorithm. Microprocess. Microsyst. **96**(104749), 104749 (2023). https://doi.org/10.1016/j.micpro.2022.104749
6. Kaur, L., Kaur, R.: A survey on energy efficient routing techniques in WSNs focusing IoT applications and enhancing fog computing paradigm. Glob. Trans. Proc. **2**(2), 520–529 (2021). https://doi.org/10.1016/j.gltp.2021.08.001
7. Kumar, D.S., Saravana Sundaram, S., Prakash, S., Selvaperumal, S.: Effective aggregate data collection and enhanced network lifetime using energy efficient aggregation data convening routing in wireless sensor network. Wirel. Pers. Commun. (2023). https://doi.org/10.1007/s11277-023-10419-y
8. Pedditi, R.B., Debasis, K.: Energy efficient routing protocol for an IoT-based WSN system to detect forest fires. Appl. Sci. (Basel, Switzerland) **13**(5), 3026 (2023). https://doi.org/10.3390/app13053026

9. Prasad, V.K.H., Periyasamy, S.: Energy optimization-based clustering protocols in wireless sensor networks and Internet of Things-survey. Int. J. Distrib. Sens. Netw. **2023**, 1–18 (2023). https://doi.org/10.1155/2023/1362417
10. Abidoye, A.P., Kabaso, B.: Energy-efficient hierarchical routing in wireless sensor networks based on fog computing. EURASIP J. Wirel. Commun. Netw. **2021**(1) (2021). https://doi.org/10.1186/s13638-020-01835-w
11. Rehman, A.U., et al.: A survey on MAC-based physical layer security over wireless sensor network. Electronics **11**(16), 2529 (2022). https://doi.org/10.3390/electronics11162529
12. Raut, A.R., Khandait, S.P., Dongre, S.S.: Time-critical data transmission scheme in wireless sensor networks using machine learning approach. Int. J. Softw. Innov. **10**(1), 1–11 (2022). https://doi.org/10.4018/ijsi.303586
13. Kaur, T., Kumar, D.: QoS mechanisms for MAC protocols in wireless sensor networks: a survey. IET Commun. **13**(14), 2045–2062 (2019). https://doi.org/10.1049/iet-com.2018.5110
14. Abid, K., Lakhlef, H., Bouabdallah, A.: A survey on recent contention-free MAC protocols for static and mobile wireless decentralized networks in IoT. Comput. Netw. **201**(108583), 108583 (2021). https://doi.org/10.1016/j.comnet.2021.108583
15. Afroz, F., Braun, R.: Energy-efficient MAC protocols for wireless sensor networks: a survey. Int. J. Sens. Netw. **32**(3), 150 (2020). https://doi.org/10.1504/ijsnet.2020.105563
16. Mathew, K.D., Jones, A.: Survey: energy efficient protocols using radio scheduling in wireless sensor network. Int. J. Electr. Comput. Eng. (IJECE) **10**(2), 1296 (2020). https://doi.org/10.11591/ijece.v10i2.pp1296-1307
17. Sakib, A.N., Drieberg, M., Sarang, S., Aziz, A.A., Hang, N.T.T., Stojanović, G.M.: Energy-aware QoS MAC protocol based on prioritized-data and multi-hop routing for wireless sensor networks. Sensors (Basel, Switzerland) **22**(7), 2598 (2022). https://doi.org/10.3390/s22072598
18. Hai, T., Zhou, J., Padmavathy, T.V., Md, A.Q., Jawawi, D.N.A., Aksoy, M.: Design and validation of lifetime extension low latency MAC protocol (LELLMAC) for wireless sensor networks using a hybrid algorithm. Sustainability **14**(23), 15547 (2022). https://doi.org/10.3390/su142315547
19. Tomovic, S., Radusinovic, I.: DR-ALOHA-Q: a Q-learning-based adaptive MAC protocol for underwater acoustic sensor networks. Sensors (Basel, Switzerland) **23**(9), 4474 (2023). https://doi.org/10.3390/s23094474
20. Fu, X., Kim, J.G.: Deep-Q-network-based packet scheduling in an IoT environment. Sensors (Basel, Switzerland) **23**(3), 1339 (2023). https://doi.org/10.3390/s23031339
21. Sarang, S., Stojanović, G.M., Drieberg, M., Stankovski, S., Bingi, K., Jeoti, V.: Machine learning prediction based adaptive duty cycle MAC protocol for solar energy harvesting wireless sensor networks. IEEE Access Pract. Innov. Open Solut. **11**, 17536–17554 (2023). https://doi.org/10.1109/access.2023.3246108
22. Kherbache, M., Sobirov, O., Maimour, M., Rondeau, E., Benyahia, A.: Reinforcement learning TDMA-based MAC scheduling in the industrial internet of things: a survey. IFAC-PapersOnLine **55**(8), 83–88 (2022). https://doi.org/10.1016/j.ifacol.2022.08.014
23. Miśkowicz, M.: Unfairness of random access with collision avoidance in industrial internet of things networks. Sensors (Basel, Switzerland) **21**(21), 7135 (2021). https://doi.org/10.3390/s21217135
24. Su, H., Pan, M.-S., Chen, H., Liu, X.: MDP-based MAC protocol for WBANs in edge-enabled eHealth systems. Electronics **12**(4), 947 (2023). https://doi.org/10.3390/electronics12040947
25. C, T., MG, J.: An enhancement for IEEE 802.11p to provision quality of service with context aware channel access for the forward collision avoidance application in vehicular ad hoc network. Sensors (Basel, Switzerland) **21**(20), 6937 (2021). https://doi.org/10.3390/s21206937

26. Khun, A.T.P., Shan, L., Lim, Y., Tan, Y.: MCST scheme for UAV systems over LoRa networks. Drones **7**(6), 371 (2023). https://doi.org/10.3390/drones7060371
27. Sadeq, A.S., Hassan, R., Sallehudin, H., Aman, A.H.M., Ibrahim, A.H.: Conceptual framework for future WSN-MAC protocol to achieve energy consumption enhancement. Sensors (Basel, Switzerland) **22**(6) (2022). https://doi.org/10.3390/s22062129
28. Zhang, H., Wang, F.: MAC protocol analysis for wireless sensor networks. J. Inf. Technol. Res. **15**(1), 1–12 (2022). https://doi.org/10.4018/jitr.298617
29. Uthayakumar, G.S., Dappuri, B., Vanitha, M., Suganthi, R., Savithiri, V., Kamatchi, S.: Design criteria for enhanced energy constraint MAC protocol for WSN. Meas. Sens. **25**(100642), 100642 (2023). https://doi.org/10.1016/j.measen.2022.100642
30. Pradhan, N., Chaudhari, B.S.: Traffic-aware autonomous scheduling for 6TiSCH networks. Int. J. Comput. Appl. **44**(11), 1039–1046 (2022). https://doi.org/10.1080/1206212x.2022.2103889
31. Subramanyam, R., Bala, G.J., Perattur, N., Kanaga, E.G.M.: Energy efficient MAC with variable duty cycle for wireless sensor networks. Int. J. Electron. **109**(3), 367–390 (2022). https://doi.org/10.1080/00207217.2021.1892202
32. Thippun, P., Sasiwat, Y., Buranapanichkit, D., Booranawong, A., Jindapetch, N., Saito, H.: Implementation and experimental evaluation of dynamic capabilities in wireless body area networks: different setting parameters and environments. J. Eng. Appl. Sci. **70**(1) (2023). https://doi.org/10.1186/s44147-022-00171-8
33. Khalifeh, A., Tanash, R., AlQudah, M., Al-Agtash, S.: Enhancing energy efficiency of IEEE 802.15.4- based industrial wireless sensor networks. J. Ind. Inf. Integr. **33**(100460), 100460 (2023). https://doi.org/10.1016/j.jii.2023.100460
34. Nisha, N., Manikandan, A., Venkataramanan, C., Dhanapal, R.: A score based link delay aware routing protocol to improve energy optimization in wireless sensor network. J. Eng. Res. **100115**, 100115 (2023). https://doi.org/10.1016/j.jer.2023.100115
35. Zhong, C., Springer, A.: Design and evaluation of innovative protocols for LoRa. IET Wirel. Sens. Syst. **12**(1), 12–20 (2022). https://doi.org/10.1049/wss2.12033
36. Mathew, K.D., Jones, A.: CSRS-MAC: cluster based synchronous radio scheduling MAC protocol using carrier sense multiple access for wireless sensor network. Wirel. Pers. Commun. **126**(1), 209–229 (2022). https://doi.org/10.1007/s11277-022



A Novel Approach of Intrusion Detection System for IoT Against Modern Attacks Using Deep Learning

A. Durga Bhavani[1(✉)] and Neha Mangla[2]

[1] BMS Institute of Technology and Management, Bangalore, Karnataka, India
durga842004@bmsit.in
[2] Atria Institute of Technology, Bangalore, Karnataka, India

Abstract. Network intrusion detection is important for protecting computer networks from malicious attacks. However, class imbalance in network traffic data can make it difficult to detect intrusions accurately. To address this challenge, we propose a novel deep learning model called the deep regularizer learning model (DRLM). DRLM uses sample similarity across categories to enhance its ability to learn from imbalanced data. To enhance the representation ability of the neural network, DRLM also employs a feature extraction encoder consisting, LayerNorm and Skip-connection units. DRLM uses a adaptive contrastive loss function to optimize the model during training. The model validation is done against IoT-23 dataset, a real-time traffic data from numerous smart home IoT devices. Experimental results showed that the DRLM outperformed existing methods, demonstrating its superior generalization ability and its ability to handle class imbalance without additional pre-training or fine-tuning.

Keywords: Intrusion Detection · IoT · Security · Class Imbalance · Contrastive Learning

1 Introduction

The development of the Internet of Things (IoT) has completely changed how we interact with our surroundings and brought a new level of efficiency, productivity, and comfort. IoT is driving the fourth industrial revolution in the modern digital era and has a significant impact on many industries, including healthcare, transportation, agriculture, and industrial automation, among others [1]. The Internet of Things (IoT) paradigm imagines a hyper-connected world in which common things, integrated with electronics, sensors, and network connectivity, exchange information without the need for human-to-human or human-to-computer interaction. IoT devices are anticipated to reach a whopping 75 billion by 2025, according to a Statista projection, demonstrating the scope and potential of this technology [2].

But as the IoT environment develops, enemies are finding it to be an appealing target, which has raised serious security issues that could undermine the advantages

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 172–182, 2024.
https://doi.org/10.1007/978-3-031-53549-9_18

of this game-changing technology [3]. IoT devices frequently have poor processing and storage capacities, which makes them susceptible to several security attacks. IoT systems are vulnerable for a variety of reasons. First off, a lot of IoT devices are made with convenience and cost-effectiveness over security in mind, making them vulnerable to a range of assaults. The challenge of assuring the security of IoT networks is further complicated by the sheer diversity and size of these networks, combined with the dynamic nature of their communication patterns. IoT infrastructure is increasingly the target of complex and sophisticated cyberattacks in the modern era [5, 6]. Zero-day exploits, phishing, botnets, malware, and distributed denial-of-service (DDoS) attacks are just a few of the diverse techniques used by cybercriminals. When these attacks go unnoticed, they have the potential to do significant harm, including the interruption of services, the theft of private information, and in some circumstances, the control of actual physical systems. Furthermore, it is becoming more and more difficult to quickly recognize and respond to these assaults due to the enormous volume of network traffic data created by IoT devices [7, 8]. Due to these difficulties, Network Intrusion Detection Systems (IDS) have emerged as a crucial element in the protection of IoT networks. Traditional IDS have proven limited effectiveness against complex, contemporary threats since they largely rely on signature or anomaly-based detection [9, 10]. While effective against known threats, signature-based IDS are unable to identify fresh, undiscovered attacks. Although anomaly-based IDS can identify unidentified threats, they frequently experience a high proportion of false positives. Therefore, there is a critical need for more sophisticated and flexible intrusion detection systems that can react to the constantly shifting threat environment. Deep learning's potential for enhancing IDS capabilities is now being investigated in research [11]. Deep learning, a branch of machine learning, is particularly skilled at seeing subtle patterns in vast amounts of data. In a variety of tasks, including audio and picture recognition, natural language processing, and increasingly intrusion detection, it has demonstrated great promise.

One of the critical issues at hand is the 'class imbalance' problem. Network traffic data, being predominantly benign, typically exhibits a massive imbalance between the majority (normal behavior) and minority (attacks) classes [12, 13]. When deep learning models are trained on such data, they tend to be biased towards the majority class, resulting in a high detection rate for normal behavior but poor performance in detecting attacks, which are the minority class. This bias severely impedes the effectiveness of IDS, making it imperative to develop mechanisms to overcome this challenge.

In order to address the issue of class imbalance, this study proposes a novel method for the deep learning model's training phase: a specialized regularizer scheme. This plan intends to make the model's learning process less biased towards the dominant class, resulting in a more accurate and balanced detection system.

In the ensuing sections of this paper, we will delve into the intricacies of our proposed scheme, discuss its theoretical underpinnings, and provide a thorough empirical evaluation of its effectiveness against modern attacks in IoT networks. By contributing to the development of more intelligent, robust, and secure IDS, we hope to strengthen the defenses of IoT networks in an era of increasingly sophisticated cyber threats.

2 Related Work

Recent works in the existing literature includes various data analytics and machine learning (ML) methods for the development of network intrusion detection systems (IDS) to defend against potential security threats and attack. The purpose of this section is to briefly review some recent works, and explores their effectiveness and limitations.

Most existing research works employ supervised machine learning approaches in developing signature-based IDSs that operate on predefined attack signatures and normal traffic patterns. Such IDS systems are reported to be quite effective in scenarios where the threats are well-known, and the signatures are up-to-date. For example, the work by Alzahrani and Alenazi [14] studied potential security threats and vulnerabilities associated with future internet architecture and put their concerns necessitating the development of a reliable and dependable IDS system. The authors have implemented an advanced tree-based classification model named XgBoost to develop IDS for SDN-enabled wireless networks. The training and validation of the implemented models were conducted on the NSL-KDD dataset considering 41 features. Experimental results claim the effectiveness of the presented classification model in comparison to Decision Tree (DT), Random Forest (RF), and deep neural network (DNN) Learning models. Mebawondu et al. [15] focus on improving the performance of supervised learning-based IDS concerning higher accuracy and rapid response action. A multi-layer perceptron (MLP) model with an information gain (IG) mechanism is utilized to select the optimal feature set. The experiment with UNSW-NB15 network traffic dataset, the study outcome claims that the model is lightweight and has a higher detection rate.

Another interesting research presented by Alhajjar et al. [16] studied the vulnerabilities associated with the ML model to assess how an attacker can fool the learning model for intrusion detection. This study has generated adversarial examples using a joint evolutionary algorithm approach and swarm intelligence with a generative adversarial network (GAN). The outcome demonstrates higher false positive outcomes by different ML models when introduced with generated adversarial examples. Sarker et al. [17] focus on reducing the computational complexities associated with ML-driven IDS. In this work, a tree-based classification model is developed and trained with optimal features sorted and ranked according to their importance. This study presented a practical approach for reducing data dimension, which cuts training costs and contributes towards preciseness in feature generalization. Wang et al. [18] have attempted to address the problem of false alarms in IDS by introducing an intelligent filter. The scheme adopts edge computing devices to minimize computing loads and achieve shorter response times. Another work by Moualla et al. [19] tried to address the problem of biased IDS due to training ML models on class-imbalanced datasets. The presented scheme has adopted Synthetic Minority Oversampling (SMO) mechanism to balance the dataset, and the Gini measure is utilized to select optimal features from the UNSW-NB15 dataset for the training classification model.

Apart from the adoption of signature-based IDS, the recent literature shows a growing trend towards anomaly-based IDS using unsupervised and semi-supervised learning approaches. This trend is because building an effective IDS via supervised learning requires many training samples, complex pre-processing feature engineering, and relies on labelled datasets to train the model accurately. In this context, an alternative is the

autoencoder, an unsupervised learning approach concerned with learning compact representation of single input (normal). Any deviation from this learned pattern is identified as an intrusion. Researchers use different autoencoder models to build anomaly-based IDS [20, 21]. The work by Zavrak et al. [22] explored the application of standard and variational autoencoders trained using normal traffic flow features to identify an anomalous pattern from the network traffic flow. The authors have also implemented a binary SVM classifier for comparative analysis. The model was validated over multiple standard datasets, namely UNSW-NB15, CIDDS-001, and CICIDS2017. The result analysis proves the effectiveness of the variational autoencoder with lower false-positive rates in the classified outcomes compared to the binary SVM and standard autoencoder model.

2.1 Problem Description

This work addresses the challenge of accurately identifying and classifying malicious network activity hidden within a large volume of normal traffic in IoT (Internet of Things) devices. Specifically, it focuses on network traffic from smart home IoT devices and attempts to develop machine learning algorithms capable of detecting various types of cyber-attacks in these devices. Traditional network traffic classification models struggle to distinguish between sample similarity and class similarity, often resulting in a bias towards numerous normal traffic samples without considering the actual differences between classes. This makes it challenging to uncover hidden malicious attacks amidst the abundance of typical network traffic. These issues make it necessary to extract features at the network layer without the need for an in-depth examination of data packets, which requires innovative approaches for effective network traffic classification. The problem of class imbalance in a network dataset, or any other dataset, is typically quantified using the class imbalance ratio. This ratio is computed by dividing the number of instances in the majority class by the number of instances in the minority class. If we have two classes, A (majority) and B (minority), and n(A) denotes the number of instances in class A and n(B) denotes the number of instances in class B, the class imbalance ratio, R, can be mathematically defined as follows:

$$R = n(A)/n(B) \tag{1}$$

For balanced datasets, $R = 1$. For imbalanced datasets, $R > 1$. The greater the value of R, the more imbalanced the dataset. Biased learning by a model in the context of class imbalance can be described by the model's tendency to favor predictions towards the majority class. This bias in learning can result in higher misclassification costs for the minority class, leading to suboptimal model performance when evaluated on metrics such as precision, recall, or F1-score for the minority class. One way to numerically describe this biased learning is by comparing the model's accuracy on the majority and minority classes. Let Acc(A) and Acc(B) denote the model's accuracy on classes A and B, respectively. The bias can be quantified as:

$$\text{Bias} = \text{Acc}(A) - \text{Acc}(B) \tag{2}$$

Here, a positive Bias value means the model performs better on the majority class than on the minority class, indicating a bias towards the majority class. A Bias value close

to 0 would indicate that the model performs similarly on both classes, suggesting no or minimal bias.

3 Proposed System

This section proposes an implementation for detecting malicious network traffic involves several key steps, starting from data loading and preprocessing, feature extraction, model design, to final model training.

The initial phase of our methodology involves data loading and preprocessing. In our case, we used the IoT-23 dataset, which is a widely acknowledged dataset in the field of network security research. It includes real, labeled network traffic data from various smart home devices such as Amazon Echo, Philips HUE, and Somfy Door Lock. Loading this data into a manageable format for our model was accomplished using the Python library, Pandas. Since the raw dataset was voluminous, we divided it into 23 different portions, each representing a unique scenario, and loaded each into separate data frames. The data loading process can be represented as follows:

$$df[i] = pd.read\_csv('iot23\_scenario\_' + str(i),\ skiprows = 10,\ nrows = 100000) \tag{3}$$

where df[i] represents the data frame for the ith scenario. Next, we combined all these data frames into a new data frame:

$$df\_combined = pd.concat(df[i] for\ i\ in\ range(1, 24)) \tag{4}$$

Irrelevant columns, such as timestamps and IDs, which have no substantial impact on the model's prediction, were dropped:

$$df\_combined.drop(['ts', 'uid', 'id.orig\ h', 'id.resph', 'service', 'localorig', 'localresp', 'history'])$$

For categorical variables like 'proto' and 'conn state', we assigned dummy values, while missing values were replaced with zeros:

$$df\_combined[[/proto/, /connstate/]] = d.get\_dummies(x)$$

$$x = (df\_combined[[/proto/, /connstate/]]) \tag{5}$$

$$df\_combined.fillna(0, inplace = True) \tag{6}$$

The next phase involved the extraction of features and the design of the model architecture. The learning model, based on the MLP and Skip-Connection architecture, served as our model of choice. The encoder network in learning helps transform raw network traffic data into a representation space that can be analyzed by the model. The encoding process starts with both categorical and continuous features being normalized using one-hot encoding, dropout layers, and LayerNorm layers respectively, as outlined

in the previous response. The normalized features are then combined and passed through two DRB modules, which effectively mitigate issues related to gradient vanishing and network degradation:

$$E = MLP(DRB(Concat(DO(OH(X\_c)), LN(X\_n)))) \tag{7}$$

This encoded vector E is then fed into a multilayer perceptron to get the 512-dimensional representation vector. Next, a projection network maps this representation vector into a final vector used for loss calculation. This mapping is accomplished using a fully connected layer and a regularization step to maximize the similarity between augmentation flow projections while minimizing the similarity between different flows:

$$P = FC(E) \tag{8}$$

where P is the projected vector and FC() denotes the fully connected layer.

The final phase of our methodology involves model training and loss function definition. Inspired by the success of supervised contrastive learning in CV and NLP, we utilized the same principle to train our network flow comparison model. The model was designed to pass each network traffic sample twice to obtain different feature vectors of the same input. This process, known as "dropout twice", generated positive samples for supervised contrastive learning. A similarity matrix S, the size of the input, was obtained by calculating the similarity between different vector representations of the same input:

$$S[i][j] = Sim(P[i], P[j]) \tag{9}$$

where Sim() is the similarity function, which could be cosine similarity or any other relevant similarity measure. In this matrix, a supervised contrastive loss was defined. It aimed to maximize the sum of the similarity distances of the same type of traffic (positive pairs) and minimize the sum of the similarity distances of different classes (negative pairs). The contrastive loss can be expressed as:

$$L = Sum(\alpha) - Sum(\beta) \tag{10}$$

where,

$$\alpha = S\,[i]\,[j]\,\forall\, positive\, pairs$$

$$\beta = S\,[i][j]\,\forall\, negative\, pairs$$

Additionally, we integrated this contrastive loss with the standard cross-entropy loss in the final loss function. This hybrid approach helped mine malicious attacks hidden under a large amount of normal traffic effectively.

4 Dataset and Experiment

This section discusses the dataset, preprocessing and experiment analysis for the proposed regularized learning model for intrusion detection system.

4.1 Dataset

The IoT-23 dataset is structured into 23 separate captures, or scenarios. 20 of these captures contain malicious data, and 3 contain benign data. Each capture from an infected device is tagged with the potential malware sample executed in that specific scenario. The labels assigned to the IoT-23 dataset to categorize malware include: Attack, C&C, *C&C − FileDownload*, C&C-HeartBeat, C&C-HeartBeat- Attack, C&C-HeartBeat-FileDownload, C&C-Mirai, C&C- Torii, DDoS, FileDownload, Okiru, *Okiru − Attack*, and *PartOfA − HorizontalPortScan*. The IoT-23 dataset was processed with Zeek, a network analysis tool. The resultant *conn.log.labeled* files are derived from the *Zeek* conn.log file, which itself is generated by analyzing the original *pcap* file through Zeek. The IoT-23 dataset was processed with *Zeek*, a net *work* analysis tool. The resultant *conn.log.labeled* files are derived from the *Zeek* conn.log file, which itself is generated by analyzing the original pcap file through Zeek Table 1.

Table 1. Variables and definition for Zeek files

ts	This is the time of the first packet
uid	A unique identifier of the connection
id	The connection's 4-tuple of endpoint addresses/ports
proto	The transport layer protocol of the connection
service	An identification of an application protocol
duration	How long the connection lasted
orig_bytes	The number of payload bytes the originator sent
resp_bytes	The number of payload bytes the responder sent
conn_state	Possible connection state values
local_orig	If the connection is originated locally, this will be T
local_resp	If the connection is responded locally, this will be T
missed_bytes	Indicates the number of bytes missed in content gaps
history	Records the state history of connections as a string
orig_pkts	Number of packets that the originator sent
orig_ip_bytes	Number of IP level bytes that the originator sent
resp_pkts	Number of packets that the responder sent
resp_ip_bytes	Number of IP level bytes that the responder sent
tunnel_parents	uid values for any encapsulating parent connections
orig_l2_addr	Link-layer address of the originator

The specific variable types and definitions pertaining to the IoT-23 dataset are detailed in Table 1. Given the considerable size of the dataset, we elected to extract a portion of records from each individual dataset and amalgamate them into a new dataset. This approach allowed our computing resources to manage the workload effectively, while still maintaining the vast majority of the attack types from the original IoT-23 dataset.

4.2 Data preprocessing

The study first used the Python library Pandas to load each of the 23 IoT-23 datasets into separate data frames. The first 10 rows of each dataset were skipped, and the next 100,000

rows were read. Then, all 23 data frames were combined into a single, unified data frame. The dataset was further refined by removing variables that did not influence the results. The variables proto and *conn* state were given dummy values, and all missing values were replaced with 0. The final combined dataset was saved as the iot23_combined.csv file. For validation purposes, the combined dataset was partitioned into a training dataset and a testing dataset. The training dataset consisted of 80% of the data, and the testing dataset consisted of the remaining 20%. Due to data sparsity, limitations, privacy, and sensitivity constraints, features were extracted at the network layer (L3). These features were statistical in nature and focused on behavioral attributes. This meant that there was no need to examine the information in the data packets in detail (Table 2).

Table 2. Counts of attack types file IOT23_ COMBINED.CSV

Label	count
PartOfAHorizonatalPosrtScan	825939
Okiru	262690
Benign	197809
DDoS	138777
Attack	3915
C&C-HeartBeat	349
C&C-FileDownload	43
C&C-Torii	30
FileDownload	13
C&C-HeartBeat-FileDownload	8
C&C-Mirai	1

The proposed system discusses a technique to resist potential attacks on the IoT system using a deep learning approach. The resistivity mechanism discussed in the prior section is assessed using a simulation-based approach, while the obtained outcomes are discussed in this section. This section elaborates on the test environment considered for the study, along with illustrating the results obtained. The proposed study considers three performance metrics such as precision (P), recall (R), and F-measure (F), to evaluate the performance of the model.

- **Precision (P):** It represents true detections over the false detection rate. The higher precision leads to a lower false alarm rate.

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

- **Recall (R):** This metric presents measures of predicted intrusions vs. all intrusions presented in the dataset.

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

- **F1_Score:** This metric measures the weighted correlation of both P and R.

$$F1 = \frac{2(Recall \times Precision)}{(Recall + Precision)} \tag{13}$$

Table 3. Numerical outcome of classification models

	KNN	NB	DL	Proposed
Precision	90.371%	87.563%	91.527%	96.61%
Recall	91.975%	87.328%	90.656%	97.24%
F1score	90.261%	86.405%	91.089%	96.82%

Based on the numerical outcome shown in Table 3, it can be seen that the proposed model achieved a precision of 96.61%, a recall of 97.24%, and an F1-score of 96.82%. These results are significantly better than the results of the other models, which had precision, recall, and F1-scores of 90.37%, 87.32%, and 90.26% for KNN, 87.56%, 87.32%, and 86.40% for NB, and 91.52%, 90.65%, and 91.08% for DL.

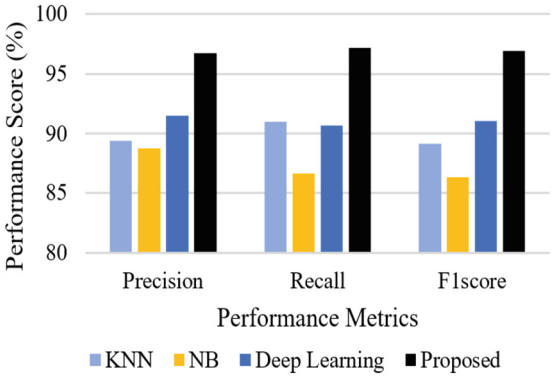


Fig. 1. Comparative analysis of the proposed system with other classification models

Figure 1 shows a comparative analysis of the proposed learning model with existing machine learning techniques KNN, Naïve Bayes, and Deep Learning (DL) with SMOTE. The graph trend exhibits that the proposed model outperforms the existing machine learning technique KNN, Naïve Byes and deep learning in precision rate, recall rate, and F1-score. The results suggest that the proposed model is the most accurate at identifying positive cases. This is likely due to the fact that the proposed model combines the strengths of the KNN and NB models. The KNN model is good at identifying positive cases, but it is also prone to generating false positives. The NB model is less prone to generating false positives, but it is not as accurate at identifying positive cases. The proposed model combines the strengths of the KNN and NB models, while minimizing the weaknesses of both models.

5 Conclusion

To overcome the difficulties caused by class imbalance in network traffic data, this research developed a supervised regularised learning strategy for network intrusion detection. The suggested learning model boosts class-imbalanced learning and increases the precision of intrusion detection by identifying similarities between samples from distinct classes and comparing them with samples from other classes. The invention of a feature ex-traction encoder that improves the neural network's capacity for representation is one of this paper's major achievements. The generation of various vector representations is made possible by the application of dropout layer randomization for data augmentation. Additionally, a powerful training framework to handle class imbalance and find covert attacks is provided by the weighted supervised contrastive loss mixed with the cross-entropy loss. By introducing regularized learning approach in training phase, this paper offers a promising solution to the class imbalance problem in network intrusion detection.

References

1. Abuadlla, Y., Kvascev, G., Gajin, S., Jovanovic, Z.: Flow-based anomaly intrusion detection system using two neural network stages. Comput. Sci. Inf. Syst. **11**(2), 601–622 (2014). https://doi.org/10.2298/CSIS130415035A
2. Yousefi-Azar, M., Varadharajan, V., Hamey, L., Tupakula, U.: Autoencoder-based feature learning for cyber security applications. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN), pp. 3854–3861, May 2017. https://doi.org/10.1109/IJCNN.2017.7966342
3. Zhang, B., Yu, Y., Li, J.: Network intrusion detection based on stacked sparse autoencoder and binary tree ensemble method. In: Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops), pp. 1–6, May 2018. https://doi.org/10.1109/ICCW.2018.8403759
4. Abolhasanzadeh, B.: Nonlinear dimensionality reduction for intrusion detection using auto-encoder bottleneck features. In: Proceedings of the 7th Conference on Information and Knowledge Technology (IKT), pp. 1–5, May 2015. https://doi.org/10.1109/IKT.2015.7288799
5. Mighan, S.N., Kahani, M.: Deep learning based latent feature extraction for intrusion detection. In: Proceedings of the Electrical Engineering (ICEE), Iranian Conference on, pp. 1511–1516, May 2018. https://doi.org/10.1109/ICEE.2018.8472418
6. Shone, N., Ngoc, T.N., Phai, V.D., Shi, Q.: A deep learning approach to network intrusion detection. IEEE Trans. Emerg. Top. Comput. Intell. **2**(1), 41–50 (2018). https://doi.org/10.1109/TETCI.2017.2772792
7. Çekmez, U., Erdem, Z., Yavuz, A.G., Sahingoz, O.K., Buldu, A.: Network anomaly detection with deep learning. In: Proceedings of the 26th Signal Processing and Communications Applications Conference (SIU), pp. 1–4, May 2018. https://doi.org/10.1109/SIU.2018.8404817
8. Aygun, R.C., Yavuz, A.G.: A stochastic data discrimination based autoencoder approach for network anomaly detection. In: Proceedings of the 25th Signal Processing and Communications Applications Conference (SIU), pp. 1–4, May 2017. https://doi.org/10.1109/SIU.2017.7960410

9. Naseer, S., et al.: Enhanced network anomaly detection based on deep neural networks. IEEE Access **6**, 48231–48246 (2018). https://doi.org/10.1109/ACCESS.2018.2863036
10. Chen, Z., Yeo, C.K., Lee, B.S., Lau, C.T.: Autoencoder-based network anomaly detection. In: Proceedings of the Wireless Telecommunication Symposium (WTS), pp. 1–5, April 2018. https://doi.org/10.1109/WTS.2018.8363930
11. KDD Cup 1999 Data. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html. Accessed 30 Mar 2019
12. Farahnakian, F., Heikkonen, J.: A deep auto-encoder based approach for intrusion detection system. In: Proceedings of the 20th International Conference on Advanced Communication Technology (ICACT), pp. 178–183, Feburary 2018. https://doi.org/10.23919/ICACT.2018.8323688
13. Mebawondu, J.O., et al.: Network intrusion detection system using supervised learning paradigm. Sci. Afr. **9**, e00497 (2020)
14. Alhajjar, E., Maxwell, P., Bastian, N.: Adversarial machine learning in network intrusion detection systems. Expert Syst. Appl. **186**, 115782 (2021)
15. Sarker, I.H., et al.: Intrudtree: a machine learning based cyber security intrusion detection model. Symmetry **12**(5), 754 (2020)
16. Wang, Y., et al.: Adaptive machine learning-based alarm reduction via edge computing for distributed intrusion detection systems. Concurr. Comput. Pract. Exp. **31**(19), e5101 (2019)
17. Moualla, S., Khorzom, K., Jafar, A.: Improving the performance of machine learning-based network intrusion detection systems on the UNSW-NB15 dataset. Comput. Intell. Neurosci. **2021**, 1–13 (2021)
18. Zavrak, S., İskefiyeli, M.: Anomaly-based intrusion detection from network flow features using variational autoencoder. IEEE Access **8**, 108346–108358 (2020)
19. Lopes, I.O., et al.: Effective network intrusion detection via representation learning: a denoising autoencoder approach. Comput. Commun. **19**, 55–65 (2022)
20. Yang, J., et al.: Conditional variational auto-encoder and extreme value theory aided two-stage learning approach for intelligent fine-grained known/unknown intrusion detection. IEEE Trans. Inf. Forensics Secur. **16**, 3538–3553 (2021)
21. Sun, J., et al.: Learning sparse representation with variational auto-encoder for anomaly detection. IEEE Access **6**, 33353–33361 (2018)
22. Osada, G., Omote, K., Nishide, T.: Network intrusion detection based on semi-supervised variational auto-encoder. In: Foley, S., Gollmann, D., Snekkenes, E. (eds.) Computer Security – ESORICS 2017. ESORICS 2017. LNCS, vol. 10493, pp. 344–361. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66399-9_19
23. Dao, T.-N., Lee, H.: Stacked autoencoder-based probabilistic feature extraction for on-device network intrusion detection. IEEE Internet Things J. **9**(16), 14438–14451 (2021)
24. Longari, S., et al: CANnolo: An anomaly detection system based on LSTM autoencoders for controller area network. IEEE Trans. Netw. Serv. Manag. **18**(2), 1913–1924 (2020)
25. Kukkala, V.K., Thiruloga, S.V., Pasricha, S.: Indra: intrusion detection using recurrent autoencoders in automotive embedded systems. IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **39**(11), 3698–3710 (2020)



Predictive Classification Framework for Software Demand Using Ensembled Machine Learning

Salma Firdose[1](✉) and Burhan Ul Islam Khan[2]

[1] School of Information Science, Presidency University, Bengaluru, India
salma.firdose@presidencyuniversity.in
[2] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University Malaya (UM), Kuala Lumpur, Malaysia
burhankhan@um.edu.my

Abstract. Software demands are the primary set of essential textual description in baseline of software engineering associated with design anticipation of clients or stakeholder which are quite challenging to be properly evaluated in case of large quantity of information. Review of existing methodologies exhibits frequent adoption of predictive approaches using manifold learning-based approaches. Hence, this manuscript presents a novel and simplified predictive methodology where ensembled machine learning is used for classifying the novel classes of software demands. The study model uses a simplified transformation approach where the input dataset of software demands is subjected to text mining and feature extraction followed by testifying the extracted features on ensembled machine learning models. The study outcome shows Random Forest to offer higher accuracy, lower processing time, and reduced memory utilization in contrast to Support Vector Machine, Decision Tree, K-Nearest Neighbour, and Naive Bayes.

Keywords: Software Demands · Machine Learning · Software engineering · Accuracy · Feature extraction · Prediction · Classification

1 Introduction

The term ‘Software demand’ relates to a textual document which captures the information from the client associated with the target design and operational features associated with the end product [1]. However, from theoretical viewpoint, there is another analogous terminology called as software requirement which is quite different from software demands [2]. Basically, software requirements relate more towards functional and non-functional requirement that further offers an insight towards interface requirements, business requirement, system requirement, and user requirements [3–5]. However, software demand relates more specifically to the precise design requirement from the viewpoint of reduced risk and minimal cost involvement with an assurance of choosing an appropriate design methodology cycle [6–10]. The information retained within software demands

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 183–195, 2024.
https://doi.org/10.1007/978-3-031-53549-9_19

are mainly textual description that offers precise information of essential design components to be involved in order to meet the actual client order [11]. However, this process is typically challenging especially in the case of analyzing a large set of information of software demands, which is humanly impossible for extracting the logical inference in least duration of time cost effectively. At present, there are various research methodology where machine learning, artificial intelligence, and deep learning has been proving its effectiveness for solving issues in software engineering [12–15]; however, there is no explicit report of any standard referential framework towards predictive classification of software demands.

The proposed paper is oriented towards predictive analysis of software demands in order to acquire a definitive classified state The organization of the manuscript is as follows: Sect. 2 discusses about related work while issues associated with related work is presented in Sect. 3. Section 4 briefs of adopted research methodology while system implementation is discussed in Sect. 5. Result analysis is carried out in Sect. 6 while Sect. 7 highlights the contribution of proposed study model.

2 Related Work

This section is a continuation of the discussion of prior related work [16–19]. Study towards realizing software demands towards the practical productive in software engineering is presented by Duarte [20]. Salais-Fierro [21] have presented a modelling of software demand prediction considering use-case of automotive industry. Similar approach for demand prediction is also carried out by Grobler-Debska et al. [22] considering use-case of enterprise resource planning. Study of software demand in perspective of varied project management methods is carried out by Mahdi et al. [23] using machine learning approach. Further adoption of machine learning approaches is reported by Piculjan and Car [24]. Bayesian method is adopted by Molinos et al. [25] towards identifying flaws in software design. Similar line of research work towards risk evaluation in software is carried out by Shan and Wang [26]. Existing study has also witnessed usage of big data analytics towards streamlining predictive analysis on manufacturing industry as noted in study of Cakir et al. [27]. The adoption of machine learning in software engineering has been witnessed in work of Naseer et al. [28] where varied approaches e.g., Random Forest, AdaBoost, Naïve Bayes, Multilayer Perceptron, J48, etc. have been evaluated. Kler et al. [29] have used machine learning along with the artificial intelligence towards predictive analysis of food industry. Adoption of deep learning towards estimating software cost is presented by Hassan et al. [30] using Constructive Cost Model (COCOMO). This technique also integrates bio-inspired algorithms for optimizing the accuracy of prediction. Deep Learning has also been used for predicting demands of musical equipment as noted in work of Zhang et al. [31]. Study of demand forecasting considering retail industry using deep learning is presented by Giri et al. [32]. Further usage of deep learning has been reported to be used for task management (Kolomvastsos et al. [33]). Adoption of deep learning in software industry is presented by Ahmed and Lee [34], Abdu et al. [35], Kalouptsoglou et al. [36], and Hoc et al. [37]. Further, there are varied hybrid approaches where machine learning is combined with deep learning as noted in work of Forootan et al. [38] and Costas et al. [39]. Hence, there are varied

approaches used for realizing software demands where both machine learning and deep learning has been frequently used. The next section discusses about issues.

3 Problem Description

The identified research problems after reviewing the existing methodologies in prior section are as follows: i) there are less studies focused on addressing issues in software demand in project management perspective while more strategies have been presented for defect analysis and cost estimation, ii) more emphasis is given towards accomplishing higher analysis on data and this is one reason towards more adoption of deep learning compared to machine learning approaches, iii) related deep learning approaches offers higher accuracy at the cost of computational resources while existing machine learning approaches offers reduced accuracy but comparatively computationally cost efficient approach, iv) there is no reported study to prove faster response time and effective memory utilization during predictive analysis.

4 Proposed Methodology

The proposed study aims towards developing a simplified and yet robust computational framework that is capable of categorizing software demands using integrated machine learning approach and text mining approach. The presented model adopts analytical research methodology whose process flow is exhibited in Fig. 1.

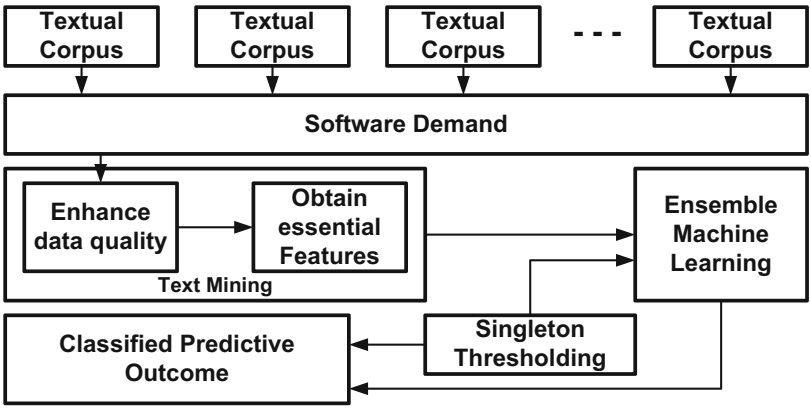


Fig. 1. Adopted Architecture

According to Fig. 1, the proposed scheme performs classification of software demands into regular and dynamic demand where the input content of software demands is maintained in textual corpus. The input data is subjected to series of transformation operation where the complexities associated with input data is normalized followed by extraction of essential feature suitable for classification. Further, an analytical test-bed is

constructed where the extracted feature is subjected to multiple machine learning models. The predictive outcome is further assessed for its convergence with the objective function that is intended to reduce the development cost by minimizing the uncertainty in software demand. The illustration of architecture is carried out in next section of system implementation.

5 System Design Implementation

The prime agenda of the proposed system design is to ensure that an effective classification of the software project demands is fulfilled by categorizing it to *standard* and *dynamic* demands. This section performs discussion of the system design with respect to formulation of research challenge, strategy towards problem solution implementation and algorithm design.

5.1 Formulating Research Challenge

In order to realize the practical perspective of software demands, considers a real-world use-case where a client furnishes n number of software demands consisting of m_1, m_2, …. m_n number of descriptive textual contents associated with the software demands. Pictorial representation of this scenario is showcased in Fig. 2.

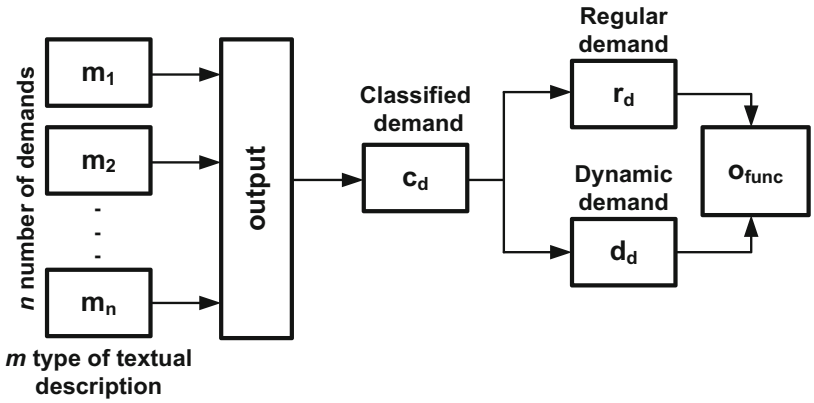


Fig. 2. Pictorial representation of Research Challenge

As per above Fig. 2, there is a need of a system which can consider the input of n number of software demands and m type of textual description of software demands in order to generate a classified data c_d. At present, this operation is manually carried out by skilled professional, however, it is beyond capability of any human if the value of n is extensively large in size. The problems towards categorizing software demands becomes much challenging if one need to classify them as regular demand r_d and dynamic demand d_d. Regular demand r_d is stated as those set of operation where the development team has already prior experiences while dynamic demand dd refers to uncertain or novel

set of operation for development team. Hence the objective function to represent this problem formulation is as shown below:

$$O_{\text{fun}} = \arg_{\max}(\mathrm{r_d}),\ \arg_{\min}(\mathrm{d_d}) \forall (n, m) \tag{1}$$

A closer look into expression (1) showcase that effective results can be only formulation of the development team adopts a working methodology that leads to maximization of r_{d} and minimization of d_{d} which will indirectly lower down the cost of development as well as reduce the possible risk of software development. Hence, an implementation strategy is developed to solve this research challenge.

5.2 Strategy Towards Implementation

One of the essential parts of proposed implementation is to reduce the d_{d} attributes, which is actually not an easy task. For this purpose, the proposed scheme carries out two essential strategic implementation steps. In the first implementation strategy, **the** proposed scheme performs a series of transformation operation on input data of software requirement that leads to final binary classification. The set of operations are streamlining data towards improving data quality, performing extraction of feature, and performing binary classification using machine learning approach. The adopted strategy of implementation is as shown in Fig. 2.

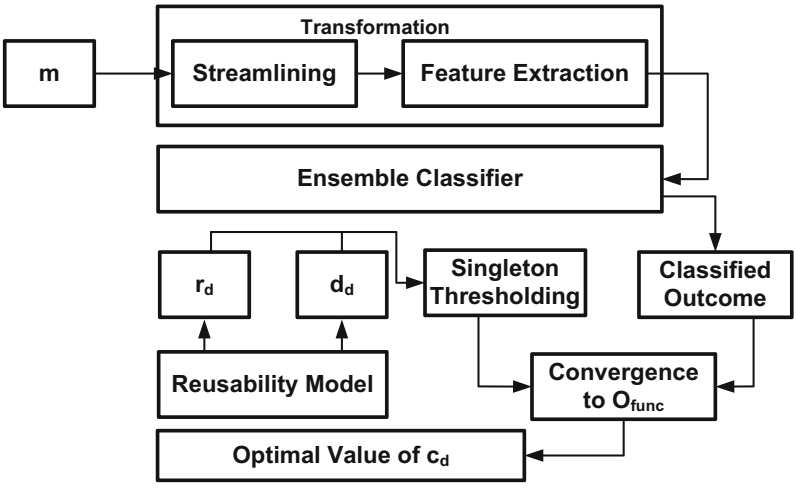


Fig. 3. Adopted Implementation Strategy

A simplified and yet novel text-mining approach is applied for this purpose whose outcome is subjected to machine learning to carry out predictive analysis of software demand types. In the second implementation strategy, the proposed scheme constructs a predefined set of regular demands $\mathrm{r_d}$ as well as dynamic demands $\mathrm{d_d}$. For this purpose, the proposed scheme utilizes the Lee et al. [40] model of software reusability on a predefined

rule set defined for r_d based on historical software project as well as development team also formulates all possible arguments of rule set towards dynamic demands d_d. This process formulates a threshold value T complying with objective function O_{fun} defined in expression (1) that is compared with outcome of binary classification leading to generation of optimal value of classified data c_d. The next section presents discussion of proposed algorithm design implementation.

5.3 Algorithm Design Implementation

The proposed algorithm implementation is based on accomplishing the objective function towards maximizing the regular demand attribute r_d and minimizing the dynamic demands d_d after classifying the software demands. In order to implement this algorithm, a standard dataset [41] is considered which consists of classified form of software demands with a pre-labelled data.

Algorithm for Classifying Software Demands

Input: n, m

Output: c_d

Start

1. **For** i=1: n
2. **For** j=1:m
3. $n_d \rightarrow f_1(m)$
4. $e_f \rightarrow f_2(n_d)$
5. $c_d \rightarrow f_3(e_f)$
6. **End**

7 **End**

End

The discussion of above-mentioned algorithmic steps is as follows: The algorithm considers input of n (number of software demands) and m (textual content of software demands) that after processing yields an outcome of c_d(classified data). After taking the input of n data (Line-1) and m number of textual content of software demands (Line-2), the algorithm executes a function $f_1(x)$ which is mainly responsible for streamlining data towards accomplishing data quality (Line-3). In this step, the textual data is subjected to cleaning process where all the articles and pronouns are deemed as irrelevant words and are eliminated. This leads to generation of normalized data n_d (Line-3) that are further subjected to another explicit function $f_2(x)$ which is responsible for carrying out feature extraction. In this operation, numerical vectors are obtained from the streamlined data that offer optimal representation of information of software demands using standard Term Frequency and Inverse Document Frequency along with Bag-of-Words.

The process generates an extracted feature e_f that are further subjected to next analytical processing (Line-4). Further, the proposed system implements its last function $f_3(x)$ in order to implement a set of machine learning approaches on extracted feature e_f in order to generate finally classified data c_d (Line-5). It is to be noted that the function $f_3(x)$ actually carry out implementation of multiple machine learning approaches where the evaluation of the classified outcome is carried out using accuracy parameters. The finally accomplished outcome is further evaluated on the basis of singleton threshold (as exhibited in Fig. 3) where it is compared with the outcome generated by machine learning models. Finally, the outcome of c_d whose value is found more than the threshold T is considered as an optimal predictive solution fulfilling the criteria stated in objective function. The function $f_3(x)$ is subjected to iteration where the stopping criterion is based on objective function i.e., to increase rd and minimize d_d values. Therefore, the proposed algorithm offers a simplified and cost-effective computational model towards software engineering framework with a flexible solution of classified software demands.

6 Results Discussion

The assessment of the proposed study model is carried out using PROMISE dataset [AR] which consists of pre-annotated data with 625 types of software demand maintained in textual form. The dataset is slightly finetuned towards representing regular and dynamic demand with 50–50 ratio. Implementing the proposed algorithm, the extracted features are subjected to varied ensembled machine learning algorithms frequently reported to be deployed viz. Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbor (KNN), and Naïve Bayes (NB). Further, the proposed system makes use of natural language toolkit for normalizing the data [41]. This section showcases the result obtained and discusses its inference.

6.1 Accomplished Results

As the proposed study introduces a technique towards predictive analysis of software demands using integrated text mining approach and machine learning methods, the core target is to obtained a highly accurate classification result. However, different from existing approaches, the analysis is carried out considering accuracy, processing time, and memory utilization to extensively put forward evidence towards its robustness. Table 1 highlights the numerical outcomes of comparative analysis.

The numerical outcome in Table 1 showcase that proposed system offers optimal performance with integrated with RF approach with respect to all the three performances metrics. The performance is also found to be better for SVM and DT while NB performance is found to be sub-optimal. This states the RF is the best alternative solution for any machine learning algorithm in order to perform predictive analysis of software demands. A closer look into the outcome shows that RF approach offers approximately 5.1% increased accuracy, 2.1% decreased processing time, and 9.7% of reduced memory utilization in contrast to averaged respective scores of SVM, DT, KNN, and NB approach.

Table 1. Numerical Outcomes

Approach	Accuracy(%)	Processing Time(s)	Memory Utilization(%)
RF	99.5	0.388	12.8
SVM	97.2	0.569	20.5
DT	96.7	0.498	18.9
KNN	91.6	0.605	19.5
NB	91.9	0.721	31.2

6.2 Discussion of Results

A deeper insight to accuracy performance in Fig. 4 shows that RF, SVM, and DT offers and accuracy in the range of 95–100%, while RF score the highest accuracy of 99.5%. SVM approach is witnessed to encounter challenge in parameter tuning when exposed to software demand dataset while DT suffers from slight instability during the feature extraction process. This leads to slightly fall of accuracy in SVM and DT. KNN and NB are found to yet not fit for high dimensional data lowering the accuracy.

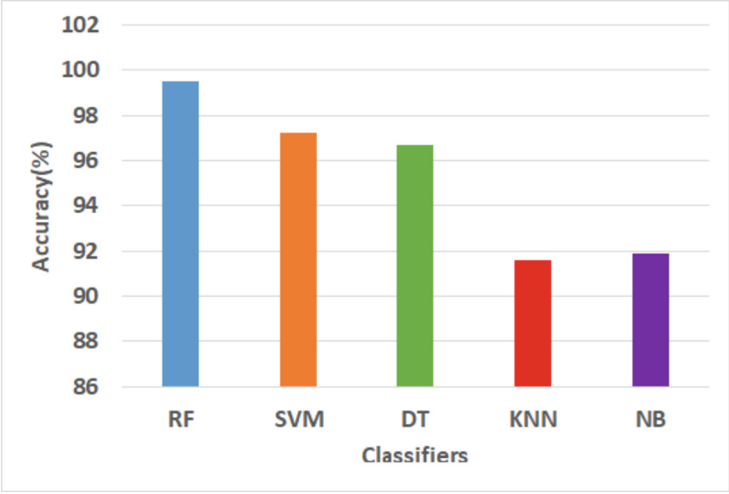


Fig. 4. Comparative Analysis of Accuracy

The trend of processing time shown in Fig. 5 showcases RF approach to offer better classification result in least time. NB approach is found to be highly sensitive to input textual data distribution of software demands while KNN takes a longer duration of time to deal with software demands with complex attributes with no consideration of underlying demand data distribution leading to more time consumption. Although, SVM and DT offers better advantage in this perspective from the viewpoint of scaling and normalization, yet they are prone to overfitting as well as unscalable. This results in increased

set of iterative operation which slightly increases the duration of processing time. On the other hand, RF not only offers higher accuracy but also capable of autonomously handle artifacts as well as minimize overfitting leading to reduced algorithm processing time.

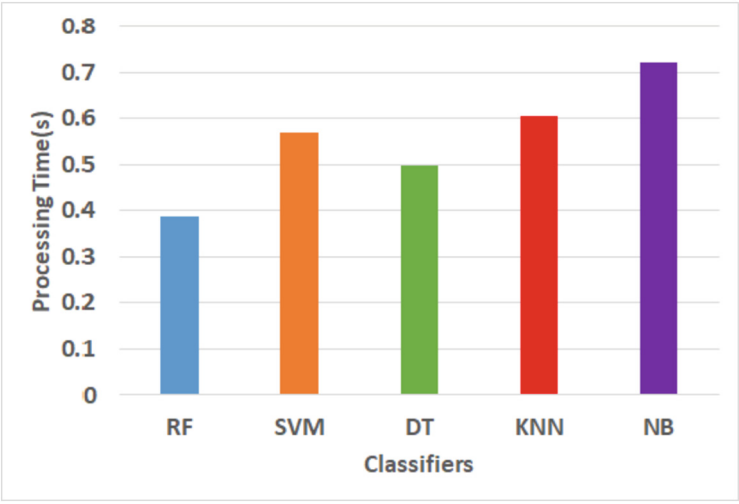


Fig. 5. Comparative Analysis of Processing Time

The proposed scheme is also evaluated with respect to memory utilization. For this purpose, allocated memory for performing training and storing the intermediate outcomes were summed up to find the final memory utilization in percentile. The current experiment has been carried out on 16 GB RAM memory where memory utilized by each process has been recorded during the observation to obtain outcome presented in Fig. 6.

According to outcome shown in Fig. 6, proposed system with RF shows lesser than 15% of memory utilization whereas the utilization for SVM, DT, and KNN ranges between 15–21% approximately while NB approach exhibits more than 30% of memory utilization. Although NB approach can successfully handle higher ranges of data yet it cannot successfully extract the complex relationship in the data. It is also noted that SVM has slightly more than 20% of memory utilization but it is noted that its training time increases with the increased data size which directly influences the heap memory spaces. DT can handle both categorical and numerical data but it is found to be quite instable and inconsistent even for same dataset in multiple iteration and hence not much considered as reliable classified on software demand extraction. Hence, RF is found as better solution from memory utilization perspective.

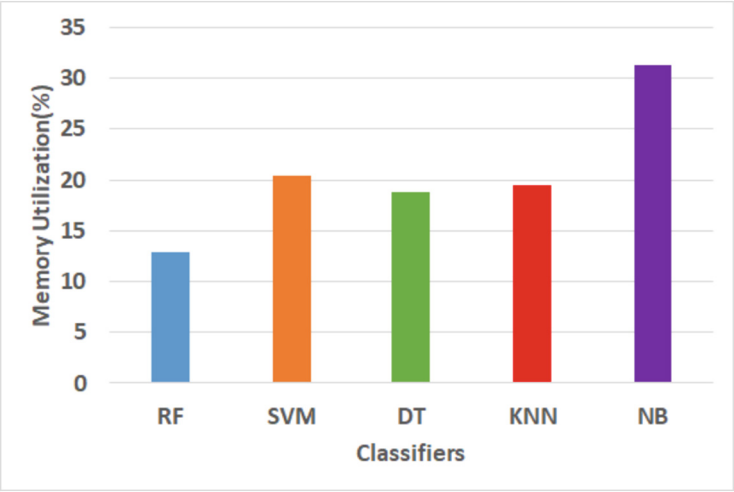


Fig. 6. Comparative Analysis of Memory Utilization

7 Conclusion

The proposed study presents a novel computational framework towards performing predictive analysis of software demands by integrating simplified text mining approach and ensembled machine learning approach. The contribution of the proposed study model are as follows: i) the proposed study model considers a binary classification of regular and dynamic data with a formulated objective function to increase the former and decrease the latter leading to cost effective software engineering methods, ii) the presented study uses a unique and simplified text vectorization method that can extract knowledge and classify the software demands on the basis of their computed significance values., iii) different from any machine learning approaches used for software engineering, the proposed scheme doesn't consider the final predictive outcome of its respective machine learning algorithms as its final result of classified demands, rather it compares the final predictive outcome with outcomes of software reusability model finetuned with binary classes with respect to singleton thresholding. This offers much fine-tuned and reliable outcomes. The future direction of proposed study will be further towards inclusion of more challenging threats to uncertainty of risk as well as towards cost evaluation process. The primary idea is to generate a highly optimal risk-free software engineering baseline model free from domain attributes during software project demands. The secondary idea is to reduce the complexity in analytical process of software engineering giving more clear and predictive results.

References

1. Hickok, M., Maslej, N.: A policy primer and roadmap on AI worker surveillance and productivity scoring tools. AI Ethics **3**(3), 673–687 (2023). https://doi.org/10.1007/s43681-023-00275-8

2. Sadiq, M., Susheela Devi, V.: Prioritization and selection of the software requirements using rough-set theory. IETE J. Res. 1–18 (2021). https://doi.org/10.1080/03772063.2021.1973593
3. Laplante, P.A., Kassab, M.H.: Requirements Engineering for Software and Systems, 4th edn. Auerbach Publications, Boca Raton (2022)
4. Hoy, Z., Xu, M.: Agile software requirements engineering challenges-solutions—a conceptual framework from systematic literature review. Information (Basel) **14**(6), 322 (2023). https://doi.org/10.3390/info14060322
5. Khatun, M.T., Hiekata, K., Takahashi, Y., Okada, I.: Design and management of software development projects under rework uncertainty: a study using system dynamics. J. Decis. Syst. **32**(2), 265–288 (2023). https://doi.org/10.1080/12460125.2021.2023257
6. Ren, T., Zeng, N., Wang, D., Cheng, S.: Effects of channel encroachment on the software and service decisions in it supply chains. J. Ind. Manag. Optim. **18**(5), 3787 (2022). https://doi.org/10.3934/jimo.2021136
7. van Dijk, J., Delacrétaz, N., Lanz, B.: Technology adoption and early network infrastructure provision in the market for electric vehicles. Environ. Resour. Econ. **83**(3), 631–679 (2022). https://doi.org/10.1007/s10640-022-00703-z
8. Marsh, E., Vallejos, E.P., Spence, A.: The digital workplace and its dark side: an integrative review. Comput. Hum. Behav. **128**(107118), 107118 (2022). https://doi.org/10.1016/j.chb.2021.107118
9. Patel, T., Bapat, H., Patel, D., van der Walt, J.D.: Identification of critical success factors (CSFs) of BIM software selection: a combined approach of FCM and Fuzzy DEMATEL. Buildings **11**(7), 311 (2021). https://doi.org/10.3390/buildings11070311
10. Savić, V., Varga, E.: Extending the SEMAT kernel with the TDD practice. IET Softw. **12**(2), 85–95 (2018). https://doi.org/10.1049/iet-sen.2016.0305
11. Ghiduk, A.S., Qahtani, A.M.: An empirical study of local-decision-making-based software customization in distributed development. IET Softw. **15**(2), 174–187 (2021). https://doi.org/10.1049/sfw2.12016
12. Wang, S., et al.: Machine/deep learning for software engineering: a systematic literature review. IEEE Trans. Softw. Eng. **49**(3), 1188–1231 (2023). https://doi.org/10.1109/tse.2022.3173346
13. Soori, M., Arezoo, B., Dastres, R.: Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cogn. Robot. **3**, 54–70 (2023). https://doi.org/10.1016/j.cogr.2023.04.001
14. Taye, M.M.: Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. Computers **12**(5), 91 (2023). https://doi.org/10.3390/computers12050091
15. Satapathy, S.C., Jena, A.K., Singh, J., Bilgaiyan, S.: Automated Software Engineering: A Deep Learning-based Approach. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-38006-9
16. Firdose, S., Rao, L.M.: PORM: predictive optimization of risk management to control uncertainty problems in software engineering. Int. J. Electr. Comput. Eng. (IJECE) **8**(6), 4735 (2018). https://doi.org/10.11591/ijece.v8i6.pp4735-4744
17. Firdose, S., Swapna Kumar, S., Gayan Narendra Meegama, R.: A novel predictive model for capturing threats for facilitating effective social distancing in COVID-19. Int. J. Electr. Comput. Eng. (IJECE) **12**(1), 596 (2022). https://doi.org/10.11591/ijece.v12i1.pp596-604
18. Firdose, S., Rao, L.M.: 3LRM-3 layer risk mitigation modelling of ICT software development projects. Int. J. Electr. Comput. Eng. (IJECE) **6**(1), 349 (2016). https://doi.org/10.11591/ijece.v6i1.pp349-356

19. Shahabadkar, R., Govinda, S., Firdose, S.: Integrated privacy preservation with novel encoding and encryption for securing video in internet-of-things. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics. CSOC 2023. LNNS, vol. 723, pp. 217–227. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_19
20. Duarte, C.H.C.: Software productivity in practice: a systematic mapping study. Software **1**(2), 164–214 (2022). https://doi.org/10.3390/software1020008
21. Salais-Fierro, T.E., Saucedo-Martinez, J.A., Rodriguez-Aguilar, R., Vela-Haro, J.M.: Demand prediction using a soft-computing approach: a case study of automotive industry. Appl. Sci. (Basel, Switzerland) **10**(3), 829 (2020). https://doi.org/10.3390/app10030829
22. Grobler-Dębska, K., Kucharska, E., Żak, B., Baranowski, J., Domagała, A.: Implementation of demand forecasting module of ERP system in mass customization industry—case studies. Appl. Sci. (Basel, Switzerland) **12**(21), 11102 (2022). https://doi.org/10.3390/app122111102
23. Mahdi, M.N., et al.: Software project management using Machine Learning technique—a review. Appl. Sci. (Basel, Switzerland) **11**(11), 5183 (2021). https://doi.org/10.3390/app11115183
24. Pičuljan, N., Car, Ž: Machine learning-based label quality assurance for object detection projects in requirements engineering. Appl. Sci. (Basel, Switzerland) **13**(10), 6234 (2023). https://doi.org/10.3390/app13106234
25. Hernández-Molinos, M.J., Sánchez-García, A.J., Barrientos-Martínez, R.E., Pérez-Arriaga, J.C., Ocharán-Hernández, J.O.: Software defect prediction with Bayesian approaches. Mathematics **11**(11), 2524 (2023). https://doi.org/10.3390/math11112524
26. Shan, J., Wang, H.: Software enterprise risk detection model based on BP neural network. Wirel. Commun. Mob. Comput. **2022**, 1–9 (2022). https://doi.org/10.1155/2022/9147090
27. Cakir, A., Akın, Ö., Deniz, H.F., Yılmaz, A.: Enabling real time big data solutions for manufacturing at scale. J. Big Data **9**(1) (2022). https://doi.org/10.1186/s40537-022-00672-6
28. Naseer, M., Zhang, W., Zhu, W.: Prediction of coding intricacy in a software engineering team through machine learning to ensure cooperative learning and sustainable education. Sustainability **12**(21), 8986 (2020). https://doi.org/10.3390/su12218986
29. Kler, R., et al.: Machine learning and artificial intelligence in the food industry: a sustainable approach. J. Food Qual. **2022**, 1–9 (2022). https://doi.org/10.1155/2022/8521236
30. ul Hassan, C.A., et al.: Optimizing deep learning model for software cost estimation using hybrid meta-heuristic algorithmic approach. Comput. Intell. Neurosci. **2022**, 1–20 (2022). https://doi.org/10.1155/2022/3145956
31. Zhang, J.: Forecasting of musical equipment demand based on a deep neural network. Mob. Inf. Syst. **2022**, 1–10 (2022). https://doi.org/10.1155/2022/6580742
32. Giri, C., Chen, Y.: Deep learning for demand forecasting in the fashion and apparel retail industry. Forecasting **4**(2), 565–581 (2022). https://doi.org/10.3390/forecast4020031
33. Kolomvatsos, K., Anagnostopoulos, C.: A deep learning model for demand-driven, proactive tasks management in pervasive computing. IoT **1**(2), 240–258 (2020). https://doi.org/10.3390/iot1020015
34. Ahmed, H.A., Lee, J.: FCP2Vec: deep learning-based approach to software change prediction by learning co-changing patterns from changelogs. Appl. Sci. (Basel, Switzerland) **13**(11), 6453 (2023). https://doi.org/10.3390/app13116453
35. Abdu, A., Zhai, Z., Algabri, R., Abdo, H.A., Hamad, K., Al-antari, M.A.: Deep learning-based software defect prediction via semantic key features of source code—systematic survey. Mathematics **10**(17), 3120 (2022). https://doi.org/10.3390/math10173120
36. Kalouptsoglou, I., Tsoukalas, D., Siavvas, M., Kehagias, D., Chatzigeorgiou, A., Ampatzoglou, A.: Time series forecasting of software vulnerabilities using statistical and deep learning models. Electronics **11**(18), 2820 (2022). https://doi.org/10.3390/electronics11182820

37. Hoc, H.T., Silhavy, R., Prokopova, Z., Silhavy, P.: Comparing stacking ensemble and deep learning for software project effort estimation. IEEE Access Pract. Innov. Open Solut. **11**, 60590–60604 (2023). https://doi.org/10.1109/access.2023.3286372
38. Forootan, M.M., Larki, I., Zahedi, R., Ahmadi, A.: Machine learning and deep learning in energy systems: a review. Sustainability **14**(8), 4832 (2022). https://doi.org/10.3390/su14084832
39. Cordeiro-Costas, M., Villanueva, D., Eguía-Oller, P., Martínez-Comesaña, M., Ramos, S.: Load forecasting with machine learning and deep learning methods. Appl. Sci. (Basel, Switzerland) **13**(13), 7933 (2023). https://doi.org/10.3390/app13137933
40. Lee, W.-T., Chen, C.-H.: Agile software development and reuse approach with Scrum and software product line engineering. Electronics **12**(15), 3291 (2023). https://doi.org/10.3390/electronics12153291
41. PROMISE Software Engineering Repository. (n.d.). Uottawa.Ca. http://promise.site.uottawa.ca/SERepository/. Accessed 31 Aug 2023



Secure Data Transmission Scheme in Wireless Sensor Network Resisting Unknown Lethal Threats

Chaya Puttaswamy[1(✉)] and Nandini Prasad Kanakapura Shivaprasad[2]

[1] Department of Information Science and Engineering, Dr. Ambedkar Institute of Technology, Bangalore, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka 590018, India
chayaneetha@gmail.com

[2] Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore, Affiliated to Visvesvaraya Technological University, Karnataka 590018 Belagavi, India
dean-fa@dsatm.edu.in

Abstract. Security data transmission has been a critical concern in Wireless Sensor Network (WSN) irrespective of secured routing protocols evolved in past decades. There are emerging new approaches towards embedding security features in WSN; however, optimal solution has yet not been evolved or found fit for practical implementation. Therefore, the proposed paper presents a novel secure data transmission scheme that is capable of identifying and resisting potential threats of unknown and uncertain form. The scheme initially designs public key-based encryption scheme to protect the data during transmission followed by introducing a novel analytical scheme that can determine the adversaries from its malicious behaviour. The study outcome shows that proposed system is not only energy efficient but also offers efficient data transmission performance along with higher coverage from unknown lethal threats.

Keywords: Public Key Encryption · Security · Energy · Secured Routing · Wireless Sensor Network · Threats

1 Introduction

Wireless Sensor Network (WSN) is characterized by an interconnected sensor nodes which are meant to capture environmental information, aggregate them, and forward them to a sink node for further utilization of such data for advance analytics or taking up certain actions [1]. However, there is a serious form of vulnerabilities noted in WSN mainly due to its distributed nature, restricted range of communication, and resource constraints [2]. Some of the prominent security issues in WSN are limited resources, node compromise, susceptible communication with tampering, interception, and eavesdropping [3]. WSN is also characterized by vulnerable location privacy [4] with highly prone to various types of attacks thereby compromising its data integrity [5]. At present, there

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 196–207, 2024.
https://doi.org/10.1007/978-3-031-53549-9_20

are various research studies being carried out towards securing the data transmission in the form of secure routing schemes [6–10]. Unfortunately, there are further loopholes in majority of the existing secure routing schemes in WSN. The most prominent issue is related to designing a light-weight security algorithm that can offer extensive resistance from lethal threats present in forms of various attacks. This is however quite computationally challenging to design as cryptography being the strongest security protocol find its infeasibility to work properly over resource-constraint sensor nodes in WSN [11]. There is also a need of approach which is capable to resist maximum forms of threats in WSN. Hence, the proposed scheme contributes towards developing a novel secure routing scheme which is capable of offering light weight encryption on its data to be propagated along with a novel solution to identify the unknown adversaries. The proposed scheme also contributes towards a novel bypassing solution towards the threat by isolating them using fake sensor broadcast to mislead the attackers. The organization of the paper is as follows: Sect. 2 discusses about existing literatures, Sect. 3 briefs of identified research problem, while adopted methodology is highlighted in Sect. 4 and system design is highlighted in Sect. 5. Result discussion is carried out in Sect. 6 while conclusion is presented in Sect. 7.

2 Related Work

Our prior work has offered discussion of various security aspects pertaining to WSN [12]. Currently, it is noted that adoption of machine learning [13, 14], deep learning [15], and blockchain technology [16] is on rise towards securing communication in WSN. Blockchain technology has been increasing in its pace of adoption towards boosting security in WSN (Hsiao & Sung [17]; Khan et al. [18]; And et al. [19]; Lee & Lee [20]). The next most frequently adopted schemes towards securing WSN is encryption-based over varied use cases. Oladipupo et al. [21] have used Elliptical Curve Cryptography in order to develop an authentication scheme along with usage of digital signature. Adoption of cryptanalysis and multiple signature scheme towards securing digital transaction is witnessed in work of Shim [22]. Zhao and Li [23] have used Rabin cryptosystem in order to develop an authentication scheme in WSN emphasizing on multi gateway system.Tan et al. [24] have presented a tamperproof mechanism for secured authentication in WSN integrated with cloud environment.Existing system has also witnessed adequate secure routing-based strategies in WSN. There are different variants explored in literatures viz. Genetic algorithm-based trust evaluation (Han et al. [25]), secure data forwarding using energy scheduling (Feng et al. [26]), heuristic-based secure data transmission (Haseeb et al. [27]), secure cooperative routing with energy efficiency (Saeed et al. [28]), secure message dissemination on optimal trust (Hu et al. [29]), integrated watermarking and homomorphic encryption based secure routing (Babaeer& Al-Ahmadi [30]), revised blowfish based routing (Alotaibi [31]). It can be noted that these approaches are mainly emphasized towards energy efficiency while performing secure routing operation. Apart from this, it is also noted that LEACH protocol has been consistently used towards incorporating energy efficient in secure communication strategies in WSN. Various equivalent approaches seen to adopt this methodology as noted in work of Aljumaie and Alhakami [32], Hussein et al. [33], Nagaraj et al. [34], Ahmad et al. [35]. Therefore, there are

various recent approaches towards secure routing scheme in WSN with claimed advantages in the perspective of various identified issues in literatures. However, they are also associated with significant levels of issues too that are highlighted in next section.

3 Problem Description

The identified research problem in existing problem solution towards secure routing in WSN are as follows: i) usage of learning and blockchain based approaches are substantially new and no benchmarked method has yet evolved to be applicable in practical ground of WSN, ii) majority of existing approaches has deployed sophisticated encryption-based strategies which is not only iterative but also consumes lot of resources of low capacity sensor nodes, iii) existing studies on secure routing doesn't showcase any potential method towards resisting unknown threats and majority of solutions are designed considering apriori information of threats, iv) there is no balance between robust security characteristics with data transmission performance for a dynamic form of network in WSN, which raises a questions of practical applicability.

4 Proposed Methodology

The prime agenda of the proposed study is to design a secured data transmission scheme in WSN in order to protect the sensors from unknown and dynamic form of threats. The implementation of this study model is carried out using analytical methodology which emphasize on balancing the energy efficiency and higher degree of resiliency from potential threats without affecting the ongoing data transmission in WSN. The adopted architecture is shown in Fig. 1.

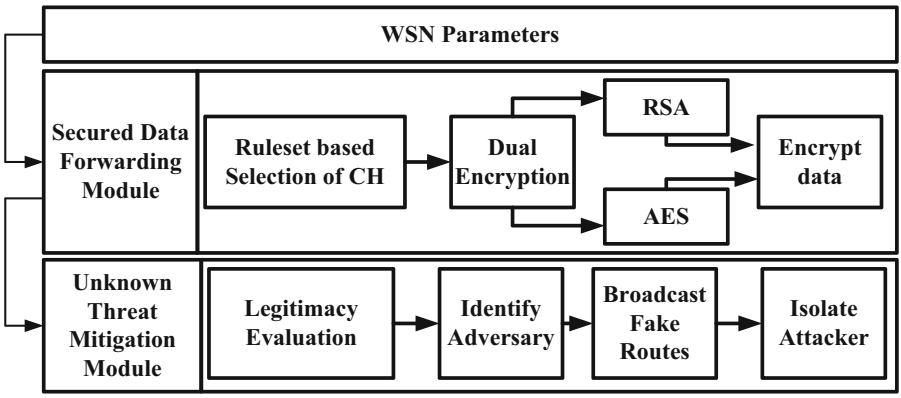


Fig. 1. Adopted Architecture for Secure Data Transmission in WSN

Figure 1 highlights that implementation of proposed study model is carried out using two sequential model which is responsible for implementing dual layer of security using public key encryption. Further, the study model also introduces a mechanism of

identifying the unknown threats on the basis of evaluation of malicious behaviour on multiple hops. The prime contribution of this methodology is also towards its mechanism of isolating the positively identified threats followed by misleading the attacker node using fake broadcasted nodes. The illustration of study model and its internal operation is given in next section of system implementation.

5 System Implementation

This section presents the core system implementation design that targets to accomplish the objective of the proposed research work towards offering an optimal security strength to WSN environment. In order to implement this, the proposed scheme introduces two discrete module of implementation whose description is as follows:

5.1 Secured Data Forwarding Module (SDFM)

This module is primary security module which is responsible for sharing/forwarding encrypted data from one to another sensor node. It is to be noted that proposed scheme uses a clustering operation where a cluster-head is selected that actually carries out data propagation. If the distance from one cluster head to sink is closer, the cluster head chooses a single hop strategy to forward the encrypted data directly to the sink node via gateway node. In case the distance from sink to current cluster head is higher than the latter transfer the encrypted data to its neighboring clustering node using multihop strategy. Further, it should be noted that probability of opting multihop transmission is more likely to happen compared to single-hop transmission strategy as not many sensors will be located near to sink location. Figure 2 highlights the internal operations carried out by SDFM:

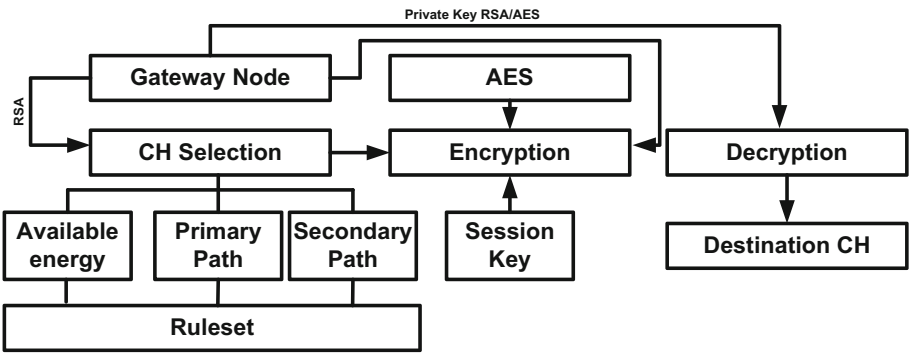


Fig. 2. Internal Operation of SDFM

According to Fig. 2, the scheme uses fuzzy inference system to construct manifold ruleset considering three essential parameters viz. Available energy, primary path, and secondary path. The primary path is basically a mean distance of shortest path between two communicating sensors in tree while secondary path is the physical spatial measure

of the source sensor showcase its availability towards facilitating data aggregation in WSN environment. The gateway node generates a secret key pair using Rivest-Shamir Algorithm (RSA) to be forwarded to all sensors where the latter uses this key to encrypt the message. A session key is used to carry out this encryption using Advanced Encryption Standard (AES) while the encrypted data is forwarded to gateway node first, which further performs decryption operation using its RSA key. The decrypted packet is either forwarded to nearest neighboring cluster head who will use AES to further decrypt it or to the sink depending on its choice of selection of single/multipath-based propagation.

5.2 Unknown Threat Mitigation Module (UTMM)

The outcome of the first module is basically a decrypted information that is accessed by neighboring cluster head. However, there is still an uncertainty to claim the legitimacy of neighboring cluster head to be a regular node. In such case, there is higher likelihood that decrypted information be accessed by a malicious node, whose identity is challenging to be ascertained. The following operation are carried out within UTMM as shown below:

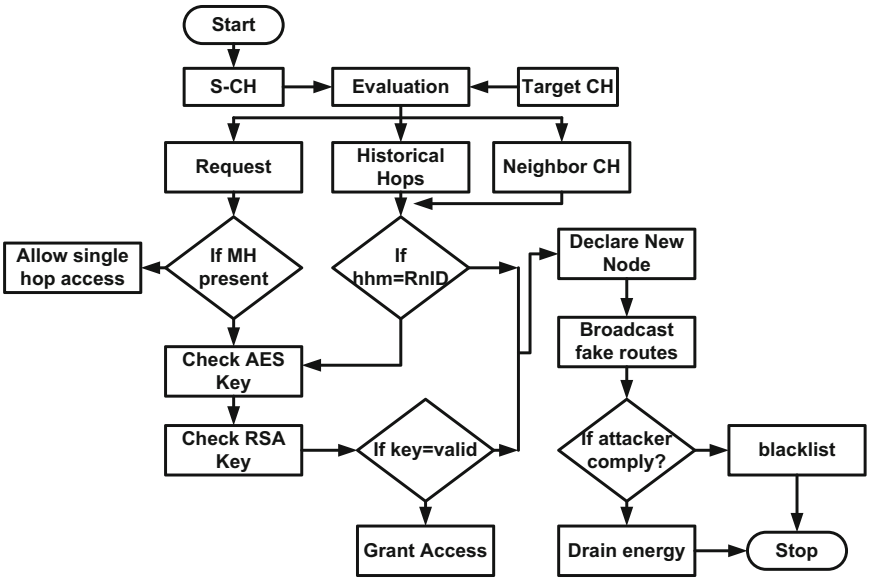


Fig. 3. Internal Operation of UTMM

According to Fig. 3, the sender cluster head (S-CH) carry out an evaluation towards confirming the legitimacy of its neighboring target (T-CH) whom it needs to forward the encrypted data. The evaluation is carried out on three perspective parameters i.e., i) evaluation of request from T-CH, ii) assessing historical hops of T-CH, and iii) adjoining CH of T-CH. In the first step, S-CH assess the incoming request from T-CH and checks if the request is towards accessing any specific nodes located in multihop. If the incoming request is found positive for accessing nodes in multihop (MH) than the S-CH assess

its AES key and RSA key (both in private key form). If these sequential validations of keys are found to be positive than S-CH grant an access to T-CH confirming it to be regular node. Further, if the incoming request of T-CH is found to access a node not in multihop than S-CH just grant an access right towards node in single hop. In such way, S-CH gets enough opportunity to monitor the activity of T-CH if they are not found to be registered node. In negative possibility, if the T-CH acts maliciously, it may intrude just one sensor while whole network will be protected. This will also cause the T-CH to get caught by S-CH which will be further blacklisted from entire network. However, if T-CH is proven to be legitimate, its monitoring will be continued until it reaches some specific threshold of trust. In the next parameter evaluation, S-CH assess the historical hops records of T-CH and formulates a logical condition to find if the historical hops bear a record of all registered sensors. In such case, similar validation of secret keys is carried out to further confirm its legitimacy just like the prior method of evaluation of incoming request. Similar evaluation strategy is continued for neighboring CH of T-CH too. If the neighboring CH of T-CH is found to be registered nodes in its hop records, similar dual secret keys evaluation is further carried out. Upon confirming presence of valid keys, the T-CH is granted access or it follows the next step of operations. It should be noted that if T-CH fails in all these three validation steps, it is quite confirmed that T-CH is a malicious node although it may bear all characteristic of regular node in data propagation perspective. In such case, the proposed scheme introduces a novel mechanism to isolate such form of uncertain and dynamic attackers. The S-CH declares a new node as an acknowledge to malicious T-CH and start broadcasting fake routes. In adherence ot non-repudiation security standard, the T-CH node cannot deny the service and will need to comply with this declaration to follow the broadcasted path. There are further two possibilities. If the malicious T-CH nodes chooses to comply with the broadcasted fake routes, they will allocate their resources to perform data transmission over fake routes and this process will gradually drain out entire resources of malicious nodes. This process will autonomously continue till the attacker node depletes its complete resources. However, if the malicious T-CH nodes chooses to violate the broadcasted fake path, then T-CH node will be permanently blacklisted and this acknowledgement is forwarded to gateway node. Further, the gateway node forwards an update information to all the underlying sensor nodes thereby protecting them from further possibilities of attacks from this identified unknown attacker.

6 Results Discussion

A simulation study was carried out using MATLAB considering 500–1000 sensors bear The implementation of the proposed study framework is carried out considering 500–1000 sensor nodes with a squared simulation area of $100x100m^2$. Each sensor is assigned with 10J of initialized energy with 50nj/bit as transmittance and receiving energy each. The control message size is 32 bits while data packet size is considered to be 2000 bits. A closer look into the existing scheme will showcase that energy trends is one of the prime indicators of intrusion and therefore, the proposed system is benchmarked with conventional standard energy-efficient algorithms applicable in WSN i.e., LEACH protocol [32] and revised LEACH protocol (RevLEACH) [32]. The comparative analysis is

carried out with respect to four performance metrics e.g., energy fluctuation, throughput, residual energy, and operational nodes. All the sensors are dispersed in simulation area in random order where a sink node is located at the far point (within the coverage of simulation area). Selection of cluster node is carried followed by random selection of one cluster head as source node. The member node senses the data using Time Division Multiple Access (TDMA) and forwards the data to its respective cluster head while the aggregated data from the cluster head is forwarded either to sink node (if the cluster head is located less than 10m of distance) or forwarded to neighbouring cluster head.

6.1 Communication-Based Analysis

A cost-effective and strong secure routing scheme must offer a better consistency in its internal operation right from encryption to proposed mechanism of validation of dual secret key as well as identifying the legitimacy of target cluster head. Hence, *energy fluctuation* parameter is used for evaluating the degree of variance in energy dissipation by each of the considered approaches in comparative analysis. The outcome for energy fluctuation is shown in Fig. 4.

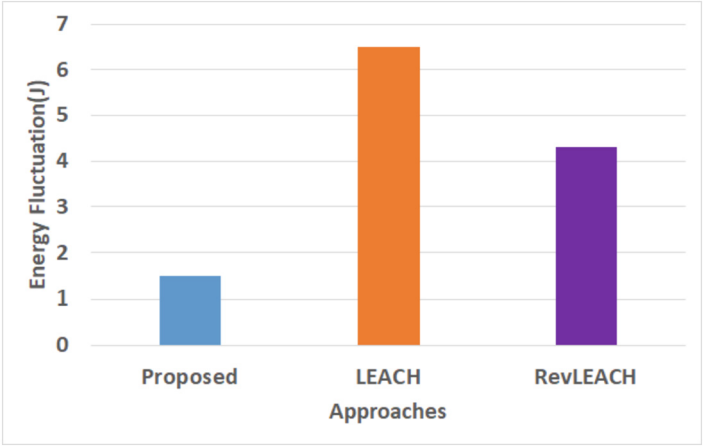


Fig. 4. Comparative Evaluation of Energy Fluctuation

The outcome in Fig. 4 exhibits that proposed scheme offers approximately 39% of reduced energy fluctuation in contrast to both existing energy-efficient approaches (i.e., LEACH and RevLEACH). The rationale are i) proposed system uses tree-based structure where evaluation of target cluster head is carried out thereby benefit not only current cluster head but also other participating neighboring cluster head leading to nearly equivalent resource consumption and hence lesser fluctuation whereas LEACH and RevLEACH only emphasize on data aggregation load on current sensors not considering its adjoining nodes leading to different scale of energy consumption dominated by uncertain traffic.

Figure 5 showcases that proposed scheme offers approximately 50% of increased throughput in comparison to existing LEACH and RevLEACH algorithm. The prime

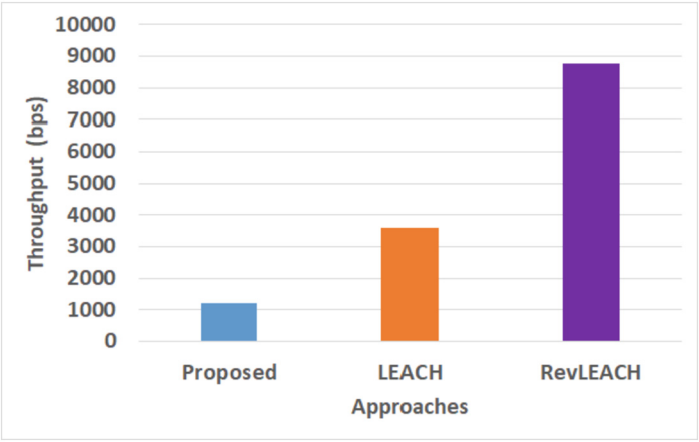


Fig. 5. Comparative Evaluation of Throughput

justification of this outcome is that proposed scheme uses a highly structured tree with spontaneous updates of blacklist node information which reduces down security checks from beginning. This form of an integrated security not only reduces the complexities of processing data during aggregation but also makes the routing path unaffected. Hence, proposed scheme presents a set up of secure network topology with an efficient shortest path attributed considered during selection of cluster head along with residual energy. This leads to more focus on data transmission without affecting a single intrinsic security operation.

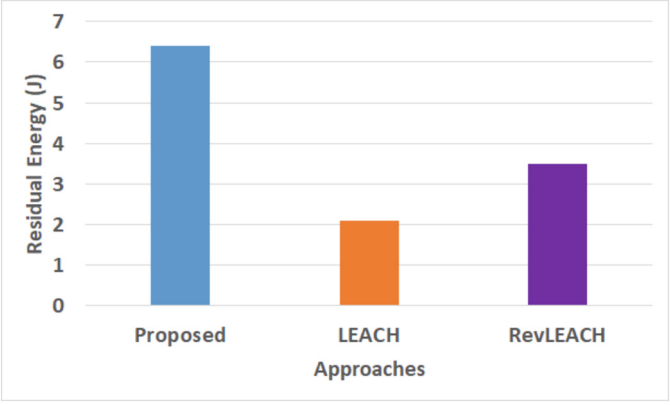


Fig. 6. Comparative Evaluation of Residual Energy

Figure 6 showcases that proposed system offers approximately 36% higher residual energy retention compared to existing energy efficient protocols in WSN. The prime reason is that unlike LEACH and RevLEACH, proposed system doesn't choose all the sensors to play a role of cluster head while it does computation based on three parameters

to select cluster head suitable for current data transmission, leading to lesser energy consumption and more residual energy.

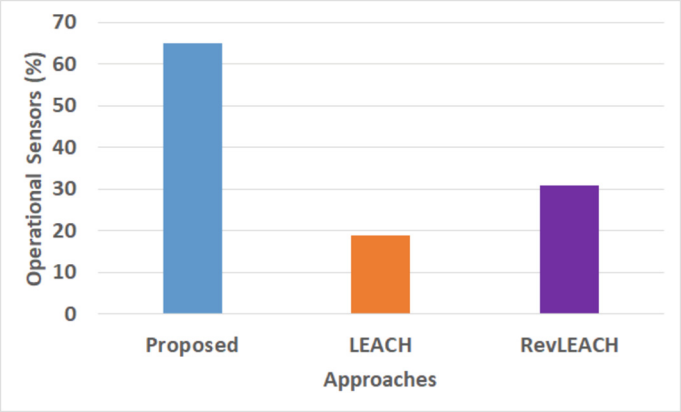


Fig. 7. Comparative Evaluation of Operational Sensor

Figure 7 highlights that proposed scheme offers approximately 40% of increased operational nodes compared to existing approaches. Basically, operational nodes are defined as sensor nodes whose residual amount is evaluated to be more than 30% of their assigned initial energy. It is noted that maximum number of sensor nodes are retained till the last round of simulation whereas sensors using LEACH and RevLEACH are found to be operational only till 5% and 12% of cumulative simulation rounds of 2000 iterations.

6.2 Security-Based Analysis

The proposed scheme is capable of resisting both active and passive forms of attacks in WSN (e.g., node replication attack, packet injection, data injection, data interception, etc.) This is because proposed scheme uses an integrated security mechanism where SDFM secures all the data to be propagated from authorized sensors only while UTMM further boost up the secure routing by identifying the malicious behaviour on the basis of manifold attributes i.e., intentions of incoming request, historical hop information, and presence of trusted neighboring cluster head. It will eventually mean that proposed scheme can resist communication attacks, node attacks, as well as data attacks too. At the same time, proposed scheme uses a simplified and non-iterative security mechanism that doesn't demand involvement of complex encryption operation which compliments its better data transmission performance too. Even if the attacker's information or strategies are unknown, proposed scheme can compute the malicious behaviour of target node using no-trust approach while this computation leads to confirmation of presence of attacker node.

7 Conclusion

This paper has presented a simplified, cost-effective, and yet robust secure data transmission scheme in WSN which is capable of identifying and bypassing the lethal forms of threats. The novelty of this study model are as following: i) the novel routing strategy is designed using dual-layer of public key encryption processed by gateway node and sensor node that is challenging to be replicated by any form of malicious nodes, ii) a simplified selection strategy of cluster head is presented on manifold criterion which is capable enough to adapt itself over dynamic scenario of WSN implementation, iii) the model uses three distinct parameters (incoming request, historical hop, and neighboring nodes) in order to perform legitimacy evaluation of target node, iv) unlike any existing security approaches in WSN, proposed scheme not only identifies the target node on the basis of calculated malicious behaviour but also presents a novel approach to isolate them from network by broadcasting fake routes. An extensive simulation study shows that proposed scheme excel better performance both from data transmission and security perspective compared to existing routing schemes in WSN.

References

1. Bagwari, A., Tomar, G.S., Bagwari, J., Barbosa, J.L.V., Sastry, M.K.: Advanced Wireless Communication and Sensor Networks: Applications and Simulations, 1st edn. Chapman and Hall/CRC, Boca Raton (2023)
2. Arab, A.: The advanced wireless sensor networks' routing protocol to detect malicious nodes and behavior. J. Chin. Inst. Eng. **46**(7), 805–812 (2023). https://doi.org/10.1080/02533839.2023.2238786
3. Singh, S., Saurabh, R., Maitra, T., Giri, D.: Security in communication for intelligent wireless sensor networks. In: Computational Intelligence for Wireless Sensor Networks, 1st edn, pp. 175–192. Chapman and Hall/CRC, Boca Raton (2022)
4. Shi, L., Li, K., Zhu, H.: Data fusion and processing technology of wireless sensor network for privacy protection. J. Appl. Math. **2023**, 1–10 (2023). https://doi.org/10.1155/2023/1046050
5. Urooj, S., Lata, S., Ahmad, S., Mehfuz, S., Kalathil, S.: Cryptographic data security for reliable wireless sensor network. Alex. Eng. J. **72**, 37–50 (2023). https://doi.org/10.1016/j.aej.2023.03.061
6. Sankar, S.U., Christo, M.S., Priyadarsini, P.U.: Secure and energy concise route revamp technique in wireless sensor networks. Intell. Autom. Soft Comput. **35**(2), 2337–2351 (2023). https://doi.org/10.32604/iasc.2023.030278
7. Sudha, G., Tharini, C.: Trust-based clustering and best route selection strategy for energy efficient wireless sensor networks. Automatika **64**(3), 634–641 (2023). https://doi.org/10.1080/00051144.2023.2208462
8. Verma, P., et al.: Security Issues for Wireless Sensor Networks, 1st edn. CRC Press, Boca Raton (2022)
9. Gebremariam, G.G., Panda, J., Indu, S.: Design of advanced intrusion detection systems based on hybrid machine learning techniques in hierarchically wireless sensor networks. Connect. Sci. **35**(1) (2023). https://doi.org/10.1080/09540091.2023.2246703
10. Deebak, B.D., Al-Turjman, F.: A hybrid secure routing and monitoring mechanism in IoT-based wireless sensor networks. Ad Hoc Netw. **97**(102022), 102022 (2020). https://doi.org/10.1016/j.adhoc.2019.102022

11. Sen, J.: A survey of cryptography and key management schemes for wireless sensor networks. In: Sen, P.J., Yi, D.M., Niu, D.F., Wu, A.P.H. (eds.), Wireless Sensor Networks - Research Issues and Effective Smart Solutions [Working Title]. IntechOpen (2023)
12. Chaya, P., Prasad, N.: Investigation of threats, vulnerabilities, attacks, and approaches in wireless sensor networks. In: Advances in IT Standards and Standardization Research, pp. 147–162. IGI Global (2022)
13. Ali, Y., Khan, H.U., Khalid, M.: Engineering the advances of the artificial neural networks (ANNs) for the security requirements of internet of things: a systematic review. J. Big Data, **10**(1) (2023). https://doi.org/10.1186/s40537-023-00805-5
14. Ahmad, R., Wazirali, R., Abu-Ain, T.: Machine learning for wireless sensor networks security: an overview of challenges and issues. Sensors (Basel, Switzerland) **22**(13), 4730 (2022). https://doi.org/10.3390/s22134730
15. Salmi, S., Oughdir, L.: Performance evaluation of deep learning techniques for DoS attacks detection in wireless sensor network. J. Big Data **10**(1) (2023). https://doi.org/10.1186/s40537-023-00692-w
16. Ismail, S., Dawoud, D.W., Reza, H.: Securing wireless sensor networks using machine learning and blockchain: a review. Futur. Internet **15**(6), 200 (2023). https://doi.org/10.3390/fi15060200
17. Hsiao, S.-J., Sung, W.-T.: Employing blockchain technology to strengthen security of wireless sensor networks. IEEE Access Pract. Innov. Open Solut. **9**, 72326–72341 (2021). https://doi.org/10.1109/access.2021.3079708
18. Khan, Z.A., Amjad, S., Ahmed, F., Almasoud, A.M., Imran, M., Javaid, N.: A blockchain-based deep-learning-driven architecture for quality routing in wireless sensor networks. IEEE Access Pract. Innov. Open Solut. **11**, 31036–31051 (2023). https://doi.org/10.1109/access.2023.3259982
19. And, I.A.A.E.-M., Darwish, S.M.: Towards designing a trusted routing scheme in wireless sensor networks: a new deep blockchain approach. IEEE Access Pract. Innov. Open Solut. **9**, 103822–103834 (2021). https://doi.org/10.1109/access.2021.3098933
20. Lee, J., Lee, M.: A blockchain system for TDMA-based tactical wireless networks with constrained resources. IEEE Access Pract. Innov. Open Solut. **11**, 6857–6866 (2023). https://doi.org/10.1109/access.2023.3237828
21. Oladipupo, E.T., et al.: An efficient authenticated elliptic curve cryptography scheme for multicore wireless sensor networks. IEEE Access: Practical Innovations, Open Solutions **11**, 1306–1323 (2023). https://doi.org/10.1109/access.2022.3233632
22. Shim, K.-A.: Cryptanalysis of two signature schemes for IoT-based mobile payments and healthcare wireless medical sensor networks. IEEE Access Pract. Innov. Open Solut. **8**, 167203–167208 (2020). https://doi.org/10.1109/access.2020.3023093
23. Zhao, X., Li, D.: A lightweight user authentication scheme for multi-gateway based wireless sensor networks using Rabin cryptosystem. IEEE Access Pract. Innov. Open Solut. **11**, 79874–79889 (2023). https://doi.org/10.1109/access.2023.3300440
24. Tan, X., Zhang, J., Zhang, Y., Qin, Z., Ding, Y., Wang, X.: A PUF-based and cloud-assisted lightweight authentication for multi-hop body area network. Tsinghua Sci. Technol. **26**(1), 36–47 (2021). https://doi.org/10.26599/tst.2019.9010048
25. Han, Y., Hu, H., Guo, Y.: Energy-aware and trust-based secure routing protocol for wireless sensor networks using adaptive genetic algorithm. IEEE Access Pract. Innov. Open Solut. **10**, 11538–11550 (2022). https://doi.org/10.1109/access.2022.3144015
26. Feng, W., et al.: Joint energy-saving scheduling and secure routing for critical event reporting in wireless sensor networks. IEEE Access Pract. Innov. Open Solut. **8**, 53281–53292 (2020). https://doi.org/10.1109/access.2020.2981115

27. Haseeb, K., Almustafa, K.M., Jan, Z., Saba, T., Tariq, U.: Secure and energy-aware heuristic routing protocol for wireless sensor network. IEEE Access Pract. Innov. Open Solut. **8**, 163962–163974 (2020). https://doi.org/10.1109/access.2020.3022285
28. Saeed, K., Khalil, W., Ahmed, S., Ahmad, I., Khattak, M.N.K.: SEECR: secure energy efficient and cooperative routing protocol for underwater wireless sensor networks. IEEE Access Pract. Innov. Open Solut. **8**, 107419–107433 (2020). https://doi.org/10.1109/access.2020.3000863
29. Hu, H., Han, Y., Yao, M., Song, X.: Trust based secure and energy efficient routing protocol for wireless sensor networks. IEEE Access Pract. Innov. Open Solut. **10**, 10585–10596 (2022). https://doi.org/10.1109/access.2021.3075959
30. Babaeer, H.A., AL-ahmadi, S.A.: Efficient and secure data transmission and sinkhole detection in a multi-clustering wireless sensor network based on homomorphic encryption and watermarking. IEEE Access Pract. Innov. Open Solut. **8**, 1 (2020). https://doi.org/10.1109/access.2020.2994587
31. Alotaibi, M.: Improved blowfish algorithm-based secure routing technique in IoT-based WSN. IEEE Access Pract. Innov. Open Solut. **9**, 159187–159197 (2021). https://doi.org/10.1109/access.2021.3130005
32. Aljumaie, G.S., Alhakami, W.: A secure LEACH-PRO protocol based on blockchain. Sensors (Basel, Switzerland) **22**(21), 8431 (2022). https://doi.org/10.3390/s22218431
33. Hussein, S.M., López Ramos, J.A., Ashir, A.M.: A secure and efficient method to protect communications and energy consumption in IoT wireless sensor networks. Electronics **11**(17), 2721 (2022). https://doi.org/10.3390/electronics11172721
34. Nagaraj, S., et al.: Improved secure encryption with energy optimization using random permutation pseudo algorithm based on Internet of thing in wireless sensor networks. Energies **16**(1), 8 (2022). https://doi.org/10.3390/en16010008
35. Ahmad, R., Wazirali, R., Abu-Ain, T., Almohamad, T.A.: Adaptive trust-based framework for securing and reducing cost in low-cost 6LoWPAN wireless sensor networks. Appl. Sci. (Basel, Switzerland) **12**(17), 8605 (2022). https://doi.org/10.3390/app12178605



An IoT-Based Cloud Data Platform with Real-Time Connecting Maritime Autonomous Surface Ships

Hyoseong Hwang[1(✉)] and Inwhee Joe[2]

[1] The Department of Computer and Software, Hanyang University, Seoul 04763, Republic of Korea
hshwang@hyundai-gs.com

[2] The Department of Computer and Software, Hanyang University, Seoul 04763, Republic of Korea
iwjoe@hanyang.ac.kr

Abstract. Maritime Autonomous Surface Ships (MASS) pose numerous challenges and complexities regarding risk, stability, and implementation. To overcome these obstacles and realize the vision of MASS, the development of an onshore data platform capable of real-time monitoring and support for ship conditions becomes imperative. In this paper, we design a cloud-based data platform that facilitates seamless interaction between ships and onshore through the utilization of IoT technology. To achieve this, we introduce a comprehensive six-layer structure that serves as the foundation of the platform, enabling effective communication and data exchange at each layer. Furthermore, we propose a methodology that leverages data sharing among neighboring ships through the platform, thereby enhancing collision avoidance performance. To implement this methodology, we introduce the architecture of Edge in the structure of MSA and propose an MQTT-based communication method designed to enhance real-time under unstable network conditions. To validate our approach, we implement and test the data platform using AWS Cloud. The simulation results show that the proposed MQTT communication method effectively limits real-time data delay and reduces total data transmission time by 34.6% compared to conventional methods.

Keywords: Maritime Autonomous Surface Ships · Cloud · Edge · MQTT · Navigation · Internet of Things · AWS IoT

1 Introduction

Recent advancements in automation systems and the integration of Internet of Things(IoT) technology are revolutionizing maritime operations. This has enabled the development of new types of ship operations, including smart ships, remote operating ships, and digital twin ships [1, 2].

Recognizing these technological strides, the International Maritime Organization(IMO) has been actively engaging in discussions around the safe operation of

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 208–220, 2024.
https://doi.org/10.1007/978-3-031-53549-9_21

autonomous ships. In 2018, the IMO introduced the concept of Maritime Autonomous Surface Ships(MASS), defining them as vessels capable of navigating surface waters with minimal or no human intervention. They further categorized these into four stages, based on the degree of automation [3]. This recognition signifies the growing interest in leveraging autonomous technologies for enhancing maritime operations.

As the MASS concept gains traction, addressing the legal and operational safety regulations for autonomous ship operations becomes imperative. The IMO has begun discussions to establish comprehensive guidelines to ensure the safe and responsible deployment of these vessels. Such regulations play a crucial role in creating a robust framework for autonomous ship operations, while also mitigating potential risks and concerns [4].

In this paper, designing platforms that bridge data and functions between ships and onshore is crucial. Such platforms allow efficient data management, real-time monitoring, and informed decision-making, thereby improving the safety, operation-al efficiency, and performance of autonomous maritime operations [5].

The wireless communication technology of ships is being developed in various forms, from satellite communication to short-range communication [6]. However, due to the nature of the ship's wireless communication, communication coverage problems or Line of Sight(LoS) problems may cause unstable communication connections. We propose an MQTT-based communication method, designed to ensure real-time reconnection after periods of communication instability between ships and onshore.

2 Related Work

The recent decade has seen a considerable focus on the fusion of Information Technology(IT) and Internet of Things(IoT) technologies in the maritime field, primarily aimed at automating ship navigation and machinery control. An essential aspect of autonomous ship operation is collision avoidance. Consequently, numerous studies have strived to develop robust algorithms, guidance systems, and controls to address this challenge [7, 8]. Furthermore, the sector is leveraging IoT technology to share and analyze data, aiming to tackle prevalent issues in maritime operations [9].

2.1 Ship Automation System

Ship automation systems can be broadly divided into two main components: The Integrated Navigation System(INS) responsible for navigation and the Control and Monitoring System(CAMS) responsible for control and monitoring of various systems of the ship. Figure 1 illustrates the conceptual structure where ship automation systems are linked with the cloud. The IEC 61924–2 standard defines the requirements and test methods for the operation and performance of INS [10].

CAMS, functions as a ship's Distributed Control System (DCS), is crucial for the comprehensive control and monitoring of the ship's control systems. In the past, CAMS on container carriers were relatively straightforward, focusing mainly on navigation and cargo control. However, with the advent of multifunctional eco-friendly vessels, such

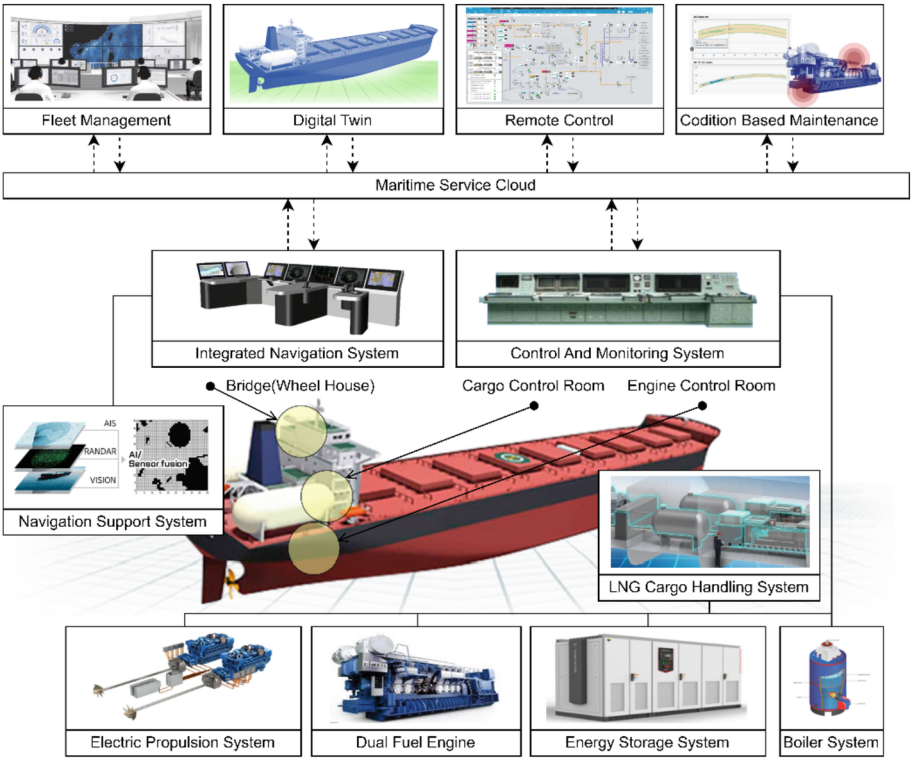


Fig. 1. Conceptual diagram of ship automation system and cloud service

as LNG carriers, electric propulsion ships, and Dual Fuel engines, the control duties of CAMS have become increasingly complex and diverse [11].

To cope with these advancements, research efforts have been devoted to establishing an environment that enables the rapid and secure development of these sophisticated ships and control technologies using digital twin technology [12]. Additionally, studies are underway to integrate and optimize systems through virtual simulation environments [13]. For these research activities to yield practical results and be applied to real ships, a real-time data platform capable of seamlessly connecting and managing ship data is indispensable.

2.2 Ship Data Platform

Various studies have been carried out to contribute to the development of autonomous ships by focusing on the collection of ship data onshore and the establishment of a data platform. In [14], research investigates the application of IoT technologies in unmanned ships, specifically intelligent recognition, sensor fusion, and communication. This study proposes an E-Navigation structure that interlinks the ship with onshore systems.

In [15], suggests a Machine-Type Communication Framework, facilitating communication between peripheral ships or clouds through short-range or broadband connections. This is directed at developing the functional elements of MASS.

The field of ship IoT technologies has advanced significantly, allowing for research to be conducted on implementing autonomous ship functions based on IoT technologies. However, the actualization of fully autonomous ships necessitates detailed investigations on how ship systems will operate within scalable cloud-based platforms that use IoT technologies. The research in [16] illustrates the complexities involved in deploying cloud-based ship IoT solutions. Cloud platforms are extensively utilized to develop large-scale IoT environments, offering advantages such as scalability, flexibility, security, and availability.

In [17], major cloud-based IoT platforms like AWS IoT, Azure IoT, Watson IoT, PTC ThingWorx, and Google IoT are identified. The study confirms that these public clouds provide the necessary functions to build large-scale IoT environments. Among these platforms, AWS IoT is the most commonly used public cloud and has extensive references and ongoing research focusing on designing and implementing structures based on specific applications and purposes [18, 19].

Rules for accessing ship machinery and equipment data are currently being standardized and are in the draft document for the NP stage [20]. However, a wider data collection is necessary for MASS. As per [21], the Ship Software Logging System(SSLS), which documents the ship software version and operational status, has been standardized. The standardization of ship data management will continue to be explored on a cloud basis [22]. In [23] emphasizes the crucial role of Software Quality Assurance(SQA) performed alongside the development of ship technology.

3 Design of Data Platform

The design of a cloud-based data platform holds significant importance in facilitating efficient data management and analysis in MASS. In this section, we will discuss the proposed cloud-based autonomous shipping platform, which comprises six layers. Moreover, we will focus on the design elements, which are instrumental in building the proposed envisioned platform.

3.1 IoT-Based Cloud Data Platform

This paper proposes a six-layered system architecture for a cloud-based data platform: The Application Layer, Data Exchange Layer, Data Process Layer, Integration Layer, Process Layer, and Instrument Layer, as depicted in Fig. 2. The system architecture design aims to achieve integrated collection and management of ship data from various sources. It provides users and machines with professional and accurate data services necessary for autonomous operations through data analysis, ship-to-ship data exchange, and the utilization of digital twin technology.

The Application Layer serves as a configuration for data collection, analysis, and display. The Data Exchange Layer manages secure transmission of data between ships and onshore or between ships using certificate-based encrypted MQTT communication.

The Data Process Layer handles the storage and distribution of data collected from lower layers or generated by upper layers. Onshore management ensures scalability for storing data from multiple ships and maintains the data's lifecycle for sustainability. The Integration Layer integrates and operates the data generated and processed by each module, enabling the fusion of functions crucial for ship operations. Onshore acts as the INS and CAMS creating a digital twin environment that accurately describes and predicts the ship's operating conditions based on real and simulated data.

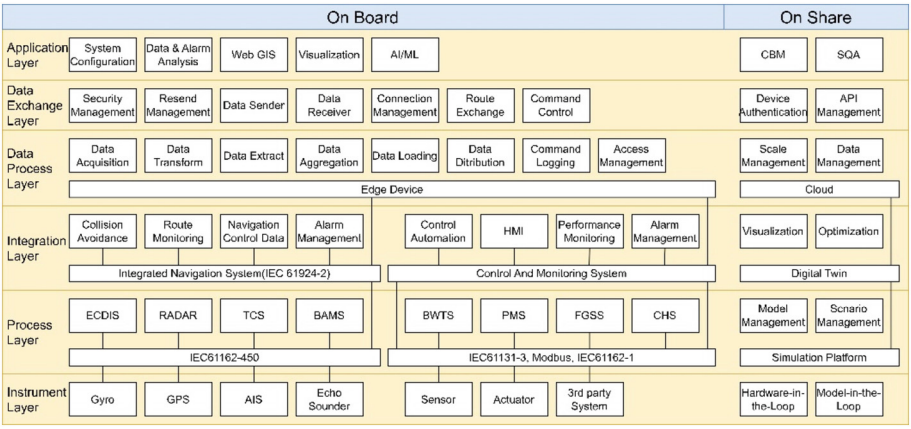


Fig. 2. Six-tier architecture for ship data and service interworking

The Process Layer consists of modules that perform ship functions by interacting with sensors or actuators in the Instrument Layer. Onshore should generate equivalent data as the ship's Process Layer modules through the Simulation Platform. Harmonizing with the Hardware-in-the-Loop Simulation (HILS) or Model-in-the-Loop Simulation (MILS) of the Instrument Layer, environments, models, and scenarios are configured to generate data similar to the actual ship environment. Compliance with established rules is essential for interconnecting the platform and data within the Integration Layer.

3.2 Ship Collision Avoidance

The Integration Layer in the ship's architecture is configured based on the INS and CAMS. Figure 3 the operational flowchart shows the real-time collection of planned route and navigation data from nearby ships using a data platform. The use of tools and technologies such as AIS, RADAR, GPS, cameras, Lidar, and others facilitates environment recognition to ensure collision avoidance.

In the past, only a single planned route was used, but in autonomous ships, various routes are generated, including collision-avoidance regional routes to prevent collisions and economically optimized routes created onshore or by remote sources. ECDIS must prioritize delivering these routes to the TCS.

Since the utmost priority is to operate without accidents, the collision avoidance route is given precedence, followed by the planned route determined by the navigator,

and then the economically optimized route generated from onshore sources. ECDIS determines which route to operate through the Route Priority Check, collecting data to detect collision risks and assess potential collisions based on the chosen route. Because MASS is currently in the research phase, the Route Priority Check should be flexible according to the degree of autonomous.

The data platform plays an instrumental role in providing planned routes and real-time navigation data from surrounding ships. The presence of nearby ships is identified by receiving Automatic Identification System(AIS) data, which is a legal requirement for vessels. Ships publish their planned navigation routes and navigation data to the data platform using the Maritime Mobile Service Identity(MMSI) as a topic. This allows the platform to collect and provide this information to subscribed ships simultaneously.

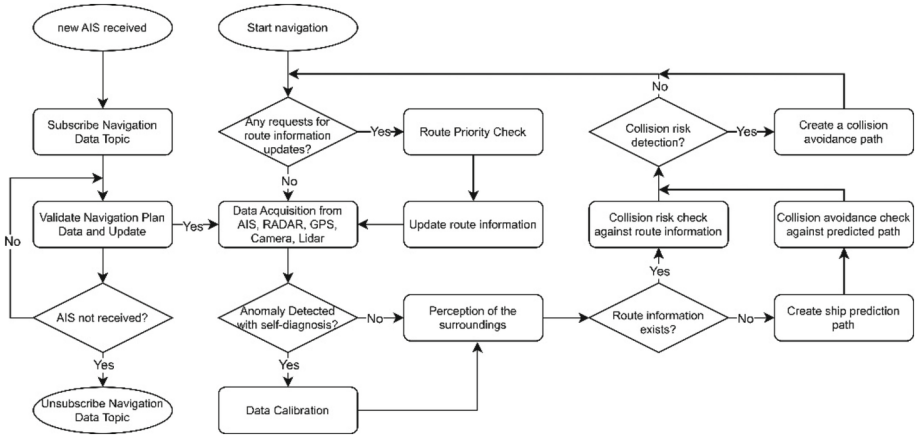


Fig. 3. Collision avoidance flowchart of INS equipment associated with data platforms

3.3 Architecture of Onboard Edge

The edge devices installed on ships fulfill the roles of the Application Layer, Data Exchange Layer, and Data Process Layer. Figure 4 demonstrates a module and data flowchart for an edge device responsible for these functions. The Transform component plays a vital role in converting the data model used on ships into a format compatible with the platform's time series data model. Due to variations in equipment manufacturers and data types across different ship types, this conversion process is necessary as it ensures uniformity in accessing data from the platform. The transformed time series data is both displayed to users through the onboard Frontend and stored in a database. The Data Receiver component is responsible for receiving the data and forwarding it to the Integration Layer, while the Command Control component scrutinizes and delivers control commands based on authority and the prevailing conditions.

Each module within the edge device is designed following the Micro Services Architecture(MSA) structure, allowing selective usage of onboard web services based on the

ship's specific requirements. This MSA architecture offers flexibility in utilizing different hardware configurations for the edge device [24].

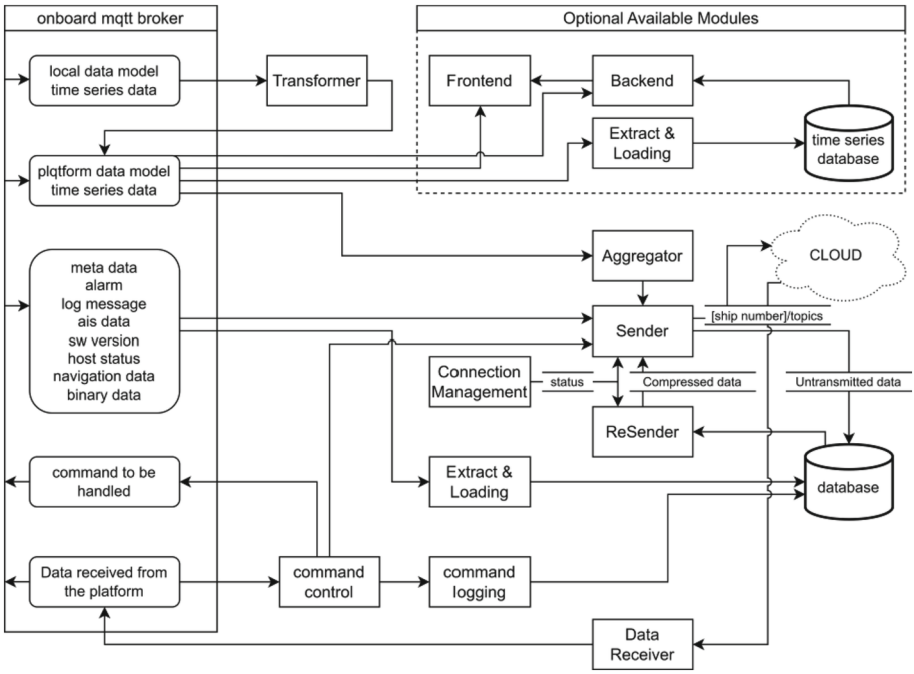


Fig. 4. Edge architecture to interface with the cloud

3.4 MQTT-Based Retransmission Method

To ensure real-time transmission of periodically generated ship data to the cloud, certificate-based encrypted MQTT communication is employed. However, due to the use of satellite communication by ships, there may be instances where the connection is unstable, and communication may be interrupted for several hours when entering shaded areas. To address this issue, the MQTT protocol offers QoS (Quality of Service) and Persistence session functions.

QoS ensures reliable transmission of data. QoS0 transmits data once, QoS1 sends data at least once, and QoS2 ensures data is sent exactly once. QoS1 and QoS2 messages that occur during disconnection are stored in a queue, and upon reconnection, transmission resumes based on the QoS mechanism.

Nevertheless, it's important to note that if the communication is disrupted for an extended period, real-time data transmission may be delayed due to the lengthy retransmission time required by QoS1 and QoS2 to ensure ordered transmission. For periodically generated time series data, the total delivery delay time TD^{tot} to the cloud service

can be defined as follows:

$$TD^{tot} = TD^{tr} + TD^{ag} + TD^{se} + TD^{cl} \tag{1}$$

- TD^{tr}: Transformer delay, which involves data model transformation, is relatively small as it mostly consists of processing time.
- TD^{ag}: Aggregator delay is determined by user-defined period P^{ud}, as data is calculated and transmitted periodically based on aggregation functions.
- TD^{se}: Sender delay represents the time required to transmit data to the cloud using the QoS1 mechanism.
- TD^{cl}: Cloud delay represents the time needed to deliver the service based on the cloud service configuration.

As the Transformer delay and Cloud delay are related to wired communication and processing time, they can be approximated as Eq. 2 when communication is unstable. The user-defined period P^{ud} can be set in seconds, and D^{us} is data stored without being transmitted during a period of communication loss. Equation 3 explains the time delay with QoS1 of sender for sending data at time t, where N^{us} refers to the total number of unsent data at that time and α means processing time for handling queue. If the message queue contains many unsent messages due to prolonged communication interruption, it may take a significant amount of time for all messages to be transmitted.

$$TD^{tot} = TD^{ag}\left(P^{ud}\right) + TD^{se}\left(D^{us}\right) \tag{2}$$

$$TD^{se} = \sum_{n=1}^{n=N^{us}} \left(RTT\left(D_{t-n}\right) + \alpha\right) \tag{3}$$

In this study, we address potential delays in data transmission as follows: QoS1 is employed for data transmission until a disconnection is detected. Upon detecting a disconnection, the application stores the unsent data and sends it separately to prevent any adverse effect on real-time data transmission.

Figure 5 shows the sequence diagram of the proposed packet slicing transmission method. To optimize the transmission process, the ReSender component updates the maximum payload function f^{mp} based on the time taken by previously sent payloads. This allows for the calculation of the maximum payload that can be sent within a given time interval T^{in} using the f^{mp} function.

Given that time-series data is generated periodically, the size of the payload can be minimized. This is achieved by representing multi-time domain data using only the start time and period information for each distinct time interval. After the real-time data D_t is transmitted, the remaining time interval T^{in} is utilized to transmit the multi-time data D^{mt}.

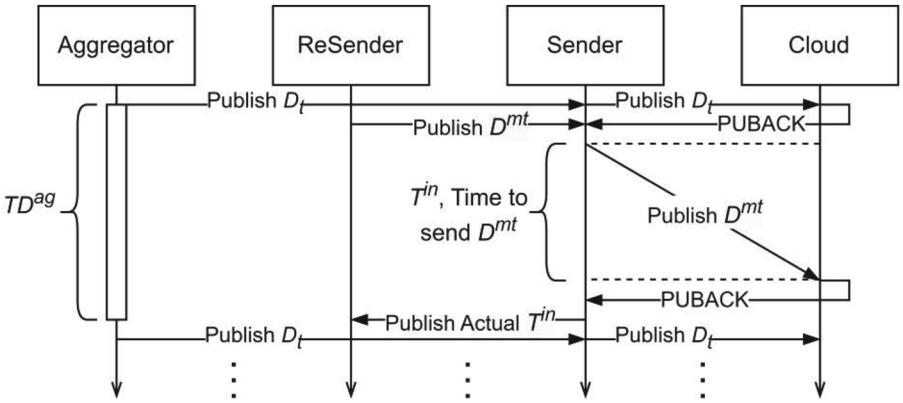


Fig. 5. Sequence diagram of packet slicing transmission method

Algorithm 1

Input : Aggregator time delay TD^{ag}, unsent data D^{us}
Initialize Number of unsent data N^{us}
while
Calculate max payload within T^{in} by using f^{mp}. $P^{max} = f^{mp}(T^{in})$
Compress data to multi-domain data as much as max payload. $D^{mt} \leftarrow D^{us}$
Publish and Update f^{mp} by using D^{mt} and Actual T^{in}.
Until(N^{us}==0)

4 Implementation and Simulation

4.1 Implementation with AWS Cloud Architecture

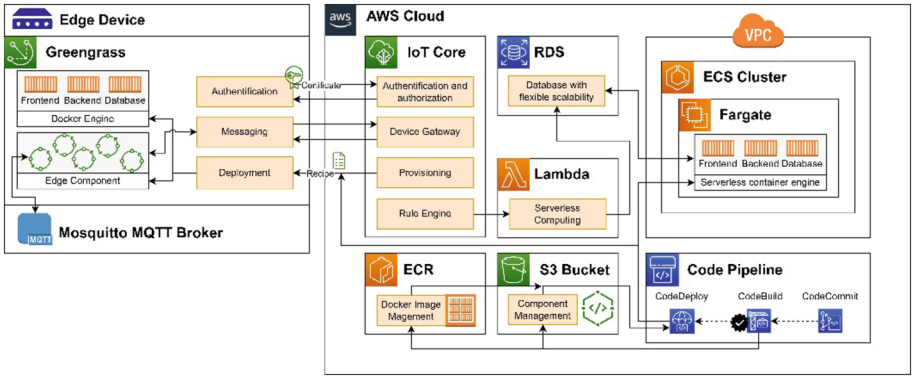


Fig. 6. Platform architecture consisting of AWS Component

The architecture, shown in Fig. 6, was designed using various AWS Cloud components. The proposed Edge Server structure was implemented on Edge devices using AWS Greengrass, an edge runtime environment.

We containerized the web service frontend and backend, and utilized Elastic Container Registry(ECR), a managed container image registry service, for registration and distribution. This allowed for improved scalability and flexibility in the cloud, with operations conducted on the serverless container engine service, Fargate, and on the Edge, employing the Docker engine. Real-time data collected through IoT Core was processed using Lambda, a serverless computing service, and directed to the Rule Engine based on the specified topic for Extract, Transform, Load(ETL) processing.

To validate the consistency of data collection between the ship and the onshore, we employed a ship operation simulator. By simulating ship operations and generating periodically changing simulation data on the edge, we could verify the coherence of the collected data. Figure 7 illustrates the interface where ship navigation data is collected and displayed on the WEB GIS service. Through this setup, we could collect and monitor real-time AIS information, track the location and navigation details of ships. Figure 8 presents the results of the simulation data collection, highlighting data consistency. The left side portrays the data trends accessed through Edge Web Services, while the right side illustrates the data trends accessed and confirmed through the cloud-based web service.



Fig. 7. WEB GIS Interface for Vessel Navigation Data Verification

4.2 Simulation of Retransmission Method

In order to evaluate the performance of the proposed retransmission method, we utilized end-to-end delay values for QoS1 in a 3G wireless environment, as reported in [24], as our simulation data. The simulations revealed a delay of approximately 0.5 s for a payload of 1000 and around 0.8 s for a payload of 16000. Considering that each tag has a length of about 10 bytes, we assumed the aggregation and generation of 100 tags per second, resulting in a total of 60 s required to detect network disconnection. The simulation results are presented in Fig. 9.

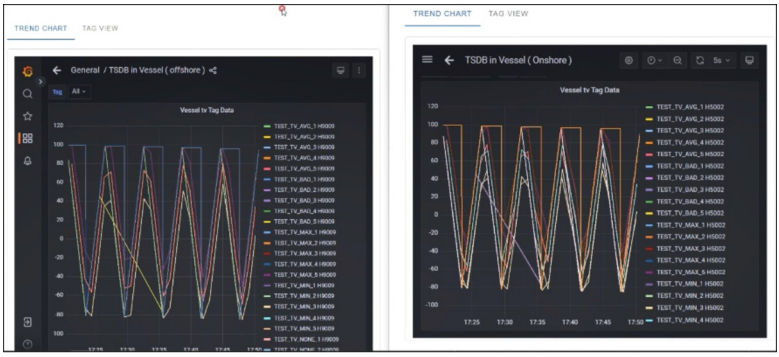


Fig. 8. Data consistency check interface between ship and onshore

The conventional method, which solely relied on MQTT QoS1, demonstrated a linear increase in delay in line with the FIFO principle, as the quantity of unsent data piled up in the queue with the duration of the disconnection. On the other hand, the proposed method utilizes the QoS1 queue only until the timeout is recognized. This strategy effectively curtails the accumulation of messages in the queue, thereby managing the delay. Furthermore, by employing a compression technique that converts multiple data into a multi-time domain with a compression rate of approximately 70%, we can send three compressed D^{mt} from T^{in}. This approach enables a reduction of approximately 34.6% in the time required to transmit all data, compared to the conventional method.

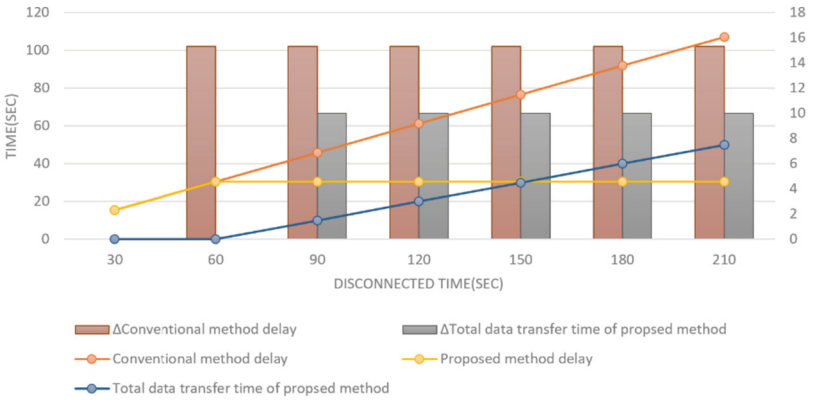


Fig. 9. Simulation results of existing and proposed methods

5 Conclusion

The paper highlights the significance of establishing platforms for MASS and propose 6-layer architecture. These platforms facilitate efficient data management, real-time monitoring, and well-informed decision-making, all of which contribute significantly to the

overall success of autonomous maritime operations. Additionally, the paper proposes an MQTT-based communication method to assure reliable and real-time connectivity between vessels and the onshore platform. This method improves real-time by mitigating data delay issues in unstable wireless communication environments. The implementation and evaluation of the proposed data platform and MQTT communication method were conducted using AWS Cloud, demonstrating their feasibility and effectiveness.

By addressing crucial aspects such as data integration, collision avoidance, real-time connectivity, and communication resilience, this paper contributes to the development and implementation of comprehensive solutions aimed at enhancing MASS. In future work, we will continue to investigate the implementation and performance evaluation of essential functions required for MASS via the platform.

Acknowledgements. This research is a part of 'AI-based heavy cargo ship logistics platform demonstration project' hosted by the Ulsan ICT Promotion Agency, supported by the National IT Industry Promotion Agency and the Ministry of Science and ICT (Project number: S1510–22-1001).

References

1. Wróbel, K., Gil, M., Montewka, J.: Identifying research directions of a remotely-controlled merchant ship by revisiting her system-theoretic safety control structure. Safety Sci. **129** (2020)
2. Mauro, F., Kana, A.A.: Digital twin for ship life-cycle: a critical systematic review. Ocean Eng. **269** (2023)
3. IMO Working group report in 100th session of IMO Maritime Safety Committee for the regulatory scoping exercise for the use of maritime autonomous surface ships (MASS). Maritime Safety Committee 100th session, MSC 100/ WP.8 (2018)
4. Fan, C., Wróbel, K., Montewka, J., Gil, M., Wan, C., Zhang, D.: A framework to identify factors influencing navigational risk for Maritime autonomous Surface Ships. Ocean Eng. **202** (2020)
5. Martelli, M., Virdis, A., Gotta, A., Cassarà, P., Di Summa, M.: An outlook on the future marine traffic management system for autonomous ships. IEEE Access **9** (2021)
6. Alqurashi, F.S., Trichili, A., Saeed, N., Ooi, B.S., Alouini, M.-S.: Maritime communications: a survey on enabling technologies, opportunities, and challenges. IEEE Internet of Things J. **10** (2023)
7. Cho, Y., Han, J., Kim, J.: Efficient COLREG-compliant collision avoidance in multi-ship encounter situations. IEEE Trans. Intell. Trans. Syst. **23** (2022)
8. Akdağ, M., Solnør, P., Johansen, T.A.: Collaborative collision avoidance for maritime autonomous surface ships: a review. Ocean Eng. **250** (2022)
9. Nielsen, R.E., Papageorgiou, D., Nalpantidis, L., Jensen, B.T., Blanke, M.: Machine learning enhancement of manoeuvring prediction for ship Digital Twin using full-scale recordings. Ocean Eng. **257** (2022)
10. IEC Standard.: Maritime navigation and radiocommunication equipment and systems - Integrated navigation systems (INS) (2021)
11. Jeong, B., Kim, M., Park, C.: Decarbonization trend in international shipping sector. J. Inter. Maritime Safety, Environ. Affairs Shipping **6** (2022)
12. Madusanka, N.S., Fan, Y., Yang, S., Xiang, X.: Digital twin in the maritime domain: a review and emerging trends. J. Marine Sci. Eng. (2023)

13. CIMAC Guideline From CIMAC WG20 System Integration: Virtual system integration & simulation - a performance-oriented approach for guiding system Simulation in the Field of Hybrid Marine Applications (2023)
14. Wang, J., Xiao, Y., Li, T., Philip Chen, C.L.: A survey of technologies for unmanned merchant ships. IEEE Access **8** (2020)
15. Zhang, J., Wang, M.M., You, X.: Maritime autonomous surface shipping from a machine-type communication perspective. IEEE Commun. Mag. (2023)
16. Cankar, M., Stanovnik, S., Cankar, M., Stanovnik, S.: In: 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (2018)
17. Barros, T.G.F., et al.: IEEE Access **10** (2022)
18. Kodali, R.K., Sabu, A.C.: Aqua monitoring system using AWS. In: 2022 International Conference on Computer Communication and Informatics (ICCCI) (2022)
19. Battula, S., Kumar, M.N.V.S.S., Panda, S.K., Rao, U.M., Laveti, G., Mouli, B.: Online Ocean Monitoring using Edge IoT. Online Ocean Monitoring using Edge IoT (2020)
20. Ando, H.: Activities of smart ship application platform (SSAP) project. JSMEA (2020)
21. ISO.: ISO 24060:2021 Ships and marine technology — Ship software logging system for operational technology (2021)
22. Lim, J.-H., Kim, J.-H., Huh, J.-H.: Recent trends and proposed response strategies of international standards related to shipbuilding equipment big data integration platform. Quality Quantity **57**, 863–884 (2023)
23. Kang, J., Ryu, D., Baik, J.: Predicting just-in-time software defects to reduce post-release quality costs in the maritime industry. Softw. Pract. Exper. 1–24 (2020). https://doi.org/10.1002/spe.2927
24. Liu, H., Jurdana, I., Lopac, N., Wakabayashi, N.: BlueNavi: a microservices architecture-styled platform providing maritime information. Sustainability **14**, 2173 (2022). https://doi.org/10.3390/su14042173
25. Lee, S., Kim, H., Hong, D.-k., Ju, H.: Correlation analysis of MQTT loss and delay according to QoS level. In: The International Conference on Information Networking 2013 (ICOIN), Bangkok, Thailand, pp. 714–717 (2013). https://doi.org/10.1109/ICOIN.2013.6496715

Deep Learning-Based Tag Mapping Automation of Ship Data Models with Natural Language Processing

Jiawei Huang, Hyoseong Hwang(✉), and Inwhee Joe(✉)

Department of Computer Science, Hanynag University, 222 Wangsimni-ro, Seongdong-gu, Seoul, Korea
hjw52237@gmail.com, hshwang@hd.com, iwjoe@hanyang.ac.kr

Abstract. The current Maritime Autonomous Surface Ship (MASS) presents many challenges and complexities in terms of risk, stability and implementation. To overcome these obstacles and realize the MASS vision, it has become imperative to develop a shore-based data platform capable of monitoring and supporting ship conditions in real time. In the data platform, a tag mapping operation is required to collect ship data onshore. The traditional approach is usually for the designer to perform the mapping manually by looking at the description and performance of the ship's I/O list. Such an approach usually requires a lot of design M/H and may lead to human errors, wasting a lot of time and resources. With the rapid development of deep learning, it becomes possible to realize the tag mapping automation task by converting it into a natural language classification task. Therefore, we propose a deep learning-based framework for tag mapping automation of ship data. Specifically, we consider the tag mapping task as a natural language classification task as follows: Classify the ship data into the corresponding platform data by a natural language classification model, and select the corresponding rules to achieve the tag mapping according to the classification result. Our proposed method reduces the manual involvement in the label mapping operation, minimizes the risk of manual errors, and saves resources. *abstract* environment.

Keywords: Tag Mapping Automation · Deep Learning · Natural Language Processing

1 Introduction

Recently, the rapid development of automatic monitoring systems incorporating Internet of Things (IoT) technology and the application in the ship sector have led to more efficient and safe ship operations. Combined with these new technologies, new types of ship operations such as smart ships, remote operating ships, and digital twin ships are constantly being developed [6,14]. Research on the automation of ship navigation and machinery control by combining information technology and IoT technologies has been a major topic in the maritime

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 221–232, 2024.
https://doi.org/10.1007/978-3-031-53549-9_22

field for the last decade. One of them is the integrated navigation system (INS) responsible for navigation, which is an important component of the ship automation system. The ship integrated navigation system (INS) can collect, store and display ship navigation data through edge servers on board and cloud servers on shore. Collecting and storing ship data onshore requires a tag mapping operation, where we need to map the ship data tags to platform data tags. The traditional solution is usually for the designer to do the tag mapping operation manually by looking at the description and performance of the ship's I/O list.

However, manual methods usually require a lot of design M/H and can be subject to human error, wasting a lot of time and resources. The composition of labels is based on natural language text, so we combine relevant techniques in the field of natural language processing to implement tag mapping, In the field of natural language processing, sequence to sequence model [12] is often used to deal with machine translation tasks, and machine translation tasks [9] can usually be regarded as mapping tasks of different languages. Specifically, sequence to sequence model [12] contains an encoder and a decoder, the text sequence is encoded into feature representation by the encoder, and then decoded by the decoder to get the target vector. Therefore, inspired by the sequence to sequence based machine translation, it is feasible to consider the tag mapping automation of ship data as a text generation task. Ideally, the ship data labels are input into the sequence to sequence model and then the target labels are generated directly, however, this method requires a large amount of data, and in our task dataset, the style of the data has a certain style, and it is difficult to standardize the style of the labels generated by the generative method.

In natural language processing, the accuracy of text classification task [7] has been greatly improved due to the emergence of Pre-training model Bert [1]. Bert is a pre-trained model for language encoding based on the Transformer. We mainly use its encoder part. Specifically the text sequence can be categorized by text classification model. Inspired by the text classification task, we can input the tag of ship data as text sequences into the model, and then consider the corresponding mapping targets as categories, and categorize each label through the text classification model, so as to automate the tag mapping. Specifically our approach has two stages,1) we use a pre-trained language model Bert [1] to learn the deep-level features of the tags, and then initially categorize the tags through the text classification model, 2) pre-define the specific generation rules for each target category, and then find the corresponding rules for generating the target labels based on the classification results. Due to the specificity and specialization of the label mapping task for ship data, the specific rules for generating target labels need to be designed by human beings. Combining AI technology and pre-defined rules, we propose a solution to automate the tag mapping of ship data, which can greatly reduce the human involvement in the ship tag mapping operation, reduce the possibility of human errors and save resources.

Specifically, we take the onshore data as the target labels and each label as a category. Next, the descriptive labels of the ship data will be encoded by the pre-trained language model in NLP technology, and then the probability of each

category of the target labels will be output by the decoder, and the category with the highest probability is the mapped target label. This cleverly converts the label mapping task into a classification task. We then predefine the label generation rules to standardize the format of the generated target labels, thus automating the label mapping of ship data. In summary, our main contributions are as follows:

- Sorted out the process of automating ship data mapping in the Integrated Ship Navigation System (INS).
- Proposed AI technology combined with a rule-based approach to implement ship data tag mapping operations.

2 Related Work

2.1 Tag Mapping Automation

The tag mapping task is to map source tags to target tags, and the traditional solution is usually for designers to manually map each source tag to a target tag, which is inefficient and prone to human error. With the development of AI technology and deep learning, it is feasible to automate the tag mapping task. Usually, the solution methods for tag mapping tasks can be divided into generation methods and classification methods. The generative approach is to view the tag mapping task as a translation task, and translate the source tags into the target tags through a translation model. The commonly used model is sequence to sequence model [12], and nowadays, research commonly uses pre-trained language models with sequence to sequence architecture such as T5 [8], GPT-2 [3] to implement machine translation [16] and generation tasks. The classification approach is to view the mapped tags as a classification task and the target tags as individual classes, and categorize each source tag into each target tag class by the classification model. Generative methods usually require a large amount of data for training, and classification methods are usually used to obtain better results in low-resource situations.

2.2 Natural Language Processing

Since the labels of ship data are composed of natural language text, the techniques related to natural language processing can be well used on the automation task of ship data tag mapping. In the past, in order to solve text classification problems, traditional machine learning methods, such as rule-based models, decision tree models [11], or probability-based models, naive Bayes classifiers [10], were often used. However, with the development of deep learning, in recent years, more deep learning models are used to solve text classification problems. LSTM [2] is commonly used to process text sequences, and Transformer [13] introduces a self-attentive mechanism to solve the information loss problem caused by transforming long sequences to fixed-length vectors. Bert [1] is a powerful pre-trained language representation model based on Transformer for embedding text

sequences into representation vectors. Recent research has typically introduced new layers at the end of the Bert model for use in downstream tasks in different domains, such as introducing MLPs at the last layer of bert for classification tasks. In the latest study, deep learning techniques like XLNet[15] and RoBERTa [4] have attained some of the biggest performance jumps for text classification problems. Our task is to use Bert as the basic backbone and use another layer for classification, thus initially implementing tag mapping.

3 Dataset

We use the I/O lists provided by ships as the training set. The dataset we use mainly consists of two parts: ship data as the source data and shore data as the target data. The ship data consists of tag name and description, and the shore data consists of LOCAL THING and Local property. There are 4470 data items in the dataset, among which there are 79 categories of LOCAL THING. The detailed construction of the I/O list is shown in the Table 1, where Tag name and Description from the ship data are used as inputs to the model, and the Label corresponding to the shore data are **LOCAL THING** and **LOCAL PROPERTY**. We want to map the ship data, which is difficult for people to use on shore, to the shore data, which is easier for people to use. We use 80% of the data as training data and 20% of the data as test data. An example of the dataset is shown in the Table 2.

Table 1. Dataset Construction

Object	Total number	Number of Categories
Tag Name	4470	-
Description	4470	-
LOCAL THING	4470	81
LOCAL PROPERTY	4470	85

4 Method

Our proposed method is used to automate the tag mapping of ship data. We first standardize the operation flow of tag mapping automation. We then detail a deep learning and rule-based approach to tag mapping automation.

Table 2. Dataset Example

Object	Example
Tag Name	BL_AC_160
Description	NO.1 BALLAST, BILGE, FIRE & GS PP RUNNING
LOCAL THING	BallastBilgeFireGSPump1
LOCAL PROPERTY	RunningState

4.1 Tag Mapping Automation for Ship Data

In order to better apply our method to the field of Maritime Autonomous Surface Ship (MASS), we have formulated the tag mapping automation process. We standardized the interaction flow between the user, the database and the tag mapping automation module. As shown in Fig. 1. First, the user calls the mapping service through the WEB Service UI by entering the Tag and running the automatic mapping function. The tag information is passed through the tag mapping automation module, which returns the desired target tags, LOCAL THING and LOCAL PROERTY, to the user through the WEB Service UI, and the user confirms the returned target tags and corrects any errors. If the tag cannot be mapped accurately, the user will do the mapping by manual operation. As shown in the Table 3. There are three types of target labels returned through the Label Mapping Automation Module,(1)Recommended local thing/property,(2)Cannot recommend due to lack of mapping information,(3)Existing mapping relationship cannot be judged due to lack of similarity. Depending on the type of the returned target label the user performs a confirmation pass, corrects the returned target label, and manually adds the target label. Next, the user downloads the mapping data through the WEB Service UI and downloads it to the ship's edge through the backend. The mapping data is also stored to the database. The new mapping data can also be used as training data for further training of the tag mapping automation model, thus improving the model performance.

Table 3. Three types of target tags returned by the tag mapping automation module with corresponding user actions

Type for Return Label	User action
Recommended local thing/property	Confirm the recommendation, if it is correct, proceed to the next step.
Cannot recommend due to lack of mapping information	Manual input of target labels
Existing mapping relationship cannot be judged due to lack of similarity	After manual confirmation, make corrections or proceed to the next step

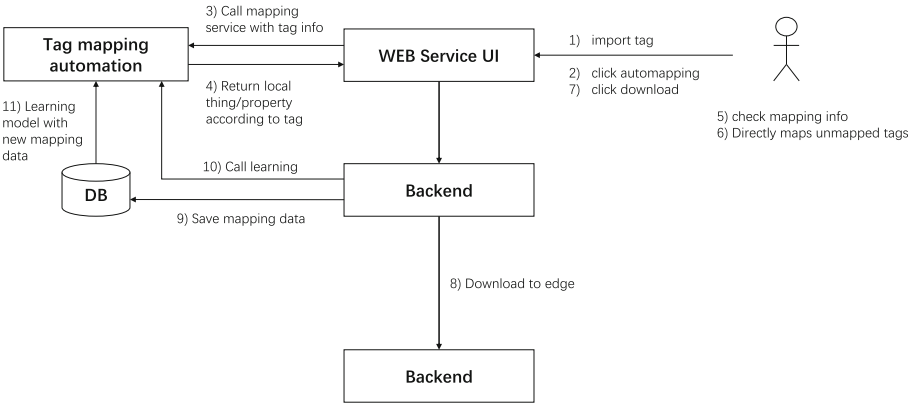


Fig. 1. The overview of Ship Data Tag Mapping Automation Framework

4.2 Tag Mapping Automation Module

Rencently, mapping ship data to shore data is mostly performed manually according to the I/O list, however, such a way has the problems of low efficiency and waste of resources. In order to improve the efficiency of label mapping, we propose a method based on deep learning and combined with rules to automatically complete the tag mapping. Our approach is divided into two steps. First the source labels (tag name and description) pass through a deep learning classification model to get the Local Thing and the generation type of Local Property. Next, according to the generation type of Local Thing to select the Local Property generation rules, and then based on the rules to generate Local Property, so as to realize the mapping from the source label (label name and description) to the target label (Local Thing and Local Property). Details are introduced in the subsections and the overview of Tag Mapping Automation are shown in Fig. 2.

Preliminary Mapping Based on Classification. We use two classification models to implement the preliminary mapping. One classification model is used to get the Local Thing of the target label, and one classification model is used to get the generation type of the target label Local Property. We first pre-process the Local Thing by converting the category of Local Thing to the corresponding Local_Thing_ID. The Local Thing_ID is used as the labels of the classification model. At the same time, we summarize the format of the Local Property and get the target label Local Property's generation type as the labels of the classification model. The generation rules of Local Property can be selected according to the generation type of the target label Local Property. The correspondence between Local Thing and Local Thing ID is stored in the dictionary ID_to_Thing is shown in the Table 4. The correspondence between Local Property generation rule and generation type is stored in the dictionary Type_to_Rule is shown in the Table 5.

We use the Bert [1] pre-trained model as the classification model, and input the description of the source label into the classification model, and the model outputs the Local Thing_ID. We input the description of the source label into another classification model, and the model outputs the generation type of Local Property.

$$\begin{aligned} Local\_Thing\_ID &= Bert_1(Description), \\ Local\_Property\_Type &= Bert_2(Description). \end{aligned} \tag{1}$$

The formulas are shown above, where $Bert_1$ denotes the classification model used to categorize Local Thing, and $Bert_2$ denotes the classification model used to categorize the Local Property generation type.

Table 4. The overview of dictionary ID_to_Thing

Local Thing	Local Thing ID
BallastBilgeFireGSPump1	1
BOGThing	2
CargoHandling	3
.....	.....
Stbd_SteerGear2	79
None_Recommend	80
Not_Sure	81
Total Local Thing	81

Table 5. The overview of Local Property and dictionary Type_to_Rules

Local Property	Generation Rules	Generation Type
RunningState	running state + (null/FWD/REV)	1
CT1_P_Level	CT + number + (P/S) + Level	2
CT + number + (P/S) + Level	CT + number + (P/S) + Volume	3
.....	.....	.....
StoppingState	StoppingState	84
FBOG	Human Input	85
Total	85	85

Rule-Based Local Property Generation. The Preliminary mapping of the classification model allows us to obtain Local Thing ID and the generation types of Local Property. According to the dictionary Thing_to_ID, which shown in

Table 4, we can get Local Thing by Local Thing_ID, and according to the dictionary Type_to_Rules, which shown in Table 5, you can find the generation rules of Local Property by generation_Type, so that we can generate Local Property based on the generation rules, then we generate Local Property based on the rule, which taking Tag_Name and Descripition as inputs.

$$\begin{aligned} Local\_Thing &= Dict\_ID\_to\_T(Local\_Thing\_ID), \\ Local\_Property\_Rule &= Dict\_T\_to\_R(Local\_Property\_Type). \\ Local\_Property &= Local\_Property\_Rule(Tag\_Name + Description). \end{aligned} \tag{2}$$

Where Dict_ID_to_T denotes dictionary ID_to_Thing and Dict_T_to_R denotes dictionary Type_to_Rules.

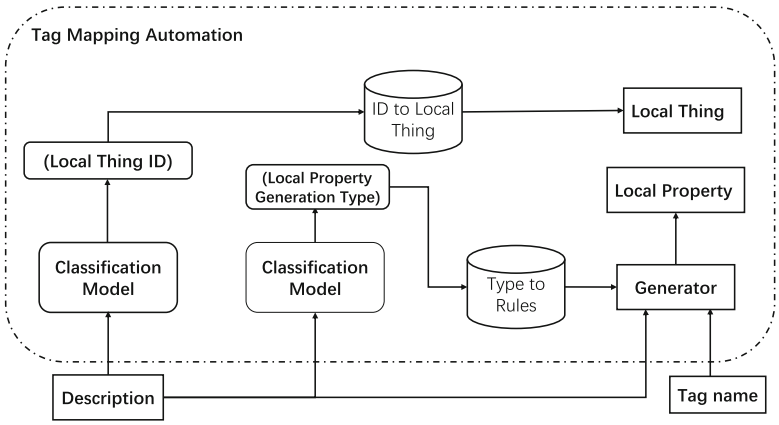


Fig. 2. The overview of Tag Mapping Automation

5 Experiments

5.1 Implementation Details

Since the success rate of the implementation of the rule-based part is 100%, our experiments only focus on the classification part. We use Bert as the encoder and a neural layer as the classifier for training, LOCAL THING and LOCAL PROPERTY are trained with different Bert and classifier for each target label. The Batch size is set to 64, 40 epochs are trained, the learning rate is set to 5e-5, and the optimizer is chosen to use AdamW [5]. We train the same task five times each, and the average of the results of the five tests is used as the final result. In order to prove the validity of our method, We also tested randomized

classification. We test five times randomly using random assignment and take the average value as the final test result for comparison.

5.2 Experimental Results

The results are shown in the Table 6. From the experimental results, our method is 0.7 higher than the random Local Thing_ID in Local Thing_ID classification accuracy and 0.59 higher than the random Local Property_Type in Local Property_Type classification accuracy, which proves the effectiveness of our method. We represent the comparison of classification accuracy in Fig. 3, showing that our method greatly outperforms the randomized method in terms of classification accuracy.

Table 6. Comparison of experimental results between the randomized method and our method

Object	Accuracy	Average Loss
Random LOCAL THING_ID	0.01	949.7
Random LOCAL PROPERTY_Type	0.01	1365.9
LOCAL THING_ID	0.95	0.34
LOCAL PROPERTY_Type	0.93	0.36

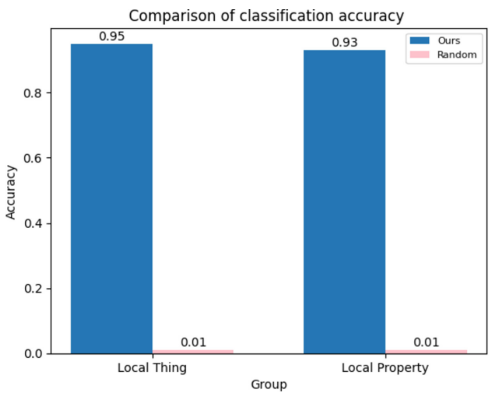


Fig. 3. The comparison of classification accuracy

5.3 Loss and Accuracy Trend Analysis

We plotted the loss curve and the accuracy curve during training. The top two plots show the loss and accuracy changes during training for the classification model trained for Local Thing. The bottom two plots are the loss and accuracy changes during training for the classification model trained for Local Property. As shown in the Fig. 4, the two classification models in the training process, with the increase of epoch, the Loss decreases rapidly in the first 15 epochs, and between 15 and 40 epochs, it decreases slowly, but the decrease is not obvious, and the decrease of the Loss tends to stabilize. Meanwhile Accuracy increases rapidly in the first 15 epochs and slowly between 15 and 40 epochs, but the result tends to stabilize. Through the trend of Loss and Accuracy during the training process, we can see that our model

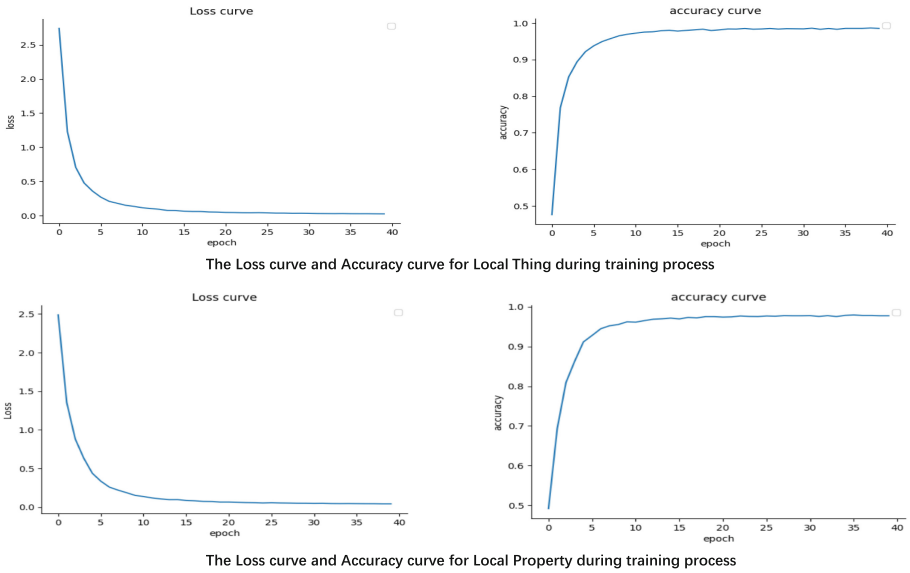


Fig. 4. The Loss and Accuracy curve for Local Thing and Local Property

5.4 Trend Analysis of Training Time

For the analysis of training time trends, we measure the time spent on each epoch and plot it on a Fig. 5. The figure shows the trend of training time, in which all epochs take less than 22 s to train, the epoch with the least training time is at 19.37 s, and there is not much difference in training time between each epoch, which shows that our model has good training efficiency.

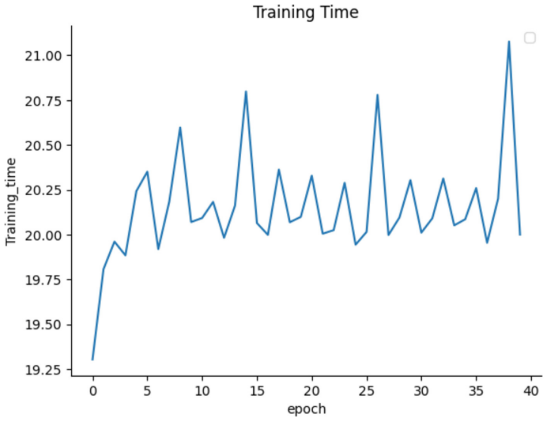


Fig. 5. Training Time

6 Conclusion

We present a framework for automating tag mapping of ship data and platform data for the first time. We first standardize the process for automating tag mapping of ship data and platform data. At the same time, our proposed method greatly reduces the human involvement in the tag mapping operation of ship data by converting the tag mapping task into a classification task, and then reducing the possibility of human errors and the human and time resources required for the tag mapping operation by deep learning model with rules.

Acknowledgements. This research is a part of 'AI-based heavy cargo ship logistics platform demonstration project' hosted by the Ulsan ICT Promotion Agency, supported by the National IT Industry Promotion Agency and the Ministry of Science and ICT (Project number: S1510-22-1001).

References

1. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
2. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
3. Lagler, K., Schindelegger, M., Böhm, J., Krásná, H., Nilsson, T.: Gpt2: empirical slant delay model for radio space geodetic techniques. Geophys. Res. Lett. **40**(6), 1069–1073 (2013)
4. Liu, Y., et al.: Roberta: a robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692 (2019)
5. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101 (2017)

6. Mauro, F., Kana, A.A.: Digital twin for ship life-cycle: a critical systematic review. Ocean Eng. **269**, 113479 (2023)
7. Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., Gao, J.: Deep learning-based text classification: a comprehensive review. ACM Comput. Surv. (CSUR) **54**(3), 1–40 (2021)
8. Raffel, C., et al.: Exploring the limits of transfer learning with a unified text-to-text transformer. J. Mach. Learn. Res. **21**(1), 5485–5551 (2020)
9. Ramachandran, P., Liu, P.J., Le, Q.V.: Unsupervised pretraining for sequence to sequence learning. arXiv preprint arXiv:1611.02683 (2016)
10. Rish, I., et al.: An empirical study of the naive bayes classifier. In: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, vol. 3, pp. 41–46 (2001)
11. Song, Y.-Y., Ying, L.U.: Decision tree methods: applications for classification and prediction. Shanghai Arch. Psychiatry **27**(2), 130 (2015)
12. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks. Advances in neural information processing systems, 27 (2014)
13. Vaswani, A., et al.: Attention is all you need. Advances in neural information processing systems, 30 (2017)
14. Wróbel, K., Gil, M., Montewka, J.: Identifying research directions of a remotely-controlled merchant ship by revisiting her system-theoretic safety control structure. Saf. Sci. **129**, 104797 (2020)
15. Yang, Z., et al.: Xlnet: generalized autoregressive pretraining for language understanding. Advances in neural information processing systems, 32 (2019)
16. Zhang, J., Zong, C., et al.: Deep neural networks in machine translation: an overview. IEEE Intell. Syst. **30**(5), 16–25 (2015)



Deep Reinforcement Learning-Based Task Offloading in Multi-access Edge Computing for Marine IoT

Ducsun Lim(✉) and Dongkyun Lim

Department of Computer Engineering, Hanyang Cyber University, Seoul 04763, Republic of Korea
{4180032,eiger07}@hycu.ac.kr

Abstract. The recent surge in maritime activities has led to a significant demand for marine Internet of Things (M-IoT) devices. These devices are responsible for meeting strict requirements in resource-limited and complex maritime network environments. Advanced communication and computing solutions are imperative for addressing these challenges. Leveraging 6G-based multi-access edge computing (MEC) offers the potential to effectively process large-scale marine data, thereby meeting diverse needs across various marine application scenarios. This paper introduces a latency-sensitive and energy-efficient task offloading scheme (DLES) anchored on deep reinforcement learning (DRL). This scheme was designed to optimize task offloading in a maritime MEC setting, aiming to reduce delays and energy consumption. Simulation results validated the superior performance of the proposed scheme in terms of reduced latency and enhanced energy efficiency.

Keywords: Multi-access Edge Computing · Maritime Internet of Things · Task Offloading · Deep Reinforcement Learning

1 Introduction

Various information services in the marine domain, based on the Internet of Things (IoT), demand high-quality network transmission coupled with robust data processing and analysis [1]. This enormous quantity of ocean data is instrumental in environmental conservation, oceanographic research, natural disaster prevention, mineral exploration, and military applications. It is also important for ocean monitoring [2–4]. Marine weather data, imagery, and other pertinent information require swift analysis, processing, and transmission to accurately assess the magnitude, trajectory, and real-time status of disasters. Precise and extensive measurements are vital for the continuous monitoring of various physical phenomena, such as acoustics, vibrations, and visuals, in marine networks. As technology advances, vast amounts of real-time monitoring data amassed from terminals during maritime search and rescue endeavors can be promptly processed through the multi-access edge computing (MEC) framework. Consequently, maritime task processing can be executed quickly and accurately. However, because

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 233–244, 2024.
https://doi.org/10.1007/978-3-031-53549-9_23

of the absence of terrestrial infrastructure support, the architecture and computational resources of maritime communication networks are significantly constrained. Therefore, relying solely on existing technologies to accomplish communication and computational tasks is challenging.

With the widespread deployment of 5G communication networks, both industry and academia have initiated research on 6G to enhance network capabilities and extend its application scope [5]. A cornerstone of 6G research is the creation of a globally interconnected network encompassing space, air, land, and sea domains to offer unbroken global coverage [6, 7]. This research trajectory is instrumental in the rapid evolution of marine IoT (M-IoT). The proliferation of maritime endeavors underscores the growing anticipation of M-IoT-driven marine services. However, considering the intricate marine environment and finite resources, devising strategies to process vast oceanic data volumes efficiently with minimal cost and energy consumption remains a significant challenge. Edge intelligence, or MEC technology, within the 6G paradigm [8, 9] allows resource-intensive computations to be offloaded from centralized units to MEC servers, aligning well with maritime network prerequisites. However, most existing MEC studies have focused on terrestrial networks, leaving maritime communications comparatively uncharted.

This study delineates an MEC network architecture tailored for the M-IoT within the context of a 6G-integrated space-air-ground communication framework. Specifically, we focus on the task offloading aspect of the MEC within a maritime network milieu. We introduce a potent task offloading strategy coupled with a deep reinforcement learning (DRL)-driven offloading algorithm that amplifies energy efficiency while curtailing the processing delays of diverse tasks. The paper is structured as follows: Sect. 2 surveys and details the pertinent antecedent studies. Section 3 describes the overarching system architecture and model, delving into the mathematical methodologies pertinent to local computations and task offloading. Section 4 elucidates the mathematical modeling of the problem, algorithmic design, and operational tenets. Section 5 describes the performance assessment and presents the empirical findings. Section 6 provides a holistic recapitulation of the study, considering prospective research trajectories and potential enhancement domains.

2 Related Works

In a 6G-based IoT environment, the sheer volume of devices and the resulting data traffic are expanding rapidly. This burgeoning growth can strain communication networks that already grapple with limited resources. IoT core technologies play a significant role in this context. By offloading and processing computational tasks and requesting information from the MEC server, they can enhance the user experience by amplifying energy efficiency and curtailing transmission delays. Academia and industry are currently engaged in this research, which probes a plethora of potential solutions [10–20].

To date, research has predominantly focused on terrestrial communication. Specifically, MEC-related studies have predominantly targeted terrestrial networks. For instance, Zheng et al. [10] employed game theory to analyze dynamic computation offloading in MEC networks and unveiled a technique for allocating wireless channels

among users. Wang et al. [11] explored a single-user wireless MEC system in which a multi-antenna energy transmitter wirelessly powered the user via energy beamforming technology. Cheng et al. [12] proposed an enhanced learning methodology anchored in the Markov decision process for SAGIN to achieve swift convergence with minimal energy consumption.

It is worth noting that the bulk of MEC-oriented research has focused on terrestrial networks. Wei et al. [13] introduced MVR, a groundbreaking framework that facilitates computational offloading in MEC landscapes. He et al. [14] crafted diverse transmission strategies to reduce latency. Meanwhile, Liu et al. [15] postulated a technique based on continuous convex approximation coupled with a decomposition and iteration method to reduce the overall energy consumption of UAVs. Kim and Kim [16] designed and implemented an edge-update-centric motor-state-estimation algorithm, whereas Trinh et al. [17] envisioned low-latency processing for high-resolution visual data via MEC.

Venturing into the realm of 6G, Rodrigues et al. [18] synthesized the features of 6G networks using IoT and proposed a machine-learning-anchored server construction strategy within the 6G IoT framework. Their findings underscored superior performance compared with existing methodologies. Yang et al. [19] proposed a dynamic resource allocation architecture by embedding a rapid heuristic-based incremental allocation mechanism for real-time resource allocation. Yang et al. [20] highlighted the importance of maritime MEC and proposed a framework to enhance performance and energy efficiency. Their methodology employed a dual-stage offloading strategy to balance energy conservation and latency optimization.

Although these investigations have undeniably advanced the MEC domain, gaps remain, particularly in the case of M-IoT-based task offloading. To address this void, this study pioneers the task-offloading dimension within the M-IoT arena. Within M-IoT frameworks, premium MEC servers strategically stationed at the network's frontier empower users to tap into a rich tapestry of services. This is achieved by leveraging offloading technology that spans computing, storage, and random access technologies.

3 System Model

Maritime autonomous surface ships (MASS) are designed to establish and monitor collaborations at sea. Each MASS is equipped with sensors to capture and analyze images and videos of varying resolutions. However, given the processing capacity and energy constraints of MASS, certain task processing must be offloaded to MEC servers (such as satellites and large vessels) that possess superior computing capabilities or can connect to distant cloud data. This underscores the importance of enhancing task-processing efficiency. To address the demands of this scenario, we introduce an edge-computing network and propose an optimization algorithm aimed at improving energy efficiency and curtailing latency for MASS.

The integrated framework for the space-air-land-maritime communication network is illustrated in Fig. 1. The backbone network primarily comprises heterogeneous networks, and tasks that require extensive computing are processed in cloud-computing data centers. MEC servers are positioned in the peripheral areas of maritime networks,

such as UAVs and nearby ships. MASS serves as an end device responsible for monitoring and tracking tasks. In this context, we propose a delay-sensitive and energy-efficient computing offloading strategy that utilizes DRL in maritime IoT environments.

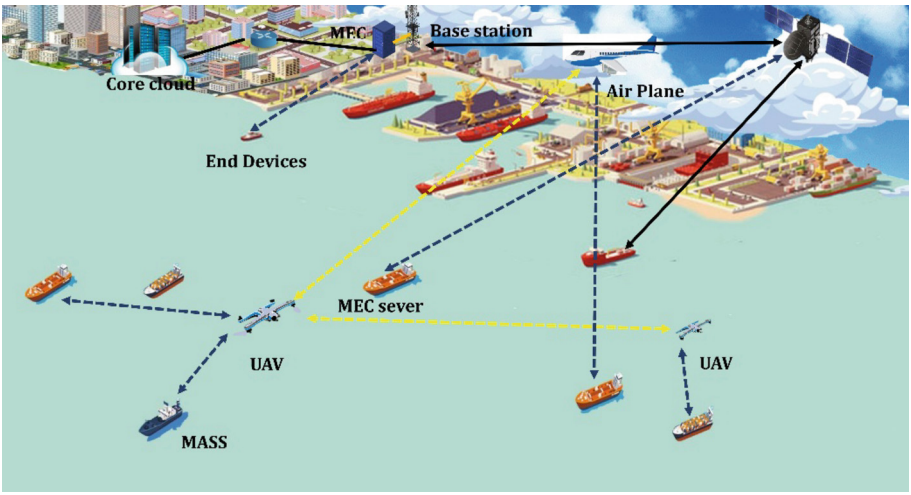


Fig. 1. Framework of maritime multi-access edge computing networks.

In a maritime edge computing network, we consider a scenario where devices are connected to a single access point (AP). The set of MASS is denoted as $M = \{m_1, m_2, \ldots, m_n\}$. Given that each AP houses an MEC server responsible for computation offloading, a MASS can offload tasks to this server via a wireless link. Each task of a MASS can be represented as $\tau_i = \{s_i, c_i, d_i\}$, where s_i indicates the data size, c_i specifies the CPU cycles needed to complete the task, and d_i denotes the task's deadline. In this scenario, under the premise that each task is indivisible and can only be processed by a single device, tasks can either be processed locally or offloaded to a superior edge server or to cloud computing. When multiple tasks need to be offloaded simultaneously, the MEC server in the AP determines the spectrum and computes the resource allocation for each MASS.

3.1 Communication Model

If a MASS cannot execute a computing task owing to resource constraints, the task is offloaded to either an MEC server or a remote cloud server. All MASS are within the communication range of the edge server, and any communication interference among the MASS can be disregarded. When a MASS decides to upload a task, the MEC server allocates bandwidth to it. $\sigma_i \in [0, 1]$ indicates the proportion of bandwidth allocated to each MASS. If σ_i is 1, the MEC server dedicates the full bandwidth to m_i. If σ_i is 0, no bandwidth is allocated. Using the Shannon formula, the upload and download speeds r_{m_i} of MASS m_i can be determined based on the upload and download speeds.

$$r_{m_i}^{ul} = \omega_i B_{m_i}^{e} log_2(1 + \frac{\rho_{m_i} g_{m_i}^{ul}}{\sigma^2}) \tag{1}$$

$$r_{m_i}^{dl} = \omega_i B_{m_i}^{e} \log_2(1 + \frac{\rho_{m_i} g_{m_i}^{dl}}{\sigma^2}) \tag{2}$$

In this context, $B_{m_i}^{e}$ represents the available channel bandwidth between m_i and the MEC server. ρ_{m_i} denotes the transmission power of the MASS. Meanwhile, $g_{m_i}^{ul}$ and $g_{m_i}^{dl}$ indicate the channel gains for upload and download, respectively, and σ^2 stands for the power of additive white Gaussian noise.

3.2 Computing Model

The maritime MEC architecture is designed to offload all tasks based on the network link status, with a specific emphasis on achieving acceptable latency. This design seeks a balance between delay and power consumption. Depending on the scenario, the MEC server determines which device should process the task. Tasks that cannot be processed locally are offloaded to the MEC server. If the MEC server does not have the necessary resources, the tasks are offloaded to the cloud. Parameters $\alpha_i, \beta_i, and \gamma_i \in [0, 1]$ are introduced to specify the processing locations for each task. The sum of these parameters always equals 1, meaning $\alpha_i + \beta_i + \gamma_i = 1$. If the task is processed locally, $\alpha_i = 1$. If it is processed on the MEC server, $\beta_i = 1$. If it is executed in the cloud, $\gamma_i = 1$. The definitions of the three computational models are as follows:

- Local computing: When α_i is set to 1, task τ_i is processed locally within MASS. The required clock cycle and clock frequency for completing the task in local MASS are denoted as $c_{m_i}^{l}$ and $f_{m_i}^{l}$, respectively. Computational time and energy consumption are calculated based on these parameters.

$$L_{m_i}^{loc} = \frac{c_{m_i}^{loc}}{f_{m_i}^{loc}} \tag{3}$$

$$E_{m_i}^{loc} = \phi (f_{m_i}^{loc})^2 c_{m_i}^{loc} \tag{4}$$

where ϕ represents the effective capacity set, and its value is given as 10^{-9} [21].

- MEC: When β_i is 1, the task τ_i is offloaded to the AP and processed on the MEC server. Once the MEC server completes the task, it returns the result to the MASS. The processing delay during this period primarily comprises the task processing time on the MEC server, the transmission delay from the MASS to the AP, and the time taken to send the processed results back to the MASS. Energy consumption includes the energy used for computation on the MEC server, the energy required for transmission from the MASS to the AP, and the energy expended when receiving the results. The energy and delay calculations are defined as follows:

$$L_{m_i}^{mec} = L_{m_i}^{mec_{tran}} + L_{m_i}^{mec_{proc}} + L_{m_i}^{mec_{recv}} = \frac{\tau_i}{r_{m_i}^{ul}} + \frac{c_{m_i}^{mec}}{\chi_{m_i} f_{m_i}^{mec}} + \frac{\tau_i}{r_{m_i}^{dl}} \tag{5}$$

$$E_{m_i}^{mec} = \frac{\rho_{m_i}\tau_i}{r_{m_i}^{ul}} + \phi(\chi_{m_i} f_{m_i}^{mec})^2 c_{m_i}^{mec} + \frac{\rho_{m_i}\tau_i}{r_{m_i}^{dl}} \tag{6}$$

$\chi_{m_i} \in \{0, 1\}$ represents the ratio of CPU allocated to different MASS. $c_{m_i}^{mec}$ Indicates the number of CPU cycles required to process the task, and $f_{m_i}^{mec}$ ⭤ denotes the clock frequency of the MEC server.

- Cloud computing: Given the limited energy and computing resources of MEC servers, the cloud is introduced as a supportive offloading platform. Tasks that require extensive CPU cycles and are challenging to handle solely on the MEC server are offloaded to the cloud server. Analogous to MEC, processing latency and energy consumption are defined as follows:

$$L_{m_i}^{cld} = L_{m_i}^{cld_{tran}} + L_{m_i}^{cld_{proc}} + L_{m_i}^{cld_{recv}} + \psi\tau_i = \frac{\tau_i}{r_{m_i}^{ul}} + \frac{c_{m_i}^{cld}}{\chi_{m_i} f_{m_i}^{cld}} + \frac{\tau_i}{r_{m_i}^{dl}} + \psi\tau_i, \tag{7}$$

$$E_{m_i}^{cld} = \phi(\chi_{m_i} f_{m_i}^{mec})^2 c_{m_i}^{mec} + \rho_{m_i}\left(\frac{\tau_i}{r_{m_i}^{ul}} + \frac{\tau_i}{r_{m_i}^{dl}} + \psi\tau_i\right), \tag{8}$$

where $c_{m_i}^{cld}$ and $f_{m_i}^{cld}$ denote the CPU cycle required to process the task and the allocated CPU clock frequency of the cloud server, respectively. ψ expresses the propagation delay parameter from the MASS to the cloud [22].

$$Cost_{m_i}^{tot} = x_i L_{m_i}^{tot} + y_i L_{m_i}^{tot}, x_i + y_i = 1, x_i, y_i \in [0, 1], \tag{9}$$

$$L_{m_i}^{tot} = \alpha_i L_{m_i}^{loc} + \beta_i L_{m_i}^{mec} + \gamma_i L_{m_i}^{cld}. \tag{10}$$

$$E_{m_i}^{tot} = \alpha_i E_{m_i}^{loc} + \beta_i E_{m_i}^{mec} + \gamma_i E_{m_i}^{cld}. \tag{11}$$

To achieve a balance between delay and energy consumption, the total consumption, denoted as $Cost_{m_i}^{tot}$, can be calculated using the aforementioned formulas, where $x_i$and $y_i$are the time and energy consumption weight coefficients, respectively. The larger the values of these parameters, the more significant their impact on the performance indicator [23]. The offloading optimization problem can be formulated as min $Cost_{m_i}^{tot}$.

$$minCost_{m_i}^{tot} = x_i L_{m_i}^{tot} + y_i L_{m_i}^{tot}, \tag{12}$$

$$\text{subject to: } x_i + y_i = 1,$$

$$x_i, y_i \in [0, 1],$$

$$\alpha_i, \beta_i, \gamma_i \in [0, 1].$$

$$\alpha_i + \beta_i + \gamma_i = 1.$$

We defined the problem of optimizing the equations for the total cost, delay, and energy consumption, along with their associated constraints.

4 Problem Formulation

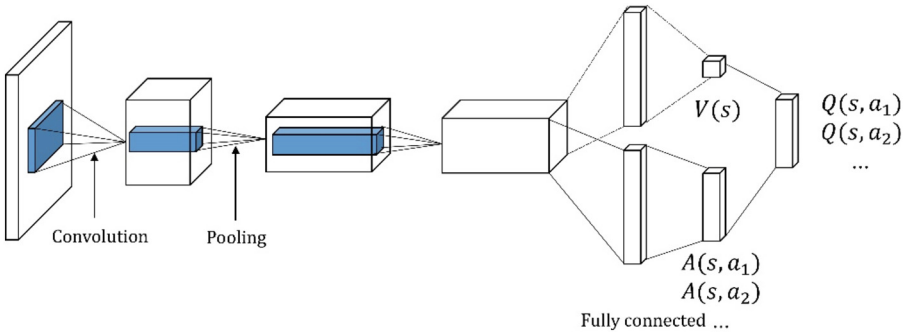


Fig. 2. Structure of dueling DQN.

We proposed a DRL-based, latency-sensitive, and energy-efficient task-offloading scheme (DLES) for maritime scenarios. Figure 2 illustrates the structure of the dueling deep Q-network (DQN). This problem can be modeled as a finite Markov decision process (MDP). In this setup, the MEC server functions as an agent and selects a task for each timeslot. It receives a reward function, denoted as $f_r\left(s_{t_{m_i}}, a_{t_{m_i}}\right)$, based on state $s_{t_{m_i}}$ and action $a_{t_{m_i}}$ in each time slot. The state space is expressed as follows:

$$s_{t_{m_i}} = \{m_i^{sel}, D, E_c, BW, CPU_{load}, s_i\} \tag{13}$$

m_i^{sel} indicates the offloading strategy for the current task and determines where the task should be processed. D is the deadline, and E_c is the energy consumed to complete the task. BW is the communication bandwidth for offloading, and CPU_{load} indicates the current CPU load status of the MEC server. Furthermore, s_i represents the task size. Action space $a_t = \{\alpha_i, \beta_i, \gamma_i\}$ is an option indicating where to process the current task. The immediate reward function $f_r(s_t, a_t)$ is the overall cost and should be designed to minimize the sum of latency and energy consumption. Therefore, the compensation function is defined as Eq. (13).

$$f_r(s_t, a_t) = -\iota_1 \times latency + \iota_2 \times (E_{max} - E_c) \tag{14}$$

where ι_1 and ι_2 serve as weights, highlighting the significance of latency and energy consumption, respectively. E_{max} denotes the maximum permissible energy consumption. Subsequently, the state function is updated using a recursive approach. The iterative formula to estimate the optimal Q value based on the dueling DQN method is as follows [24]:

$$\mathrm{Q(s, a)} = \mathrm{V(s)} + \left(A(s, a) - \frac{1}{|A|} \sum_{a\prime} A(s, a\prime) \right) \tag{15}$$

where $A(s, a)$ 는 denotes the advantage of choosing a particular action in the state. $\alpha \in (0, 1)$ represents the learning rate, and γ represents the discount factor. Because these variables may affect the Q-value update, the details of the proposed DRL-based algorithm are presented in Algorithm 1.

DLES Algorithm 1

Initialize: state space s, action space a, Q-values Q(s, a)
  for each episode do
    Initialize current state s
   while s is not terminal do
      Choose action a from s using policy derived from Q
      Take action a, observe reward r, next state s'
      Update Q-value using Equation (15)
      s = s'
    end while
end for

5 Performance Evaluation

To demonstrate the superiority of the proposed delay-sensitive and energy-efficient task offloading method using the proposed DRL, we compared the performance of the DLES scheme using two basic strategies: executing tasks locally and offloading tasks to the MEC. The experimental settings are as follows: The channel's bandwidth is set at 4 MHz, and σ^2 is -30 dBm. Task sizes range from 0.1 MB to 4 MB. The clock frequencies for the local MASS, MEC, and cloud servers were 0.5–1 GHz, 10 GHz, and 100 GHz, respectively. The learning rate was set to 0.6, the discount factor was set to 0.85, and the values of parameters $x_i$and $y_i$were adjusted based on various task types.

Figure 3 illustrates the convergence analysis of the dueling DQN. In the early stages of the learning process, the episodes exhibited high variability as they primarily focused on exploring the environment. As the episodes progressed, the variability decreased, and the model approached the optimal policy. In the later stages, the stabilization of the average reward confirms the convergence of the algorithm.

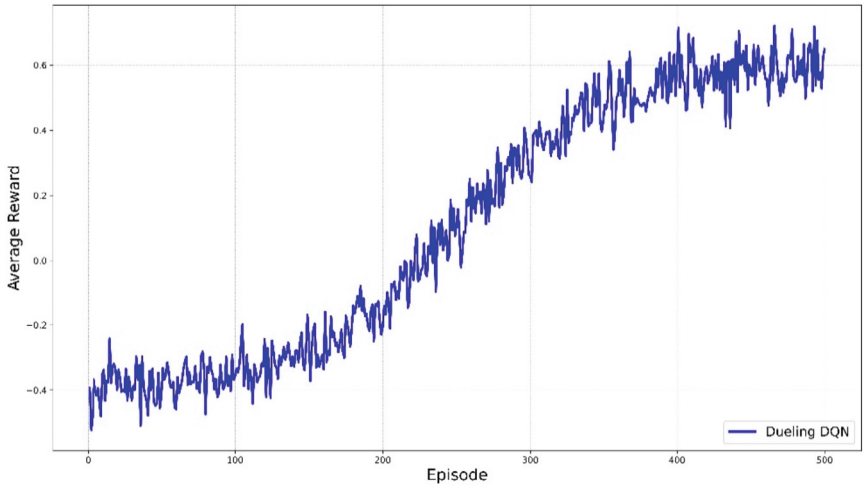


Fig. 3. Convergence analysis of dueling DQN.

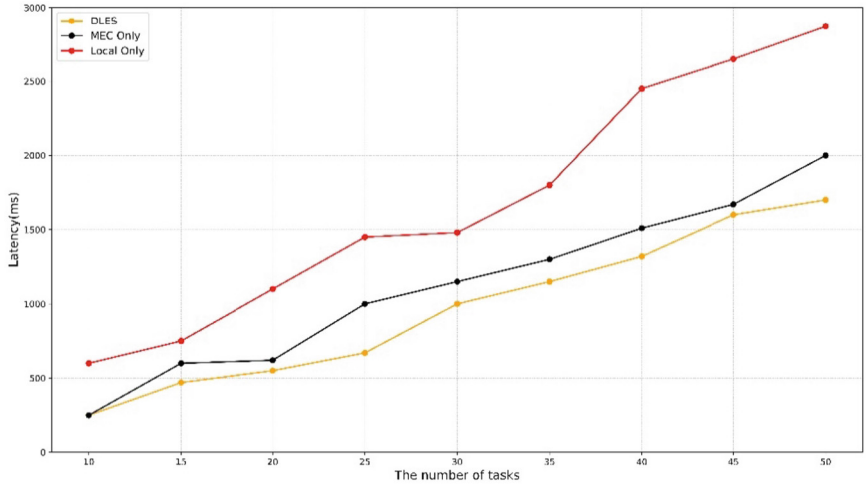


Fig. 4. Latency with the number of tasks.

Figures 4 and 5 illustrate the trends in delay time and energy consumption, respectively, in relation to the number of tasks. As expected, both the delay time and energy consumption increased with an increasing number of tasks. Notably, the proposed DLES scheme demonstrated superior performance by effectively minimizing both delay and energy consumption.

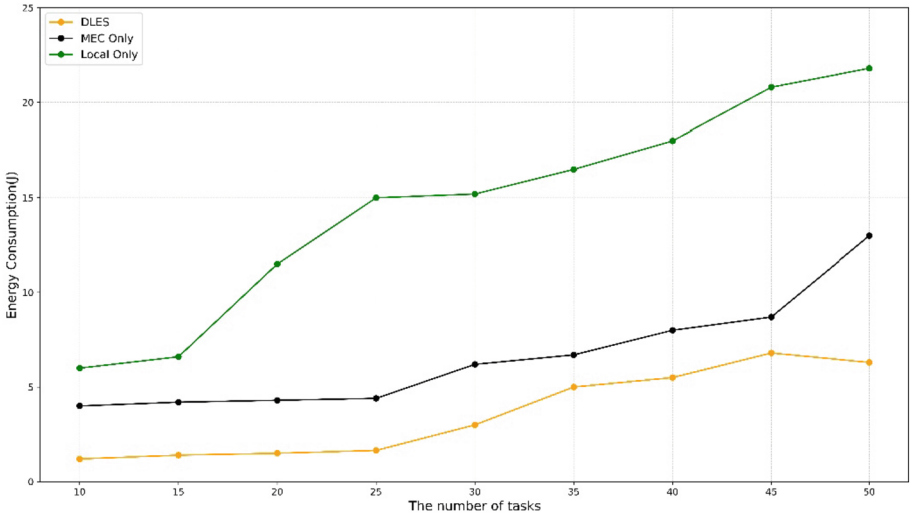


Fig. 5 Energy consumption with the number of tasks.

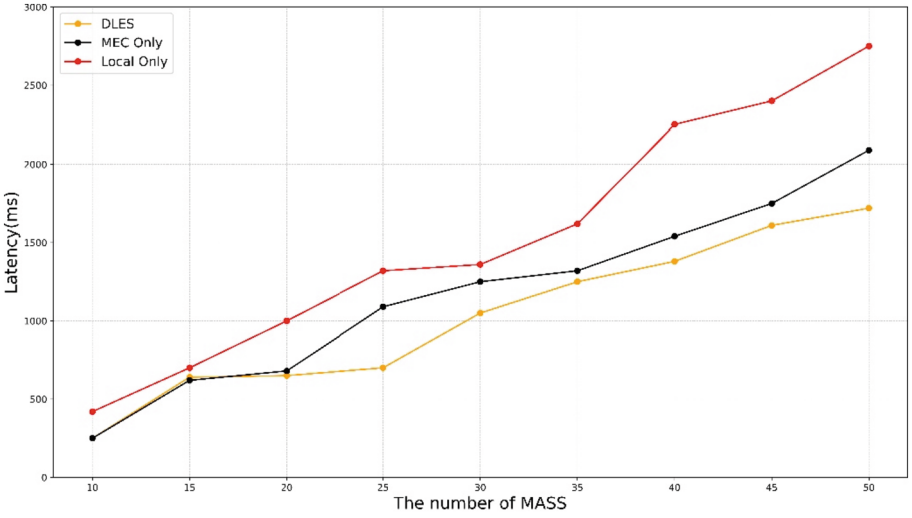


Fig. 6 Latency with the number of MASS.

Figures 6 and 7 illustrate the trends in time delay and energy consumption, respectively, based on varying numbers of MASS. While the performance varies with changes in the number of MASS, the proposed DRL-based algorithm consistently minimizes both the delay and energy consumption, outperforming other methods.

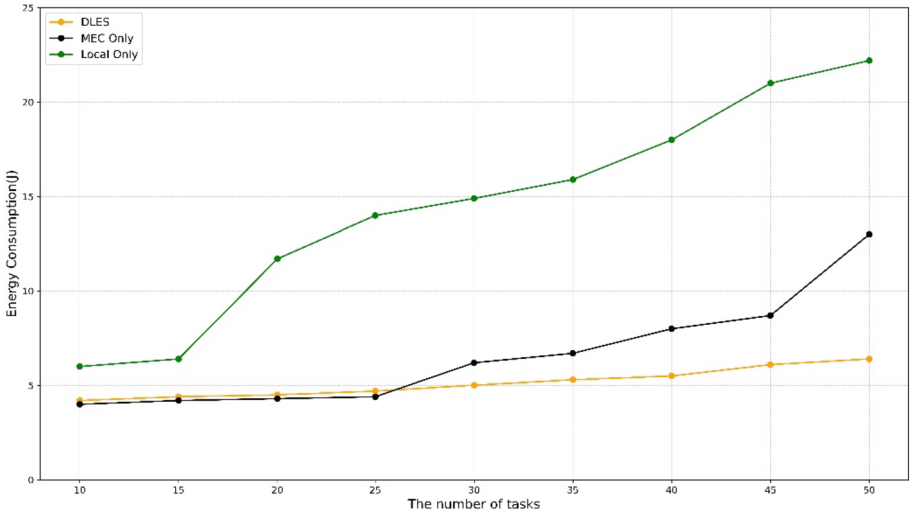


Fig. 7 Energy consumption with MASS.

6 Conclusion

In this paper, we introduce a DRL-based task offloading scheme designed to enhance execution efficiency by minimizing time delays and energy consumption in maritime MEC networks. We employed the DRL algorithm to make intelligent offloading decisions and select an optimal processing strategy. Through a series of simulations, the delay and energy efficiency of the proposed method were evaluated and compared with those of the two prevalent offloading methods. The results confirm that the proposed DLES scheme excels in the unique communication and computing conditions of the maritime environment.

References

1. Liu, R.W., Nie, J., Garg, S., Xiong, Z., Zhang, Y., Hossain, M.S.: Data-driven trajectory quality improvement for promoting intelligent vessel traffic services in 6G-enabled maritime IoT systems. IEEE Internet Things J. **8**(7), 5374–5385 (2020)
2. Yang, J., Wen, J., Wang, Y., Jiang, B., Wang, H., Song, H.: Fog-based marine environmental information monitoring toward ocean of things. IEEE Internet Things J. **7**(5), 4238–4247 (2020)
3. Hu, C., Pu, Y., Yang, F., Zhao, R., Alrawais, A., Xiang, T.: Secure and efficient data collection and storage of IoT in smart ocean. IEEE Internet Things J. **7**(10), 9980–9994 (2020)
4. Park, S.-H., Yoo, J., Son, D., Kim, J., Jung, H.-S.: Improved calibration of wind estimates from advanced scatterometer MetOp-B in Korean seas using deep neural network. Korean J. Remote Sens. **13**(20), 4164 (2021)
5. Saad, W., Bennis, M., Chen, M.: A vision of 6G wireless systems: applications, trends, technologies, and open research problems. IEEE Netw. **34**(3), 134–142 (2019)
6. Sun, W., Zhang, H., Wang, R., Zhang, Y.: Reducing offloading latency for digital twin edge networks in 6G. IEEE Trans. Veh. Technol. **69**(10), 12240–12251 (2020)

7. Jamil, S.U., Arif Khan, M., Rehman, S.U.: Intelligent task offLoading and resource allocation for 6G smart city environment. In: 2020 IEEE 45th Conference on Local Computer Networks (LCN), Sydney, NSW, Australia, pp. 441–444 (2020)
8. Cao, J., Feng, W., Ge, N., Lu, J.: Delay characterization of mobile-edge computing for 6G time-sensitive services. IEEE Internet Things J. **8**(5), 3758–3773 (2020)
9. Peltonen, E., et al.: 6G white paper on edge intelligence. arXiv preprint arXiv:2004.14850 (2020)
10. Zheng, J., Cai, Y., Wu, Y., Shen, X.: Dynamic computation offloading for mobile cloud computing: A stochastic game-theoretic approach. IEEE Trans. Mobile Comput. **18**(4), 771–786 (2019)
11. Wang, F., Xu, J., Cui, S.: Optimal energy allocation and task offloading policy for wireless powered mobile edge computing systems. IEEE Trans. Wireless Commun. **19**(4), 2443–2459 (2020)
12. Cheng, N., et al.: Space/aerial-assisted computing offloading for IoT applications: a learning-based approach. IEEE J. Sel. Areas Commun. **37**(5), 1117–1129 (2019)
13. Wei, X., et al.: MVR: An architecture for computation offloading in mobile edge computing. In: 2017 IEEE international conference on edge computing (EDGE), pp. 232–235 (2017)
14. He, S., et al.: Cloud-edge coordinated processing: Low-latency multicasting transmission. IEEE J. Sel. Areas Commun. **37**(5), 1144–1158 (2019)
15. Liu, Y., Xiong, K., Ni, Q., Fan, P., Letaief, K.B.: UAV-Assisted wireless powered cooperative mobile edge computing: joint offloading, cpu control, and trajectory optimization. IEEE Internet of Things J. **7**(4), 2777–2790 (2020)
16. Kim, J., Kim, B.K.: Development of precise encoder edge-based state estimation for motors. IEEE Trans. Ind. Electron. **63**(6), 3648–3655 (2016)
17. Trinh, H., et al.: Energy-aware mobile edge computing and routing for low-latency visual data processing. IEEE Trans. Multimedia **20**(10), 2562–2577 (2018)
18. Rodrigues, T.K., Suto, K., Kato, N.: Edge cloud server deployment with transmission power control through machine learning for 6G Internet of Things. IEEE Trans. Emerg. Top. Comput.Comput. **9**(4), 2099–2108 (2019)
19. Yang, B., Chai, W.K., Xu, Z., Katsaros, K.V., Pavlou, G.: Costefficient NFV-enabled mobile edge-cloud for low latency mobile applications. IEEE Trans. Netw. Service Manag. **15**(1), 475–488 (2018)
20. Yang, T., et al.: Two-stage offloading optimization for energy–latency tradeoff with mobile edge computing in maritime internet of things. IEEE Internet of Things J. **7**(7), 5954–5963 (2020)
21. Lim, D., Lee, W., Kim, W.T., Joe, I.: DRL-OS: a deep reinforcement learning-based offloading scheduler in mobile edge computing. Sensors **22**(23), 9212 (2022)
22. Ji, J., Zhu, K., Yi, C., Niyato, D.: Energy consumption minimization in UAV-assisted mobile-edge computing systems: Joint resource allocation and trajectory design. IEEE Internet Things J. **8**(10), 8570–8584 (2021)
23. Lim, D., Joe, I.: A delay and energy-aware task offloading and resource optimization in mobile edge computing. In: Computer Science On-line Conference, pp. 259–268 (2023)
24. Wang, Z., Schaul, T., Hessel, M., Hasselt, H., Lanctot, M., Freitas, N.: Dueling network architectures for deep reinforcement learning. In: International Conference on Machine Learning, pp. 1995–2003 (2016)

Neural Network Development for Quality Analysis of ERP Systems

A. D. Selyutin[1,2], V. A. Kushnikov[1,2,4], A. S. Bogomolov[1,2], A. F. Rezchikov[3], V. A. Ivashchenko[1], E. V. Berdnova[5], T. V. Pakhomova[5], O. I. Dranko[3], I. A. Stepanovskaya[3], A. A. Kositzyn[1], and A. A. Dnekeshev[6(✉)]

[1] Federal State Budgetary Institution Federal Research Centre, «Saratov Scientific Centre of the Russian Academy of Sciences», 24 Rabochaya Street, Saratov 410028, Russia
aseliutin@ya99.ru

[2] Saratov State University, 83 Astrakhanskaya Street, Saratov 410012, Russia

[3] V. A. Trapeznikov Institute of Control Sciences of RAS, 65 Profsoyuznaya Street, Moscow 117997, Russia

[4] Yuri Gagarin State Technical University of Saratov, 77 Politechnicheskaya Street, Saratov 410054, Russia

[5] Saratov State Vavilov Agrarian University, 1 Teatral'naya Pl. Street, Saratov 410012, Russia

[6] Zhangir Khan West Kazakhstan Agrarian Technical University, 51 Zhangir Khan Street, Uralsk 090009, Kazakhstan
dnekeshev1991@gmail.com

Abstract. In the modern world, information systems it's important tool in corporate structures and business processes. ERP systems are widely used for automating business processes and improving communication between employees in companies. However, like any other system, ERP systems can vary in quality, which in turn affects the efficiency and feasibility of business processes. This article describes a development an information system that uses a neural network to analyze the quality of ERP systems. The input parameters of the neural network are a subset of 12 quality parameters defined in ISO 25010. The developed neural network is based on a multilayer perceptron architecture. The newly created reduction methodology using the mean of L1 and L2 regularization allowed reducing the connections in the neural network and increasing the model accuracy. As a result, the neural network has a high accuracy, sufficient for practical use. A web application has been developed for the neural network, which is a graphical user interface that allows interacting with the neural network and evaluating the quality of ERP systems. Thus, the created information system is capable of significantly simplifying and speeding up the process of analyzing the quality of ERP systems.

Keywords: ERP Systems · Multilayer Perceptron · ISO25010 · Neural Network Reduction · Flask

1 Introduction

The problem of analyzing ERP (Enterprise Resource Planning) systems quality is relevant in case developing new ERP and in case changing functions of existing ERP.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 245–253, 2024.
https://doi.org/10.1007/978-3-031-53549-9_24

The research [1] compared the effectiveness of several machine learning methods, including neural networks, to predict the level of operability of the ERP system being developed. The results of article showed low accuracy of predictions for neural network.

The article [2] found that neural network models do not provide a sufficient level of accuracy in predicting the quality of an ERP system. The authors highlight the problem of insufficient data for neural network training.

In the article [3], the authors used neural networks to predict the success of ERP projects on the example of state-owned enterprises in Taiwan. As a result of the research, it was found that the accuracy of prediction using neural networks was 62%, which is a low accuracy.

Thus, most studies highlight the problems of using neural networks in analyzing the quality of ERP systems associated with low prediction accuracy. Creating an effective neural network requires a large amount of data, the right choice of neural network architecture and configuration parameters for training. The purpose of this article is development a neural network for analyzing the quality of ERP systems, which allows to increase the accuracy of the forecast.

It is necessary to develop a model for analyzing the quality of ERP systems, which will take into account the quality parameters of the software according to ISO 25010 [4, 5]. This will allow for a more complete analysis of the quality of ERP systems. The key steps in solving this issue are to determine the quality characteristics for ERP systems and determine data collection methods.

2 Methods and Materials

The research of the quality of ERP systems requires a systematic analysis of large structured data. One of the solutions to achieve this goal it's neural networks [6–10]. Neural network is well suited for processing structured data, as they are able to find patterns and reduce the dimension of input variables.

One of the most common neural network architectures is a multilayer perceptron. This means that with the help of a multilayer perceptron, any continuous function can be approximated at a given interval. This property makes the multilayer perceptron applicable for solving various tasks in the field of data analysis, including quality analysis of ERP systems.

We use the Python programming language [11] and the Keras library [12] to train neural network models. The graphical user interface for the neural network is developed via Flask [13].

3 Dataset Creation

To develop a neural network for analyzing the quality of ERP systems, data collection work was carried out. As part of the research, a questionnaire (web form) was created and sent to 100 leading experts and developers in the field of ERP systems. The questionnaire offered to evaluate 12 quality parameters corresponding to the ISO 25010. These parameters are: x_1 - functionality, x_2 - suitability, x_3 - correctness, x_4 - consistency, x_5 - security, x_6 - reliability, x_7 - stability, x_8 - recoverability, x_9 - practicality, x_{10} - clarity,

x_{11} - simplicity, x_{12} - efficiency. Respondents had to rate each parameter on a scale from 1 (minimal) to 5 (very high).

As a result, 3013 records of the quality parameters for various ERP systems were collected. For further research, the data was divided into training, test and validation dataset. The training dataset contained 70% of records (2110), the test dataset - 20% (603), the validation dataset - 10% (300).

The data distribution between datasets is shown on Fig. 1.

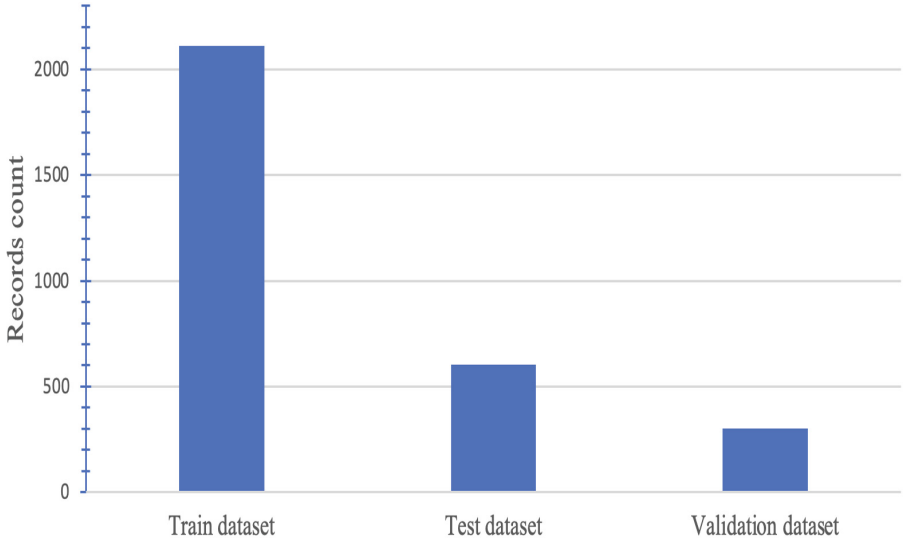


Fig. 1. Data distribution between datasets.

The collected data used for training a neural network model.

4 Neural Network Training

Before the neural network training is start, data normalization was carried out. This was necessary so that the neural network could correctly interpret the values of the input parameters. Normalization is reduced to convert the values of the input data into a range from 0 to 1.

To train the neural network, we used a laptop with 16 GB of RAM and GPU NVIDIA GeForce GTX 1050. This configuration made it possible to quickly train a neural network on a large amount of data.

To train the neural network, a multilayer perceptron with 12 input parameters, 6 hidden layer neurons and one neuron on the output layer was used.

In the training process of neural network, we used a minibatches method. This means that the training dataset is divided into tuples of 20 records, and the weights common to the entire model are adjusted at the end of the iteration. There were 400 such iterations for the first stage of training.

After the first stage of neural network training, the accuracy of the model on the validation dataset was 82%, which is a pretty good result. However, for a more accurate assessment of the quality of the system, it is possible to reduce unnecessary dependencies from the neural network.

The structure of the multilayer perceptron used in our neural network is shown in Figure 2.

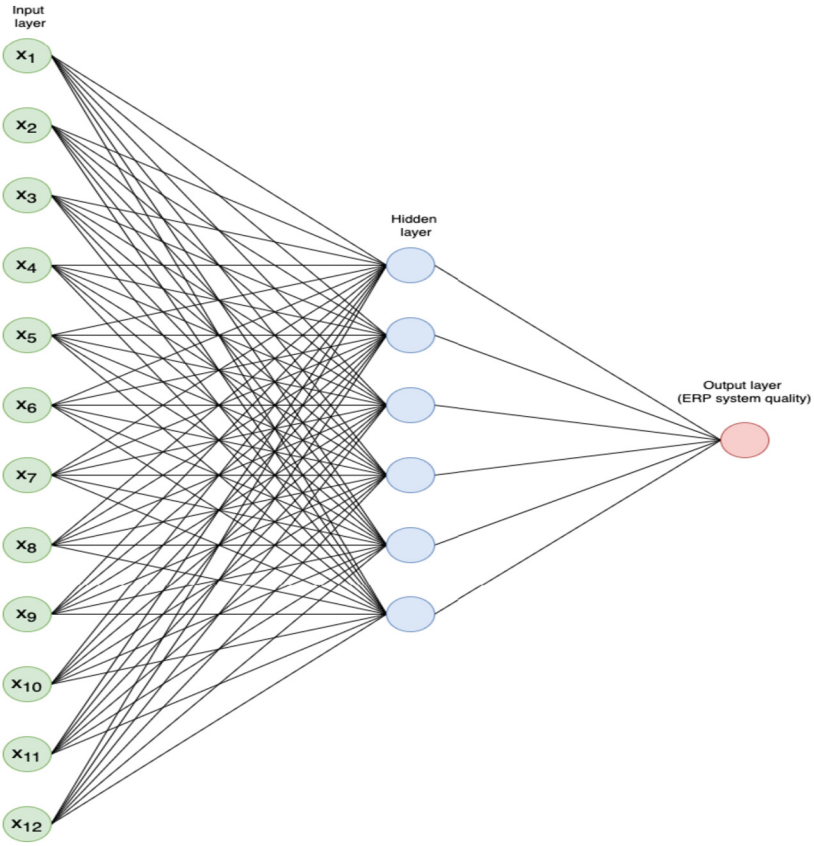


Fig. 2. The structure of a multilayer perceptron.

4.1 Neural Network Reduction

Reduction of a multilayer perceptron is a process of decrease the dimension of the output data [14, 15]. Multilayer perceptron reduction may be necessary to improve the generalizing ability of the model and lessening the risk of retraining.

Multilayer perceptron reduction algorithms are used to reduce the dimension of the model layers, their parameters, or a combination of both. The following reduction methods are most often used: spatial reduction, spatial-temporal reduction, spectral reduction, as well as methods based on data compression.

One of the new reduction methods described in this article it's average values of L1 and L2 regularizations. It allows us to find optimal values for the regularization coefficients, which leads to an improvement in the quality of the model and its ability to generalize.

L1 regularization method uses the sum of absolute component values of the model weights matrix $\|x\|_1 = \sum_i^n |x_i|$, and L2 regularization uses the sum of components squares of the model matrices $\|x\|_2 = \sqrt{\sum_i^n (x_i)^2}$. The difference between L1 and L2 regularization is shown in Fig. 3.

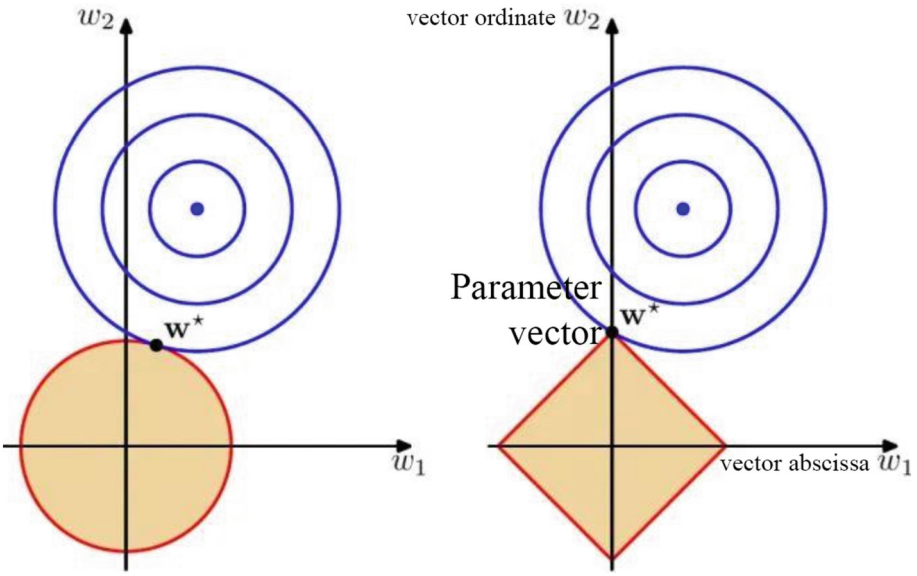


Fig. 3. Difference between L1 and L2 regularization, respectively.

The combination of L1 and L2 regularization methods makes it possible to obtain a solution that takes into account both the scale of the features and their contribution to the overall accuracy of the neural network.

4.2 Algorithm of Average L1 and L2 Regularization

The algorithm based on averages using L1 and L2 regularization described below:

1. Start;
2. Choose the values of regularization constants for L1 and L2, for example, 0.001 and 0.01, respectively;
3. Find the optimal values of the model weights, taking into account the regularization requirements of L1 and L2;
4. Recalculate the average values for each layer of the model using the new values of the weights;

5. Repeat steps 3–4 until the stop criterion is reached;
6. End.

Reduction makes it possible to preserve neural network generalization ability and increase prediction accuracy. Based on the developed method, a multilayer perceptron designed for quality analysis of ERP systems was reduced. The accuracy changes of the neural network during the reduction process is shown on the Fig. 4.

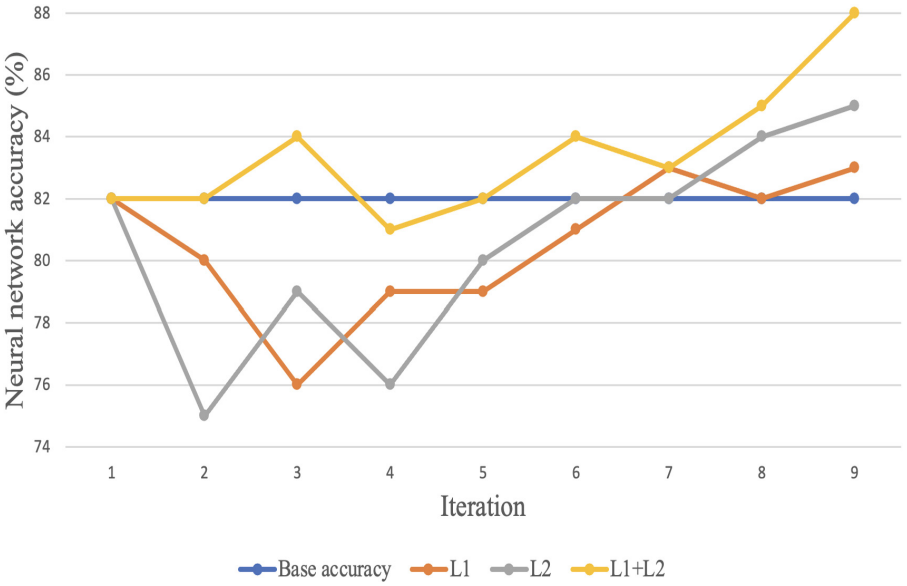


Fig. 4. Neural network accuracy changing in the reduction process.

The figure shows the base accuracy (blue) of the neural network at 82%, which does not change over all iterations. The L1 regularization method (orange) showed a maximum accuracy of 84% on the 9th iteration. The L2 regularization method (gray) showed a maximum accuracy of 85% also on the 9th iteration. The new method of combining L1 and L2 regularization showed an accuracy of 88%. This accuracy will be enough to analyze the quality of ERP systems. However, the neural network model cannot take into account all the rules of the subject area due to the subjectivity of expert. In future studies, it is planned to change the configuration of the neural network and increase its accuracy.

5 Modelling

For the developed neural network model, a web interface was created that allows us to input data of the quality parameters for the current ERP system and get a result with a percentage level of quality. The developed web application was published on the Heroku [16].

To show the possibilities of the developed mathematical model we evaluated the Odoo system quality. We conducted a comparative analysis of the work of the created neural network model and an expert assessment. This ERP system was chosen for comparison because it is one of the most popular ERP systems on the market at the moment.

We evaluated the quality of Odoo based on the described input parameters $x_1 \ldots x_{12}$ via our web application. The input parameters coincided with the expert's assessment. However, the final value of the quality of the ERP system was different. The neural network gave an average quality score of 0.89, and the expert 0.63 (the average of 5 surveys). Figure 5 shows the difference between Odoo quality assessment by a neural network and an expert.

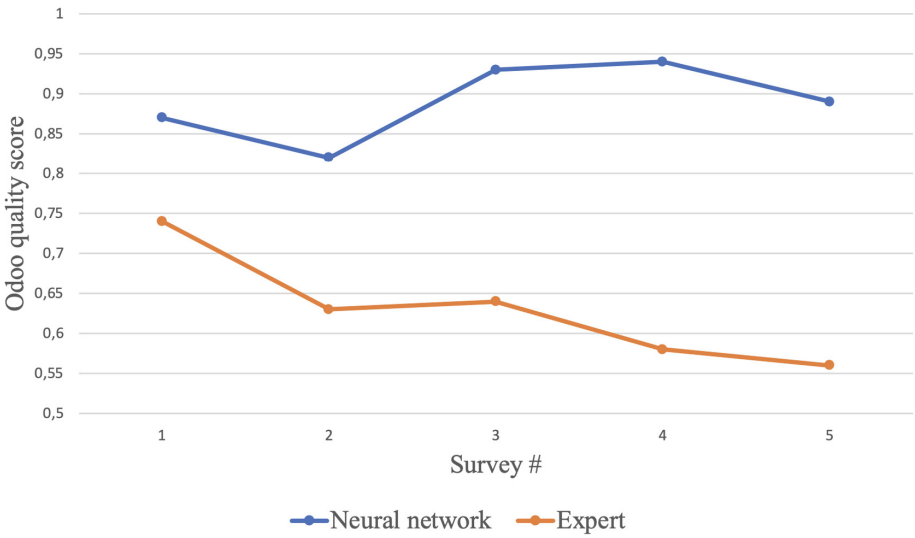


Fig. 5. Odoo quality score difference by neural network and expert.

The expert justified his conclusion based on the functionality and effectiveness of Odoo. However, the expert could not justify the assessment of consistency and recoverability. This may be due to the limited expertise and experience of an expert in this subject area.

The developed mathematical software can be used in conjunction with already known methods for assessing the quality of ERP systems [17–19].

Based on this, we can conclude that the use of a neural network model can be used to analyze the quality of ERP systems. However, it is not recommended to replace expert evaluation completely with neural network models, due to the inability of the model to cover all the rules of the subject area, but instead, it is proposed to use 2 methods together.

6 Conclusion

A methodology for quality assessment of ERP systems has been developed. To check the quality, a neural network model of a multilayer perceptron was developed to analyze the quality of ERP systems. The model has 12 input parameters based on the ISO 25010, 6 neurons on the hidden layer and 1 on the output layer. It differs from existing models in that neural network was reduced using a new algorithm based on average of L1 and L2 regularization, which allows the model to more accurately assess the quality of ERP systems.

The results obtained show that the model has a high accuracy (88%). However, when comparing the predictions of the model with expert estimates, differences were found, which may indicate an incomplete assimilation of the rules by the subject area neural network. In further articles, the optimization of the multilayer perceptron will be carried out.

This research can be used to improve the functioning of both existing and prospective ERP systems.

References

1. Kohli, M.: Supplier evaluation model on SAP ERP application using machine learning algorithms. Inter. J. Eng. Technol. **7** (2018). https://doi.org/10.14419/ijet.v7i2.28.12951
2. Rouhani, S., Zareravasan, A.: ERP success prediction: an artificial neural network approach. Scientia Iranica **20** (2013). https://doi.org/10.1016/j.scient.2012.12.006
3. Zareravasan, A., Rouhani, S.: An Expert System for Predicting ERP Post-Implementation Benefits Using Artificial Neural Network (2017). https://doi.org/10.4018/978-1-5225-3909-4.ch036
4. Estdale, J., Georgiadou, E.: Applying the ISO, IEC 25010 Quality Models to Software Product: 25th European Conference, EuroSPI,: Bilbao, Spain, 5–7 September 2018. Proceedings (2018). https://doi.org/10.1007/978-3-319-97925-0_42
5. Hussain, A., Mkpojiogu, E.: An application of the ISO/IEC 25010 standard in the quality-in-use assessment of an online health awareness system. Jurnal Teknologi **77** (2015). https://doi.org/10.11113/jt.v77.6107
6. Reddy, P., Pham, V., Deepa, N., Dev, et al.: Industry 5.0: a survey on enabling technologies and potential applications. J. Indus. Inform. Integr. **26** (2021). https://doi.org/10.1016/j.jii.2021.100257
7. Pande, M.: Industrial Safety Equipment Market Size, Trends, Growth, Forecast and Global Industry Analysis (2023)
8. Furqon, R., Lukitasari, L.: Valuasi saham perusahaan manufaktur menggunakan guideline publicly traded company method, vol. 3, pp. 329–342 (2023). https://doi.org/10.32493/j.perkusi.v3i2.29584
9. Auad, A., Hilal, M., Khalaf, N.: Best approximation of unbounded functions by modulus of smoothness. Eur. J. Pure Appl. Math. **16**. 944–952 (2023). https://doi.org/10.29020/nybg.ejpam.v16i2.4730
10. Pasias, S.: Boundary behavior of analytic functions and Approximation Theory (2023)
11. Oliphant, T.: Python for scientific computing. Comput. Sci. Eng. **9**, 10–20 (2007). https://doi.org/10.1109/MCSE.2007.58
12. John, J., Ferdin, J., Nonsiri, S., Monsakul, A.: Keras and TensorFlow: A Hands-On Experience (2021). https://doi.org/10.1007/978-3-030-66519-7_4

13. Kirby, G., Connor, R., Cutts, Q., Morrison, R., Munro, D., Scheuerl, S.: Using the Flask Architecture to Build Distributed Applications (1996)
14. Korobevnikov, I., Selyutin, A., Kushnikova, E.: Models, algorithms and software complexes for operational monitoring of the software quality of the compressor station of an industrial enterprise. In: 15th International Conference Management of Large-Scale System Development (MLSD), Moscow, Russian Federation 2022, 1–5 (2022). https://doi.org/10.1109/MLSD55143.2022.9934446
15. Kushnikova, E., Selyutin, A., Stepanovskaya, I.: Improving the quality of control systems according to the criterion of minimum loss from the effects of atmospheric pollutants. In: 15th International Conference Management of Large-Scale System Development (MLSD), Moscow, Russian Federation 2022, pp. 1–4 (2022). https://doi.org/10.1109/MLSD55143.2022.9934572
16. Tsvirkun, A., Kushnikov, V., Bogomolov, A., Selyutin, A., Kushnikova E.: Management of compressed air production at an industrial enterprise according to technical and economic criteria. IFAC-PapersOnLine, vol. 55(9), pp. 181–186 (2022) ISSN 2405–8963
17. Selyutin, A. et al.: Models and algorithms for analysis the software quality of the system of automatic segmentation and pathology analysis of the Lumbar spine MRI images. In: Silhavy, R. (eds.) Software Engineering Perspectives in Systems. CSOC 2022. LNNS, vol. 501. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09070-7_37
18. Stapper, C.: An Analysis of Heroku and AWS for Growing Startups (2017)
19. Felderer, M., Tanriverdi, E., Low, S., Breu, R.: A Quality Analysis Procedure for RequestData of ERP Systems (2013). https://doi.org/10.1007/978-3-642-37021-2_18

Designing an Algorithm for Recognizing the Kazakh-Latin Alphabet in an Image

Zhumazhan Kulmagambetova[1](✉), Damir Murzagulov[1], Ulmeken Smailova[2], Gulmira Shangytbayeva[3], and Bazargul Kulzhagarova[4]

[1] Zhubanov Regional University, Aktobe, Kazakhstan
kulma_zh@mail.ru
[2] Center of Excellence, AEO "Nazarbayeb Intellectual ISchools" , Astana, Kazakhstan
[3] Zhubanov Regional University, Aktobe, Kazakhstan
[4] Yesenov University, Aktau, Kazakhstan

Abstract. This paper presents and investigates a new algorithm for recognition of the Kazakh-Latin alphabet in an image using convolutional neural network. Alphabet recognition in images is an important task in the field of computer vision and image processing, especially for languages with special characters and graphic features, such as the Kazakh-Latin alphabet. The aim of the research was to develop an efficient algorithm capable of automatically recognizing the characters of the Kazakh-Latin alphabet in images and representing them in text format. The paper proposed a new approach based on the use of deep convolutional neural networks. #COMESYSO1120.

Keywords: Algorithm · Text · Symbol · Recognition · Neural Network

1 Introduction

The human brain contains billions of neurons - highly specialized cells designed to receive, process, store, transmit and release information from outside by means of electrical and chemical signals [1].

An artificial neural network has been constructed based on the neurons of the human brain. Artificial neural network is a mathematical model based on the principle of organization and operation of biological neural networks of a living organism, including its software implementation. The basic notion of artificial neural networks is an artificial neuron, which itself is a connector of all incoming signals.

The main architectures of neural networks are direct propagation networks and perseptrons. These neural networks are very simple. Information about them is provided directly from input to output. The cells in this layer of networks are unconnected, unlike neighboring layers, which are usually fully connected. Directly coupled neural networks are usually trained on errors with back propagation. In this type of learning, the neural network receives a lot of data both as input and output. This process is called learning with a teacher [2].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 254–262, 2024.
https://doi.org/10.1007/978-3-031-53549-9_25

There are also convolutional neural networks. This type of neural network is very different from the others. Convergent neural networks are usually used for image classification. The main feature of this neural network is a "scanner" that reads the image in parts, and a "convolution" process that reduces the size of the image feature matrix.

More advanced configurations of neural networks are neural networks that include repetition. The peculiarity of these neural networks is that they receive information not only from the previous layer, but also from the previous path. Because of this feature the order of data input becomes important for neural network training. The difficulty in using this type of neural network is the vanishing gradient problem, which is determined by the rapid loss of information over time. This problem only affects the weights during model training. Neural networks of this type are often used for automatic information augmentation [3].

2 Implementation of the Convolutional Neural Network

The main stages for the implementation of the program were identified as the following:

- Training a neural network to recognize the letters of the Kazakh Latin alphabet.
- Identify the letters on the image.
- Divide the image into parts with letters.
- Recognize the letters in the individual images.
- Make a word out of the obtained letters.
- Arrange the spaces between the different words.

To train the model a convolutional neural network was chosen which was trained to distinguish between handwritten English letters and numbers. The choice of the convolutional neural network is due to the fact that at the moment this kind of neural networks has one of the best algorithms for image recognition and classification. Compared to the fully connected neural network, it has a much smaller number of adjustable weights. The main feature of convolutional neural network is "convolution". The process of "convolution" is a reduction in the size of the feature matrix of the input image. To obtain a matrix cell of reduced size, the elements of the original matrix in a certain area are multiplied by a weight, followed by the summation of all the elements in this area. To obtain the next reduced matrix cell, the area is shifted and the same actions are performed.

The sequence of operations described above can be formulated as follows:

$$(I * K)_{xy} = \sum_{i=1}^{h} \sum_{i=1}^{w} K_{ij} * I_{x+i-1,y+k-1},$$

where I is the original feature matrix, K - weight matrix, x and y - indices of the chosen block, h and w - height and width [4].

Figure 1 shows the process of convolution of the feature matrix of dimension.

$7 * 7 = 49$ features, which is denoted by the letter I. The "grid" of the input matrix I of size 3 by 3 "cells" is multiplied by the elements of the weight matrix K, after which the matrix values obtained are summed up and the obtained value is entered into the

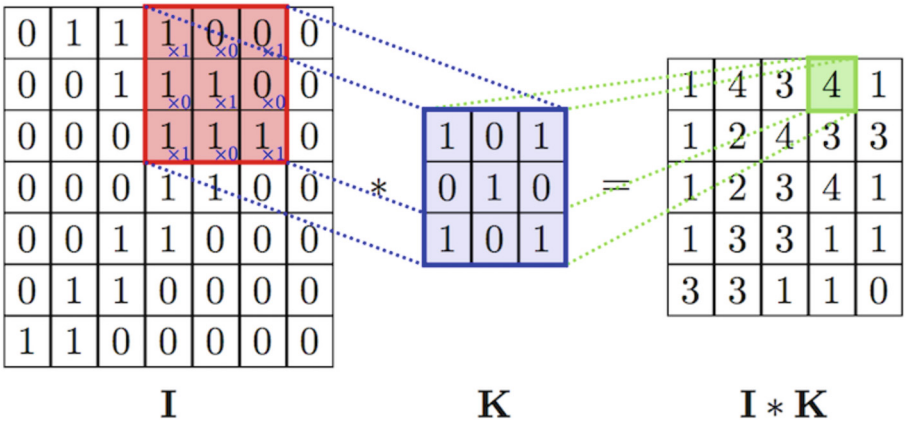


Fig. 1. The process of "convolution".

cell of the output matrix. Then the "grid" is shifted and the above described actions are repeated.

Depending on the type of project, using different datasets to train neural networks can be very useful. There are several popular classifications of datasets that can be used in various machine learning tasks.

One such dataset is MNIST. It is a collection of images of handwritten digits from 0 to 9. MNIST is widely used to train neural networks in number recognition tasks.

Two other popular datasets are CIFAR-10 and CIFAR-100. They contain images divided into 10 and 100 classes, respectively. These datasets are widely used in the field of image classification.

One of the most extensive datasets is ImageNet. It includes more than 1.4 million images belonging to more than 1,000 classes. ImageNet is one of the most popular datasets for training deep neural networks in image classification.

When it comes to the task of object detection, the COCO dataset is widely used. It contains more than 330,000 images illustrating different objects in different contexts.

For machine translation tasks, the WMT dataset, which is a collection of parallel texts in several languages, is often used [5].

However, we sometimes need to create our own dataset. In the course of our work, we encountered a situation where there was no suitable ready-made database for the Kazakh-Latin alphabet. As a result, we developed our own dataset consisting of 99736 images generated using 410 different fonts. This dataset can be very useful in training of neural networks in tasks related to Kazakh-Latin alphabet.

Stages of neural network model creation and training are presented below. Before the construction of the neural network model it is necessary to perform the preprocessing of the data set (Fig. 2).

First of all, we bring the images from the finished database to a certain size. In our case, the images have been resized to 50 × 50 pixels. It is important to note that the input images can be in color. Therefore, before using them in the model, we convert them from the RGB color model (red, green, blue) to gradient gray (Fig. 3).

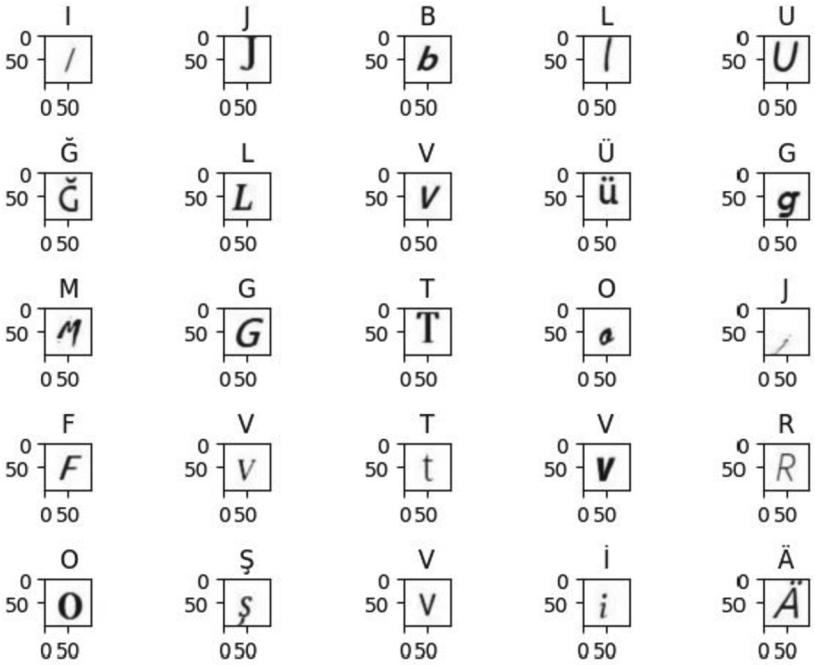


Fig. 2. Data set.

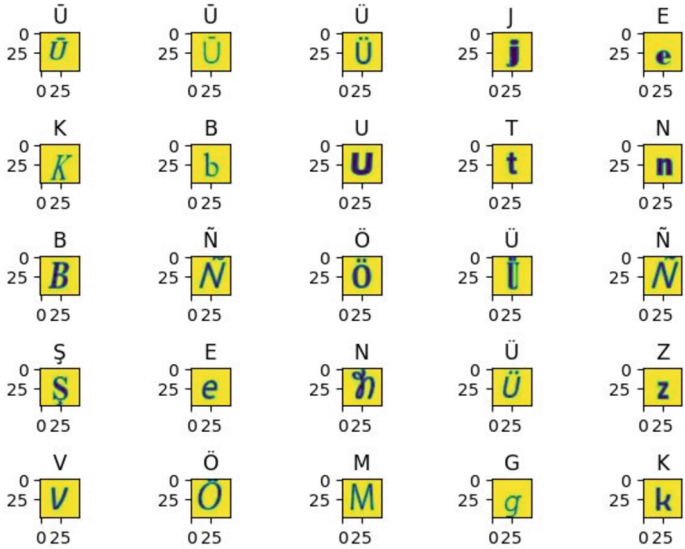


Fig. 3. Processed dataset.

Each pixel in the processed dataset has a value that ranges from 0 to 255. In order to provide more efficient training of the neural network, we perform a process of data normalization. The normalization of the neural network is to reduce the input data to a standard normalized form. This improves the performance of the neural network and speeds up its convergence to the optimal solution. Usually neural network data has a different scale and range of values. This can cause problems in training because large input values can lead to saturation of the activation function and slow down the neural network convergence. In addition, large differences in the scale of the input data can reduce the effectiveness of the gradient descent in the learning process. Normalization solves these problems by converting the input data to a standard range of values, such as 0 to 1 or by a standard normal distribution with a mean of 0 and a standard deviation of 1. This approach makes gradient descent more efficient and faster [6].

After completing the data normalization process, we move on to the important step of dividing the data set into three parts: the training set, the validation set, and the test set. This step is necessary to build a robust neural network model, allowing us to evaluate the quality of the model's performance on independent data and avoid the problem of overtraining.

The training dataset is the set on which the model will be trained. The test dataset is used to evaluate and control the quality of the model during training. Finally, the test dataset is for the final evaluation of the model's performance after the training process is complete. There are several strategies for dividing the data set into these three parts, but one of the most common approaches is random assignment. In our case, we used a ratio of 70% training set, 20% validation set, and 10% test set.

In the early stages of the study, we analyzed different types and architectures of neural networks. As a result of this analysis the most effective architecture for our Kazakh-Latin alphabet recognition algorithm was selected - convolutional architecture (Fig. 4).

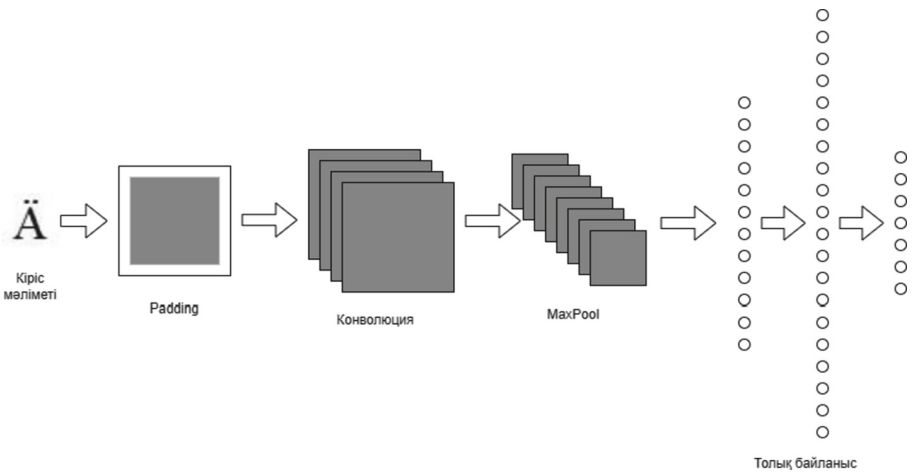


Fig. 4. Neural network architecture.

The next step in our work is to create a model of the neural network. We built a model consisting of several layers, and for this we used the Keras library.

```
model = Sequential()
```

```
model.add(ZeroPadding2D(padding=(1,1)))
model.add(Conv2D(8,(3,3),1, activation="relu", input_shape=(50,50,1)))
model.add(MaxPooling2D())

model.add(ZeroPadding2D(padding=(1,1)))
model.add(Conv2D(16,(3,3),1, activation="relu"))
model.add(MaxPooling2D())

model.add(ZeroPadding2D(padding=(1,1)))
model.add(Conv2D(32,(3,3),1, activation="relu"))
model.add(MaxPooling2D())

model.add(Flatten())

model.add(Dense(248, activation="relu"))
model.add(Dense(128, activation="relu"))
model.add(Dense(31, activation="softmax"))
```

```
model.compile("adam", loss=tf.losses.CategoricalCrossentropy(), metrics=["accuracy"])
```

Fig. 5. Program for creating the model.

Using the Sequential class in the Keras library, we can create a sequential collection of layers that go one after another (Fig. 5). In our model we add the following parameters:

Layer Convolution2D. This layer allows to create a convolution layer with the kernel and get the output tensor. We can define the size of the kernel using the kernel_size parameter, where the first parameter indicates the height and the second parameter indicates the width of the convolution window.

Input_shape layer. This function converts an input two-dimensional image of 28 x 28 pixels into a one-dimensional array of 784 pixels. Activation 'relu' represents the activation function that is used to calculate the output of the neuron. The activation function is needed to introduce nonlinearity into the model.

Layer ReLU. Rectified Linear Unit (ReLU) represents the activation function that is applied to the output of the previous layer. The formula for the ReLU function is as follows:

$$f(x) = \max(0, x).$$

The ReLU function helps in modeling nonlinear relationships and provides a more flexible model (Fig. 6).

Layer ZeroPadding2D. Convex neural networks use zeropadding to control the size of the output after the convolution operation. During convolution, the dimensionality of the data is usually reduced. However, this can lead to loss of information around the edges of the images and limit the visibility of the neural network. To avoid this problem, we use the ZeroPadding2D layer, which adds zeros around the input data before applying the convolution operation. This preserves the dimensionality of the data and prevents narrowing of the field of view.

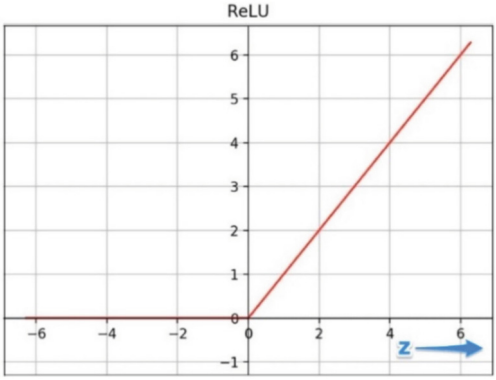


Fig. 6. "ReLu" activation graph.

The MaxPooling2D layer is used in neural networks to process two-dimensional data, such as images. It performs feature space dimensionality reduction by selecting the largest value from each subarea of the input data (usually 2×2 or 3×3 in size) and moving through the image in a given step (2 or 3). This reduces the number of parameters in the model, simplifies the calculations, and avoids overtraining by highlighting the most important features of the image.

The Flatten layer is an intermediate layer that converts multidimensional input tensor to one-dimensional tensor. Usually this layer is placed before the fully connected layers (Dense) of the neural network. For example, if the input tensor size is (size_package, height, width, channels), then after applying the Flatten layer, a one-dimensional tensor of (size_package, height * width * channels) dimension is obtained. This layer allows the output of the previous layers to be converted into a form that can be used by fully connected layers that expect a one-dimensional input tensor. In addition, the Flatten layer can reduce the dimensionality of the input data, which helps to simplify the model and reduce computational complexity.

The Dense layer, also known as a fully connected layer, is a type of layer in a neural network where each neuron in the input layer is connected to each neuron in the output layer. This means that all input layer neurons are connected to all output layer neurons. During neural network training, the Dense layer performs a linear transformation of the input data by weighting each input with an appropriate function, and then applies a non-linear activation function (e.g., ReLU, Sigmoid, Tanh) to produce the output data.

The softmax activation function represented by the following formula is widely used in classification problems of multiple classes [7]:

$$softmax(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}}.$$

It takes an input vector z and calculates probabilities for each class i, scaling the index z_i with respect to the sum of all indices. Thus, the softmax function converts output values into probabilities, where each value represents the probability of belonging to the corresponding class.

The number of classes is denoted by the variable k. This is very important, because Softmax function allows to get probabilities for each class, where sum of all probabilities equals 1. This is useful, because the neural network can predict probabilities for each class, not only to determine the most likely class. When training a neural network using the Softmax activation function, the cross-entropy loss is usually minimized, which measures the discrepancy between predicted probabilities and actual class labels.

When we compile our model, we define parameters that determine how the neural network will be trained. In particular, the compilation method takes the following important parameters:

Optimizer. This is the optimization algorithm used during training. Here we can choose from different optimizers such as Adam, SGD (stochastic gradient descent) and others. Each optimizer has its own characteristics and may be more or less suitable for the specific task of training the neural network.

Loss function. This is a function that measures the mismatch between the output values of the neural network and the target values. For example, for a binary classification task we may choose "binary_crossentropy" loss function, and for a multi-class classification task we may choose "categorical_crossentropy". Choosing the appropriate loss function is important for successful model training.

Metrics. These are the metrics used to evaluate the performance of the model during training. For example, we can choose the metric "accuracy" which measures the proportion of correctly classified samples, or other metrics such as "precision", "recall", etc. The choice of appropriate metrics depends on the specific task and model requirements.

Adam (Adaptive Moment Estimation) is an effective way to train neural networks. It combines the advantages of two other methods - AdaGrad and RMSProp. The Adam algorithm uses the first and second moment gradient to adjust the learning rate and momentum. With Adam, we compute the gradient of the loss function using weights, update the first and second moments of the gradient using an exponential moving average, compute the learning rate and time at the current optimization step, and update the weights with the computed learning rate and momentum. Adam has the advantage that it can vary the learning rate in different parts of the loss function, allowing optimal values of the loss function to be achieved. This method is considered one of the most effective methods for optimizing neural networks.

Categorical cross-entropy is a loss function widely used in multi-class classification problems. In neural networks it is used to measure the difference between predicted class probabilities and actual class labels. When the neural network model predicts the probabilities of belonging to each class, the categorical cross-entropy calculates the sum of the logarithms of the predicted negative true class probabilities. This allows to estimate the classification costs and minimize them [8].

3 Conclusion

The study analyzed existing methods of text recognition and basic machine learning methods. A neural network based on convolutional neural network architecture was implemented for recognition of the Kazakh Latin alphabet. This task required the preliminary processing and preparation of data, as well as the selection of a suitable neural

network architecture. Convergent neural networks are powerful tools for working with images and text data, so they were chosen for this task.

References

1. Omelianenko, Y.: Evolutionary neural networks in the Python language. Moscow (2020)
2. Cook, D.: Machine learning using library (2018)
3. Chardin, B., Massaron, L., Boschetti, A.: Large-scale machine learning with Python (2018)
4. What is a convolutional neural network. https://habr.com/ru/post/309508/, (Accessed 6 June 2023)
5. Thevenoux, R., et al.: Image based species identification of Globodera quarantine nematodes using computer vision and deep learning. Comput. Elect. Agricult. **186**, 106058 (2021). https://doi.org/10.1016/j.compag.2021.106058
6. Sukanya, J., Rajiv Gandhi, K., Palanisamy, V.: An assessment of machine learning algorithms for healthcare analysis based on improved MapReduce. Adv. Eng. Softw. **173**, 103285 (2022)
7. Arkhangelskaya, E., Kadurin, A., Nikolenko, S: Deep Learning, Moscow (2022)
8. Goodfellow, J., Bengio, I., Courville, A.: Deep Learning: A Practical Guide, Moscow (2020)

Influence of TMDC Layers on the Optical Properties of Silicon Nanoparticles

Denis Kislov[1] and Vjaceslavs Bobrovs[2(✉)]

[1] Center for Photonics and 2D Materials, Moscow Institute of Physics and Technology, Dolgoprudny 141700, Russia
denis.a.kislov@gmail.com
[2] Institute of Telecommunications, Riga Technical University, Riga 1048, Latvia
vjaceslavs.bobrovs@rtu.lv

Abstract. An analytical model is implemented based on the angular spectral representation of the near field of electric and magnetic dipoles located near a flat substrate, taking into account the presence of an active layer of a two-dimensional material (TMDC). This approach is applicable only for the case of planar structures, but knowledge of the angular spectra of dipole fields is important for designing nanophotonic systems based on subwavelength dielectric nanoantennas with a high refractive index.

Keywords: TMDC · Green functions · Angular spectrum · Multipole decomposition

1 Introduction

Optical properties of semiconductor metasurfaces can be controlled using active two-dimensional materials such as graphene or layers of transition metal dichalcogenides (TMDC). These materials have unique optical properties that can be used to modify the characteristics of metasurfaces.

One way to control the optical properties of a nanoantenna is to change the refractive index of the active material. The refractive index determines the speed of light propagation in the material, and its change can lead to a change in the resonance wavelength of the nanoantennas that make up the metasurface. This allows you to tune the resonant frequency of the nanoantenna and control its interaction with electromagnetic radiation [1–3].

Also, active two-dimensional materials can be used to change the quality factor of nanoantennas. The quality factor determines the efficiency of absorption and emission of light by a nanoantenna. Changing the quality factor allows you to control the efficiency of the nanoantenna and optimize its functionality [4].

Another way to control the optical properties of a silicon nanoantenna is to use active two-dimensional materials to create so-called metascatters. Metascatters are structures capable of controlling the phase and amplitude of scattered light. By controlling these parameters, it is possible to change the direction and intensity of the radiation of the nanoantenna [5–10].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 263–270, 2024.
https://doi.org/10.1007/978-3-031-53549-9_26

Thus, the interaction of a silicon nanoantenna and two-dimensional transition metal dichalcogenides can have a wide range of applications, in various practical applications such as as lab-on-a-chip [11–15], photonic computing [16, 17], metamaterials [18–20] and metasurfaces [21–24], and others [25–27]. These effects can be used in optical devices, optical communications, sensors and other areas.

Analytically, the interaction of active two-dimensional materials with a nanoantenna located on it can be described in terms of reflected and incident electromagnetic fields from individual multipoles excited in the nanoantenna.

2 Methods

Two-dimensional TMDC films were chosen as an optically active material that affects the optical response of subwavelength semiconductor nanoparticles in this work.

A spherical silicon nanoparticle lying on a glass substrate coated with a monatomic TMDC layer is considered. The particle size is chosen so that only dipole modes (p, m) are predominantly excited (see Fig. 1).

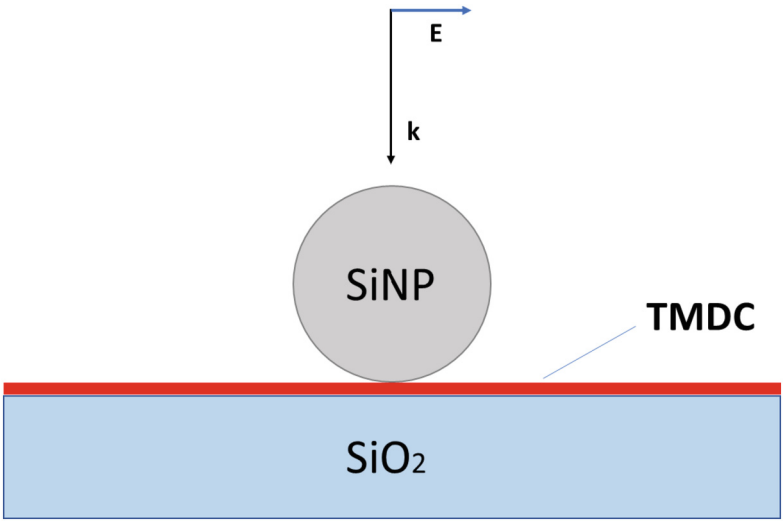


Fig. 1. The geometry of the problem. Silicon particle on a substrate with a two-dimensional active material. A plane linearly polarized electromagnetic wave is incident on the system at zero angle.

We used the model of the effective permittivity of two-dimensional TMDC layers. For this, experimentally measured spectral dependences of the complex refractive index were taken. At the same time, the model has the ability to specify a TMDC coating with a thickness of one or several layers (1, 2, or 3). For MoS2 and MoSe2 materials, it was possible to model the substrate as a bulk material.

Curves of the real and imaginary parts of the refractive index were plotted for four different TMDC materials. On these graphs, one can observe the evolution of the dielectric properties of the studied materials with an increase in the number of layers of a two-dimensional material (see Fig. 2).

In the case of TMDC, the model of effective permittivity has proven itself well, where for each material its own effective thickness is introduced

$$d_{MoS_2} = 6.15\,\text{Å} \quad d_{MoSe_2} = 6.46\,\text{Å} \quad d_{WS_2} = 6.18\,\text{Å} \quad d_{WSe_2} = 6.49\,\text{Å}$$

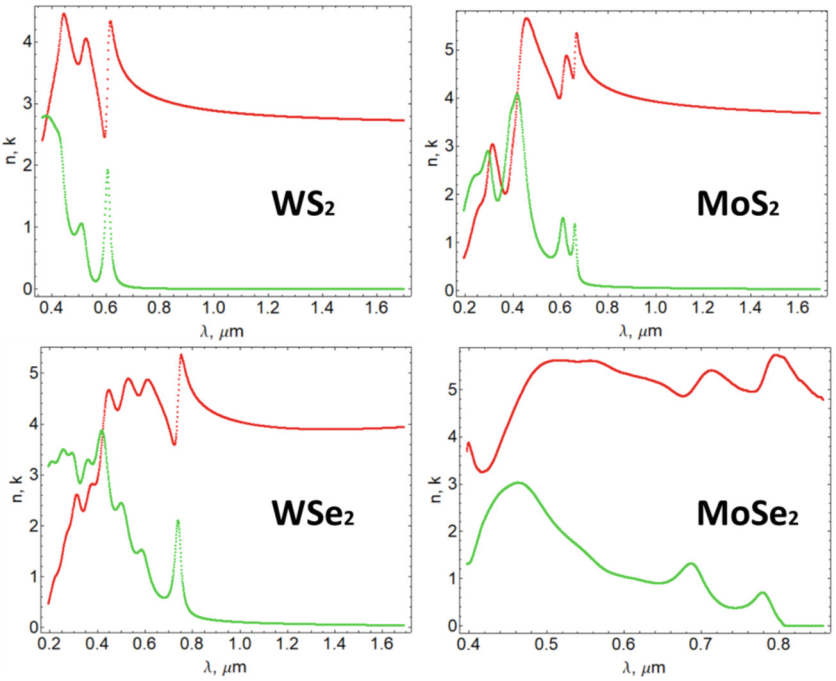


Fig. 2. Real and imaginary parts of the refractive index of various TMDC materials. Differences in the dielectric properties are shown with a change in the number of layers of a 2D material.

3 Results

In the presence of a substrate, the total field can be represented as a superposition:

$$\begin{cases} \mathbf{E}_{Tot} = \mathbf{E}_{Inc} + \mathbf{E}^{p}_{ref} + \mathbf{E}^{m}_{ref} \\ \mathbf{H}_{Tot} = \mathbf{H}_{Inc} + \mathbf{H}^{p}_{ref} + \mathbf{H}^{m}_{ref} \end{cases} \tag{1}$$

where $\mathbf{E}_{Inc}$ and $\mathbf{H}_{Inc}$ sum of incident and reflected fields. $\mathbf{E}^{p,m}_{ref}$ и $\mathbf{H}^{p,m}_{ref}$ reflected fields of dipoles.

$$\begin{cases} \mathbf{E}^{p}_{ref} = \dfrac{\mu k^2}{\varepsilon_0} \widehat{\mathbf{G}}^{E}_{ref}(r_0, r_0)\mathbf{p} \\ \mathbf{E}^{m}_{ref} = i\omega\mu\mu_0 \left[\nabla \times \widehat{\mathbf{G}}^{H}_{ref}(r_0, r_0)\right]\mathbf{m} \end{cases} \quad \begin{cases} \mathbf{H}^{p}_{ref} = -i\omega \left[\nabla \times \widehat{\mathbf{G}}^{E}_{ref}(r_0, r_0)\right]\mathbf{p} \\ \mathbf{H}^{m}_{ref} = \varepsilon\mu k^2 \widehat{\mathbf{G}}^{H}_{ref}(r_0, r_0)\mathbf{m} \end{cases} \tag{2}$$

r_0- dipole location.

Then the effective dipole moments, taking into account the effect of the substrate, can be written in terms of the polarizability tensor:

$$\begin{cases} \mathbf{p} = \varepsilon_0 \widehat{\alpha}_{EE} \mathbf{E}_{Inc} + \widehat{\alpha}_{EM} \mathbf{H}_{Inc} \\ \mathbf{m} = \widehat{\alpha}_{MM} \mathbf{H}_{Inc} + \widehat{\alpha}_{ME} \mathbf{E}_{Inc} \end{cases} \tag{3}$$

where the elements of the polarizability tensor are expressed:

$$\begin{aligned}
\widehat{\alpha}_{EE} &= \left[\hat{I} - \varepsilon_0 \alpha_E \tfrac{\mu k^2}{\varepsilon_0} \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) - \varepsilon_0 \alpha_E i\omega\mu\mu_0 \left[\nabla \times \widehat{\mathbf{G}}^H_{ref}(r_0, r_0) \right] \left(\hat{I} - \alpha_H \varepsilon\mu k^2 \widehat{G}^H_{ref}(r_0, r_0) \right)^{-1} \alpha_H \left(-i\omega \left[\nabla \times \widehat{G}^E_{ref}(r_0, r_0) \right] \right) \right]^{-1} \alpha_E \\
\widehat{\alpha}_{MM} &= \left[\hat{\mathbf{I}} - \varepsilon_0 \alpha_E \tfrac{\mu k^2}{\varepsilon_0} \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) - \alpha_H \left(-i\omega \left[\nabla \times \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) \right] \right) \left(\hat{\mathbf{I}} - \varepsilon_0 \alpha_E \tfrac{\mu k^2}{\varepsilon_0} \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) \right)^{-1} \varepsilon_0 \alpha_E i\omega\mu\mu_0 \left[\nabla \times \widehat{\mathbf{G}}^H_{ref}(r_0, r_0) \right] \right]^{-1} \alpha_H \\
\widehat{\alpha}_{EM} &= \left[\hat{\mathbf{I}} - \varepsilon_0 \alpha_E \tfrac{\mu k^2}{\varepsilon_0} \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) - \varepsilon_0 \alpha_E i\omega\mu\mu_0 \left[\nabla \times \widehat{\mathbf{G}}^H_{ref}(r_0, r_0) \right] \left(\hat{\mathbf{I}} - \alpha_H \varepsilon\mu k^2 \widehat{\mathbf{G}}^H_{ref}(r_0, r_0) \right)^{-1} \alpha_H \left(-i\omega \left[\nabla \times \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) \right] \right) \right]^{-1} \times \\
&\times \varepsilon_0 \alpha_E i\omega\mu\mu_0 \left[\nabla \times \widehat{\mathbf{G}}^H_{ref}(r_0, r_0) \right] \left(\hat{\mathbf{I}} - \alpha_H \varepsilon\mu k^2 \widehat{\mathbf{G}}^H_{ref}(r_0, r_0) \right)^{-1} \alpha_H \\
\widehat{\alpha}_{ME} &= \left[\hat{\mathbf{I}} - \varepsilon_0 \alpha_E \tfrac{\mu k^2}{\varepsilon_0} \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) - \alpha_H \left(-i\omega \left[\nabla \times \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) \right] \right) \left(\hat{\mathbf{I}} - \varepsilon_0 \alpha_E \tfrac{\mu k^2}{\varepsilon_0} \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) \right)^{-1} \varepsilon_0 \alpha_E i\omega\mu\mu_0 \left[\nabla \times \widehat{\mathbf{G}}^H_{ref}(r_0, r_0) \right] \right]^{-1} \times \\
&\times \alpha_H \left(-i\omega \left[\nabla \times \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) \right] \right) \left(\hat{\mathbf{I}} - \varepsilon_0 \alpha_E \tfrac{\mu k^2}{\varepsilon_0} \widehat{\mathbf{G}}^E_{ref}(r_0, r_0) \right)^{-1} \varepsilon_0 \alpha_E
\end{aligned} \tag{4}$$

here $\alpha_E = \frac{6i\pi a_1}{k^3}$ $\alpha_H = \frac{6i\pi b_1}{k^3}$ Mie polarizability of magnetic and electric dipoles.

To simplify the numerical calculation procedure, we write the Green's function of the reflected field in cylindrical coordinates. This technique helps to move from double integration to single integration.

$$\begin{aligned}
\hat{\mathbf{G}}^E_{ref}(r_0, r_0) &= \frac{i}{8\pi k_0^2} \int_0^\infty \frac{k_\rho}{k_{z1}} \begin{bmatrix} k_0^2 r^s - k_{z1}^2 r^p & 0 & 0 \\ 0 & k_0^2 r^s - k_{z1}^2 r^p & 0 \\ 0 & 0 & 2k_\rho^2 r^p \end{bmatrix} e^{2ik_{z1}z_0} dk_\rho \\
\hat{\mathbf{G}}^H_{ref}(r_0, r_0) &= \frac{i}{8\pi k_0^2} \int_0^\infty \frac{k_\rho}{k_{z1}} \begin{bmatrix} k_0^2 r^p - k_{z1}^2 r^s & 0 & 0 \\ 0 & k_0^2 r^p - k_{z1}^2 r^s & 0 \\ 0 & 0 & 2k_\rho^2 r^s \end{bmatrix} e^{2ik_{z1}z_0} dk_\rho
\end{aligned} \tag{5}$$

$$\nabla \times \hat{\mathbf{G}}^E_{ref}(r_0, r_0) = -\nabla \times \hat{\mathbf{G}}^H_{ref}(r_0, r_0) = \frac{1}{8\pi} \int_0^\infty k_\rho \left(r^s - r^p \right) \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} e^{2ik_{z1}z_0} dk_\rho \tag{6}$$

Then the efficiency of nanoparticle extinction can be estimated using formula (7). The calculation results are shown in Figs. 3 and 4.

$$\sigma_{ext} = \sigma^p_{ext} + \sigma^m_{ext} = \frac{\omega}{2P_{Inc}} \mathrm{Im}\left[\mathbf{E}^*_{Inc} \mathbf{p} \right] + \frac{\omega\mu_0}{2P_{Inc}} \mathrm{Im}\left[\mathbf{H}^*_{Inc} \mathbf{m} \right] \tag{7}$$

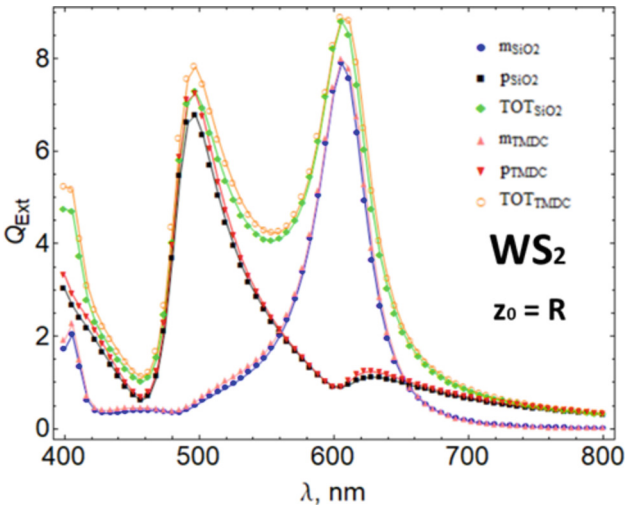


Fig. 3. Efficiency of extinction of a silicon spherical nanoparticle with a radius of 75 nm on a glass substrate and a glass substrate coated with a WS_2 layer. The total efficiency and efficiency of individual multipoles (electric and magnetic dipoles) are shown. The dipoles are located at a distance R from the substrate

4 Discussion

The extinction efficiency of a silicon spherical nanoparticle with a radius of 75 nm on a glass substrate and a glass substrate coated with a TMDC layer was calculated. As an example, in fig. Figure 3 shows the total efficiency and efficiency of individual multipoles (electric and magnetic dipoles) of such a nanoparticle on a WS2 monolayer. The multipole expansion of the extinction efficiency shows the influence of the active two-dimensional material on the optical properties of individual multipole resonances. In this system, the extinction at the electric dipole resonance increases by 15%, while for the magnetic dipole resonance the extinction does not change. This effect is associated with the spectral properties of the complex refractive index of the WS_2 monolayer.

A controlled change in the extinction coefficient of a nanoantenna requires a controlled change in the spectrum of the complex refractive index of an active two-dimensional material. One way to achieve this is to increase the number of layers of 2D material.

Calculations were carried out showing the dynamics of the influence of infinite layers of 2D materials of various thicknesses on the optical resonances of a silicon nanoparticle (see Fig. 4.).

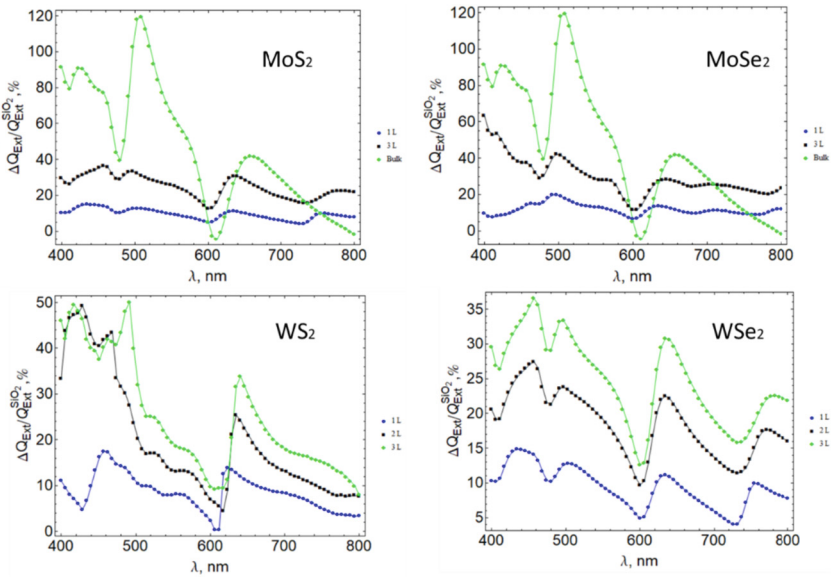


Fig. 4. Relative change in the extinction efficiency of a silicon spherical nanoparticle with a radius of 75 nm on a glass substrate and a glass substrate coated with a TMDC.

5 Conclusion

An approach based on the angular spectral representation of the near field of individual dipoles located near a flat substrate with an active TMDC monolayer is considered.

It is shown that by changing the type of TMDC material and the number of layers, one can control the optical properties of nanoantennas in contact with it.

Acknowledgements. The authors gratefully acknowledge the financial support from the Ministry of Science and Higher Education of the Russian Federation (Agreement No. № 075–15-2022–1150). V.B. acknowledges the support of the Latvian Council of Science, project: NEO-NATE, No. Lzp-2022/1–0553.

References

1. Terekhov, P.D., Evlyukhin, A.B., Redka, D., Volkov, V.S., Shalin, A.S., Karabchevsky, A.: Magnetic octupole response of dielectric quadrumers. Laser Photon Rev. **14**(4), 1900331 (2020)
2. Barhom, H., et al.: Biological kerker effect boosts light collection efficiency in plants. Nano Lett. **19**(10), 7062 (2019)
3. Vestler, D., et al.: Circular dichroism enhancement in plasmonic nanorod metamaterials. Opt. Express **26**(14), 17841–17848 (2018)
4. Canós Valero, A., Gurvitz, E.A., Benimetskiy, F.A., et al.: Theory, observation, and ultrafast response of the hybrid anapole regime in light scattering. Laser Photon Rev. **15**(10), 2100114 (2021)

5. Izmodenova, S.V., Kislov, D.A., Kucherenko, M.G.: Accelerated nonradiative electron-excitation energy transfer between molecules in aqueous pools of reverse micelles containing encapsulated silver nanoparticles. Colloid J. **76**, 683 (2014)
6. Kostina, N.A., Kislov, D.A., Ivinskaya, A.N., et al.: Nanoscale tunable optical binding mediated by hyperbolic metamaterials. ACS Photonics **7**(2), 425–433 (2020)
7. Kislov, D.: Effect of plasmonic silver nanoparticles on the photovoltaic properties of graetzel solar cells. Phys. Procedia **73**, 114 (2015)
8. Shalin, A.S., Sukhov, S.V.: Plasmonic Nanostructures as accelerators for nanoparticles: optical nanocannon. Plasmonics **8**(2), 625–629 (2013)
9. Simovski, C.R., Shalin, A.S., Voroshilov, P.M., Belov, P.A.: Photovoltaic absorption enhancement in thin-film solar cells by non-resonant beam collimation by submicron dielectric particles. J. Appl. Phys. **114**, 103104 (2013)
10. Kucherik, A., et al.: Nano-antennas based on silicon-gold nanostructures. Sci. Rep. **9**, 338 (2019)
11. Rodríguez-Ruiz, I., Tobias, N.A., Muñoz-Berbel, X., Llobera, A.: Photonic lab-on-a-chip: integration of optical spectroscopy in microfluidic systems. Anal. Chem. **88**(13), 6630 (2016)
12. Canós Valero, A., Kislov, D., Gurvitz, E.A., et al.: Nanovortex-driven all-dielectric optical diffusion boosting and sorting concept for lab-on-a-chip platforms. Adv. Sci. **7**(11), 1903049 (2020)
13. Rostamian, A., Madadi-Kandjani, E., Dalir, H., Sorger, V.J., Chen, R.T.: Towards lab-on-chip ultrasensitive ethanol detection using photonic crystal waveguide operating in the mid infrared. Nanophotonics **10**, 1675 (2021)
14. Kucherenko, M.G., Kislov, D.A.: Plasmon-activated intermolecular nonradiative energy transfer in spherical nanoreactors. J. Photochem Photobiology A: Chem. **354**, 25 (2018)
15. Kislov, D., Novitsky, D., Kadochkin, A., Redka, D., Shalin, A.S., Ginzburg, P.: Diffusion-inspired time-varying phosphorescent decay in a nanostructured environment. Phys. Rev. B **101**, 035420 (2020)
16. Li, M., Ling, J., He, Y.: Lithium niobate photonic-crystal electro-optic modulator. Nat. Commun. **11**, 4123 (2020)
17. Xu, X.Y., et al.: A scalable photonic computer solving the subset sum problem. Sci. Adv. **6**(5), eaay5853 (2020)
18. Kuznetsov, A.V., Canós Valero, A., Tarkhov, M., Bobrovs, V., Redka, D., Shalin, A.S.: Transparent hybrid anapole metasurfaces with negligible electromagnetic coupling for phase engineering. Nanophotonics **10**(17), 4385–4398 (2021)
19. Novitsky, D., Karabchevsky, A., Lavrinenko, A., Shalin, A., Novitsky, A.: PT symmetry breaking in multilayers with resonant loss and gain locks light propagation direction. Phys. Rev. B **98**, 125102 (2018)
20. Novitsky, D.V., Shalin, A.S., Redka, D., Bobrovs, V., Novitsky, A.V.: Quasibound states in the continuum induced by PT symmetry breaking. Phys. Rev. B **104**(8), 85126 (2021)
21. Shamkhi, H.K., Sayanskiy, A., Valero, A.C., et al.: Transparency and perfect absorption of all-dielectric resonant metasurfaces governed by the transverse Kerker effect. Phys. Rev. Mater. **3**(8), 85201 (2019)
22. Novitsky, D.V., Valero, A. C., Krotov, A., Salgals, T., Shalin, A.S., Novitsky, A.V.: CPA-lasing associated with the quasibound states in the continuum in asymmetric non-Hermitian structures. ACS Photonics (2022)
23. Shalin, A.S.: Broadband blooming of a medium modified by an incorporated layer of nanocavities. Jetp Lett. **91**, 636–642 (2010)
24. Shalin, A.S., Moiseev, S.G.: Optical properties of nanostructured layers on the surface of an underlying medium. Opt. Spectrosc. **106**, 916–925 (2009)
25. Kislov, D., et al.: Multipole engineering of attractive-repulsive and bending optical forces. Adv. Photonics Res., 2100082 (2021)

26. Kuznetsov, A.V., et al.: Special scattering regimes for conical all-dielectric nanoparticles. Sci. Rep. **12**, 21904 (2022)
27. Terekhov, P.D., Evlyukhin, A.B., Shalin, A.S., Karabchevsky, A.: Polarization-dependent asymmetric light scattering by silicon nanopyramids and their multipoles resonances. J. Appl. Phys. **125**, 173108 (2019)

Flexible GaNP Nanowire-Based Platform: Optical Studies

Alina Kurinnaya[1], Olga Koval[1(✉)], Alex Serov[2], Vjaceslavs Bobrovs[3], Igor Shtrom[2], and Alexey Bolshakov[1,2,4]

[1] Center for Photonics and 2D Materials, Moscow Institute of Physics and Technology, Dolgoprudny 141700, Russia
falseo.yu.koval@gmail.com
[2] Saint Petersburg State University, Saint-Petersburg 199034, Russia
[3] Institute of Telecommunications, Riga Technical University, Riga 1048, Latvia
[4] Alferov University (Formerly St. Petersburg Academic University), Saint-Petersburg 19402, Russia

Abstract. This work is devoted to studying the features of the optical properties of self-catalyzed axially heterostructured GaNP/GaP nanowires on Si(111) grown by plasma-assisted molecular beam epitaxy. Transparent polydimethylsiloxane rubber was used to determine the photoluminescent properties of the grown NW arrays and the parasitic layer separately. Low-temperature photoluminescence studies demonstrate the transition to a quasi-direct bandgap in nanowires, which is characteristic feature of diluted nitrides with a low nitrogen content. The bright photoluminescent response at room temperature demonstrates the potential application of nanowires/rubber membranes in flexible optoelectronic devices. #COMESYSO1120.

Keywords: GaP · GaNP · Nanowire · NW Membrane: III-V on Si · Axially Heterostructure · diluted nitride

1 Introduction

The integration of A3B5 semiconductor compounds and silicon is one of the promising ways for the development of modern optoelectronics, as it makes it possible to combine the capabilities of silicon technology. Among the family of A3B5 semiconductor compounds, the least lattice mismatch with silicon (0.36%) has an indirect-gap semiconductor material is gallium phosphide. However, the efficiency of light-emitting devices and photoelectric converters based on GaP is limited due to the indirect gap nature [1].

The using of Ga(N,P) type ternary nitrogen-containing solid alloys makes it possible to expand the area of functional application of heterostructures based on GaP/Si. Thus, the addition of a small nitrogen fraction to Ga(N,P) compounds leads to a decrease in the band gap by ~100 meV/% [2], which is accompanied by a decrease in the lattice constant. At concentrations of N as early as 0.4.. 0.5% in Ga(N,P), an indirect to a direct bandgap transition is observed [3, 4].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 271–277, 2024.
https://doi.org/10.1007/978-3-031-53549-9_27

The quasi-one-dimensional morphology can contribute to the optical radiation localization into these structures [5, 6] and the manifestation of resonance and waveguide properties [7, 8], determining the radiation pattern and the efficiency of optical radiation input and output from the structure active layers, and also opens up possibilities in a huge variety of other applications [9–14]. The relaxation of elastic strains on the developed NW free surface makes it possible to form lattice-mismatched heterojunctions in axial and radial geometries [15]. In this case, each NW can act as a separate functional device, and the epitaxial array itself, due to the small area of the NW heterointerface with the substrate, can be separated from it. Thus, materials based on heterostructured NW Ga(N,P) solid alloys can be considered as a material platform for energy-efficient light-emitting devices in the visible range.

2 Methods

The GaP and GaNP/GaP heterostructured NW arrays were grown on vicinal Si(111) substrates misoriented on 4° within a vapor-liquid-solid (VLS) self-catalyzed method by plasma-assisted molecular beam epitaxy (PA-MBE) using Veeco GEN-III MBE machine (Plainview, NY, USA). The MBE machine was equipped with the Riber radio-frequency plasma assisted source (13.56 MHz) of activated nitrogen and phosphorus valved cracker.

The growth temperature was measured by both thermocouple and pyrometer and was equal 710 °C during the growth process. Ga flux was kept constant and equivalent to the 0.4 monolayer per second growth rate of planar GaP on Si(111), which corresponds to Ga beam equivalent pressure as 8×10^{-8} Torr.

The grown samples represent pure GaP and GaPN/GaP heterostructured NW arrays, last of all consisting of GaNP segment grown on GaP stem. The growth conditions of the GaP stem and GaNP segment was equal. P/Ga beam equivalent pressure ratio was kept constant during the growth of GaP and GaNP and equal 18. The growth was not interrupted during the plasma activation.

In order to determine the photoluminescence properties of the synthesized NW arrays and the parasitic layer separately, we studied the released membrane and as-grown NW array also separately. The “Sylgard 184” polydimethylsiloxane (PDMS) was used as a transparent rubber source. NW arrays were encapsulated into a PDMS using the G-coating method [16]. The PDMS NW membrane was released from the growth Si(111) substrate using a thick PDMS supported layer (~100 μm), labelled as a cap-film in Fig. 1. Necessity of this layer described by maintaining of the GaP and GaNP/GaP NW-encapsulated membrane mechanical stability.

The NW morphology was studied using scanning electron microscopy (SEM) (Zeiss SUPRA 25) (D-73446, Oberkochen, Germany).

The measurements from as-grown NW arrays and individual NWs were performed at room temperature (RT, 300K) using a Horiba LabRAM HR800 spectrometer (Kyoto, Japan) and an optical microscope for excitation and collection in backscattering geometry. The measurements were carried out with a $\lambda = 532$ nm diode-pumped solid-state (DPSS) YAG: Nd (neodymium-doped yttrium aluminum garnet) continuous wave laser providing slightly higher bandgap value of GaP and nitrogen containing GaNP solid alloys [17]. The local photoluminescence (μ-PL) measurements from individual NWs

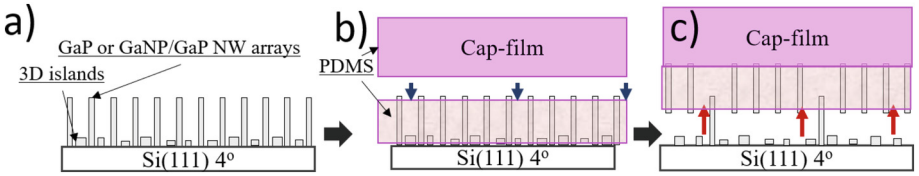


Fig. 1. The schematically presentation of the GaP and heterostructured GaNP/GaP NW_PDMS membrane release

were obtained using an × 100 collecting objective. The excitation beam was focused on the sample to a spot of ≈ 1 μm diameter.

The optical measurements from NW arrays at the temperature of 5 K were performed with the Melles Griot laser Model: 85-GLS-301 (Rochester, NY, USA) an excitation wavelength of $\lambda = 507$ nm using a closed cycle He cryostat. The spectrometer was based on an MDR-204–2 monochromator LOMO FOTONIKA (Saint-Petersburg, Russia) and a Hamamatsu R298 photomultiplier tube (Hamamatsu, Japan).

3 Results

To determine the effect of the nitrogen atom embeddability on the optical properties of ternary nitrogen diluted solid alloys GaNP, a set of samples with and without nitrogen impurities was synthesized in this work. As clearly seen in Fig. 2. The GaP and GaNP/GaP NW array morphology was almost identical. We conclude that nitrogen atom embedding into the zincblend GaP does not affect the nanostructure morphology and preserves self-catalyzed growth mechanism. Remarkable, the remaining Ga- catalytic metal droplets on the NW tips is the evidence of the self-catalyzed growth mechanism [18].

According to SEM analysis, NW are hexagonal in cross-sections, it means that they are following the Si (111) substrate crystallographic orientation. The average diameter 90 nm, and average length 5,8 and 6,1 μm for GaP and GaNP/GaP NW arrays, correspondently. On the silicon growth substrate surface observe the existence of the parasitic 3D islands, which are spontaneously appear during the MBE- growth.

4 Discussions

It is obvious that the analysis of the grown GaP and GaNP/GaP NW arrays using integral evaluation techniques is difficult due to the presence of parasitic islands. Optical studies of individual NW mechanically transferred to the transparent quartz substrate as a schematically shown in Fig. 3. b) inset were performed.

Figure 3. Presents PL spectra both from GaP and GaNP/GaP NW arrays measured at backscattering geometry. The sharp lines located in low-wavelength region in the room-temperature PL spectra, which demonstrated in the Fig. 3. Can be attributed to the GaP and Si first-order and second-order Raman scattering signal of the investigated samples [19].

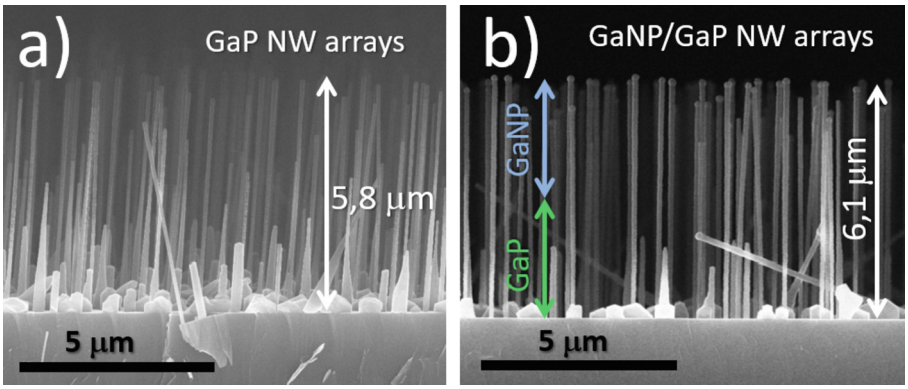


Fig. 2. a) SEM images of the pure GaP NWs grown homogeneously on Si[111] substrates, where: b) SEM images GaNP/GaP NW array

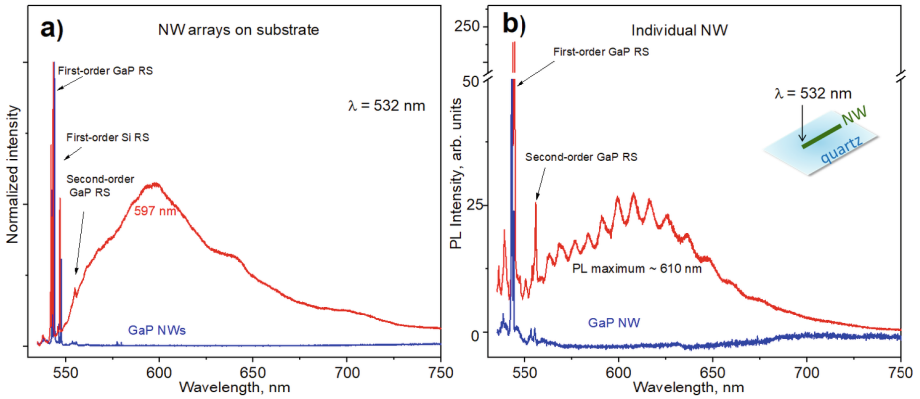


Fig. 3. Room-temperature PL spectra a) pure GaP and heterostructured GaNP/GaP NW arrays, and b) individual GaP and heterostructured GaNP/GaP NWs, inset: experiment geometry.

The absence of the PL signal on the PL spectra of pure GaP NW array and individual GaP NW is evidence of the indirect bandgap structure [2]. The appearance of PL response into the heterostructured GaNP/GaP demonstrate the conversion to the direct gap into the ternary GaNP-like solid alloys [17]. Broad PL signal is common for the highly mismatched alloys such as a GaNP, the maximum PL position is located at 597 nm and to evidence of nitrogen concentration is about ~0.5% throughout the volume of the as-prepared GaNP/GaP NW array on growth substrate [20].

The difference in the position of the room-temperature photoluminescence signal maximum for the as-grown GaNP/GaP NW array and the individual GaNP/GaP NW indicates the unequal nitrogen atom embedding into the NW and parasitic island volume onto the as-grown studied sample.

To quantitatively determine the embedded nitrogen atom concentration in the NW and parasitic island volume, the low-temperature photoluminescent response of the GaP

and GaNP/GaP NW arrays was studied. As demonstrated earlier, the parasitic 3D layer is practically not transferred into the rubber membrane [21], that allows us to state that the signal from the NW/PDMS membrane was collected only from NW arrays.

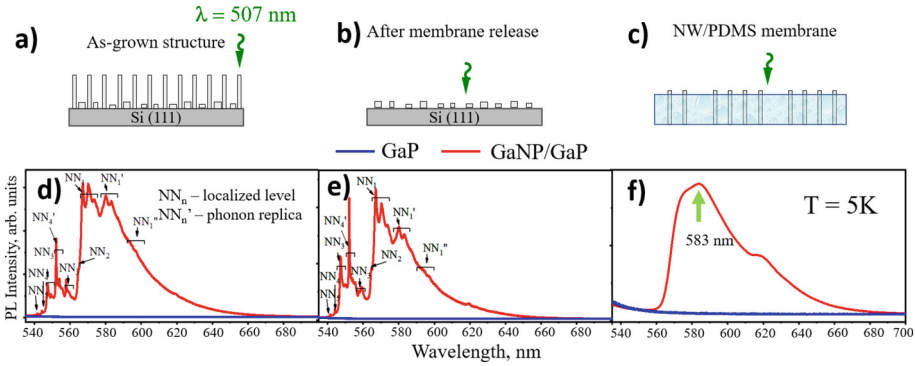


Fig. 4. a)–c) schematic representation of the optical experiment; low-temperature PL response measured on: d) the as-grown GaNP/GaP and GaP NW arrays on Si(111), e) the growth substrate after NW/PDMS membrane release, f) the released NW/PDMS membrane

From the shape and the intensity of NN_n localized levels, described in detail by Thomas [22], of the low-temperature PL spectra one can judge the approximate concentration of embedded nitrogen in a ternary GaNP solid alloys. According to the PL spectrum shape, which shown in Fig. 4. e), the concentration of nitrogen impurities into the parasitic layer can be estimated as 0.3% [23]. The broadband PL response and the PL maximum position with redshift indicate that the concentration of diluted nitrogen atom in GaNP/GaP NWs is about 0.9% [24], which is higher than the evaluated value for parasitic islands. The main source of emission is parasitic islands, which were grown on the substrate surface.

5 Conclusion

The photoluminescence studies show the appearance of a broad PL signal at room temperature in the spectral range between 550 and 700 nm, demonstrating perspectives for creation of the yellow-green flexible light emitters. The oscillation of modulated PL signal is found to be due to Fabry–Pérot-like resonances between the opposite facets of the individual NWs allowing the creation of nano-sized waveguides for future photonic applications. It was demonstrated that nitrogen atoms embed 3 times more efficiently into NWs in comparison with the parasitic 3D layer.

Acknowledgements. The authors gratefully acknowledge the financial support from The Ministry of Science and Higher Education of the Russian Federation (Grant FSRM-2023–0009; agreement 075–03-2023–106, project FSMG-2021–0005). V.B. acknowledges the support of the Latvian Council of Science, project: NEO-NATE, No. Lzp-2022/1–0553.

References

1. Lorenz, M.R., Pettit, G.D., Taylor, R.C.: Band gap of gallium phosphide from 0 to 900°k and light emission from diodes at high temperatures. Phys. Rev. **171**, 876–881 (1968)
2. Sukrittanon, S. et al.: Growth and characterization of dilute nitride GaNxP1−x nanowires and GaNxP1−x/GaNyP1−y core/shell nanowires on Si (111) by gas source molecular beam epitaxy. Appl. Phys. Lett. **105** (2014)
3. Amin, M.N., Faisal, M.: Highly nonlinear polarization-maintaining photonic crystal fiber with nanoscale GaP strips. Appl. Opt. **55**, 10030 (2016)
4. Lan, Y., et al.: Free-standing self-assemblies of gallium nitride nanoparticles: a review. Micromachines **7**, 121 (2016)
5. Kuznetsov, A., et al.: Self-assembled photonic structure: a Ga optical antenna on GaP nanowires. Nanoscale **15**, 2332–2339 (2023)
6. Bolshakov, A.D., et al.: Single GaP nanowire nonlinear characterization with the aid of an optical trap. Nanoscale **14**, 993–1000 (2022)
7. Kuznetsov, A. et al.: Elastic gallium phosphide nanowire optical waveguides—versatile subwavelength platform for integrated photonics. Small **19** (2023)
8. Anikina, M.A., et al.: Numerical study of gap nanowires: individual and coupled optical waveguides and resonant phenomena. Nanomaterials **13**, 56 (2022)
9. Kondratev, V.M., et al.: Silicon nanowire-based room-temperature multi-environment ammonia detection. ACS Appl. Nano Mater. **5**, 9940–9949 (2022)
10. Kondratev, V.M., et al.: Si nanowire-based schottky sensors for selective sensing of NH 3 and HCl via impedance spectroscopy. ACS Appl. Nano Mater. **6**, 11513–11523 (2023)
11. Kadinskaya, S.A., et al.: Hydrothermal zinc oxide nanostructures: geometry control and narrow band UV emission. J. Phys. Conf. Ser. **2227**, 012007 (2022)
12. Canós Valero, A., et al.: Theory, observation, and ultrafast response of the hybrid anapole regime in light scattering. Laser Photon. Rev. **15**, 2100114 (2021)
13. Terekhov, P.D., Evlyukhin, A.B., Shalin, A.S., Karabchevsky, A.: Polarization-dependent asymmetric light scattering by silicon nanopyramids and their multipoles resonances. J. Appl. Phys. **125**, 173108 (2019)
14. Kostina, N.A., et al.: Nanoscale tunable optical binding mediated by hyperbolic metamaterials. ACS Photonics **7**, 425–433 (2020)
15. Sibirev, N.V., Berdnikov, Y.S., Fedorov, V.V., Shtrom, I.V., Bolshakov, A.D.: Parameter-free model of the self-catalyzed growth of ga(As, P) nanowires. Semiconductors **56**, 14–17 (2022)
16. Neplokh, V., et al.: Modified silicone rubber for fabrication and contacting of flexible suspended membranes of n-/p-GaP nanowires with a single-walled carbon nanotube transparent contact. J. Mater. Chem. C **8**, 3764–3772 (2020)
17. Bellaiche, L., Wei, S.-H., Zunger, A.: Band gaps of GaPN and GaAsN alloys. Appl. Phys. Lett. **70**, 3558–3560 (1997)
18. Bolshakov, A.D. et al.: Growth and Characterization of GaP/GaPAs Nanowire Heterostructures with Controllable Composition. Phys. status solidi – Rapid Res. Lett. **13**, 1900350 (2019)
19. da Silva, B.C., et al.: Optical absorption exhibits pseudo-direct band gap of wurtzite gallium phosphide. Sci. Rep. **10**, 7904 (2020)
20. Buyanova, I.A., Rudko, G.Y., Chen, W.M., Xin, H.P., Tu, C.W.: Radiative recombination mechanism in GaNxP1−x alloys. Appl. Phys. Lett. **80**, 1740–1742 (2002)
21. Koval, O.Y., et al.: Structural and optical properties of self-catalyzed axially heterostructured GaPN/GaP nanowires embedded into a flexible silicone membrane. Nanomaterials **10**, 2110 (2020)

22. Thomas, D.G., Hopfield, J.J., Frosch, C.J.: Isoelectronic traps due to nitrogen in gallium phosphide. Phys. Rev. Lett. **15**, 857–860 (1965)
23. Buyanova, I.A., et al.: Direct experimental evidence for unusual effects of hydrogen on the electronic and vibrational properties of GaNxP1−x alloys. Phys. Rev. B **70**, 245215 (2004)
24. Baillargeon, J.N., Cheng, K.Y., Hofler, G.E., Pearah, P.J., Hsieh, K.C.: Luminescence quenching and the formation of the GaP1− x N x alloy in GaP with increasing nitrogen content. Appl. Phys. Lett. **60**, 2540–2542 (1992)

Transverse Kerker Effects in All-Dielectric Conical Nanoparticles

Alexey V. Kuznetsov[1,2(✉)] and Vjaceslavs Bobrovs[2]

[1] Moscow Institute of Physics and Technology, Dolgoprudny, Russia
alexey.kuznetsov98@gmail.com
[2] Riga Technical University, Riga, Latvia

Abstract. A lot of new effects and phenomena based on dielectric nanoscatterers are becoming possible through research in the field of dielectric nanophotonics. These resonators show exceptional optical characteristics, encompassing both electric and magnetic responses within the visible spectrum. The conical shape introduces additional degrees of freedom, paving the way for fresh opportunities to tailor light-matter interactions at the nanoscale.In this research, we study truncated cone resonators, which demonstrate transverse Kerker effects.

Keywords: Transverse Kerker effects · Truncated Cone · All-dielectric nanophotonics

1 Introduction

In recent years, interest in studying the optical properties of subwavelength dielectric nanoparticles has been constantly growing. These nanoparticles represent a promising basis for creating photonic devices that outperform their counterparts. They suggest several undeniable advantages [1], such as minimal losses, high efficiency [2–4], and the capability to manipulate both the electric and magnetic components of light simultaneously [5]. As a result of these remarkable properties, nanolasers [6], nanoantennas [7], ultrathin lenses [8], sensors, detectors, metamaterials [9], metasurfaces [10], and numerous other emerging applications have been developed to harness the distinctive effects facilitated by these resonators [11–13].

Remarkable progress in high-index dielectric nanophotonics has been achieved through the efficient use of versatile manipulation of multipole excitations in subwavelength scatterers [14, 15]. This contributed to a deeper understanding and effective control of various phenomena through the integration and interaction of multipole moments. An important result of this achievement is the transverse Kerker effect, which entails the suppression of both backward and forward scattering of nanoparticles. Originally conceived for a hypothetical sphere with equal epsilon and mu [16], the Kerker effect now holds great significance in dielectric nanophotonics. It has paved the way for the development of fully transparent phase-tailoring “Huygens” metasurfaces, offering a wide range of possibilities in this field [8, 17].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 278–282, 2024.
https://doi.org/10.1007/978-3-031-53549-9_28

Truncated cones are increasingly preferred as components in photonic structures due to their numerous advantages [18]. An important benefit lies in their capacity to adjust one of the radii, introducing an additional level of flexibility. Unlike the more commonly used cylindrical nanoscatterers, nanocones provide precise control over the modal composition of the scatterer. This advanced control allows for improved customization options and effects that were previously unavailable.

This research primarily focuses on the transverse Kerker effects exhibited in silicon truncated nanocones. With the increasing importance of the transverse Kerker effect in the rapidly developing field of dielectric nanophotonics, there is a need to investigate and conceptualize these effects in the context of different scatterer geometries.

2 Transverse Kerker Effects in Cone Silicon Particles

As the Kerker effect has been studied, various combinations of multipole moments and phase differences have been discovered, leading to a broader understanding of the Kerker effect. In addition to the generalized Kerker effect, another crucial scattering phenomenon is the transverse Kerker effect, which was first described in [19, 20]. The transverse Kerker effect primarily arises from the interplay of multipole interactions, leading to scattering that predominantly occurs sideways, with only a minor portion involving forward scattering, in line with the optical theorem. This effect can be realized through straightforward combinations of two multipoles, as well as through more complex configurations (see Fig. 1).

Figure 1 shows the transverse Kerker effects achieved through a combination of numerical and analytical approaches. The term ‘transverse’ is derived from the shape of the scattering patterns. By adjusting the phases and amplitudes of the multipole components, it is possible to induce lateral scattering while minimizing forward and backward scattering.

This work shows various manifestations of the transverse Kerker effects. The radiation patterns of these cones generated using COMSOL Multiphysics closely resemble the expected Kerker-type patterns obtained from ideal point calculations. However, minor differences arise due to practical limitations in implementing the ideal combination of required multipoles in real systems.

In addition, the results highlight that higher order multipoles are redundant, as evidenced by the good agreement between summation of multipole scattering cross sections and calculations obtained from integration of the Poynting vector over a closed surface within the long range. This agreement is evident from the gray and orange lines shown in Fig. 1a–e.

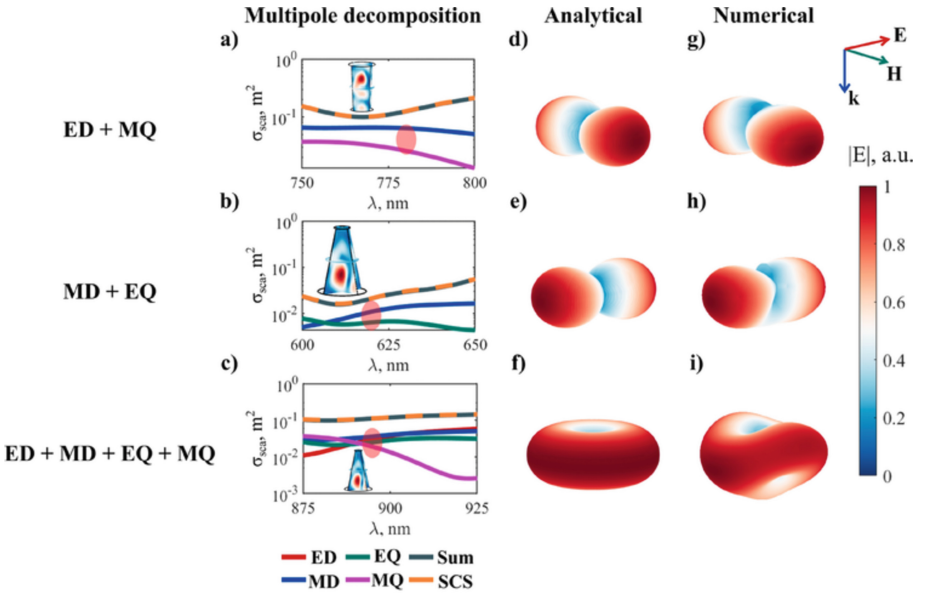


Fig. 1. Transverse Kerker effects in truncated conical nanoparticles with different geometries. Figures (a–c) provide a visual representation of the multipolar decomposition, while figures (d–f) display the far-field sum of multipoles calculated analytically for a point. Figures (g–i) illustrate the numerically calculated far-field distribution for conical nanoparticles of different sizes at the spectral points indicated in (a–c).

3 Conclusion

In this study, we have demonstrated transverse Kerker effects within individual nanoscatterers of the same geometry shape. This makes nanocones a versatile and easy-to-fabricate platform for realizing new photonic devices that use a wide range of effects. This platform provides the flexibility to customize the scattering pattern.

Moreover, our study reveals conical geometry as a versatile platform for achieving numerous important optical effects in the field of nanophotonics. We achieved transverse Kerker effects on a single scatterer shape while accounting for realistic refractive index dispersion of silicon. This research represents a significant advance in the field of nanophotonics, especially in the context of more complex shapes, offering the ability to precisely control the effects achievable with a single nanoscatterer shape. Such advances significantly reduce the costs associated with developing photonic devices and open new opportunities for practical applications in next-generation photonics.

The results of this study have wide implications in various fields of research. In addition, the proposed approach can be used in the development of metasurfaces, opening a number of optical effects that were previously unattainable.

Acknowledgments. The authors gratefully acknowledge the financial support from the Ministry of Science and Higher Education of the Russian Federation (Agreement No. № 075–15-2022–1150). V.B. acknowledges the support of the Latvian Council of Science, project: DNSSN, No. Lzp-2021/1–0048.

References

1. Shalin, A.S., Sukhov, S.V.: Plasmonic nanostructures as accelerators for nanoparticles: optical nanocannon. Plasmonics **8**, 625–629 (2013)
2. Simovski, C.R., Shalin, A.S., Voroshilov, P.M., Belov, P.A.: Photovoltaic absorption enhancement in thin-film solar cells by non-resonant beam collimation by submicron dielectric particles. J. Appl. Phys. **114** (2013)
3. Shalin, A.S.: Broadband blooming of a medium modified by an incorporated layer of nanocavities. JETP Lett. **91**, 636–642 (2010)
4. Shalin, A.S., Moiseev, S.G.: Optical properties of nanostructured layers on the surface of an underlying medium. Opt. Spectrosc. **106**, 916–925 (2009)
5. Canós Valero, A., et al.: Theory, observation, and ultrafast response of the hybrid anapole regime in light scattering. Laser Photon Rev. **15**, 2100114 (2021)
6. Novitsky, D.V., Valero, A.C., Krotov, A., Salgals, T., Shalin, A.S., Novitsky, A.V.: CPA-lasing associated with the quasibound states in the continuum in asymmetric non-hermitian structures. ACS Photonics **9**, 3035–3042 (2022)
7. Kucherik, A., et al.: Nano-antennas based on silicon-gold nanostructures. Sci. Rep. **9**, 338 (2019)
8. Tian, Y., Li, Z., Xu, Z., Wei, Y., Wu, F.: High transmission focusing lenses based on ultrathin all-dielectric Huygens' metasurfaces. Opt. Mater. (Amst). **109**, 110358 (2020)
9. Vestler, D., et al.: Circular dichroism enhancement in plasmonic nanorod metamaterials. Opt. Express **26**, 17841 (2018)
10. Kuznetsov, A.V., Canós Valero, A., Tarkhov, M., Bobrovs, V., Redka, D., Shalin, A.S.: Transparent hybrid anapole metasurfaces with negligible electromagnetic coupling for phase engineering. Nanophotonics. **10**, 4385–4398 (2021)
11. Canós Valero, A., et al.: Nanovortex-driven all-dielectric optical diffusion boosting and sorting concept for lab-on-a-chip platforms. Adv. Sci. **7**, 1903049 (2020)
12. Novitsky, D.V., Karabchevsky, A., Lavrinenko, A.V., Shalin, A.S., Novitsky, A.V.: PT symmetry breaking in multilayers with resonant loss and gain locks light propagation direction. Phys. Rev. B **98**, 125102 (2018)
13. Kostina, N.A., et al.: Nanoscale tunable optical binding mediated by hyperbolic metamaterials. ACS Photonics **7**, 425–433 (2020)
14. Terekhov, P.D., Evlyukhin, A.B., Redka, D., Volkov, V.S., Shalin, A.S., Karabchevsky, A.: Magnetic octupole response of dielectric quadrumers. Laser Photon Rev. **14**, 1900331 (2020)
15. Terekhov, P.D., Evlyukhin, A.B., Shalin, A.S., Karabchevsky, A.: Polarization-dependent asymmetric light scattering by silicon nanopyramids and their multipoles resonances. J. Appl. Phys. **125**, 173108 (2019)
16. Kerker, M., Wang, D.S., Giles, C.L.: Electromagnetic scattering by magnetic spheres. J. Opt. Soc. Am. **73**, 765–767 (1983). https://doi.org/10.1364/JOSA.73.000765
17. Pfeiffer, C., Grbic, A.: Metamaterial Huygens' surfaces: Tailoring wave fronts with reflectionless sheets. Phys. Rev. Lett. **110**, 1–5 (2013)
18. Kuznetsov, A.V., et al.: Special scattering regimes for conical all-dielectric nanoparticles. Sci. Rep. **12**, 21904 (2022)

19. Shamkhi, H.K., et al.: Transparency and perfect absorption of all-dielectric resonant metasurfaces governed by the transverse Kerker effect. Phys Rev Mater. **3**, 1–10 (2019)
20. Shamkhi, H.K., et al.: Transverse scattering and generalized kerker effects in all-dielectric mie-resonant metaoptics. Phys. Rev. Lett. **122**, 193905 (2019)

Generalized Kerker Effects in All-Dielectric Conical Nanoparticles

Alexey V. Kuznetsov[1,2] and Vjaceslavs Bobrovs[2(✉)]

[1] Moscow Institute of Physics and Technology, Dolgoprudny, Russia
[2] Riga Technical University, Riga, Latvia
vjaceslavs.bobrovs@rtu.lv

Abstract. All-dielectric nanophotonics allows the study of new optical phenomena and different scattering characteristics from dielectric nanoparticles. These resonators have unique optical properties spanning both electrical and magnetic responses in the visible spectrum, offering promising applications in nano-optics, biology, sensing, and various other fields. Our research focuses on truncated cone resonators, which demonstrate the full range of generalized Kerker effects. This achievement is attributed to an inherent characteristic of cones—their asymmetry along the primary axis. Through a simple geometric approach, these effects enable the optimization of the fabrication of photonic devices with diverse functionalities.

Keywords: Generalized Kerker effects · Truncated Cone · All-dielectric nanophotonics

1 Introduction

Recently, there has been significant interest in studying the optical properties of subwavelength dielectric and semiconductor nanoparticles. These particles hold significant promise, potentially surpassing current technologies and even creating entirely new classes of devices. They offer several undeniable advantages [1], including low losses, high efficiency [2–4], and the ability to manipulate both the electric and magnetic components of light simultaneously [5]. These properties allow to develop of nanolasers [6], nanoantennas [7], ultrathin lenses [8], sensors [9], metamaterials [10, 11], metasurfaces [12], and other emerging applications that harness the unique effects enabled by these resonators [13, 14].

Significant progress has been made in high-index dielectric nanophotonics by successfully achieving flexible control over multipole excitations in subwavelength scatterers [15]. This achievement has enabled a deeper understanding and effective manipulation of various phenomena by combining multipole moments. One notable outcome of this advancement is the realization of the Kerker effect, which allows for the suppression of backward or forward scattering from a nanoparticle [16]. Kerker effect now holds immense importance in the field of dielectric nanophotonics. It has paved the way for the development of fully transparent Huygens metasurfaces that can precisely tailor the phase of light, offering a wide range of possibilities in this area of research [17, 18].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 283–287, 2024.
https://doi.org/10.1007/978-3-031-53549-9_29

Truncated cones are becoming increasingly popular as components of photonic structures due to their undeniable advantages [19]. A key benefit is that one of the radii can be adjusted, providing an additional level of flexibility. Unlike the frequently employed cylindrical nano-scattering structures, nanocones provide finer control over the mode of the scatterer. This increased level of control results in improved customization capabilities, allowing you to achieve effects that were previously unattainable.

Notably, the truncated-cone geometry is still relatively underexplored when compared to more 'conventional' shapes like spheres, cubes, and cylinders [20, 21]. In this study, we explore the generalized Kerker effects displayed by truncated silicon nanocones. The Kerker effect has found wide application in the rapidly developing field of dielectric nanophotonics. Therefore, it is critical to delve deeper and understand these effects in the context of different scatterer geometries.

2 Generalized Kerker Effects in Cone Silicon Particles

The Kerker effect was originally discovered in spherical particles characterized by identical electrical permeabilities and magnetic permeabilities. In such particles, when the electric and magnetic dipoles have identical field amplitudes and a certain phase shift, only forward or backward scattering is observed. Numerous studies have experimentally confirmed the manifestation of this phenomenon in dielectric nanoparticles. As research progressed, additional important combinations of multipole moments and phase differences were discovered, leading to the broader concept of the Kerker effect. Currently, the term "generalized Kerker effect" is used to characterize situations where there is significant forward or backward scattering. (see Fig. 1).

Figure 1 presents a study of the Kerker effects using both numerical (Comsol Multiphysics) and analytical calculations. Figure 1(a-e) shows the dependence of the multipole contributions to the scattering cross section.

The Kerker effect points are highlighted in red. Corresponding analytical representations (f-j) and numerical simulations (k-o) of the far-field patterns are provided for each case. Although the scattering shapes are similar, small deviations occur in the numerical calculations due to the small contributions of other multipoles. The strong agreement between the summation of scattering cross-sections from the multipoles and the calculations derived by integrating the Pointing vector over a closed surface in the far-field region, normalized to the incident field intensity, further supports the notion that higher-order multipoles are not required. This agreement is showed by the gray and orange lines in Fig. 1(a-e).

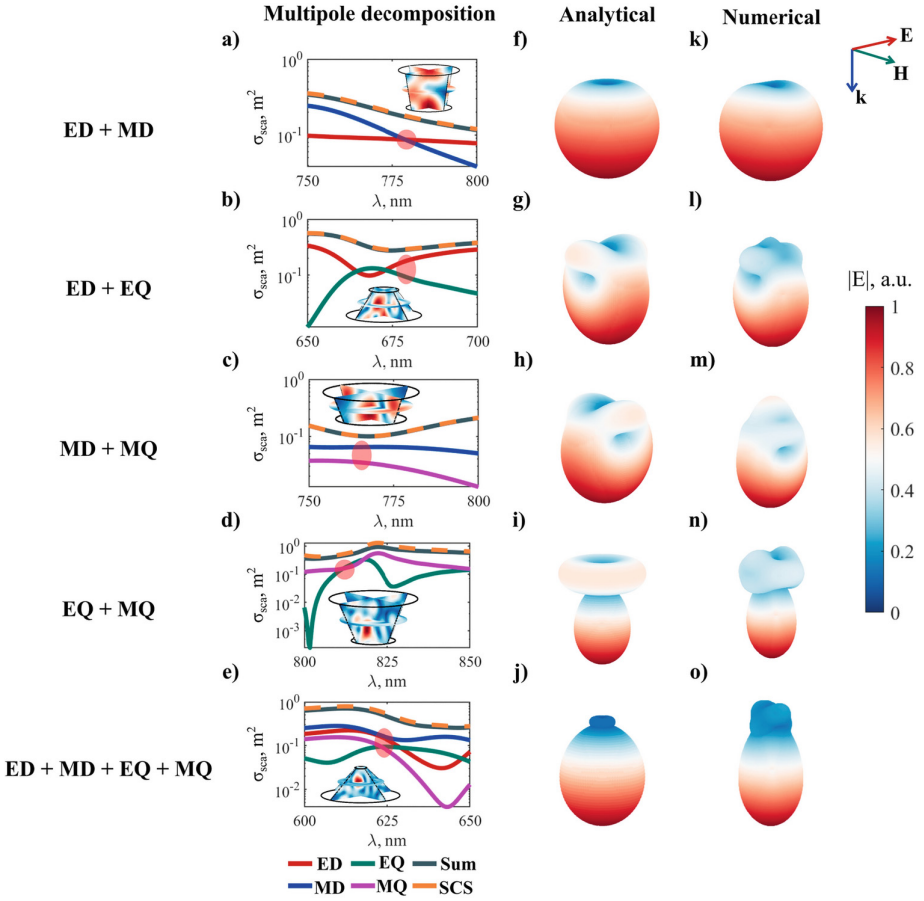


Fig. 1. Generalized Kerker effects in truncated conical nanoparticles. Figures (a-e) illustrate the decomposition of multipolar contributions, while figures (f-j) present the analytical summation of multipoles in the far field for a point. Figures (k-o) show the numerically computed far-field distributions for conical nanoparticles with varying geometries at the spectral points indicated red dots in (a-e). The incident plane wave illuminates the nanoparticles from the top of cone.

3 Conclusion

In this study, we have showed the complete range of established generalized Kerker effects for individual nanoscatterers with the same geometric shapes, representing a pioneering achievement. This underscores the adaptability associated with nanocones as a foundation for introducing novel photonic devices that demonstrate a broad spectrum of multipolar interference phenomena. These effects provide versatile control over scattering patterns.

Furthermore, we propose the conical geometry as a universal platform for achieving all known generalized Kerker effects within a single scatterer shape, while considering the realistic dispersion of the refractive index of silicon. This research marks particularly

advantages concerning more intricate shapes. The use of nanoscatterers of the same shape significantly reduces the cost of developing photonic devices and paves the way for fresh prospects in the next generation of photonics.

Acknowledgments. The authors gratefully acknowledge the financial support from the Ministry of Science and Higher Education of the Russian Federation (Agreement No. № 075–15-2022–1150). V.B. acknowledges the support of the Latvian Council of Science, project: DNSSN, No. Lzp-2021/1–0048.

References

1. Shalin, A.S., Sukhov, S.V.: Plasmonic nanostructures as accelerators for nanoparticles: optical nanocannon. Plasmonics **8**, 625–629 (2013)
2. Simovski, C.R., Shalin, A.S., Voroshilov, P.M., Belov, P.A.: Photovoltaic absorption enhancement in thin-film solar cells by non-resonant beam collimation by submicron dielectric particles. J. Appl. Phys. **114** (2013)
3. Shalin, A.S.: Broadband blooming of a medium modified by an incorporated layer of nanocavities. JETP Lett. **91**, 636–642 (2010)
4. Shalin, A.S., Moiseev, S.G.: Optical properties of nanostructured layers on the surface of an underlying medium. Opt. Spectrosc. **106**, 916–925 (2009)
5. Terekhov, P.D., Babicheva, V.E., Baryshnikova, K.V., Shalin, A.S., Karabchevsky, A., Evlyukhin, A.B.: Multipole analysis of dielectric metasurfaces composed of nonspherical nanoparticles and lattice invisibility effect. Phys. Rev. B. **99** (2019)
6. Novitsky, D.V., Valero, A.C., Krotov, A., Salgals, T., Shalin, A.S., Novitsky, A.V.: CPA-lasing associated with the quasibound states in the continuum in asymmetric non-hermitian structures. ACS Photonics **9**, 3035–3042 (2022)
7. Kucherik, A., et al.: Nano-antennas based on silicon-gold nanostructures. Sci. Rep. **9**, 338 (2019). https://doi.org/10.1038/s41598-018-36851-w
8. Tian, Y., Li, Z., Xu, Z., Wei, Y., Wu, F.: High transmission focusing lenses based on ultrathin all-dielectric Huygens' metasurfaces. Opt. Mater. (Amst). **109**, 110358 (2020)
9. Kuznetsov, A.V., Canós Valero, A., Tarkhov, M., Bobrovs, V., Redka, D., Shalin, A.S.: Transparent hybrid anapole metasurfaces with negligible electromagnetic coupling for phase engineering. Nanophotonics. **10**, 4385–4398 (2021)
10. Vestler, D., et al.: Circular dichroism enhancement in plasmonic nanorod metamaterials. Opt. Express **26**, 17841 (2018)
11. Kostina, N.A., et al.: Nanoscale tunable optical binding mediated by hyperbolic metamaterials. ACS Photonics **7**, 425–433 (2020)
12. Canós Valero, A., et al.: Theory, observation, and ultrafast response of the hybrid anapole regime in light scattering. Laser Photon Rev. **15**, 2100114 (2021)
13. Canós Valero, A., et al.: Nanovortex-driven all-dielectric optical diffusion boosting and sorting concept for lab-on-a-chip platforms. Adv. Sci. **7**, 1903049 (2020)
14. Novitsky, D.V., Karabchevsky, A., Lavrinenko, A.V., Shalin, A.S., Novitsky, A.V.: PT symmetry breaking in multilayers with resonant loss and gain locks light propagation direction. Phys. Rev. B **98**, 125102 (2018)
15. Terekhov, P.D., Evlyukhin, A.B., Redka, D., Volkov, V.S., Shalin, A.S., Karabchevsky, A.: Magnetic octupole response of dielectric quadrumers. Laser Photon Rev. **14**, 1900331 (2020)
16. Kerker, M., Wang, D.S., Giles, C.L.: Electromagnetic scattering by magnetic spheres. J. Opt. Soc. Am. **73**, 765–767 (1983)

17. Rahimzadegan, A., et al.: Beyond dipolar Huygens' metasurfaces for full-phase coverage and unity transmittance. Nanophotonics. **9**, 75–82 (2020)
18. Chen, M., Kim, M., Wong, A.M.H., Eleftheriades, G.V.: Huygens' metasurfaces from microwaves to optics: a review. Nanophotonics. **7**, 1207–1231 (2018)
19. Kuznetsov, A.V., et al.: Special scattering regimes for conical all-dielectric nanoparticles. Sci. Rep. **12**, 21904 (2022)
20. Terekhov, P.D., Evlyukhin, A.B., Shalin, A.S., Karabchevsky, A.: Polarization-dependent asymmetric light scattering by silicon nanopyramids and their multipoles resonances. J. Appl. Phys. **125**, 173108 (2019)
21. Terekhov, P.D., et al.: Broadband forward scattering from dielectric cubic nanoantenna in lossless media. Opt. Express **27**, 10924 (2019)

Ultrashort Pulse Generation in Spaser Through Nonlinear Regime

Morteza A. Sharif[1], Mehdi Borjkhani[2], and Vjaceslavs Bobrovs[3](✉)

[1] Photonics Lab, Faculty of Electrical Engineering, Urmia University of Technology, Band Road, Urmia, Iran
m.abdolahisharif@ee.uut.ac.ir

[2] ICTER, Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, Warsaw 01-224, Poland

[3] Institute of Telecommunications, Riga Technical University, Riga 1048, Latvia
vjaceslavs.bobrovs@rtu.lv

Abstract. A complete Jacobi Elliptic Functions (JEFs)-based model is developed for describing the dynamical behavior of Surface Plasmons (SPs) in a plasmonic waveguide and nanolaser (spaser). JEFs are arranged in two Down Asymptotic (DA) and Up Asymptotic (UA) solutions. DA stand for the large absorption coefficient and subsequently, the linear state while the UA are ascribed to the large nonlinearity. The large absorption coefficient causes decaying feature of SPs as anticipated. The huge nonlinearity can lead to the generation of ultrafast pulses in a SP nanolaser (spaser). It is thus learnt that one can study the quantum effects of a spaser system on a mesoscopic scale, instead of having to treat SPs as individual quasi-particles.

Keywords: Spaser · Jacobi Elliptic Functions · ultrashort pulses

1 Introduction

Today, nanophotonics attracts researchers due to the novel emerging aspects [1–10]. Surface Plasmons (SPs) are oscillating electrons formed at an interface of dielectric/metal layers. SPs modified dispersion relation reveals a field confinement beyond the light diffraction limit promising for the future nanophotonic devices [11–15]. However, high decaying rate of SPs propagation restricts their applicability. One important approach is to amplify the SPs using a nanocavity and active medium like that occurs in a laser known as SPs nanolaser (spaser). Multiple studies have investigated dynamical mechanism of the spaser, lots of them using a quantum-based approach. Jacobi Elliptic Functions (JEFs)-based quasi-classical model has been used to describe the nonlinear dynamics in plasmonic waveguides [16, 17]. Analysis of the nonlinear behavior of the spaser is important due to the strong interaction of SPs and nanomaterial. Furthermore, a spaser may exhibit instability and complex chaotic behavior inspired by what transpires in a laser cavity. JEFs-based model is a simple but still rigorous method to involve the mentioned requirement as we have previously shown in our papers. Based on our developed model for spaser dynamics, we show an ultrafast switching between the decaying and amplifying behavior that can result in ultrashort pulse generation.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 288–291, 2024.
https://doi.org/10.1007/978-3-031-53549-9_30

2 Theory and Model

Introducing $k_1 = k_\omega|_{\omega=\omega_0} = 1/v_g(\omega_0)$ (v_g is the group velocity), $k_2 = k_{\omega\omega}|_{\omega=\omega_0}$ and γ as the nonlinear coefficient, SPs spatiotemporal evolution of the carrier amplitude Ψ can be given in Eq. (1) [16, 17].

$$i\Psi_z - k_2/2\Psi_{tt} + \gamma|\Psi|^2\Psi + i\eta\Psi = 0. \tag{1}$$

where η is the absorption coefficient. Assuming $\gamma' = -2\gamma/k_2$, $\eta' = -2i\eta/k_2$ and writing $\Psi = \mathrm{Z}(z)\mathrm{T}(t)$, Eq. (1) reduces to Eq. (2).

$$\mathrm{T_{tt}} + \gamma'|\mathrm{T}|^2\mathrm{T} + \eta'\mathrm{T} = 0. \tag{2}$$

Equation (2) is a quasi-classical formalism to investigate the temporal dynamics of T and thus, the energy eigenvalues of the plasmonic system. It allows to study of a quantum system on a mesoscopic scale, instead of having to deal with a complicated many-body quantum problem. Two Down Asymptotic solutions (DA) for $\eta' \gg \gamma'$ (Eq. 3) and Up Asymptotic (UA) solutions for $\gamma' \gg \eta'$ (Eq. 4) can be distinguished.

$$\mathrm{T}_L \simeq c_1 \cos \Pi^{(R)} t - i \sin \Pi^{(R)} t \tanh \Pi^{(i)} t / \mathrm{sech}\Pi^{(i)} t, \tag{3}$$

$$\mathrm{T}_{NL} \simeq c_1 \mathrm{sech}\Pi^{(R)} t \Big(\cos \Pi^{(i)} t - i \tanh \Pi^{(R)} t \sin \Pi^{(i)} t\Big) / \Big(1 - \mathrm{sech}^2 \Pi^{(R)} t \sin^2 \Pi^{(i)} t\Big), \tag{4}$$

where cn, sn and dn are the JEFs; $\Pi = \left(\eta' + c_1^2\gamma'\right)^{1/2}$; $\Theta = \left(c_1^2\gamma'/2\left(\eta' + c_1^2\gamma'\right)\right)^{1/2}$ and c_1 and c_2 are the constant should be obtained from the initial conditions [16, 17].

3 Simulation Results and Discussion

DA solutions (Eq. (3)) reveals a quasi-periodic behavior. Spatiotemporal evolutions are shown in Fig. 1(a) and Fig. 1(c) (for 10 times larger group velocities). Figure 1(b) and Fig. 1(d) respectively depict the corresponding attractor. Temporal evolutions of Eq. (4) are depicted in Fig. 2(a) and Fig. 2(c) (for 10 times larger group velocity). Figure 2(b) an Fig. 2(d) respectively show their phase portrait. There is a clear sequence of temporally spaced pulses. The amplitude T can be either decaying or amplifying depending on η' real and imaginary forms. In the presence of huge nonlinearity, an ultrafast switching between the decaying and amplifying features can be resulted appearing as ultrashort pulse generation. This pulsation effect is formed through an unstable and chaotic regime [16, 17]. For a real plasmonic system of graphene-Si waveguide, thes model predict pulse durations ~10^{-15} s and tunable pulse rate.

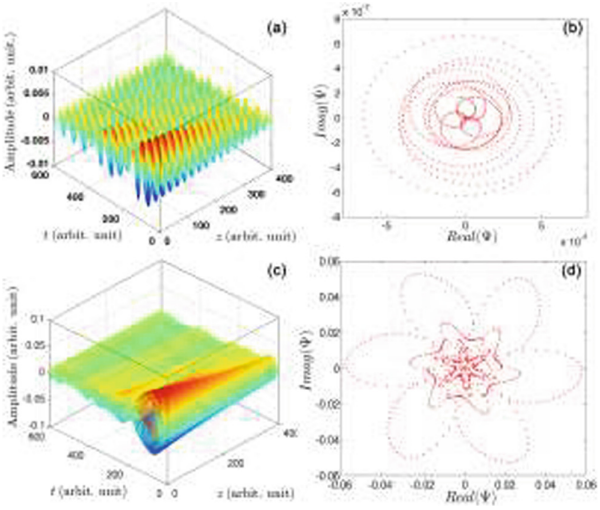


Fig. 1. (a) DA solutions, (b) strange attractor, (c) & (d) same for 10 times larger group velocity.

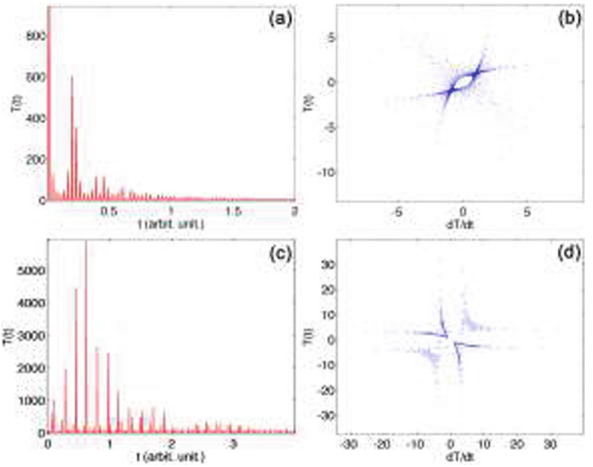


Fig. 2. (a) UA solutions, (b) phase portrait, (c) & (d) same for 10 times larger group velocity

4 Conclusion

We have shown that the generation of ultrashort pulses in a spaser can be assigned to the nonlinear dynamical regime. We have described the formation of ultrashort pulses based on the JEFs-based model.

Acknowledgements. The authors gratefully acknowledge the financial support from the Ministry of Science and Higher Education of the Russian Federation (Agreement No. № 075–15-2022–1150). V.B. acknowledges the support of the Latvian Council of Science, project: DNSSN, No. Lzp-2021/1–0048.

References

1. Simovski, C.R., Shalin, A.S., Voroshilov, P.M., Belov, P.A.: Photovoltaic absorption enhancement in thin-film solar cells by non-resonant beam collimation by submicron dielectric particles. J. Appl. Phys. **114**(10), 103104 (2013)
2. Shalin, A.S.: Broadband blooming of a medium modified by an incorporated layer of nanocavities. JETP Lett. **91**, 636–642 (2010)
3. Shalin, A.S., Moiseev, S.G.: Optical properties of nanostructured layers on the surface of an underlying medium. Opt. Spectrosc. **106**, 916–925 (2009)
4. Terekhov, P.D., Evlyukhin, A.B., Redka, D., Volkov, V.S., Shalin, A.S., Karabchevsky, A.: Magnetic Octupole Response of Dielectric Oligomers. arXiv preprint arXiv:1910.04538 (2019)
5. Novitsky, D.V., Karabchevsky, A., Lavrinenko, A.V., Shalin, A.S., Novitsky, A.V.: PT symmetry breaking in multilayers with resonant loss and gain locks light propagation direction. Phys. Rev. B **98**(12), 125102 (2018)
6. Canós Valero, A., et al.: Theory, observation, and ultrafast response of the hybrid anapole regime in light scattering. Laser Photonics Rev. **15**(10), 2100114 (2021)
7. Terekhov, P.D., Evlyukhin, A.B., Shalin, A.S., Karabchevsky, A.: Polarization-dependent asymmetric light scattering by silicon nanopyramids and their multipoles resonances. J. Appl. Phys. **125**(17), 173108 (2019)
8. Kuznetsov, A.V., Canós Valero, A., Tarkhov, M., Bobrovs, V., Redka, D., Shalin, A.S.: Transparent hybrid anapole metasurfaces with negligible electromagnetic coupling for phase engineering. Nanophotonics **10**(17), 4385–4398 (2021)
9. Novitsky, D.V., Valero, A.C., Krotov, A., Salgals, T., Shalin, A.S., Novitsky, A.V.: CPA-lasing associated with the quasibound states in the continuum in asymmetric non-Hermitian structures. ACS Photonics **9**(9), 3035–3042 (2022)
10. Kuznetsov, A.V., et al.: Special scattering regimes for conical all-dielectric nanoparticles. Sci. Rep. **12**(1), 21904 (2022)
11. Shalin, A.S., Sukhov, S.V.: Plasmonic nanostructures as accelerators for nanoparticles: optical nanocannon. Plasmonics **8**(2), 625–629 (2013)
12. Canós Valero, A., et al.: Nanovortex-Driven all-dielectric optical diffusion boosting and sorting concept for lab-on-a-chip platforms. Adv. Sci. **7**(11), 1903049 (2020)
13. Vestler, D., et al.: Circular dichroism enhancement in plasmonic nanorod metamaterials. Optics Express **26**(14), 17841–17848 (2018)
14. Kostina, N. A., et al.: Nanoscale tunable optical binding mediated by hyperbolic metamaterials. ACS Photonics **7**(2), 425–433 (2019)
15. Kucherik, A., et al.: Nano-antennas based on silicon-gold nanostructures. Sci. Reports **9**(1), 338 (2019)
16. Sharif, M.A.: Spatio-temporal modulation instability of surface plasmon polaritons in graphene-dielectric heterostructure. Physica E **105**, 174–181 (2019)
17. Sharif, M.A., Ashabi, K.: A Quasi-classical model for delineation of dynamical states and chaotic maps in a spaser. Plasmonics **16**(1), 97–105 (2021)

Numerical Solution of Mass Transfer Resistances Problem in an Electrolysis Process

Ever Peralta-Reyes[1], Iris C. Valdez-Dominguez[1], Alejandro Regalado-Méndez[1](✉), Reyna Natividad[2], Edson E. Robles-Gómez[1], Hugo Pérez-Pastenes[3], and Rubi Romero[2]

[1] Universidad del Mar, 70902 Puerto Ángel, Oaxaca, Mexico
alejandro.regalado33@gmail.com
[2] Universidad Autónoma del Estado de México, 50200 Toluca, Estado de México, Mexico
[3] Universidad Veracruzana, 96538 Coatzacoalcos, Veracruz, Mexico

Abstract. The aim of this work was to establish the effect of mass transfer resistances in an electrochemical reaction carried out on the surface of an electrode by using mathematical modeling. For this purpose, it was assumed that the reaction occurs instantaneously on the surface of the electrode. This mathematical model considers the transport phenomena in the Nernst diffusion layer. Numerical predictions of the model show that when the Biot number (Bi) is small, the external resistance to the mass transfer is large, while the amount of the electroactive species reaching the electrode's surface is small, which indicates that the process is controlled by mass transfer. In contrast, for large values of Biot number, the external resistance decreases, as the amount of electroactive species transported from bulk liquid to the electrode's surface is large, indicating that the process is controlled by the reaction on the electrode's surface. As a result of the application of this model, it has been found that the required dimensionless time (τ) to reach a steady state for Bi = [100 1.0 0.1] are 0.5, 1.5, and 3.0, respectively.

Keywords: Electrolysis process · Mass transfer resistances · Numerical solution

1 Introduction

Whenever a heterogeneous electrochemical reaction takes place, it is necessary to promote the contact between the reacting species and electrodes since it is widely acknowledged that heterogeneous reactions occur at the solution/electrode interface. This implies a constant species transport from the bulk to the reacting solution-electrode interface in order to ensure the feasibility and continuity of the electrochemical reaction. The transport rate of these species from solution to electrode or vice versa (when dealing with reaction products) depends on the concentration gradient through the diffusion layer and consequently on the parameters affecting both of them [1]. One factor that directly affects the concentration gradient in the diffusion layer is the reaction rate at the electrodes surface. Let us consider the following simple reaction,

$$O + ne^- \leftrightarrow R \tag{1}$$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 292–302, 2024.
https://doi.org/10.1007/978-3-031-53549-9_31

Reaction 1 (Eq. (1)) will only proceed if the "O" specie is transferred from bulk solution to the electrodes surface. This mass transfer step is inherent to any multiphase or electrochemical reaction mechanism and is governed by Fick's diffusion law [2] and Nernst-Planck equation [3]. Also, this mass transfer step may limit the overall reaction process whenever the reaction at the electrode surface is relatively fast [4]. Therefore, it is rather important to characterize and maximize such a mass transport phenomenon. To theoretically describe it, a simple one-direction electrochemical model for static solutions has been considered [5]. In this scenario, ion transport occurs by diffusion and migration so electrochemical electrode reactions only take place on the electrode/electrolytic solution interface.

In electrochemical systems where both mass transport mechanisms, convection and diffusion, are present, the electroactive solution is usually divided into two regions: (a) a diffusion layer (characterized by a δ thickness) constituted by a hypothetical completely stagnant liquid next to the electrode surface and where chemical species are transported only by diffusion and (b) relatively far away from this diffusion layer, a region where chemical species are strongly and homogeneously transported by convection towards the electrode, the bulk. This transport slows down as the electrode is approached and fully stops on the surface of the electrode itself [6]. Under these circumstances and when the occurrence of homogeneous reactions can be neglected, the mass balance for the electroactive species charge is represented by the accumulation term, convection transport, and migration transport due to electrical field, and by diffusion transport. The problem can be simplified if the number of electroactive species is assumed to be small, and also by adding an electrolytic support to diminish the electric field. In this manner, the transport by migration can be disregarded [7, 8], and the mass balance equation can be reduced to the accumulation and diffusion terms. This implies that the reaction is diffusion-controlled [9]. The reaction is controlled by charge if intensity of the applied current does not surpass intensity of the limited current. On the contrary, if the applied current is higher than limited current, then the process is mass transport-controlled [10, 11].

This study presents a mathematical model that allows the analysis of the effect of mass transfer resistances in an electrochemical reaction when the mass transfer at the diffusion layer controls the process, that is, when applied current surpasses the limited current.

2 Process Modeling

2.1 Mathematical Model

In this section a mathematical model that describes the dynamical behavior of the mass transport in the electrode-liquid interface is presented. Model constraints and boundary layers are also specified. The electrochemical reaction system under study is heterogeneous, and therefore homogeneous reactions in the bulk are not considered. An example of this type of systems is the in situ electrochemical production of hydrogen peroxide [12].

The area of analysis (electrode-liquid interface) is depicted in Fig. 1. As can be seen in this figure, the element of analysis consists of a stagnant Nernst diffusion layer with

thickness δ that develops at the electrode surface. Electro-active species diffuse through this layer to reach the electrode and react. It is worth pointing out that, for the theoretical analysis, it is assumed that the flow regime was already fully developed. This implies that the main parameters such as Nernst diffusion layer thickness [13] (δ), mass transfer coefficient (k_f) and limited current (IL), are uniform [14].

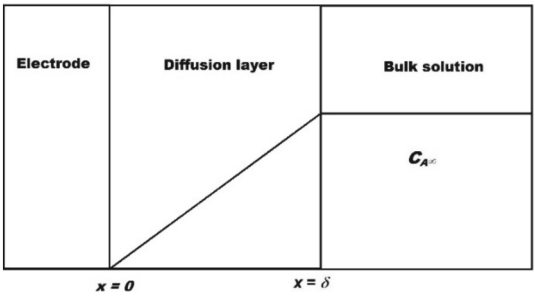


Fig. 1. Schematic representation of electrode-solution interface.

The electrochemical system without reaction in the bulk liquid is represented by the mass balance of the electroactive species (Eq. (2)) [4],

$$\frac{\partial C_A}{\partial t} = \nabla \cdot D\nabla C_A + z_A F \nabla \cdot (u_{mA} C_A \nabla \Phi) - \mathbf{u} \cdot \nabla C_A \tag{2}$$

where C_A is the concentration of the electroactive specie in the bulk, D is molecular diffusion, z_A is the charge of the electroactive species, F is the Faraday constant, u_{mA} is the ion mobility, Φ is the potential, t is the time and u is the convective velocity.

In this manner, the change in concentration of the electroactive species with respect to time is dependent upon diffusion (first term on the right), migration (second term to the right), and convection (last term). In consequence, the concentration of electroactive species will be affected by all operation parameters that impact the aforementioned transport terms. These operation parameters include applied current, electrolyte concentration, stirring speed and electrodes area. The following assumptions have been made to solve Eq. (2).

1. Solution contains an excess of support electrolyte and therefore migration term can be dismissed.
2. Bulk concentration of electroactive species is uniform.
3. Convection can be neglected at stagnant Nernst diffusion layer, and its thickness can be estimated from movement beyond the boundary layer.
4. Concentration of the electroactive species in diffusion layer is a function of one-dimensional molecular diffusion and time.
5. Physical properties remain unchanged.
6. Process is isothermal.
7. Reaction occurs instantaneously on the surface of the electrode.

When only the electrochemical zone or diffusion layer is considered, Eq. (2) can be reduced to a partial differential equation (Eq. (3)). In this zone, it is supposed that electroactive species concentration has a value ranging from concentration at the chemical or bulk zone and concentration at the electrodes surface [15]. Rate of mass exchange between electrochemical and bulk zones is obtained from equaling diffusion mass transport and convection mass transport (Eq. (6)).

$$\frac{\partial C_A}{\partial t} = D\frac{\partial^2 C_A}{\partial x^2}, \quad 0 \leq x \leq \delta \tag{3}$$

Equation (3) is subject to the following initial conditions,

$$t = 0, \quad C_A = C_{A\infty}, \quad 0 \leq x \leq \delta \tag{4}$$

The first boundary condition for Eq. (3) considers that electroactive species can react instantaneously on the electrode's surface,

$$x = 0, \quad C_A = 0, \quad \forall\, t \geq 0 \tag{5}$$

The second boundary condition, on diffusion layer's end, where diffusion mass transport is equal to convection mass transport, is,

$$x = \delta, \quad -D\frac{\partial C_A}{\partial x} = k_f (C_A - C_{A\infty}) \quad \forall\, t \geq 0 \tag{6}$$

Above equations describe the mass transport in the diffusion layer-fluid interface. Where $C_{A\infty}$ is the bulk concentration, x is the distance, k_f is the mass transfer coefficient and δ is the diffusion layer thickness. Before solving this equation, we define the following dimensionless variables.

2.2 Dimensional Equations

In order to formulate a general problem, first, it is convenient to describe a dynamic system in dimensionless form, which requires considering the following reduced variables,

$$U = \frac{C_A}{C_{A\infty}}, \quad \tau = \frac{tD}{\delta^2}, \quad \chi = \frac{x}{\delta} \tag{7}$$

where U is the dimensionless concentration, τ is the dimensionless time, and χ is the dimensionless distance.

Thus, the dimensionless form of the problem defined by Eq. (3) is given by the following differential equation,

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial \chi^2} \;\forall\; 0 \leq \chi \leq 1 \tag{8}$$

which is subject to the next initial and boundary conditions in dimensionless form (Eqs. (9)–(10)),

$$\tau = 0, \quad U = 1 \;\forall\; 0 \leq \chi \leq 1 \tag{9}$$

$$\chi = 0, \quad U = 0 \tag{10}$$

$$\chi = 1, \quad \frac{\partial U}{\partial \chi} = Bi(1 - U) \tag{11}$$

The dimensionless parameter known as mass Biot number (Bi) [16, 17], which appears on Eq. (11), is defined as follows,

$$Bi = \frac{\frac{\delta}{D}}{\frac{1}{k_f}} \tag{12}$$

Hence, the mathematical model that describes the mass transfer of electroactive species through the stagnant layer depicted in Fig. 1 is given by the set of Eqs. (8)–(11)) and requires a numerical solution. For this case, a difference finite method has been used.

3 Results and Discussions

This section presents the results of solving the electroactive species mass transport model through the diffusion layer in both dynamic and steady state conditions.

3.1 Steady State Solution

At steady state conditions, the analytical solution of the proposed model is given by Eq. (13). This equation allows establishing the electroactive species concentration profiles through the diffusion layer (Fig. 2). In this figure, the effect of Biot number is also included. Equation (13) predicts that electroactive species concentration at the electrodes surface is close to the concentration of that species in bulk solution as Biot number value increases, which might be due to the unconstrained migration of the electroactive species from bulk solution to the electrodes surface when external resistance to mass transport decreases. As long as Biot number is kept small, the number of electroactive species reaching the electrodes surface is also small because external resistance to mass transport is relatively large.

$$U = \frac{Bi\chi}{1 + Bi\chi} \tag{13}$$

3.2 Steady State Solution

The numerical solution was performed by a finite difference method in MATLAB® 2017a software package. Centered finite difference and forward finite difference were used for diffusion and accumulation terms, respectively.

The effect of Biot number variation on electroactive species concentration on electrodes surface for a transient-state condition is shown in Figs. 3, 4, 5 and 6. It can be noticed that, when Biot number is small, the number of electroactive species reaching

the electrode's surface is also small, which prolongs time until steady state is reached. On the contrary, when Biot number increases, that is, when resistances to external mass transport are small, the number of electroactive species that reach the electrode's surface approaches to the amount of that species in bulk solution in a short period. Steady state is reached easily when the Bi value is large, as shown in Fig. 7.

3.3 Discussion

Solution proposed by the mathematical model in steady state produces different concentration profiles for electroactive species for several Biot number values as shown in Fig. 2. It can be noted that, for large Biot number values (small mass transport resistances), diffusion layer is small, so the number of electroactive species that reaches the electrode's surface is large. This can be because of convection, which decreases diffusion layer thickness. It can also be observed in the same figure that a reduction in Biot number, which implies an increase in mass transport resistance, produces an increment in diffusion layer thickness. As Biot number decreases, the number of electroactive species that reaches the electrode's surface also diminishes.

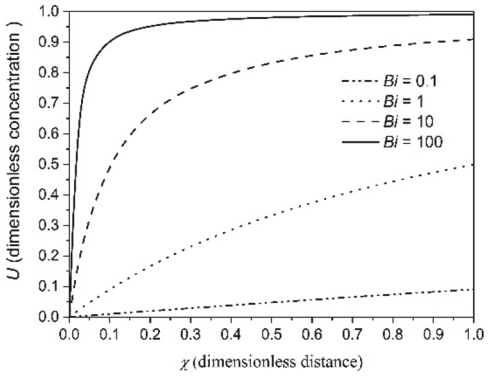


Fig. 2. Concentration profiles for several values of Biot's number (Bi) in steady state.

Solution of the mathematical model for transient-state conditions is depicted as concentration profiles of electroactive species versus distance for different Biot number values on Figs. 3, 4, 5 and 6. As can be seen, Biot number has a significant effect on each result. Figure 3 shows results obtained for Bi $= 0.1$. For $\tau = 0$ (before reaction takes place), concentration of the electroactive species at the electrode's surface is equal to that found in bulk solution. As reaction time increases, concentration profile changes in function of time (τ) and distance (χ), such as diffusion layer thickness (δ). For small time increases, δ is also small, and for large time increases, δ gets thicker. Steady state is reached after a long period (in this case, dimensionless time $\tau = 3$). As shown in Fig. 3, once steady state has been reached, the number of electroactive species that arrives to the electrode's surface is small.

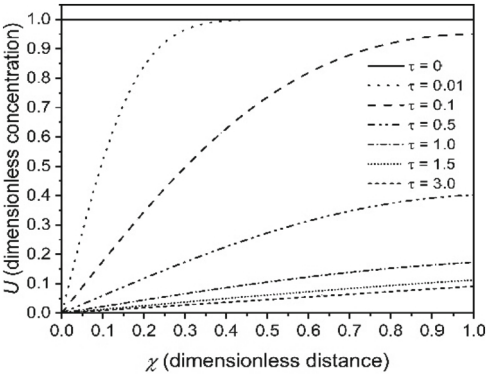


Fig. 3. Concentration profile for Bi = 0.1.

Concentration profile of electroactive species for Bi = 1 is shown in Fig. 4. For $\tau = 0$, concentration of the electroactive species at the electrode's surface is equal to that at the bulk solution. As time increases, concentration at the electrode's surface decreases gradually as reaction is taking place, whereas diffusion layer enlarges. For this Biot number value, it takes less time to reach steady state than for Bi = 0.1. This occurs when $\tau = 1.5$. In this moment, the number of electroactive species that comes into contact with electrode surface is higher than the one when Bi = 0.1.

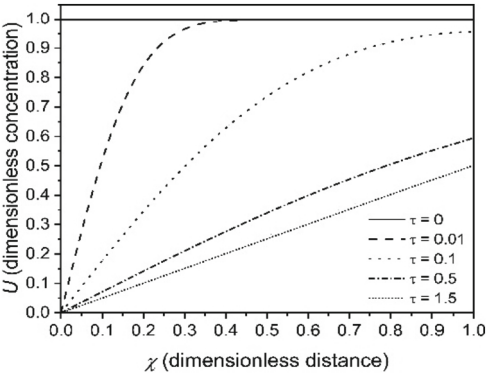


Fig. 4. Concentration profile for Bi = 1.0.

Figure 5 presents concentration profiles of the electroactive species for several reaction dimensionless times obtained when Bi = 10. When $\tau = 0$, the concentration of the electroactive species on the electrode's surface equals the concentration of that species at bulk solution. As reaction time increases, the concentration of the electroactive species on the electrode's surface diminishes due to reactions on electrode's surface until steady state is reached in a dimensionless time of 1, which is less than that obtained for Bi values of 0.1 and 1.0. For this time value, the amount of the electroactive species in contact with electrode's surface is higher since the mass transfer resistance is lower.

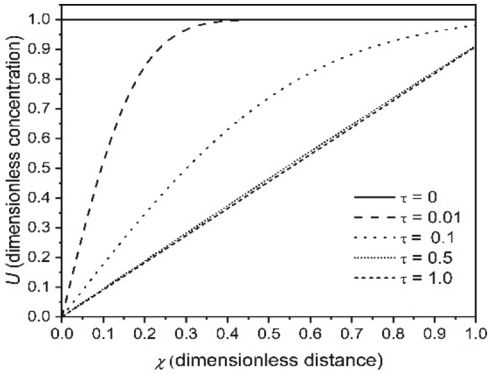


Fig. 5. Concentration profile for Bi = 10.0.

Results obtained for Bi = 100 are shown in Fig. 6. In this case, at the beginning of the electrochemical reaction, concentration of the electroactive species at the electrode's surface is also equal to the concentration at bulk solution, but as time goes on, this concentration decreases as electroactive species are consumed by reaction on the electrode's surface until steady state is reached at a dimensionless time of 0.5. This time is lower than that required when Bi = 0.1, 1.0 or 10.0, due to a higher convection. Consequently, diffusion layer decreases, which allows that electroactive species gets to the electrode more readily, replenishing the reaction, as shown in Fig. 2.

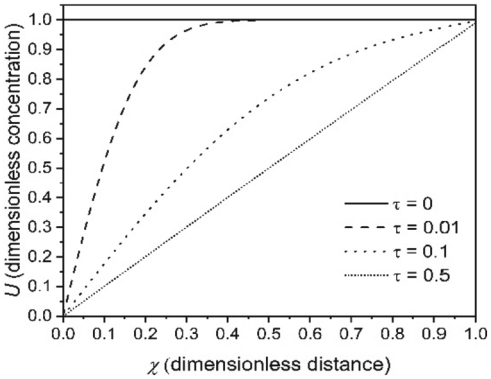


Fig. 6. Concentration profile for Bi = 100.0.

Results of the model allowed predicting that the effect of the mass transfer resistance is very significant. Thus, for large mass transfer resistances, the number of electroactive species that reaches the electrode's surface is small, and tends to be consumed fast, whereas it takes six-times longer to reach steady state when Bi = 0.1, three-times longer when Bi = 1, and two-times longer when Bi = 10 with respect to the situation when Bi = 100.

The dependence of τ with Bi is depicted in Fig. 7. It can be clearly seen that steady state is reached sooner for Bi larger values than for smaller Bi values.

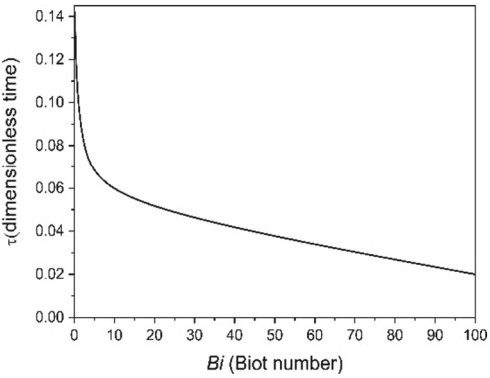


Fig. 7. Effect of Bi on τ.

Therefore, these results suggest that it is better to work in a turbulent regime since steady state is reached in a shorter period, which becomes important for electrochemical reactor operation. Furthermore, it is highly desirable in reactor design to reduce the effect of mass transfer resistances and to maintain concentration of the electroactive species at the electrode's surface as close as possible to the concentration of the species of interest at the bulk solution to prevent unwanted reactions.

4 Conclusions

A mathematical model was employed to obtain several concentration profiles of electroactive species on the diffusion layer (called the electrochemical zone) on the electrode's surface where electrochemical reactions were taking place. This model considers the effect of mass transfer resistances, which are important in mass transport-controlled reactions (diffusion). For these reactions, it is observed that, for large Biot number values (small mass transfer resistances), the amount of electroactive species reaching the electrode's surface is large, and is not depleted by reaction, whereas for small Biot number values (large mass transfer resistances), the amount of electroactive species that gets to the electrode's surface is small, and is depleted as soon as it arrives, which is not indicated to operate at industrial scale. On the other hand, the equation obtained from modelling can be used to determine the effect of mass transfer resistances in transient regime experiments, and to obtain plots of situations where transient state is important, especially when data collection of steady state experiments is necessary but time-consuming, in which case a transient solution might become an acceptable alternative because only one experiment is required.

Acknowledgements. This work has been financially supported by Universidad del Mar through the internal project (2021–2022) with CUP: 2II2104. Also, Alejandro Regalado, Ever Peralta, Reyna Natividad, Edson Robles, and Rubi Romero are grateful to CONAHCYT for the financial support through the Investigators National System (SNI) program.

Nomenclature

List of symbol

$Bi = \frac{k_f \delta}{D}$	Biot number, resistance to mass transfer by diffusion/resistance to mass transfer by convection (dimensionless)
C_A	Electroactive species concentration (mol/m^3)
$C_{A\infty}$	Electroactive species concentration at bulk solution (mol/m^3)
D	Diffusivity (m^2/s)
k_f	Mass transfer coefficient (m/s)
t	Time (s)
U	Dimensionless concentration of the electroactive species (dimensionless)
x	Distance (m)

Greek symbols

δ	Diffusion layer thickness (m)
τ	Dimensionless time (dimensionless)
χ	Dimensionless distance (dimensionless)

Subindex

f	Fluid
0	Initial

References

1. Pingarron, J.M., Sánchez Batanero, P.: Química Electroanalítica: Fundamentos y Aplicaciones, 1st edn. Síntesis, Madrid (1999)
2. Fick, A.V.: On liquid diffusion. London Edinburgh Dublin Philos. Mag. J. Sci. **10**(63), 30–39 (1855)
3. Gulati, R., Rudraraju, S.: Spatio-temporal modeling of saltatory conduction in neurons using Poisson–Nernst–Planck treatment and estimation of conduction velocity. Brain Multiphys. **4**, 100061 (2023)
4. Ciobanu, M., Wilburn, J.P., Krim, M.L., Cliffel, D.E.: Fundamentals. In: Zoski, C.G. (ed.) Handbook of Electrochemistry, pp. 3–29. Elsevier Science, Amsterdam (2007)
5. Faulkner, L.R.: Understanding electrochemistry: some distinctive concepts. J. Chem. Educ. **60**(4), 262–264 (1983)
6. Pletcher, D.: A First Course in Electrode Processes, 2nd edn. Royal Society of Chemistry, Cambridge (2009)

7. Deconinck, J.: The current distribution in electrochemical systems. In: Deconinck, J. (ed.) Current Distributions and Electrode Shape Changes in Electrochemical Systems. Lecture Notes in Engineering, vol. 75, pp. 1–55. Springer, Heidelberg (1992). https://doi.org/10.1007/978-3-642-84716-5_1
8. Newman, J., Balsara, N.P.: Electrochemical Systems, 4th edn. Wiley, New York (2021)
9. Wang, J.: Controlled-potential techniques. In: Wang, J. (ed.) Analytical Electrochemistry, pp. 67–114. Wiley, New York (2006)
10. Montilla, F., Michaud, P.A., Morallón, E., Vázquez, J.L., Comninellis, C.: Electrochemical oxidation of benzoic acid at boron-doped diamond electrodes. Electrochim. Acta **47**(21), 3509–3513 (2002)
11. Rodrigo, M.A., Michaud, P.A., Duo, I., Panizza, M., Cerisola, G., Comninellis, C.: Oxidation of 4-chlorophenol at boron-doped diamond electrode for wastewater treatment. J. Electrochem. Soc. **148**(5), D60 (2001)
12. Peralta, E., Natividad, R., Roa, G., Marín, R., Romero, R., Pavón, T.: A comparative study on the electrochemical production of h2o2 between bdd and graphite cathodes. Sustain. Environ. Res. **23**(4), 259–266 (2013)
13. Amatore, C., Szunerits, S., Thouin, L., Warkocz, J.S.: The real meaning of nernst's steady diffusion layer concept under non-forced hydrodynamic conditions. a simple model based on Levich's seminal view of convection. J. Electroanal. Chem. **500**(1–2), 62–70 (2001)
14. Vázquez, L., Alvarez-Gallegos, A., Sierra, F.Z., de León, C.P., Walsh, F.C.: CFD evaluation of internal manifold effects on mass transport distribution in a laboratory filter-press flow cell. J. Appl. Electrochem. **43**(4), 453–465 (2013)
15. Cañizares, P., García-Gómez, J., Lobato, J.A. Rodrigo, M.: Modeling of wastewater electrooxidation processes part I. General description and application to inactive electrodes. Ind. Eng. Chem. Res. **43**(9), 1915–1922 (2004)
16. Huang, J.C.: Effect of mass transfer resistance on electrochemical bioensors with asymetrical sandwirch membranes. Int. J. Polym. Mater. Polym. Biomater. **53**(7), 577–586 (2004)
17. Kanevce, G.H., Kanevce, L.P., Dulikravich, G.S.: Estimation of thermophysical properties of a drying body at high mass transfer BIoT number. In: International Symposium on Inverse Problems in Engineering Mechanics 2003 (ISIP 2003), pp. 13–20. Elsevier Science, Japan (2003)

Trust in Electronic Record Management System: Insights from Islamic-Based Professional and Moral Engagement-Based Digital Archive

Miftachul Huda[1(✉)], Reda Owis Hassan Serour[2], Mukhamad Hadi Musolin[2], Mohd Azman[3], Andi Muhammad Yauri[4], Abu Bakar[4], Muhammad Zuhri[4], Mujahidin[4], and Uswatun Hasanah[4]

[1] Universiti Pendidikan Sultan Idris, Tanjong Malim, Malaysia
halimelhuda@gmail.com
[2] Sultan Abdul Halim Mu'adzam Shah International Islamic University, Kuala Ketil, Malaysia
[3] Sultan Ibrahim Johor Islamic University College, Johor Bahru, Malaysia
[4] State Institute for Islamic Studies, Bone, South Sulawesi, Indonesia

Abstract. This paper aims to examine the strategic attempt on empowering the trust system in recordkeeping management. The further assessment was conducted in looking into detail about the critical phase in resulting to the trust system in record arrangement as the initiative to sustain information stability. The critical review from the recent related literatures was employed from the data base engine, google scholar. The finding revealed that strategic attempts on empowering trust system in recordkeeping management are in line with building the digital archive with professional and moral engagement. The further details are pointed out with the following three phases. Those include Enhancing emerging trust as strategic discipline for electronic records management system (ERMS), empowering professional and ethical balance for digital-based recordkeeping responsibilities in ERMS, strengthening social and personal development on professional and ethical balance in ERMS.

Keywords: Trust System · Recordkeeping Management · Electronic Records Management System (ERMS) · Critical Insights · Islamic-based Professional and Moral Strategic Engagement · Digital Archive

1 Introduction

In the past few decades, the particular extent of responsibility awareness of examining the structure of archival record in focusing on the security levels in detail plays a key role in giving insights into the wide context of records management [36, 37]. In this view, the record types should be conceived in providing the strategic system of recordkeeping initiative involved with the training end users set up into the context of organisation structure to work with the parts of the records plan in underlying the archival records [1]. Moreover, the wide range of file structures and the record types with other big task in the system is focused on the functional levels of the record plan in the context of

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 303–315, 2024.
https://doi.org/10.1007/978-3-031-53549-9_32

organisation in the sense that would require the whole records in the reworking each time of plan [38, 39]. In the attempts to incorporate this achievement referring to the planning and implementing of recordkeeping, the functional basis in facilitating the management concerned with creating the quality performance of archival records has to be maintained into the setting of functions in providing the facilities management to tackle of issues on records management [2]. In particular, the case of physical resources in the context of engineering issues refers to enhance the accommodation attempt with managing the space planning based on making the records plan concerned into the particular physical objects [40, 41]. Towards the point of view, regulating the professionalism engaged into the ethical concern is necessary to expand the facilities with empowerment into recordkeeping in the sense that can determine the records plan maintained into the extent of function to build the committed strength in linking the extensive basis of professional and ethical balance to have personalised cooperation [42, 43].

Through adapting the awareness quality on social and personal engagement with appropriate media technology, attempts to take account as the way in preventing negative impact in records from lack of authenticity to give insightful influences should manage in providing the extensive guidelines considered to enhance the positive feedback through empowering recordkeeping initiatives [3]. In particular, the extensive inquiry in considering moral dimensions to give the feedback in underlying the current practice of recordkeeping appropriately within the current situation needs to provide an entire platform in providing the clear understanding about the importance of archival records amidst the emerging trends of technological advancement with its significances [44–46]. As a result, the theoretical based proposed guideline to give the feedback with entire reflection in making the important goal of records management become more reflective with ethical regulations which may become more widely valuable in recognizing and articulating the practice amidst the participants and communicators [4]. It is necessary to note that important quality to address the practical stage to conduct appropriately with ethical concern and professional skills plays an entire integrity to let the process run wisely referring to the point of value in the visible features on records management landscape.

In addition, the proper arrangement is required to elaborate the value on building professional and moral sufficiency reflected into the considering the strategic attempts on on responsible awareness on records management with a proper manner [47–49]. This initiative refers to enhance the very important quality in underlying the recordkeeping along through implementing the systemic approach in incorporating the records management [5]. Attempts to accomplish the task in providing the accountability in information sharing through this initiative would need to entirely take a benefit of the feedback appropriately from technology innovation [50–52]. In dealing with the key instruments to promoting the recordkeeping with its significant role, it have a significant impact on contemporary life styles as increasing number of people spend valuable amount of time on watching television, playing records, listening to radio, reading electronic newspapers and magazines [6]. With engaging the sufficient and effective skills to enhance the recordkeeping initiative to apply for the information data, the variety of outstanding performance value such as critical thinking skills with professional enhancement and ethical engagement to develop the wide range of potential value to underlie the records

management becomes an entirely effective way to determine the sources and abundant issues among the users to solve in particular engagement [53–55].

2 Literature Review

The essences of emerging trust in recordkeeping management along with the recent development of technological side yielded the featured management in professional and ethical balance engagement [56, 57]. Such orientation is facilitated widely in expanding the interaction basis through the extent of experiential learning basis requires a further setting up amongst the users in the digital recordkeeping. The professional skills in the records management has to be involved with use of media technology in the context of utilizing the number of appropriate methods in giving insights into the instruction which necessitates the pivotal guidance appropriately with engaging the morally-based enhancement in considering the wide expansion on the technology adoption in recordkeeping arrangement with its beneficial value [7, 58, 59]. In particular, the quality sense of determining the inclusiveness about the technology skills in transmitting the experiential-based learning in the recordkeeping initiative should be more widely engaged into the users' records in terms of building both recreation and entertainment basis [60, 61]. With this regard, the particular inquiry on the technology integrated skills in recordkeeping compliance, for instance, has to do with the commercial and recreational devices linked into the learning materials [8]. As a result, it could create the adoption in various communication applications of like web based learning, for instance, has to do usually with the entertainment which media can support the growth and learning process in providing the extent of powerful models in the technology adoption within the cultural orientation [9].

Considered to be the wide characteristics to let the quality of interpersonal relationships, the mediated systems in the collaborative initiation within the recordkeeping contexts should provide the strategic mechanisms to foster the quality of interpersonal trust within the relationship between the process and maintenance of records [10]. In particular, the commitment of ties towards the number of files contained in the system should be strengthened in particular through building the structural integrity within the organizational setting [62, 63]. The context of consistency to develop the quality of information set in the document file should bring along with situating the trust in demonstrating the extent of reliability over time in enabling the users to be more consistently in the regards to providing the access of information data quickly [11]. In terms of elaborating the reputation systems aggregated into the quick access in the particular data among the users, the extent of information in the behavioural substance should be appropriately explored with referring to the consequence in giving the value about the quality of access [12]. As a result of accumulating the quality in exploring the wide systems of emerging trust, this potential dynamic in ensuring the activities together with enhancing the extent of motivations among the stakeholders would give a clear picture in fostering the trust quality within the systemic approach into the greater insights of trust commitment.

In the attempts to allow the users in both schools and larger society level, the number of ways which expand in transmitting the technology acceptance model to adopt amidst

the people's daily activities in the basis of track record keeping management system [13]. It is important to point out the valuable insights into the skills development of technology use included into the digital records activities in the sense that can be determined into the trend of technology support emerged into widely variable applications [14]. In terms of the settings expanded into the professional skills with moral enhancement, the potency to shape users' skills and experience on records management should be managed in the technology application within the influential feedback set out from the stage of skills acquisition in the process to develop in promoting such initiative within the organisational culture orientation context [15]. Along with the growing-up on the advancement of technology, the experiential basis of acceptable and natural part in the exposure of record keeping initiative referring to the number of application on the digital system provided amidst home and community settings in the ways, which are worthwhile to accord the high value to information as evidence.

3 Analysis and Discussion

3.1 Enhancing Emerging Trust as Strategic Discipline for Electronic Records Management System (ERMS)

The emerging trust has the potential value in giving the accuracy of information involved in ensuring the quality and reliability [16]. In the context of the recordkeeping management, attempts to build the frameworks with archiving the information data whether it is from the document and digital version would need to consider the engagement of accountability among the users in fostering recordkeeping process, both in organization and amidst the public or government sector [64, 65]. The need to maintain the emerging trust determined to the key performance value should bring along with addressing the concern about societal and cultural heritage issues [17]. In terms of the data collection along with documenting the skilful scenarios required to work with collecting the information data, the accurate consent should be prioritised to verify in recollecting both innovations and implications in shaping the human experience increasingly affect media entertainment and information for connectivity and networking [18]. In particular, the scenario of crowdsourcing assigned into the participatory sensing to enable the users to collect appropriately about the data whether they are in both document file and digital version refers to enhance the integrity along with getting the beneficial value [66, 67]. In this digital era, the records management should be maintained with the flexibility to engage into the societal level at large in the sense that could give feedback to the decision making mainly to the global community to access.

In line with the attempts to help foster emerging trust in the context of records management, providing the groups with a greater insight in understanding the organization need based on their demand which the stakeholders can perform has to do with maintain the focus on the way of the users to have the significant view in seeing the phenomena mainly in the digital context [19]. With inculcating the perceptions to see the system with its dynamics to ensure the accessibility during the use of systems, the particular characteristic of trustworthiness could be entangled with pointing out the interpersonal quality along with integrating the technology adoption into the records management [20]. As the key role to deliver the initial value of trustworthiness, the information quality with its

significant role to expand the goodness enlarging the technical property systems aligned into the platform in focusing the essence of development of standards on security and privacy arrangement [68, 69]. Regarded as key to records management, the wide system of trustworthiness should take into consideration in supplying the identity together with the personal information.

3.2 Empowering Professional and Ethical Balance for Digital-Based Recordkeeping Responsibilities in Electronic Records Management System (ERMS)

The exposure on engaging the professional skills and ethical values in the basis of record management should bring along with adapting the promotion in underpinning the wide range of recognitions from technology use. As a result, the record management would need to expand in deepening the potentials about the trustworthiness in ensuring the process to run well with incorporating the integrity to give insightful value to enhance the recordkeeping discipline [21]. In this view, attempts to incorporate the system of records management refer to enhance the connection between the capability of users and information process which can be transmitted in managing the professional skills appropriately within the instructional standard [70, 71]. In this view, the particular guidance in giving insightful view to determine the point of value has to point out the essence of adaptive potentials in the way which can be maintained within the analysis process to recognise the entire structure conducted in the recordkeeping [22]. Conceiving the particular structures in records management, the necessity to transmit the professional skills in the way to apply it appropriately into the comprehensive basis of the guidelines has to be strengthened with underpinning the metadata analysis.

The attempts to elaborate the exercise among the users through the activity history data in meeting their daily calorie burning goals would need to anticipate the worldwide of shortfall which can be transformed into the activity when the number of users could have chance to reach their aims [23]. Moreover, the way to enable them in getting accessed through gauging the wide range of trustworthiness in supplying the quality of consistency about the data sources to enhance the level of trustworthiness overall in determining the system with its prerequisite to face the number of challenges could be set up at the minimum level [24, 72, 73]. In particular, the management of getting access into the number of list of data sources determined to exert in the system's calculations needs to embark the formalized vetting process in developing the recommender system for data sources to become more helpful to gain the insightful value in ensuring the trustworthiness [74, 75]. This wide performance could be engaged with associating the listed sources already recorded in the system within the dominant power structures, in the sense that can maintains the characteristics like dominant, oppositional and negotiated readings of the media [25, 76, 77]. Attempts to creating the records management through strengthening the production of social media aims at giving the insights among the users in creating their own understanding of recordkeeping initiation [25]. It is essential to expand the purposeful relation in determining the opportunities and meeting the challenges to solve appropriately with the truth in living the conditions of individuals, groups and broader society.

3.3 Strengthening Social and Personal Development on Professional and Ethical Balance in Electronic Records Management System (ERMS)

The wide attempts to enhance the particular initiation in records management are necessary to point out the valuable insights of social and personal concern in the sense that can be determined with professional and ethical engagement [78–80]. In this regard, the wide range of purposeful concern in the way to expand the skills and values to give insights into fostering the recordkeeping management should be taken into consideration with adapting the professional skills and ethical values in both digital version and document basis [26]. In particular, the worldwide of pervasive growth of technology enhancement leading to the managerial competitiveness refers to enhance the initiative aims with instructional guidance to carry out the appropriately ethical considerations in strengthening the awareness of social and personal basis to deliver the number of information assigned into the knowledge understanding in records management with its beneficial significance [81–83]. Moreover, providing the creative ability engaged into the awareness in strengthening the quality level transformed in disseminating the recordkeeping skills between social and personal development is widely the pivotal role to enhance in maximizing their capacity building on the implementation stage [84–86]. With this regard, the quality extension of the combination between professional and ethical engagement would need to provide the entire approach to deliver knowledge understanding to be shared into the electronic basis.

With the common sense to disseminating the knowledge to focus on complementing the process of technology use appropriately and wisely referring to ensuring the positive role of recordkeeping through providing the archival access amidst the contemporary society [87–89]. As a result of additional complementary concern to inform the users to fully support the users to applicate the records management basis, attempts to strengthen attitudes with professional ability should be embedded with ethical engagement through continuous interaction with adoption process in the recordkeeping [28]. Moreover, the proposed approach in addressing the use and customization of innovative technology in the records management is widely more systematic and robust with enhancing the quality extension of its attractiveness to give the appropriate solution towards some challenges which could be solved [90–92]. The expansion of initial value in solving some behavioural issues should bring along with enhancing the wide collective adaption on the diffusion of digital recordkeeping, particularly in promoting the entire strength to enable the worldwide with the adaptive use skills [29]. Thus, it could be adopted into the context of digital recordkeeping management through addressing feedback to get benefit and analysing the challenges to have tackled in solving.

4 Implications and Future Directions

Regarded to include the essential component in the recordkeeping initiation, professional skills are needed to align with the basis of information culture associated with the flows of effective communication [30, 93, 94]. In this view, the partnerships circumstance in cross-organizational context is widely the engagement along with working in the cooperative practices, which should be aligned into the open access aligned into the relevant information referring to the need and demand [31]. As a result of setting up

the information data with the particular systems of recordkeeping management, the various stages of strategy, and clear guidelines and documentation in accordance with illustrating the significance of trustworthiness could be oriented to run the archival records into working collaboratively in creating and managing information quality with being committed to the management of trust [32, 95]. For the information data, the wide characteristic which can be managed into the more recently point of value in giving the forefront to adapt amidst organizational context [96, 97]. This is important to consider in talking the challenges of mistrust, for instance, in order to ensure the limit of characterized and differentiated information in the context of cultural orientation [33].

In addition, the professional skills in managing the records management with strengthening the potential value of traceability in ensuring the quality assurance of information data transmitted into the archival records should point out ultimately forcing trustworthiness which can be recommended to perform the reliance to possess the accuracy of data sources [34]. In order to make sure in guarantee about the accuracy of data source through filtering the pollution data, the attempts to enhance the initiative of aggregating process derived from the wide range of numbers of data sources should be managed in a particular way to get the less problematic on the inaccurate information [35]. In particular, the quality of information data in transmitting the wide accuracy to make the system trustworthy enough has to be involved into estimating the number of sources possibly needed with the record-keeping initiation [98, 99]. As a result of the entire attempts in the collaboration of the number of steps undertaken, the professional activity needs to significantly engage in recommending the necessary exercise enough to meet the demand on the accuracy of information data [35]. It is necessary to note that important quality to address the practical stage to conduct appropriately with ethical concern and professional skills plays an entire integrity to let the process run wisely referring to the point of value in the visible features on records management landscape.

5 Conclusion

In the attempts to respond the urgent demand about the challenging issues, the initiative on gaining the careful engagement on adopting the required procedures and practices in the records management should be paid a particular attention on considering the systematic arrangement. It means that the strategic enhancement on controlling process about the information archive basis is required to address both effectivity and efficiency in generating the trust in underlying the digital records. In order to have a serious concern in empowering the proper application on the records of information, the transmission process needs to enhance the proper manner in elaborating the strategic initiative of both professional and moral engagement as the values in making a balance in generating it into the reality. As a result, the significant instrument in guiding process with an insightful value is strategically enhanced to deliver in undertaking the arranged procedure on systematic control. The main point of this study revealed that strategic attempts on empowering trust system in recordkeeping management are in line with building the digital archive with professional and moral engagement. The further details are pointed out with the following three phases. Those include Enhancing emerging trust as strategic discipline for electronic records management system (ERMS), empowering professional and ethical balance for digital-based recordkeeping responsibilities in ERMS,

strengthening social and personal development on professional and ethical balance in ERMS.

References

1. Ashok, M., Madan, R., Joha, A., Sivarajah, U.: Ethical framework for artificial intelligence and digital technologies. Int. J. Inf. Manage. **62**, 102433 (2022)
2. Bellini, E., Iraqi, Y., Damiani, E.: Blockchain-based distributed trust and reputation management systems: a survey. IEEE Access **8**, 21127–21151 (2020)
3. Berdik, D., Otoum, S., Schmidt, N., Porter, D., Jararweh, Y.: A survey on blockchain for information systems management and security. Inf. Process. Manage. **58**(1), 102397 (2021)
4. Biswas, M., Md. Niamat Ullah Akhund, T., Kaiser, M.S.: DRLAS: digital record keeping in land administration system relying on blockchain. In: Yang, X.S., Sherratt, S., Dey, N., Joshi, A. (eds.) Proceedings of Sixth International Congress on Information and Communication Technology. LNNS, vol. 216, pp. 965–973. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-1781-2_83
5. Cerchione, R., Centobelli, P., Riccio, E., Abbate, S., Oropallo, E.: Blockchain's coming to hospital to digitalize healthcare services: designing a distributed electronic health record ecosystem. Technovation **120**, 102480 (2023)
6. Duranti, L., Rogers, C.: Trust in digital records: an increasingly cloudy legal area. Comput. Law Secur. Rev. **28**(5), 522–531 (2012)
7. El-Amin, A., Marks, Q.L.: Organizational stakeholder management develops intentional education ethos. In: Elevating Intentional Education Practice in Graduate Programs, pp. 241–252. IGI Global (2023)
8. Farazmand, A. (ed.): Global Encyclopedia of Public Administration, Public Policy, and Governance. Springer, Cham (2023)
9. Forster, D., Maxwell, B.: Using codes of professional ethics and conduct in teacher education: pitfalls and best practice. In: Eaton, S.E., Khan, Z.R. (eds.) Ethics and Integrity in Teacher Education. Ethics and Integrity in Educational Contexts, vol. 3, pp. 25–42. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-16922-9_3
10. Foscarini, F., Krucker, M., Golick, D.: Meeting technologies and recordkeeping: a preliminary study. Rec. Manag. J. **32**(1), 96–109 (2022)
11. Frings-Hessami, V., Oliver, G.: Recordkeeping culture in Switzerland: the impact of language and communication on the success of recordkeeping initiatives. Rec. Manage. J. (2022)
12. Graskemper, J.P.: Professional Responsibility in Dentistry: A Practical Guide to Law and Ethics. Wiley, Hoboken (2023)
13. Hofman, D., Katuu, S.: Law and record keeping: a tale of four African countries. In: Managing Digital Records in Africa. Taylor & Francis (2023)
14. Huda, M.: Empowering professional and ethical balance in digital record management. Organ. Cybersecur. J.: Pract. Process People **2**(1), 60–73 (2022)
15. Huda, M., Rosman, A.S., Mohamed, A.K., Marni, N.: Empowering adaptive learning technology: practical insights from social network site (SNS). In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) CoMeSySo 2021. LNNS, vol. 232, pp. 410–420. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-90318-3_34
16. Jo, E. S., Gebru, T.: Lessons from archives: strategies for collecting sociocultural data in machine learning. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp. 306–316 (2020)
17. Katuu, S.: Tapestry of the education and training landscape for archives and records management in Africa. In: Managing Digital Records in Africa. Taylor & Francis (2023)

18. Keshta, I., Odeh, A.: Security and privacy of electronic health records: concerns and challenges. Egypt. Inform. J. **22**(2), 177–183 (2021)
19. Luepker, E.T.: Record Keeping in Psychotherapy and Counseling: Ethics, Practice and Supervision. Routledge (2022)
20. Lustgarten, S.D., Garrison, Y.L., Sinnard, M.T., Flynn, A.W.: Digital privacy in mental healthcare: current issues and recommendations for technology use. Curr. Opin. Psychol. **36**, 25–31 (2020)
21. Malhan, A., Manuj, I., Pelton, L., Pavur, R.: Electronic health records using a resource advantage theory perspective: an interdisciplinary literature review. Rec. Manag. J. **32**(2), 126–150 (2022)
22. Mendyk, R.W., Block, J.: Record keeping as an aid to captive care. In: Warwick, C., Arena, P.C., Burghardt, G.M. (eds.) Health and Welfare of Captive Reptiles, pp. 535–559. Springer, Cham (2023). https://doi.org/10.1007/978-3-030-86012-7_17
23. Modiba, M.: User perception on the utilisation of artificial intelligence for the management of records at the council for scientific and industrial research.Collect. Curat. (ahead-of-print) (2023)
24. Mukred, M., Yusof, Z.M., Al-Moallemi, W.A., Mokhtar, U.A.A., Hawash, B.: Electronic records management systems and the competency of educational institutions: evidence from Yemen. Inf. Dev. **38**(1), 125–148 (2022)
25. Ngoepe, M., Jacobs, L., Mojapelo, M.: Inclusion of digital records in the archives and records management curricula in a comprehensive open distance e-learning environment. Inf. Dev. 02666669221081812 (2022)
26. Örtegren, A.: Digital citizenship and professional digital competence—Swedish subject teacher education in a postdigital era. Postdigit. Sci. Educ. **4**(2), 467–493 (2022)
27. Rahmatullah, A.S., Mulyasa, E., Syahrani, S., Pongpalilu, F., Putri, R.E.: Digital era 4.0: the contribution to education and student psychology. Linguist. Cult. Rev. **6**(S3), 89–107 (2022)
28. Subramanian, M., Shanmuga Vadivel, K., Hatamleh, W.A., Alnuaim, A.A., Abdelhady, M., Sathiskumar, V.E.: The role of contemporary digital tools and technologies in Covid-19 crisis: an exploratory analysis. Expert. Syst. **39**(6), e12834 (2022)
29. Tiwari, S., Dhanda, N., Dev, H.: A real time secured medical management system based on blockchain and internet of things. Meas.: Sens. **25**, 100630 (2023)
30. Tseng, E., Sabet, M., Bellini, R., Sodhi, H.K., Ristenpart, T., Dell, N.: Care infrastructures for digital security in intimate partner violence. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, pp. 1–20 (2022)
31. Uunona, G.N., Goosen, L.: Leveraging ethical standards in artificial intelligence technologies: a guideline for responsible teaching and learning applications. In: Handbook of Research on Instructional Technologies in Health Education and Allied Disciplines, pp. 310–330. IGI Global (2023)
32. Van de Poel, I., Royakkers, L.: Ethics, Technology, and Engineering: An Introduction. Wiley, Hoboken (2023)
33. Wendelken, S.: Pandemic recordkeeping–the New Zealand experience. Rec. Manage. J. (2023)
34. Xie, S.L., Siyi, L., Han, R.: Competing with artificial intelligence–can the records and information management profession withstand the challenge?. Rec. Manage. J. (2022)
35. Zhang, D.: Design of information teaching management system integrating association rule mining algorithm. In: Jan, M.A., Khan, F. (eds.) BigIoT-EDU 2022, Part I. LNICST, vol. 465, pp. 439–445. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-23950-2_48
36. Huda, M., et al.: From digital ethics to digital partnership skills: driving a safety strategy to expand the digital community? In: Digital Transformation for Business and Society: Contemporary Issues and Applications in Asia. Taylor & Francis, UK (2023a)

37. Huda, M., et al.: Towards digital servant leadership for organisational stability: driving processes in the pandemic age? In: Digital Transformation for Business and Society: Contemporary Issues and Applications in Asia. Taylor & Francis, UK (2023b)
38. Zainuri, A., Huda, M.: Empowering cooperative teamwork for community service sustainability: insights from service learning. Sustainability **15**(5), 4551 (2023)
39. Huda, M.: Trust as a key element for quality communication and information management: insights into developing safe cyber-organisational sustainability. Int. J. Organ. Anal. (2023). https://doi.org/10.1108/IJOA-12-2022-3532
40. Syofiarti, S., Saputra, R., Lahmi, A., Rahmi, R.: The use of audiovisual media in learning and its impact on learning outcomes of islamic cultural history at Madrasah Tsanawiyah Negeri 4 Pasaman. Firdaus J. **1**(1), 36–44 (2021). https://doi.org/10.37134/firdaus.vol1.1.4.2021
41. Huda, M.: Towards digital access during pandemic age: better learning service or adaptation struggling? Foresight **25**(1), 82–107 (2023). https://doi.org/10.1108/FS-09-2021-0184
42. Huda, M.: Digital marketplace for tourism resilience in the pandemic age: voices from budget hotel customers. Int. J. Organ. Anal. **31**(1), 149–167 (2023). https://doi.org/10.1108/IJOA-10-2021-2987
43. Zainul, Z., Rasyid, A., Nasrul, W.: Distance learning model for Islamic religious education subjects in non internet server provider (ISP) areas. Firdaus J. **1**(1), 45–53 (2021). https://doi.org/10.37134/firdaus.vol1.1.5.2021
44. Huda, M.: Between accessibility and adaptability of digital platform: investigating learners' perspectives on digital learning infrastructure. High. Educ. Skills Work-Based Learn. (2023). https://doi.org/10.1108/HESWBL-03-2022-0069
45. Huda, M., et al.: Strategic role of trust in digital communication: Critical insights into building organizational sustainability. In: Arai, K. (ed.) FTC 2023. LNNS, vol. 815, pp. 387–403. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-47457-6_25
46. Abdul Aziz, N.A., Mohd Razali, F., Saari, C.Z.: Penggunaaan media sosial dari perspektif psiko spiritual Islam. Firdaus J. **2**(1), 65–75 (2022). https://doi.org/10.37134/firdaus.vol2.1.6.2022
47. Hanafi, H.F., et al.: A review of learner's model for programming in teaching and learning. J. Adv. Res. Appl. Sci. Eng. Technol. **33**(3), 169–184 (2024)
48. Huda, M., Bakar, A.: Culturally responsive and communicative teaching for multicultural integration: qualitative analysis from public secondary school. Qual. Res. J. (2024). https://doi.org/10.1108/QRJ-07-2023-0123
49. Syafri, N., Ali, A.H., Ramli, S.: Penggunaan kaedah Inovasi Sambung dan Baca Bahasa Arab (SaBBAr) dalam meningkatkan kemahiran membaca perkataan Bahasa Arab murid di Sekolah Rendah. Firdaus J. **2**(2), 62–71 (2022). https://doi.org/10.37134/firdaus.vol2.2.6.2022
50. Huda, M., et al.: Enhancing digital leadership direction: insight into empowering gender violence prevention. In: Mishra, D., Le Ngoc, A., McDowell, Z. (eds.) Communication Technology and Gender Violence. Signals and Communication Technology, pp. 147–164. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-45237-6_13
51. Ab Rahim, N.M.Z., Saari, Z., Mohamad, A.M., Rashid, M.H., Mohamad Norzilan, N.I.: Konsep Ulul Albab dalam Al-Quran dan hubungannya dengan pembelajaran kursus Sains, Teknologi dan Manusia di UTM Kuala Lumpur. Firdaus J. **2**(2), 72–78 (2022). https://doi.org/10.37134/firdaus.vol2.2.7.2022
52. Huda, M., et al.: Understanding of digital ethics for information trust: a critical insight into gender violence anticipation. In: Mishra, D., Le Ngoc, A., McDowell, Z. (eds.) Communication Technology and Gender Violence. Signals and Communication Technology, pp. 165–181. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-45237-6_14

53. Hanafi, H.F., Wahab, M.H.A., Selamat, A.Z., Masnan, A.H., Huda, M.: A systematic review of augmented reality in multimedia learning outcomes in education. In: Singh, M., Kang, D.-K., Lee, J.-H., Tiwary, U.S., Singh, D., Chung, W.-Y. (eds.) IHCI 2020, Part II. LNCS, vol. 12616, pp. 63–72. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68452-5_7
54. Khalili, Huda, M., Rosman, A.S., Mohamed, A.K., Marni, N.: Digital learning enhancement through social network site (SNS). In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) CoMeSySo 2021. LNNS, vol. 232, pp. 421–431. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-90318-3_35
55. Leh, F.C., Anduroh, A., Huda, M.: Level of knowledge, skills and attitude of trainee teachers on Web 2.0 applications in teaching geography in Malaysia schools. Heliyon **7**(12) (2021)
56. Sukadari, S., Huda, M.: Culture sustainability through co-curricular learning program: learning Batik cross review. Educ. Sci. **11**(11), 736 (2021)
57. Huda, M., Gusmian, I., Mulyo, M.T.: Towards eco-friendly responsibilities: Indonesia field school model cross review. J. Comp. Asian Dev. (JCAD) **18**(2), 1–12 (2021)
58. Rachman, A., Oktoviani, I., Manurung, P.: Digitization of Zakat and charity BAZNAS tangerang city through crowdfunding platform tangerangsedekah. Id. Firdaus J. **3**(1), 96–106 (2023). https://doi.org/10.37134/firdaus.vol3.1.9.2023
59. Huda, M., Hashim, A.: Towards professional and ethical balance: insights into application strategy on media literacy education. Kybernetes **51**(3), 1280–1300 (2022)
60. Mohamad Shokri, S.S., Salihan, S.: Modul pembangunan program Huffaz Profesional Universiti tenaga nasional: satu pemerhatian kepada konstruk pembangunan Al-Quran. Firdaus J. **3**(2), 12–23 (2023). https://doi.org/10.37134/firdaus.vol3.2.2.2023
61. Berlian, Z., Huda, M.: Reflecting culturally responsive and communicative teaching (CRCT) through partnership commitment. Educ. Sci. **12**(5), 295 (2022)
62. Huda, M.: Towards an adaptive ethics on social networking sites (SNS): a critical reflection. J. Inf. Commun. Ethics Soc. **20**(2), 273–290 (2022)
63. Omar, M.N.: Inovasi pengajaran & pemudahcaraan menggunakan Aplikasi Ezi-Maq (MAHARAT AL-QURAN) untuk menarik minat pelajar menguasai Ilmu Tajwid. Firdaus J. **2**(2), 79–87 (2022). https://doi.org/10.37134/firdaus.vol2.2.8.2022
64. Huda, M., Sutopo, L., Liberty, Febrianto, Mustafa, M.C.: Digital information transparency for cyber security: critical points in social media trends. In: Arai, K. (ed.) FICC 2022. LNNS, vol. 439, pp. 814–831. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98015-3_55
65. Borham, A.H., et al.: Information and communication ethics in social media for indigenous people's religious understanding: a critical review. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer, Cham (2024)
66. Ali, A.H.: Pelajar Khalifah profesional tempaan ulul albab sorotan penerapannya berasaskan komponen QEI di UPSI. Firdaus J. **3**(2), 51–63 (2023). https://doi.org/10.37134/firdaus.vol3.2.5.2023
67. Jusoh, A., Huda, M., Abdullah, R., Lee, N.: Development of digital heritage for archaeo-visit tourism resilience: evidences from E-lenggong web portal. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer, Cham (2024)
68. Rahim, M.M.A., Huda, M., Borham, A.H., Kasim, A.Y., Othman, M.S.: Managing information leadership for learning performance: an empirical study among public school educators. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer, Cham (2024)
69. Hehsan, A., et al.: Digital Muhadathah: framework model development for digital Arabic language learning. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer, Cham (2024)

70. Muharom, F., Farhan, M., Athoillah, S., Rozihan, R., Muflihin, A., Huda, M.: Digital technology skills for professional development: insights into quality instruction performance. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer, Cham (2024)
71. Muhamad, N., Huda, M., Hashim, A., Tabrani Z.A., Maarif, M.A.: Managing technology integration for teaching strategy: public school educators' beliefs and practices. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer, Cham (2024)
72. Wahid, A., et al.: Digital technology for indigenous people's knowledge acquisition process: insights from empirical literature analysis. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer, Cham (2024)
73. Huda, M., Taisin, J.N., Muhamad, M., Kiting, R., Yusuf, R.: Digital technology adoption for instruction aids: insight into teaching material content. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer, Cham (2024)
74. Alwi, A.S.Q., Ibrahim, R.: Isu terhadap Penggunaan teknologi media digital dalam kalangan guru pelatih jurusan Pendidikan Khas. Firdaus J. **2**(2), 88–93 (2022). https://doi.org/10.37134/firdaus.vol2.2.9.2022
75. Zamri, F.A., Muhamad, N., Huda, M., Hashim, A.: Social media adoption for digital learning innovation: insights into building learning support. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer, Cham (2024)
76. Huda, M., et al.: Digital record management in Islamic education institution: current trends on enhancing process and effectiveness through learning technology. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
77. Yahya, S.F., Othman, M.A.: Penggunaan video dalam Pengajaran dan Pembelajaran Pendidikan Moral Tingkatan 2. Firdaus J. **2**(2), 94–105 (2022). https://doi.org/10.37134/firdaus.vol2.2.10.2022
78. Huda, M., et al.: Empowering Islamic-based digital competence and skills: how to drive it into reconstructing safety strategy from gender violence. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
79. Cita Sari, D., et al.: Transformation of artificial intelligence in Islamic edu with Ulul Albab value (global challenge perespective). Firdaus J. **3**(1), 1–9 (2023). https://doi.org/10.37134/firdaus.vol3.1.1.2023
80. Huda, M., et al.: From digital ethics to digital community: an Islamic principle on strengthening safety strategy on information. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
81. Zamri, F.A., Muhamad, N., Huda, M.: Information and communication technology skills for instruction performance: beliefs and experiences from public school educators. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
82. Masud, A., Borham, A.H., Huda, M., Rahim, M.M.A., Husain, H.: Managing information quality for learning instruction: Insights from public administration officers' experiences and practices. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
83. Rahman, M.H.A., Jaafar, J., Huda, M.: Information and communication skills for higher learners competence model. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)

84. Mohd Nawi, M.Z.: Media variations in education in Malaysia: a 21st century paradigm. Firdaus J. **3**(1), 77–95 (2023)
85. Huda, M., et al.: From technology adaptation to technology adoption: An insight into public school administrative management. In: ICICT 2024. LNNS. Springer, Singapore (2024)
86. Tan, A.A., Huda, M., Rohim, M.A., Hassan, T.R.R., Ismail, A.: Chat GPT in supporting education instruction sector: an empirical literature review. In: ICICT 2024. LNNS. Springer, Singapore (2024)
87. Wahid, A., et al.: Augmented reality model in supporting instruction process: a critical review. In: ICICT 2024. LNNS. Springer, Singapore (2024)
88. Musolin, M.H., Serour, R.O.H., Hamid, S.A., Ismail, A., Huda, M., Rohim, M.A.: Developing personalized Islamic learning in digital age: pedagogical and technological integration for open learning resources (OLR). In: ICICT 2024. LNNS. Springer, Singapore (2024)
89. Musolin, M.H., Ismail, M.H., Huda, M., Hassan, T.R.R., Ismail, A.: Towards an Islamic education administration system: a critical contribution from technology adoption. In: ICICT 2024. LNNS. Springer, Singapore (2024)
90. Musolin, M.H., Ismail, M.H., Farhan, M., Rois, N., Huda, M., Rohim, M.A.: Understanding of artificial intelligence for Islamic education support and service: insights from empirical literature review. In: ICICT 2024. LNNS. Springer, Singapore (2024)
91. Susilowati, T., et al.: Getting parents involved in child's school: using attendance application system based on SMS gateway. Int. J. Eng. Technol. **7**(2.27), 167–174 (2018)
92. Susilowati, T., et al.: Learning application of Lampung language based on multimedia software. Int. J. Eng. Technol. **7**(2.27), 175–181 (2018)
93. Abadi, S., et al.: Application model of k-means clustering: insights into promotion strategy of vocational high school. Int. J. Eng. Technol. **7**(2.27), 182–187 (2018)
94. Aminudin, N., et al.: The family hope program using AHP method. Int. J. Eng. Technol. **7**(2.27), 188–193 (2018)
95. Wulandari, et al.: Design of library application system. Int. J. Eng. Technol. **7**(2.27), 199–204 (2018)
96. Aminudin, N., et al.: Higher education selection using simple additive weighting. Int. J. Eng. Technol. **7**(2.27), 211–217 (2018)
97. Huda, M., Teh, K.S.M.: Empowering professional and ethical competence on reflective teaching practice in digital era. In: Dikilitas, K., Mede, E., Atay D. (eds.) Mentorship Strategies in Teacher Education, pp. 136–152. IGI Global, Hershey (2018). https://doi.org/10.4018/978-1-5225-4050-2.ch007
98. Huda, M., Teh, K.S.M., Nor, N.H.M., Nor, M.B.M.: Transmitting leadership based civic responsibility: insights from service learning. Int. J. Ethics Syst. **34**(1), 20–31 (2018). https://doi.org/10.1108/IJOES-05-2017-0079
99. Huda, M., Mulyadi, D., Hananto, A.L., Nor Muhamad, N.H., Mat Teh, K.S., Don, A.G.: Empowering corporate social responsibility (CSR): insights from service learning. Soc. Responsib. J. **14**(4), 875–894 (2018)



Digital Record Management in Islamic Education Institution: Current Trends on Enhancing Process and Effectiveness Through Learning Technology

Miftachul Huda[1(✉)], Mukhamad Hadi Musolin[2], Reda Owis Hassan Serour[2], Mohd Azman[3], Andi Muhammad Yauri[4], Abu Bakar[4], Muhammad Zuhri[4], Mujahidin[4], and Uswatun Hasanah[4]

[1] Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Malaysia
halimelhuda@gmail.com
[2] Sultan Abdul Halim Mu'adzam Shah International Islamic University, Kuala Ketil, Malaysia
[3] Sultan Ibrahim Johor Islamic University College, Johor Bahru, Malaysia
[4] State Institute for Islamic Studies, Bone, South Sulawesi, Indonesia

Abstract. The significant role of technological advancement offered in the society has been emerged consistently into the education sector. It is the reality that the digital record management is being the transitional pathway in conveying the sources, materials and also data in the education purpose. This paper aims to examine the actual practice of digital record management strategically employed in underlying the education teaching and learning process. The further detail is to look into detail about the current trends on enhancing the practice in the context on developing the learning technology in achieving the effectiveness. The critical review from the recent literatures, including referred journals, books, chapters and conference proceeding related to the field was conducted. The finding revealed that the strategic attempts on understanding of digital record management in education sector are significantly proposed with building the enhancement of both process and effectiveness in achieving the learning objective attainment. The main point of this paper refers to give an insightful value on strengthening the process, procedure and pathway in underlying the technology-based digital learning in education sector.

Keywords: digital record management · education · current trend · process · effectiveness · learning technology

1 Introduction

The role play of record digitation referring to affecting wide range of information should be underlined in giving the value of technology use skills as the transforming bridge in archival records. The concern among the scholarly attention in this issue has to be involved with embarking with particular reference to records management in the sense

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 316–333, 2024.
https://doi.org/10.1007/978-3-031-53549-9_33

that can determine moral foundation engaged into the professional skills [39, 40]. In terms of proposing grounded foundation on professional and ethical balance, records management could be entirely explored within providing the positive benefits to give the value of moral based instructional suggestion [1]. As a result, it is necessary to look into the commitment aligned with professional and ethical concern needed to give an insightful value in the level of individuals and society to have linkages on records management in providing systematic essence to control over records through the wide range of kinds of approaches like the document of transactions on business processes which gives insights into the decision-making process in the organizational context [41, 42]. It is highly supplied to have the wide regulation to direct the organizations including public and private sector to determine the significance of records management skills among the workers and staffs [2, 43, 44]. In this view, the wide regulations would be maintained through the requirement on legislating the assurance effectively managed, in order to successfully protect themselves from litigation and public scrutiny.

Determined as the key role of fortunate characteristics in allowing the users to get access the information data as they need, the quality exposure of trust should be strengthened to obtain the easy way to perform in the context of reliability with lack of the single failure [45, 46]. In this regard, the necessary part of valuable insights in providing the significant guidelines in the records management process refers to implementing the systematic approach with reasonable expectation which the users can adopt in taking place of the existing part of body developed through working with improving the intelligibility of data systems [3, 47, 48]. As a result, attempts to applying further in supplying the performance quality within the records management remains the certain way to let the process run well in the systematic stage planned into the intelligibility [49, 50]. This aims to offer the route in increasing the emerging trust explicitly with explaining the behavioural substance among the users expected on the techniques in interfacing the design which enables them to have such acts with ethical quality concern [4, 51, 52]. In this view, the need to pay serious attention to have both professional and ethical balance should be taken into consideration to underlie the online interaction wisely and appropriately [53, 54].

With regards to underlining the procedural stage committed into the emerging trust, the trend of information data in the context of daily needs, media messages and also the extent of productions is entirely the valuable insight to embark in providing the level of enhancing creative and captivating innovation [55, 56]. In ensuring the shape of the record which the users need based on the efficient and effectiveness to access, the quality of networking in underlining the trustworthy used in the records management in the way which is creative and critical to potentially result in engaging the services with increasing the content in the media platforms [5]. As the key insights into the records management, the behavioural substance emerged into the interpersonal quality in implementing the systems should be widely strengthened to enhance the affection in the context of individuals, organizations,and societies [57, 58]. Towards the attribution of hyper-connections in the digital records, building the worthiness in the communication style for instance might also play a significant role in enabling the process to run well with the effective medium in the attempts to take handling the issues about the archives [6]. In particular, the issues of lack of information quality would be the essential feature

in ensuring the factorisation about the archival records with adopting the technological development system. It is potential to promote the focus on the engagement process with the extent of distractive common practice among the users in addressing the variety of records management.

2 Literature Review

2.1 Digital Record Management in Islamic Education Sector

The strategic line in engaging the essential value of functioning the moral aspect refers to give an ultimate pathway in isolating of understanding the benefit of history records about any activity among the users. The various ways to recognise the platform of social media's information data for instance would need to expand the literacy skills and archival records expertise in enhancing the number of outstanding skills subject [59, 60]. The clear example could be seen such as language arts, social studies, health, science, and other subjects in ensuring the wide ability to get access in the way to examine and evaluate the complex information with educational standards of nations manifested in enabling the users to gain the better understanding [7, 61]. It aims at engaging the sufficient and effective skills to enhance the recordkeeping initiative to apply for the information data. On this view, the variety of outstanding performance value such as critical thinking skills with professional enhancement and ethical engagement to develop the wide range of potential value to underlie the records management becomes an entirely effective way to determine the sources and abundant issues among the users to solve in particular [8, 62, 63]. With this regard, the potential value to determine information data refers to give insights into the users to get accessibility assigned with taking place about the number of conditions.

In particular, the framework in reflecting the superficial level with featuring the unique into the particular organisation should be implied the wide range of characteristics in the trust-based information architecture, which can be managed amidst the organisational information systems [64, 65]. In terms of the entire attempts to make the assistance of application stage in the recordkeeping management, the quality of ethical engagement in directing the process wisely with less misconduct in the basis of digital identities in helping the appropriate recognizance of the way to draw the reference with the wide range of records management [9]. It is important to point out considering the potential value in fostering the positive feedback to meet the challenges with increasing the awareness of potential issues. It is necessary to enhance the strategic attempts to highlighting the particular effects possibly occurred when the raise of records management in addressing the consequence to propose the theoretical guideline in managing the potential value of professional and ethical balance transmitted into the instructions in recordkeeping [10]. As a result, the point of view in determining the good management skills in the way to empower the users within professional and ethical balance should be engaged into giving the necessary feedback associated with setting up the system itself.

2.2 Professional and Moral Balance on Digital Record Management Process

The professional attribution in the attempts to determine initial value of continuing the committed integrity in documenting the records through widening to manage the mandate

with a more coherent to the basis of corporate information needs to gain the extensive profession with better understanding about ethical manners [66–68]. This would lead to enhance the scale of recordkeeping processes assigned into skills and capability in harnessing technology utility to achieve the proactive support knowledge practices [11]. Both expectations of records management and capabilities to operationalize could be designed to offer the initiative to ensure the service in the quality of reflecting of the long term strategic information needs which the organizations could achieve for being accountable, transparent and compliant [69, 70]. In terms of the records management tools which can be adopted along with the organization's purposeful plan, the emphasis should focus on recordkeeping and information management initiation [12, 71, 72]. As a result, professional and ethical balance needs to be strengthened in the sense that can be transmitted into organizational expectations and its purpose within records management.

Moreover, the wide range of this process will lead to the decisions making initiative made from the strategic system in setting up the connection between the use and maintenance. Through demonstrating the way on the management performance which can be adopted among the users in recordkeeping, significant contribution could be obtained through making it appropriately and widely along with the procedural stage well deigned in terms of demand and need among the organizational basis [13, 73]. As a result, the key point of view determined to be the initiative to provide an ultimate application guideline as a counter measure against the emerging challenges of the dynamic records management system needs to bring along with urging for an appropriate professional and ethical empowerment [74, 75]. Such arrangement is required to elaborate particularly across the procedural stage proposed referring to the demand and the response with the express purpose of promoting appropriate and wise usage for the sustainable positive benefit of responsibilities on recording management initial arrangement for solving identification [14, 76, 77]. In an attempt to support this approach, this paper proposes the theoretical framework or guideline for empowering both professional and ethical foundations associated with recordkeeping initiative.

3 Analysis and Discussion

3.1 Re-actualising Digital Record Management with Privacy and Security Concern

The emerging trust explained into the contextual attribution in terms of privacy and security about the information data plays a key role in giving insights to enable the users to poses professional and ethical engagement as the quality in the records management [15]. At this point of view, the level of confidence which performs the system in protecting the information data within the archival records should be enlarged to expand the essential attribution of exploring the particular guideline in enabling the users to adopt it appropriately and wisely [78, 79]. In the extent of degree into the responsibility in managing the records management, trustworthy essence might need a comprehensive system to create the intention of making sufficiently with regards to supporting the process to explore in search for the information and utilising the systems within the skilful ability [16, 80, 81]. It is necessary to make sure in the development process of the extent in utilizing the technology in underlying the individuals to have the opportunities

with behaving the critical understanding in terms of techniques, nature and feedback of records and archival documentation.

In terms of underlining the recordkeeping system, the significant essence to enhance this particular attribution to represent the potential value in conceiving the structured data in the system would need to enhance in determining the wide trustworthiness in underlying the entire process of source of data within records management [82, 83]. As a result, the style of data aggregation with the complex analyses in enabling the interesting part of recordkeeping should bring the more data sources to possess the capability equipped into relying on data sources in assessing the quality of trustworthiness [17, 84, 85]. In the attempts to gain the value of trustworthiness, the stability of data sources in the archival records with reliability about the information data quality should be engaged into the sufficient assurances in providing the initial value of complementary engagement towards the satisfactory explanations about trustworthiness [86, 87]. In particular, the lack of trust information will lead to the failure in determining the steps which should undertake to remedy it with providing the footholds in preventing the extent quality of trust entirely being adapted to the system's reliability [18, 88, 89]. In this view, the examination process has to be involved with the entire perceptions of eradicating the comprehensive mapping of data potentially to support the distribution of chains in the context of digital infrastructure within interdependent systems.

3.2 Re-empowering Professional and Moral Skills for Digital Record Management Responsibility

The concern among the scholarly attention in this issue has to be involved with embarking with particular reference to records management in the sense that can determine moral foundation engaged into the professional skills. In terms of proposing grounded foundation on professional and ethical balance, records management could be entirely explored within providing the positive benefits to give the value of moral based instructional suggestion [19]. As a result, it is necessary to look into the commitment aligned with professional and ethical concern needed to give an insightful value in the level of individuals and society to have linkages on records management in providing systematic essence to control over records through the wide range of kinds of approaches like the document of transactions on business processes which gives insights into the decision-making process in the organizational context [90, 91]. It is highly supplied to have the wide regulation to direct the organizations including public and private sector to determine the significance of records management skills among the workers and staffs [20]. In this view, the wide regulations would be maintained through the requirement on legislating the assurance effectively managed, in order to successfully protect themselves from litigation and public scrutiny.

In line with the context of distinguished skills in the technology use to give an insightful value in the recordkeeping, the need to further exposure in particular is the process of transforming the new age-based performance skills among the staff member with long period working to update their practical experience to be more updated into the digital basis [92, 93]. This process needs the necessary part of establishing the clear diagnosis of the existing culture to be paid particular attention directed to identify and implement the appropriate strategies in the archival records [21]. The particular

attention refers to the entire part of initial diagnostic stage within the cultural orientation in the learning basis through the features of the technological environment [94, 95]. Both internal to the organisation and the broader regional capabilities should come up with the legislation and standards in the context of national, occupational and corporate cultures. It refers to the identification and evaluation of these settings in providing the framework highlighted into the relevant features, which need to be addressed in order to develop and promote the culture that is conducive to good recordkeeping [22]. In particular, attempts to strengthen the information quality assurance is assured to enhance the education process and purpose through digital record management.

Associated with the organizational information, the management framework here could be maintained with attitudes and values in order to shape information culture to meet the needs of users within the compliance requirements referring to the significant place of information quality [23]. In this view, the management of information with a fundamental process within the recordkeeping should be managed through the strategic resource with playing the significant role to embark in retrieving and storing part of wide range of sources in both public and private sector. Moreover, integrating the records management considered with the strategic initiation in being aware of the personal skills and commitment of self-efficacy refers to give a feedback with key role to fulfil the professional skills and competence in the recordkeeping management [24]. The wide range of significant quality in adopting, adapting and sustaining the way to possess the competence in particular aims to enhance the emotional aspect to create the reactions [96, 97]. There is a need to enhance in promoting the constructive engagement to apply and adopt the recordkeeping through the digital device appropriately and wisely in the basis of rational thinking into the understanding and highlighting the users about the importance and implications of online initiative [25]. Moreover, the procedural stage with the existing identities needs to be further on the human genetic engineering in understanding the way of improving and enhancing the digital human performance together with helping to better use into the general application.

3.3 Committing Awareness on Quality and Reliability in Electronic Records Management System (ERMS)

In the attempts to achieve the credibility through paying the serious attention about the recordkeeping, the point which has a pivotal role to transmit the process of managing the records management quality in strengthening the committed awareness should gather both quality and reliability which give insight into representing the mechanical issues based on the attainment of wide opportunities in accessing the procedural stage of records management [26]. Moreover, the attempts to get access technology which is affordable would enable one to continue living the users' social lives online with its attendant concerns about the security for service users and providers to abuse information and personal data stored on their sites. As the number of users together with the complexity of records, the growing number in ethical concern should begin with connecting the potential value in encouraging the moral foundation through ensuring the quality and reliability in preventing the abuse in the records management [27]. It is necessary to point out the pivotal value in encouraging the moral engagement use in cultivating the quality informing in the recordkeeping to enhance the extent of social information and

increase the level of convenience in developing social capital without cause for tension, misunderstanding, or miscommunication. In particular, the implication concerned into the credibility about the information shared in the ethical awareness in ensuring the users to get benefit equally and fairly, should be constructed with an essential value to achieve the procedural stage guideline in records management.

Moreover, the credibility of information in creating the social well-being among users is an important prerequisite to be handled with care in the sense that can be determined into the technological interactions [98, 99]. The attempts to accomplish the task in providing the accountability in information sharing would need to entirely accomplish the feedback on technology innovation [28]. It refers to the real actualisation in dealing with the key instruments to promoting the recordkeeping with its significant role to have a significant impact on contemporary life styles. The clear identification could be viewed into the recent increasing number of people spend valuable amount of time on watching television, playing records, listening to radio, reading electronic newspapers and magazines. This initiative enables individuals to attain media literacy and how to use it to promote critical understanding of the nature, techniques, impacts of media messages and productions to improve human mental and physical abilities [29, 100]. For example, smart technologies introduce various stakeholders in society in the area of business and education to improve performance and productivity with the objective of attaining customer satisfaction. Similarly, smart technologies have led to evolution in instructional strategies on using social media in formal and informal learning [30, 101]. Records management could shape the way in looking for the archives with more convenience to get benefit with building trust, care, friendship and mutual commitment to promote critical understanding of the phenomena of communication. This creates an enabling environment for individuals and groups to explore information around the world so as to maximize their potential abilities for personal development, social and epistemic practices and moral experiences.

In line with the wide range of users' activities arising from the ethical convictions nurtured and enshrined into the recordkeeping management including economic, social, business, legal, and medical aspects, an attempt to pursue the human potentials through digital-based innovative media technologies would need to provide the great opportunities to lay considerable foundation to create a balance between ethical and professional skills and ability [31, 102]. It is indeed imperative that the need for an integrated approach to apply both professional and ethical standards in routine human endeavours in personal and social orientation as corporate standards of behaviour. Thus the creation of such a balance would be a key point in nurturing the human conviction in records management [32]. As a source of human actions, there is a need to take into account how moral consideration should be applied in creating a balance for positive orientation based upon an ethically appropriate behaviour standard [103, 104]. Therefore, providing an innovative way of helping the human being adopted with considerable moral guidelines and standards is critical in transmitting it into the records management. This is intended to complement action-based approaches with value-based agenda within professional and ethical engagement in records management in giving insights into resolving the emerging issues in records.

3.4 Transmitting Records Quality Initiative into Organizational Culture

The essential value taken among the private sector organizations and government initiation refers to the quality extension on empowering professional based moral engagement. It should be emphasised in the sense that emphasizes social and personal awareness in records management, mainly in the digital basis and document in the sense that can be managed through addressing the metadata [33]. It important to enhance the digital era with emerging and innovative forms of technology through advancing the ethical way of commanding the records management utilized in the sense can be properly and wisely adopted. In particular, it refers to the different cultures across the intensive extension on fostering positive outcome of increased recordkeeping by paying serious attention on the potential challenges which may be occurred [34, 105]. Moreover, empowering moral engagement combined with professional attribution balanced to preserve the wide foundation in the way to operationalise long strategic initiation needs should be cooperated with a comprehensive coverage in theory and application. As a result, the core element of enabling individuals and groups to enhance such benefits for social and personal development would certainly contribute to the development of a wide range of recordkeeping through wide processes and application procedure stage.

In addition, it is important to note the potential value of ensuring the entire enhancement in reflecting the key performance in allowing the information data to have the pivotal role in the archival records has to be generated into the organizational culture initiative [35, 106]. Moreover, the various practices like business, transaction and learning resources are widely inseparable about the connections of the management of information, knowledge and technology more closely. As a result, the management practices in this regard play a key role in underlying the entire process of practices aligned into the organization level in the sense that can be determined into the records management to deliver along with practices amidst the organizational culture [36, 107]. In the attempts to recognise the particular decision to make in the context of information data transmitted into the archival records, this initial value to perform into the cultural stage amidst the organizational circumstance whether it is private or public sector effected systematically by the quality extent of the efficacy of records management by end-users.

At this point of view, the professional skills adapted into affecting the management quality in the records management specialists should be well-established in the initiative about the wider context on information determined into the evidence [108]. Towards the records management, the starting point of view to begin with examining the specification of the information continuum linked into the organizational culture with growing the scepticism over the current state of digital tools in providing the widely adequate support to the challenges encountered in the context of recordkeeping [34]. In terms of records management application and theory in determining the wide range of practices including, the archival records are received in the basis of creating the collective memory and the preservation of individual and community identity and history in the attempts that can determine to perform the support of administration running the business and even in the library resources in ensuring the quality of accountability of recordkeeping. In the context of enhancing the particular enhancement to transmit into the orientation of organizational culture, building the enthusiasms amidst the management construction among the staff to ensure the quality of information culture would give insights into

affecting the support and cooperation aligned with management of information assets [28]. Moreover, the entire attempts to enhance the commitment in creating the value of enthusiasm in the records management in the context of information culture becomes a pivotal role with a coherent and critical inclusiveness to let the process successful in the way which can determine the management of information assets. In particular, developing the exploratory model with its vital component should be widely engaged in nurturing the commitment within the management of information culture.

4 Implications and Future Directions

In particular, the extent of responsibility awareness of examining the structure of archival record in focusing on the security levels in detail plays a key role in giving insights into the wide context of records management. In this view, the record types should be conceived in providing the strategic system of recordkeeping initiative involved with the training end users set up into the context of organisation structure to work with the parts of the records plan in underlying the archival records [36]. Moreover, the wide range of file structures and the record types with other big task in the system is focused on the functional levels of the record plan in the context of organisation in the sense that would require the whole records in the reworking each time of plan [109]. In the attempts to incorporate this achievement referring to the planning and implementing of recordkeeping, the functional basis in facilitating the management concerned with creating the quality performance of archival records has to be maintained into the setting of functions in working with dealing into providing the facilities management to tackle of issues on records management [37]. In particular, the case of physical resources in the context of engineering issues refers to enhance the accommodation attempt with managing the space planning based on making the records plan concerned into the particular physical objects.

In addition, the strategic point on this view aims at regulating the professionalism engaged into the ethical concern [110]. It is necessary to expand the facilities with empowerment into recordkeeping in the sense that can determine the records plan maintained into the extent of function to build the committed strength in linking the extensive basis of professional and ethical balance to have personalised cooperation [29]. Through adapting the awareness quality on social and personal engagement with appropriate media technology, attempts to take account as the way in preventing negative impact in records from lack of authenticity to give insightful influences should manage in providing the extensive guidelines considered to enhance the positive feedback through empowering recordkeeping initiatives. In particular, the extensive inquiry in considering moral dimensions to give the feedback in underlying the current practice of recordkeeping appropriately within the current situation needs to provide an entire platform in providing the clear understanding about the importance of archival records amidst the emerging technological trends [38]. As a result, the theoretical based proposed guideline to give the feedback with entire reflection in making the important goal of records management become more reflective with ethical regulations which may become more widely valuable in recognizing and articulating the practice amidst the participants and communicators. It is necessary to note that important quality to address

the practical stage to conduct appropriately with ethical concern and professional skills plays an entire integrity to let the process run wisely referring to the point of value in the visible features on records management landscape.

5 Future Directions

The strategic point of records management tools in this view should be incorporated into the systematic approach in order to enable the users to appropriately use referring to its professionals needs in the attempt to meet the complex relation about system development and user practice [111]. The level of acceptance together with adopting technology acceptance and utility in a wise basis combined with strengthening strategic interaction basis has to be involved with better understanding in order to achieve the reflective initiation at considerable risk [31]. It indicates there should be an insightful value to gather the devolution adopted among the users' needs by considering professional skills and knowledge understanding among the workers. It is necessary to be considerable along with ensuring the role of records management effectively managed with understanding the implications of the decision making process within the world of organizational systems [112]. With this regard, getting aware of personal and social levels in terms of the implementation of recordkeeping has to be engaged into the moral enhancement together with professional skills [33]. The wide rules needed to give the entire support into the system in making more convenient to execute of the decision making process in ensuring the acceptance of the system with setting of some responsibility aligned into digital professional skills in implementing the document store in metadata-based database.

In addition, there should play a key role to transform the way to work with and applicate the archival records taken into the consideration into professional and ethical balance towards the emerging and dynamic trends of recordkeeping management. The framework with wide range of levels including the fundamental factors to assist in the records management needs to be arranged amongst the number of members in both organisation place and in the archival file records [37]. It is necessary to manage the quality of information in aiming to commit the accountability in supporting the ongoing activities to run business and to make the decision in the particular firm. The extent of influencing the identification and evaluation recordkeeping together manifesting the information behaviours set into the file record with assessing the wide preferences in stabilizing the fundamental values, attitudes and behaviours in the records management [38]. As a result, the need to expand the recordkeeping skills assigned into the level of knowledge and experience of employees in the context of second tier should be managed into applying the fundamental characteristics in the sense that can cover to all members of the organisation [113, 114]. Moreover, it is beyond the coverage into those with oversight of recordkeeping activities, where the extent of employer in this area to develop the appropriate strategies with tackling the isolation which would be and to underlie the level of knowledge understanding to assist the recordkeeping culture.

6 Conclusion

The significant role of technological advancement offered in the society has been emerged consistently into the education sector. It is the reality that the digital record management is being the transitional pathway in conveying the sources, materials and also data in the education purpose. Through examining the actual practice of digital record management, the strategic application to give insight into underlying the education teaching and learning process refers to look into detail about the current trends on enhancing the practice in the context on developing the learning technology in achieving the effectiveness. The strategic attempts on understanding of digital record management in education sector are significantly proposed with building the enhancement of both process and effectiveness in achieving the learning objective attainment. The main point of this paper refers to give an insightful value on strengthening the process, procedure and pathway in underlying the technology-based digital learning in education sector.

References

1. Aiusheeva, M.: Tackling a sense of insecurity: Enhancing digital literacy as a call to action for educators during the pandemic. In: Cahusac de Caux, B., Pretorius, L., Macaulay, L. (eds.) Research and Teaching in a Pandemic World, pp. 497–508. Springer, Singapore (2023). https://doi.org/10.1007/978-981-19-7757-2_33
2. AlAjmi, M.K.: The impact of digital leadership on teachers' technology integration during the COVID-19 pandemic in Kuwait. Int. J. Educ. Res. **112**, 101928 (2022)
3. Alam, A.: Platform utilising blockchain technology for elearning and online education for open sharing of academic proficiency and progress records. In: Asokan, R., Ruiz, D.P., Baig, Z.A., Piramuthu, S. (eds.) Smart Data Intelligence. Algorithms for Intelligent Systems, pp. 307–320. Springer, Singapore (2022). https://doi.org/10.1007/978-981-19-3311-0_26
4. Allayarova, S.N.: Implementation of modern information communication technologies (ICT) in higher education sector: international experience and the example of Uzbekistan. Int. J. Innov. Technol. Explor. Eng. **9**(1), 386–392 (2019)
5. Castro, R.: Blended learning in higher education: Trends and capabilities. Educ. Inf. Technol. **24**(4), 2523–2546 (2019). https://doi.org/10.1007/s10639-019-09886-3
6. Dahalan, F., Alias, N., Shaharom, M.S.N.: Gamification and game based learning for vocational education and training: a systematic literature review. Educ. Inf. Technol. 1–39 (2023)
7. de Obesso, M.D.L.M., Núñez-Canal, M., Pérez-Rivero, C.A.: How do students perceive educators' digital competence in higher education? Technol. Forecast. Soc. Chang. **188**, 122284 (2023)
8. Di Vaio, A., Latif, B., Gunarathne, N., Gupta, M., D'Adamo, I.: Digitalization and artificial knowledge for accountability in SCM: a systematic literature review. J. Enterp. Inf. Manage. (2023)
9. Dwivedi, Y.K., et al.: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. Int. J. Inf. Manage. **71**, 102642 (2023)
10. Gonçalves, M.J.A., et al.: Digital tools and methods to enhance learning: the digitools project. In: Mesquita, A., Abreu, A., Carvalho, J.V., de Mello, C.H.P. (eds.) Perspectives and Trends in Education and Technology. Smart Innovation, Systems and Technologies, vol. 320, pp. 399–413. Springer, Singapore (2023). https://doi.org/10.1007/978-981-19-6585-2_36

11. Haleem, A., Javaid, M., Qadri, M. A., Suman, R.: Understanding the role of digital technologies in education: a review Sustain. Oper. Comput. (2022)
12. Haleem, A., Javaid, M., Singh, R.P., Suman, R., Khan, S.: Management 4.0: concept, applications and advancements. Sustain. Oper. Comput. **4**, 10–21 (2023)
13. Han, H., Shiwakoti, R.K., Jarvis, R., Mordi, C., Botchie, D.: Accounting and auditing with blockchain technology and artificial Intelligence: a literature review. Int. J. Account. Inf. Syst. **48**, 100598 (2023)
14. Huda, M.: Empowering professional and ethical balance in digital record management. Organ. Cybersecur. J.: Pract. Process People **2**(1), 60–73 (2022)
15. Huda, M., Sutopo, L., Liberty, Febrianto, Mustafa, M. C.: Digital information transparency for cyber security: critical points in social media trends. In: Arai, K. (ed.) FICC 2022. LNNS, vol 439, pp. 814–831. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98015-3_55
16. Huynh-The, T., et al.: Blockchain for the metaverse: a review. Future Gener. Comput. Syst. (2023)
17. Javaid, M., Haleem, A., Singh, R. P., Suman, R., Gonzalez, E.S.: Understanding the adoption of industry 4.0 technologies in improving environmental sustainability. Sustain. Oper. Comput. (2022)
18. Kadhim, J.Q., Aljazaery, I.A., AlRikabi, H.T.S.: Enhancement of online education in engineering college based on mobile wireless communication networks and IoT. Int. J. Emerg. Technol. Learn. (Online) **18**(1), 176 (2023)
19. Kurniawan, T.A., Meidiana, C., Othman, M.H.D., Goh, H.H., Chew, K.W.: Strengthening waste recycling industry in Malang (Indonesia): lessons from waste management in the era of Industry 4.0. J. Clean. Prod. **382**, 135296 (2023)
20. Kutieshat, R., Farmanesh, P.: The impact of new human resource management practices on innovation performance during the COVID 19 crisis: a new perception on enhancing the educational sector. Sustainability **14**(5), 2872 (2022)
21. Kuzior, A., Sira, M.: A bibliometric analysis of blockchain technology research using VOSviewer. Sustainability **14**(13), 8206 (2022)
22. Lin, S., Lin, J.: How organizations leverage digital technology to develop customization and enhance customer relationship performance: an empirical investigation. Technol. Forecast. Soc. Chang. **188**, 122254 (2023)
23. Mansurjonovich, J.M.: Designing an electronic didactic environment to ensure interdisciplinary integration in the teaching of "Informatics and information technologies" during professional education. Confrencea **3**(03), 78–82 (2023)
24. Matriano, M.T.: Balancing of Skills in the digital transformation of education and employability. In: SHS Web of Conferences, vol. 156, p. 08004. EDP Sciences (2023)
25. Mhlanga, D.: Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. In: Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning (2023)
26. Mukred, M., Yusof, Z.M., Alotaibi, F.M.: Ensuring the productivity of higher learning institutions through electronic records management system (ERMS). IEEE Access **7**, 97343–97364 (2019)
27. Mukred, M., Yusof, Z.M., Alotaibi, F.M., Asma'Mokhtar, U., Fauzi, F.: The key factors in adopting an electronic records management system (ERMS) in the educational sector: a UTAUT-based framework. IEEE Access **7**, 35963–35980 (2019)
28. Mukred, M., Yusof, Z.M., Mokhtar, U.A., Manap, N.A.: Electronic records management system adoption readiness framework for higher professional education institutions in Yemen. Int. J. Adv. Sci. Eng. Inf. Technol. **6**(6), 804–811 (2016)

29. Mukred, M., Yusof, Z.M., Mokhtar, U.A., Sadiq, A.S., Hawash, B., Ahmed, W.A.: Improving the decision-making process in the higher learning institutions via electronic records management system adoption (2021)
30. Ngoepe, M., Jacobs, L., Mojapelo, M.: Inclusion of digital records in the archives and records management curricula in a comprehensive open distance e-learning environment. Inf. Dev. 02666669221081812 (2022)
31. Nguyen, L.A.T., Habók, A.: Tools for assessing teacher digital literacy: a review. J. Comput. Educ. 1–42 (2023)
32. Okoye, K., et al.: Impact of digital technologies upon teaching and learning in higher education in Latin America: an outlook on the reach, barriers, and bottlenecks. Educ. Inf. Technol. **28**(2), 2291–2360 (2023)
33. Rahardja, U., Aini, Q., Khairunisa, A., Sunarya, P.A., Millah, S.: Implementation of blockchain technology in learning management system (LMS). APTISI Trans. Manage. (ATM) **6**(2), 112–120 (2022)
34. Skitsko, V., Osypova, O.: Education 5.0 maturity index: concept and prospects for development. In: Ortiz-Rodríguez, F., Tiwari, S., Sicilia, M.A., Nikiforova, A. (eds.) EGETC 2022. CCIS, vol. 1666, pp. 95–108. Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-22950-3_8
35. Tavares, D., Lopes, A.I., Castro, C., Maia, G., Leite, L., Quintas, M.: The intersection of artificial intelligence, telemedicine, and neurophysiology: opportunities and challenges. Handb. Res. Instruct. Technol. Health Educ. Allied Discip. 130–152 (2023)
36. Tojimamatovich, J.V.: Digital transformation of educational management system. Web Semant.: Univ. J. Innov. Educ. **2**(4), 202–206 (2023)
37. Tran, V.T., Tran, N.H.: A review of smart education and lessons learned for an effective application in Binh Duong province. Vietnam. Pegem J. Educ. Instruct. **13**(1), 234–240 (2023)
38. Yousefi, S., Tosarkani, B.M.: Exploring the role of blockchain technology in improving sustainable supply chain performance: a system-analysis-based approach. IEEE Trans. Eng. Manag. (2023)
39. Huda, M.: Trust as a key element for quality communication and information management: insights into developing safe cyber-organisational sustainability. Int. J. Organ. Anal. (2023). https://doi.org/10.1108/IJOA-12-2022-3532
40. Syofiarti, S., Saputra, R., Lahmi, A., Rahmi, R.: The use of audiovisual media in learning and its impact on learning outcomes of Islamic cultural history at Madrasah Tsanawiyah Negeri 4 Pasaman. Firdaus J. **1**(1), 36–44 (2021). https://doi.org/10.37134/firdaus.vol1.1.4.2021
41. Huda, M.: Towards digital access during pandemic age: better learning service or adaptation struggling? Foresight **25**(1), 82–107 (2023). https://doi.org/10.1108/FS-09-2021-0184
42. Huda, M.: Digital marketplace for tourism resilience in the pandemic age: voices from budget hotel customers. Int. J. Organ. Anal. **31**(1), 149–167 (2023). https://doi.org/10.1108/IJOA-10-2021-2987
43. Zainul, Z., Rasyid, A., Nasrul, W.: Distance learning model for Islamic religious education subjects in non Internet Server Provider (ISP) areas. Firdaus J. **1**(1), 45–53 (2021). https://doi.org/10.37134/firdaus.vol1.1.5.2021
44. Huda, M.: Between accessibility and adaptability of digital platform: investigating learners' perspectives on digital learning infrastructure. High. Educ. Ski. Work. Based Learn. (2023). https://doi.org/10.1108/HESWBL-03-2022-0069)
45. Huda, M. et al.: Strategic role of trust in digital communication: citical insights into building organizational sustainability. In: Arai, K. (ed.) Proceedings of the Future Technologies Conference FTC 2023, vol. 3. LCNS, vol 815. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-47457-6_25

46. Abdul Aziz, N.A., Mohd Razali, F., Saari, C.Z.: Penggunaaan Media Sosial dari Perspektif Psiko Spiritual Islam . Firdaus J. **2**(1), 65–75 (2022). https://doi.org/10.37134/firdaus.vol2.1.6.2022
47. Hanafi, H.F., et al.: Review of learner's model for programming in teaching and learning. J. Adv. Res. Appl. Sci. Eng. Technol. **33**(3), 169–184 (2024)
48. Huda, M., Bakar, A.: Culturally responsive and communicative teaching for multicultural integration: qualitative analysis from public secondary school. Qual. Res. J. (2024). https://doi.org/10.1108/QRJ-07-2023-0123
49. Syafri, N., Ali, A.H., Ramli, S.: Penggunaan kaedah Inovasi Sambung dan Baca Bahasa Arab (SaBBAr) dalam meningkatkan kemahiran membaca perkataan Bahasa Arab murid di Sekolah Rendah. Firdaus J. **2**(2), 62–71 (2022). https://doi.org/10.37134/firdaus.vol2.2.6.2022
50. Huda, M., et al.: Enhancing digital leadership direction: insight into empowering gender violence prevention. In: Mishra, D., Ngoc Le, A., McDowell, Z. (eds.) Communication Technology and Gender Violence. Signals and Communication Technology. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-45237-6_13
51. Ab Rahim, N.M.Z., Saari, Z., Mohamad, A.M., Rashid, M.H., Mohamad Norzilan, N.I.: Konsep Ulul Albab dalam Al-Quran dan hubungannya dengan pembelajaran kursus Sains, Teknologi dan Manusia di UTM Kuala Lumpur. Firdaus J. **2**(2), 72–78 (2022). https://doi.org/10.37134/firdaus.vol2.2.7.2022
52. Huda, M., et al.: Understanding of digital ethics for information trust: a critical insight into gender violence anticipation. In: Mishra, D., Ngoc Le, A., McDowell, Z. (eds.) Communication Technology and Gender Violence. Signals and Communication Technology. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-45237-6_14
53. Hanafi, H.F., Wahab, M.H.A., Selamat, A.Z., Masnan, A.H., Huda, M.: A systematic review of augmented reality in multimedia learning outcomes in education. In: Singh, M., Kang, D.K., Lee, J.H., Tiwary, U.S., Singh, D., Chung, W.Y. (eds.) IHCI 2020. LNCS, vol. 12616, pp. 63–72. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68452-5_7
54. Khalili, Huda, M., Rosman, A.S., Mohamed, A.K., Marni, N.: Digital learning enhancement through social network site (SNS). In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Software Engineering Application in Informatics. CoMeSySo 2021. Lecture Notes in Networks and Systems, vol. 232. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-90318-3_35
55. Leh, F. C., Anduroh, A., Huda, M.: Level of knowledge, skills and attitude of trainee teachers on Web 2.0 applications in teaching geography in Malaysia schools. Heliyon **7**(12) (2021)
56. Sukadari, S., Huda, M.: Culture sustainability through co-curricular learning program: learning batik cross review. Educ. Sci. **11**(11), 736 (2021)
57. Huda, M., Gusmian, I., Mulyo, M.T.: Towards eco-friendly responsibilities: Indonesia field school model cross review. J. Comp. Asian Dev. (JCAD) **18**(2), 1–12 (2021)
58. Rachman, A., Oktoviani, I., Manurung, P.: Digitization of zakat and charity BAZNAS Tangerang city through crowdfunding platform tangerangsedekah. id. Firdaus J. **3**(1), 96–106 (2023). https://doi.org/10.37134/firdaus.vol3.1.9.2023
59. Huda, M., Hashim, A.: Towards professional and ethical balance: insights into application strategy on media literacy education. Kybernetes **51**(3), 1280–1300 (2022)
60. Mohamad Shokri, S.S., Salihan, S.: Modul pembangunan Program Huffaz Profesional Universiti Tenaga Nasional: satu pemerhatian kepada konstruk pembangunan Al-Quran. Firdaus J. **3**(2), 12–23 (2023). https://doi.org/10.37134/firdaus.vol3.2.2.2023
61. Berlian, Z., Huda, M.: Reflecting culturally responsive and communicative teaching (CRCT) through partnership commitment. Educ. Sci. **12**(5), 295 (2022)
62. Huda, M.: Towards an adaptive ethics on social networking sites (SNS): a critical reflection. J. Inf. Commun. Ethics Soc. **20**(2), 273–290 (2022)

63. Omar, M.N.: Inovasi pengajaran & pemudahcaraan menggunakan Aplikasi Ezi-Maq (MAHARAT AL-QURAN) untuk menarik minat pelajar menguasai Ilmu Tajwid. Firdaus J. **2**(2), 79–87 (2022). https://doi.org/10.37134/firdaus.vol2.2.8.2022
64. Huda, M., Sutopo, L., Liberty, Febrianto, Mustafa, M.C.: Digital information transparency for cyber security: critical points in social media trends. In: Arai, K. (eds.) Advances in Information and Communication. FICC 2022. LNNS. vol. 439. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98015-3_55
65. Borham, A.H., et al.: Information and communication ethics in social media for indigenous people's religious understanding: a critical review. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer (2024)
66. Ali, A.H.: Pelajar Khalifah Profesional Tempaan Ulul Albab sorotan penerapannya berasaskan komponen QEI di UPSI . Firdaus J. **3**(2), 51–63 (2023). https://doi.org/10.37134/firdaus.vol3.2.5.2023
67. Jusoh, A., Huda, M., Abdullah, R., Lee, N.: Development of digital heritage for archaeo-visit tourism resilience: evidences from E-Lenggong web portal. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer (2024)
68. Rahim, M.M.A., Huda, M., Borham, A.H., Kasim, A.Y., Othman, M.S.: Managing information leadership for learning performance: an empirical study among public school educators. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer (2024)
69. Hehsan, A., et al.: Digital muhadathah: framework model development for digital arabic language learning. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer Nature Switzerland, Cham (2024)
70. Muharom, F., Farhan, M., Athoillah, S., Rozihan, R., Muflihin, A., Huda, M.: Digital technology skills for professional development: insights into quality instruction performance. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer Nature Switzerland, Cham (2024)
71. Muhamad, N., Huda, M., Hashim, A., Tabrani, Z.A., Maarif, M.A.: Managing technology integration for teaching strategy: public school educators' beliefs and practices. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer Nature Switzerland, Cham (2024)
72. Wahid, A., et al.: Digital technology for indigenous people's knowledge acquisition process: insights from empirical literature analysis. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer Nature Switzerland, Cham (2024)
73. Huda, M., Taisin, J.N., Muhamad, M., Kiting, R., Yusuf, R.: Digital technology adoption for instruction aids: insight into teaching material content. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer Nature Switzerland, Cham (2024)
74. Alwi, A.S.Q., Ibrahim, R.: Isu terhadap Penggunaan Teknologi Media Digital dalam kalangan guru pelatih jurusan Pendidikan Khas. Firdaus J. **2**(2), 88–93 (2022). https://doi.org/10.37134/firdaus.vol2.2.9.2022
75. Zamri, F.A., Muhamad, N., Huda, M., Hashim, A.: Social media adoption for digital learning innovation: insights into building learning support. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer Nature Switzerland, Cham (2024)
76. Huda, M., et al.: Trust in electronic record management system: insights from Islamic-based professional and moral engagement-based digital archive. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)
77. Yahya, S.F., Othman, M.A.: Penggunaan video dalam Pengajaran dan Pembelajaran Pendidikan Moral Tingkatan 2. Firdaus J. **2**(2), 94–105 (2022). https://doi.org/10.37134/firdaus.vol2.2.10.2022

78. Huda, M., et al.: Empowering Islamic-based digital competence and skills: how to drive it into reconstructing safety strategy from gender violence. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)
79. Cita Sari, D., et al.: Transformation of artificial intelligence in Islamic Edu with Ulul Albab value (global challenge perspective). Firdaus J. **3**(1), 1–9 (2023). https://doi.org/10.37134/firdaus.vol3.1.1.2023
80. Huda, M., et al.: From digital ethics to digital community: an Islamic principle on strengthening safety strategy on information. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)
81. Zamri, F.A., Muhamad, N., Huda, M.: Information and communication technology skills for instruction performance: beliefs and experiences from public school educators. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023, Springer Nature Switzerland, Cham (2024)
82. Masud, A., Borham, A.H., Huda, M., Rahim, M.M.A., Husain, H.: Managing information quality for learning instruction: insights from public administration officers' experiences and practices. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023, Springer Nature Switzerland, Cham (2024)
83. Rahman, M.H.A., Jaafar, J., Huda, M.: Information and communication skills for higher learners competence model. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023, Springer Nature Switzerland, Cham (2024)
84. Mohd Nawi, M.Z.: Media variations in education in malaysia: a 21st century paradigm. Firdaus J. **3**(1), 77–95 (2023). https://doi.org/10.37134/firdaus.vol3.1.8.2023
85. Huda, M., et al.: From technology adaptation to technology adoption: an insight into public school administrative management. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. LNNS. Springer, Singapore (2024)
86. Tan, A.A., Huda, M., Rohim, M.A. Hassan, T.R.R., Ismail, A.: Chat GPT in supporting education instruction sector: an empirical literature review. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. LNNS. Springer, Singapore (2024)
87. Wahid, A., et al.: Augmented reality model in supporting instruction process: a critical review. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. LNNS. Springer, Singapore (2024)
88. Musolin, M.H., Serour, R.O.H., Hamid, S.A., Ismail, A., Huda, M., Rohim, M.A.: Developing personalized Islamic learning in digital age: pedagogical and technological integration for open learning resources (OLR). In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. LNNS. Springer, Singapore (2024)
89. Musolin, M.H., Ismail, M.H., Huda, M., Hassan, T.R.R., Ismail, A.: Towards an Islamic education administration system: a critical contribution from technology adoption. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. LNNS. Springer, Singapore (2024)
90. Musolin, M.H., Ismail, M.H., Farhan, M., Rois, N., Huda, M., Rohim, M.A.: Understanding of artificial intelligence for Islamic education support and service: insights from empirical literature review. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. LNNS. Springer, Singapore (2024)

91. Susilowati, T., et al.: Getting parents involved in child's school: using attendance application system based on SMS gateway. Int. J. Eng. Technol. **7**(2.27), 167–174 (2018)
92. Huda, M., et al.: From digital ethics to digital partnership skills: driving a safety strategy to expand the digital community? In: Digital Transformation for Business and Society: Contemporary Issues and Applications in Asia. Taylor & Francis, UK (2023a)
93. Susilowati, T., et al.: Learning application of Lampung language based on multimedia software. Int. J. Eng. Technol. **7**(2.27), 175–181 (2018)
94. Abadi, S., et al.: Application model of k-means clustering: insights into promotion strategy of vocational high school. Int. J. Eng. Technol. **7** (2.27), 182–187 (2018)
95. Huda, M., et al.: Towards digital servant leadership for organisational stability: driving processes in the pandemic age? In: Digital Transformation for Business and Society: Contemporary Issues and Applications in Asia. Taylor & Francis, UK (2023b)
96. Aminudin, N., et al.: The family hope program using AHP method. Int. J. Eng. Technol. **7**(2.27), 188–193 (2018)
97. Wulandari, et al.: Design of library application system. Int. J. Eng. Technol. **7**(2.27), 199–204 (2018)
98. Aminudin, N., et al.: Higher education selection using simple additive weighting. Int. J. Eng. Technol. **7**(2.27), 211–217 (2018)
99. Zainuri, A., Huda, M.: Empowering cooperative teamwork for community service sustainability: insights from service learning. Sustainability **15**(5), 4551 (2023)
100. Huda, M., Teh, K.S.M.: Empowering professional and ethical competence on reflective teaching practice in digital era. In: Dikilitas, K., Mede, E., Atay D. (eds.) Mentorship Strategies in Teacher Education, pp. 136–152. IGI Global, Hershey (2018). https://doi.org/10.4018/978-1-5225-4050-2.ch007
101. Huda, M., Teh, K.S.M., Nor, N.H.M., Nor, M.B.M.: Transmitting leadership based civic responsibility: insights from service learning. Int. J. Ethics Syst. **34**(1), 20–31 (2018). https://doi.org/10.1108/IJOES-05-2017-0079
102. Huda, M., Mulyadi, D., Hananto, A.L., Nor Muhamad, N.H., Mat Teh, K.S., Don, A.G.: Empowering corporate social responsibility (CSR): insights from service learning. Soc. Res. J. **14**(4), 875–894 (2018)
103. Huda, M., et al.: Innovative teaching in higher education: big data approach. Turkish Online J. Educ. Technol. **15**, 1210–1216 (2016)
104. Anshari, M., Almunawar, M.N., Shahrill, M., Wicaksono, D.K., Huda, M.: Smartphones usage in the classrooms: learning aid or interference? Educ. Inf. Technol. **22**(6), 3063–3079 (2017)
105. Huda, M., et al.: Empowering children with adaptive technology skills: careful engagement in the digital information age. Int. Electron. J. Elem. Educ. **9**(3), 693–708 (2017)
106. Huda, M., Shahrill, M., Maseleno, A., Jasmi, K.A., Mustari, I., Basiron, B.: Exploring adaptive teaching competencies in big data era. Int. J. Emerg. Technol. Learn. **12**(3), 68–83 (2017b)
107. Susilowati, T., et al.: Decision support system for determining lecturer scholarships for doctoral study using CBR (case-based reasoning) method. Int. J. Recent Technol. Eng. **8**(1), 3281–3290 (2019)
108. Huda, M., Jasmi, K.A., Alas, Y., Qodriah, S.L., Dacholfany, M.I., Jamsari, E.A.: Empowering civic responsibility: insights from service learning. In: Burton, S. (ed.) Engaged Scholarship and Civic Responsibility in Higher Education, pp. 144–165. IGI Global, Hershey (2017c). https://doi.org/10.4018/978-1-5225-3649-9.ch007
109. Huda, M., et al.: Innovative e-therapy service in higher education: mobile application design. Int. J. Inter. Mob. Technol. **11**(4), 83–94 (2017d)
110. Muslihudin, M., et al.: Decision support system in kindergarten aelection using TOPSIS method. Int. J. Recent Technol. Eng. **8**(1), 3291–3298 (2019)

111. Huda, M., et al.: From live interaction to virtual interaction: an exposure on the moral engagement in the digital era. J. Theor. Appl. Inf. Technol. **95**(19), 4964–4972 (2017e)
112. Huda, M., Maseleno, A., Jasmi, K.A., Mustari, I., Basiron, B.: Strengthening interaction from direct to virtual basis: insights from ethical and professional empowerment. Int. J. Appl. Eng. Res. **12**(17), 6901–6909 (2017)
113. Huda, M., Haron, Z., Ripin, M.N., Hehsan, A., Yaacob, A.B.C.: Exploring innovative learning environment (ILE): big data era. Int. J. Appl. Eng. Res. **12**(17), 6678–6685 (2017)
114. Maseleno, A., et al.: Combining the previous measure of evidence to educational entrance examination. J. Artif. Intell. **10**(3), 85–90 (2017)



Integrating Machine Learning Algorithms with EEG Signals to Identify Emotions Among University Students

Mohd Fahmi Mohamad Amran(✉), Venothanee Sundra Mohan, Nurhafizah Moziyana Mohd Yusop, Yuhanim Hani Yahaya, Muhammad Fairuz Abd Rauf, Noor Afiza Mat Razali, Fazilatulaili Ali, and Sharifah Aishah Syed Ali

Department of Computer Science, Faculty of Defence Science and Technology, Universiti Pertahanan Nasional Malaysia, Kuala Lumpur, Malaysia
fahmiamran@upnm.edu.my

Abstract. Emotion level of the students during their academic sessions has a significant effect on their performance and overall academic grades. The most pre-eminent way of evaluating emotion levels is by analysing the EEG signals obtained from the brain. This paper showcases the experimental study on obtaining the Electroencephalogram (EEG) signals from the students during their academic sessions and classifying it to the types of emotions using machine learning algorithms. This paper explores the method of how the data collection session is conducted and recorded. The data collected is then compiled and divided for the machine learning algorithms. Furthermore, this paper proposed a method of acquiring the emotion labels without prior inducing by using a standard normal distribution method. Finally, this paper also proposed a deep learning neural network model and machine learning models that can be used to determine the emotion level from the EEG signals.

Keywords: Machine Learning · EEG · Emotions

1 Introduction

Students undergoing their tertiary education has a much more stressful and difficult lifestyle compared to their primary and secondary counterparts as it requires more efforts in terms of diligence, time management and prioritization. Due to this stressful environment, the students' progress and achievements during their academic sessions are affected which leads to an overall loss of performance [1]. One of the main factors that affects the students' performance is the students' emotions and feelings that they have during their classes and exams [2]. Emotions can be described as sensations or perceptions that arise within a specific situation a person finds themselves in, often influenced by the interactions of those around them [3]. As the human emotions are produced inside the brain, it's difficult to determine their emotion level based on their facial expressions

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 334–342, 2024.
https://doi.org/10.1007/978-3-031-53549-9_34

by itself as the students are capable of hiding it or just giving a neutral face [4]. Therefore, the most viable option is to analyze the brain signals to identify the type of emotions that they are feeling by using an EEG device [1].

The first EEG device consisted of multiple electrodes that are connected to a heavy machine that is able to read the EEG signals. Due to the machine's size, weight and cost, EEG devices were mainly used in medical field [5]. However, technological advancement has allowed the EEG devices to be compacted into a low-cost portable device that is able to be used in other fields such as sports and education fields [5]. These low-cost portable devices are able to obtain the EEG signals using electrodes that are connected to the scalp and send it to other electronic devices such as a phone or a computer. Based on the type of devices, the signals are saved in the terms of signal itself or in numerical numbers. The EEG device is able to obtain the brainwaves from the human brain called the EEG signals. The characteristics of the EEG signals can be classified into the type of emotions using a machine learning algorithms and deep learning algorithms [6]. The model is able to identify the characteristics and patterns in the EEG signals and classify the features into the types of emotions.

From previous research, the performance level of students during their academic sessions were analyzed based on their attention level obtained from the EEG signals. However, in order to determine the emotion level, most of the past research utilized the inducing method where the participants are kept in a controlled environment and they were exposed to either audio or visual stimuli [7–9]. The stimuli are done to induce the corresponding emotion in the participants EEG signals which allows the machine learning model to classify the signals into the types of emotions based on the time of induction. This method is difficult to use as the emotions will need to be obtained during the academic sessions itself. This paper showcases the experimental study of classifying the emotions of the students during their academic session.

2 Related Work

2.1 Introduction

In terms of this research, the research articles that were chosen for analysis consisted of the portable EEG devices that is suitable for this research, articles that analyses the performance and emotion levels of the students using said portable EEG devices, articles that define the method of obtaining the emotion levels from the EEG signals without prior inducing and articles that showcase the multiple different machine learning and deep learning models that is used to classify the EEG signals into the emotions.

2.2 Portable EEG Devices

There are currently multiple different portable EEG devices in the market due to the convenience and the low cost of the devices. One of these devices is the Neurosky Mindwave mobile device which is made up of a single electrode that is positioned on the forehead and a second sensor at the ear clips that functions as a ground to filter out electronic noise [10, 11]. This device has been used for measuring performance in

academic sessions by analyzing the attention level of the students [10, 11]. The attention level is obtained from the eSense feature that is available in the Neurosky device [10, 11]. The level of attention serves as a gauge for assessing the students' performance during their learning sessions [12]. However, even though this device has been used in obtaining emotion levels, most of the research is done by inducing the participants using external stimuli before classifying the EEG signals into the type of emotions [13]. There are a smaller number of researches that was done on obtaining the emotions during the academic session itself.

2.3 Real-Time Emotion Analysis

As most of the research in emotion classification in EEG signals is done by inducing said emotions in a controlled environment, additional research will need to be done on the method of identifying the emotions in a real time situation [14]. The EEG signals are made up of multiple different waveforms known as Alpha, Beta, Gamma, Delta and Theta. Each of these signals, based on the level of concentration, indicates if the emotions are negative or positive. If the signals are too high or too low, it can be classified as having negative emotions while if the signals are in optimum level, it can be classified as positive emotions [15–17].

2.4 Machine Learning Classification Algorithm

In order to classify the EEG signals into the type of emotions, a machine learning classification model is used to identify the patterns and classify it into the emotions based on the prior research done on the real-time emotion analysis. From previous research the most common machine learning models that were used for emotion classification is the KNN model [18], Decision Tree model [7], Gaussian Naïve Bayes model [19], Logistic Regression model [20], SVM model [21] and LDA model [19]. Besides the machine learning algorithm, some research also utilizes the deep learning model that is able to handle large amount of data compared to the machine learning models. The common deep learning models that has been used to classify emotions in EEG signals are the deep feedforward neural network [22], CNN model [23], RNN model [24] and LSTM model [25]. This research will focus on the five machine learning models listed above and the deep feedforward neural network.

3 Experimental Study

3.1 Subjects

The subjects of this research are made up of 32 first-year computer science students from the Faculty of Defence Science and Technology in the National Defence University of Malaysia. The students consist of both university students and cadets that are undergoing their academic sessions. The data is collected during two semesters, and during two different classes. In the first semester, eleven students are chosen while during the second semester, 21 students were chosen.

3.2 Environment and Data Collection

The data is collected during two different environments which are known as positive or negative environment. The positive environment is used to ensure there are more positive emotions present in the data collection while the negative environment is to ensure there are more negative emotions presents in the data. In the first semester, the positive environment is the lectures while the negative environment is during the final exam. While in the second semester, the positive environment is the lectures but the negative environment is during the quiz sessions.

During the data collection session, the students were given a short briefing on consent and the outline of the research. The students were also informed on the working of the device and the time taken to complete the data collection. The students were also advised to not fiddle with the device once the device is securely fastened on their heads.

The time taken for the data collection is for 15 min with a five seconds interval. The data collection utilises an Android Application that is able to see the if there's any loss of connection or poor signals. If the application shows any interruptions, the device was promptly checked to ensure the device is properly fitted and if the device is showing a low battery level. If the data collection is interrupted for more than a minute, the device is rebooted and the data will be rerecorded. Once the data has been recorded for the required time, the data is then saved as a CSV file and transferred for the data compilation and analysis.

3.3 Data Compilation and Division

In a single environment session, the number of data collected per student is 180. Since there are a total of 32 students for each environment, there are a total of 5760 amount of data. By combining both environment sessions, there are a total of 11520 amount of data. In order for the data to be used for the machine learning classification process, the data will need to be divided into three different datasets, which is the training dataset, evaluation dataset and testing dataset. The training dataset is made up of 70% of the total dataset which leads to 16 students, 20% of the dataset is taken for the evaluation dataset which is a total number of 10 students and the testing dataset is made up of 10% of the dataset which is a total number of 6 students.

The training and testing dataset were chosen randomly while the students for the testing dataset is chosen according to an equal number of genders and a good mixture of cadets and students. By combining both scenarios of data into the dataset, the total number of EEG data in the training dataset is 5760, the evaluation dataset has 3600 data and the testing dataset has 2160 data. The data is then divided and compiled into CSV files. The datasets are then analysed to ensure there are no missing values. Due to the elimination of missing values, the total number of data is reduced compared to earlier. Table 1 shows the final division of the dataset.

3.4 Acquiring the Emotion Labels

In order to obtain the emotion labels of the EEG signals, based on the related works done in previous research, the emotion labels are divided into two categories, negative

Table 1. Dataset Division

Dataset Name	Percentage	Number of Students	Total number of EEG data in the dataset
Training Dataset	70%	16	5655
Evaluation Dataset	20%	10	3244
Testing Dataset	10%	6	2155

emotions or positive emotions. The negative emotions are obtained if the signals are too low or too high while the positive emotions are obtained if the emotions are in the average level. The best way to divide the signals into the required emotions is by utilizing the normal distribution graph. The characteristics of the normal distribution graphs allocates the average number in the peak while the low range is on the left side and the high range is on the right side of the graph. For this research, the standard normal distribution graph is chosen due to its clarity and easy understanding. The standard normal distribution graph is made up of mean values on the Y-axis while the X-axis is the standard deviation values. The Z-score is then plotted as the data points in the graph. Therefore, the standard deviation, mean and Z-score of each data signals to obtain the emotion labels according to the range of the standard normal distribution graph.

The standard deviation, mean and Z-score is obtained using the Pandas library in Python. The Pandas library already contains the formulas library required to obtain the standard deviation value, mean value and Z-score. From the standard normal distribution graph, it can be determined that the values between -1 to 1 are the average values while the others are high and low values. This is used to divide the dataset into the type of emotions. Once the dataset has been normalized in Python, the data is then labelled 1 as positive emotion, and 0 as negative emotion. However, in a single row in the dataset, there are multiple signals at a certain point of time which leads to multiple different emotion labels. To make it easier to determine the emotion label, the labels are then averaged to obtain the average emotion label at each time interval. This makes the label easier to be read by the machine learning models. Please try to avoid rasterized images for line-art diagrams and schemas. Whenever possible, use vector graphics instead (see **Error! Reference source not found.**).

3.5 Proposed Deep Feedforward Neural Network Model

A neural network is made up of multiple different neurons combined together in multiple layers of a network. The first step in designing a neural network is to determine the structure of the dataset and its availability to be used for deep learning models. The dataset needs to be ensured there are no missing data as that will cause an error in the reading later. The elimination method is used in this research to delete the data that are missing in the dataset.

The next step is to determine the key feature of the dataset and how it correlates with one another to obtain the required output. The key features of the dataset are the eight types of raw signals which then is able to correlate with one another to obtain the output

emotion which is positive or negative emotions. Once the key features and the labels are determined, the batch size and number of epochs are then determined. The batch size is the number of inputs into the model at each iteration. For this model, the batch size is 15 while the number of epochs is 200. This is to ensure that the model is able to trained itself efficiently with a high accuracy value.

Next is the number of nodes for the model, since the key feature are made up of eight signals so the number of input layer is made up of eight nodes. The input nodes are then connected to the hidden layers which is then connected to the output layer. The output layer is made up of two nodes which is either 0 for negative emotion and 1 for positive emotion. There are multiple different layers available which is the Dense layer, Convolutional Layer and Flatten layer. Each of the layers has its own purpose and method of analyzing the data. The proposed model is made up of dense layer which performs the dot product between the input and the kernel with the addition of the bias for each layer. The output of the computation is then passed on to the activation function of each layer before being passed to the next layer.

The activation functions determine which value should be the output of the neuron. There are multiple different types of activation layers which is Sigmoid function, TanH function and ReLU function. The proposed model uses the ReLU activation function for the hidden layers while the output function is made up of the SoftMax activation function.

The final proposed model is made up of eight nodes in the input layer, with a total of five hidden layers with two layers of 15 nodes, two layers of 12 nodes and a single layer of 10 nodes. The output layer is made up of two nodes. The architecture of the model is shown in Fig. 1.

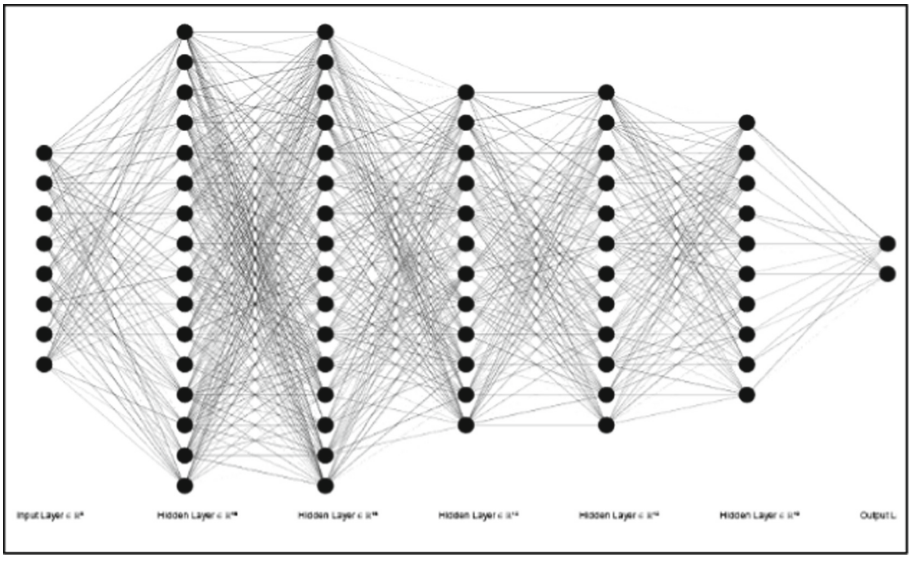


Fig. 1. Architecture of the Neural Network Model

The proposed model is then ready for the training and evaluation dataset to obtain the loss and accuracy of the model. Once the model has obtained a high accuracy with small number of loss, the model is then saved and the confusion matrix and classification table is obtained to determine the precision, accuracy and F1-Score of the model.

3.6 Proposed Machine Learning Models

Compared with the neural network model, the machine learning models are much simpler because of the convenience of the Scikit-Learn library. The Scikit-Learn library in Python provides efficient tools for machine learning algorithms including models for both the supervised and unsupervised machine learning algorithms. Due to the straightforward operation of the library, multiple models can be utilized simultaneously for the training and evaluation process of the machine learning model. Besides the Scikit-Learn library, the Pandas library in Python allows easy analysis and manipulation of datasets and data frame.

The first step is similar to the neural network model which is to analyze the dataset and remove the missing values. The key features are then identified which is the eight raw signals and the label feature. The key features are labelled as X while the label features are labelled as Y. The next step is determining the model to use for the machine learning process. From previous research, the most commonly used models are the Gaussian Naïve Bayes model, KNN model, LDA model, Decision Tree model and SVM model. All these models are available in the Scikit-Learn library.

The Scikit-Learn library allows the training data and label to be inserted into the model and is then evaluated with the evaluation data and label.fit command. The model is then compared with the evaluation labels and raw signals to determine the prediction score. Once the prediction score is obtained, Python then generates the confusion matrix, precision, recall and F1-Score of each model. The results of the models are then saved and compared with the other models to determine which model has the highest accuracy.

4 Conclusion

The main objective of this paper is to develop a new method of obtaining EEG signals to determine the emotion analysis of students during their academic sessions. This paper gives an overview of the data collection process including the subjects, environment and the data compilation and division. Furthermore, this paper proposes a method to acquire the emotion labels of the EEG signals by using standard normal distribution to divide the signals into positive and negative emotions. Finally, this paper also explains the proposed deep learning neural network model and the machine learning models that are used to determine the emotion level of the students. This model is able to classify the emotion levels of the students based on the acquired emotional label which allows future interpretation on whether the emotion level of the students is affecting their academic performance.

Acknowledgement. The authors honorably appreciate Universiti Pertahanan Nasional Malaysia for continuous support and the financial sponsorship through research grant

(UPNM/2021/GPJP/ICT/2). We also wish to thank the volunteer participants for their generous cooperation.

References

1. Konar, A.: Artificial intelligence and soft computing: behavioral and cognitive modeling of the human brain (2018)
2. Banerjee, P.A.: A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools. Cogent Educ. **3**(1), 1178441 (2016). https://doi.org/10.1080/2331186X.2016.1178441
3. Mohan, V.S., Amran, M.F.M., Yahaya, Y.H., Yusop, N.M.M., Sembok, T.M.T., Ahmad, M.A.: Emotion classifications in electroencephalogram (EEG) signals. Int. J. Recent Technol. Eng. **8**(3), 2736–2740 (2019)
4. Raval, D., Sakle, M.: A literature review on emotion recognition system using various facial expression. **17**(2), 326–329 (2015)
5. Minguillon, J., Lopez-Gordo, M., Pelayo, F.: Trends in EEG-BCI for daily-life: Requirements for artifact removal. Elsevier (2017). https://www.sciencedirect.com/science/article/pii/S1746809416301318?casa_token=uOLv-4a3CrEAAAAA:-bBt2Z1H0uPN-7nIi_RSGIVPqismBl8z6IJwcWp9HsAC4Q0IdA5RX1wnGqOlSL2SQ9cI3ULuCM-Z. Accessed 10 June 2020
6. Soroush, M.Z., Maghooli, K., Setarehdan, S.K., Nasrabadi, A.M.: A review on EEG signals based emotion recognition. **4**(4), 118–129 (2017). https://doi.org/10.15171/icnj.2017.01
7. Liu, J., Meng, H., Nandi, A., Li, M.: Emotion detection from EEG recordings. In: 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, ICNC-FSKD 2016, pp. 1722–1727 (2016). https://doi.org/10.1109/FSKD.2016.7603437
8. Bhatti, A., Majid, M., Anwar, S.,Khan, B.: Human emotion recognition and analysis in response to audio music using brain signals. Elsevier (2016). https://www.sciencedirect.com/science/article/pii/S0747563216305945?casa_token=VLJf7L_LMYgAAAAA:1LEcEnNJrdim79GEE4vZuxclmNXWkIJkfu3UynuuUqO16hiHieC2_j8-6efnQC0ximKeqbLidO2m. Accessed 10 June 2020
9. Peled-Avron, L., Levy-Gigi, E., Richter-Levin, G., Korem, N., Shamay-Tsoory, S.G.: The role of empathy in the neural responses to observed human social touch. Cogn. Affect. Behav. Neurosci. **16**(5), 802–813 (2016). https://doi.org/10.3758/s13415-016-0432-5
10. Katona, J., Farkas, I., Ujbanyi, T., Dukan, P., Kovari, A.: Evaluation of the NeuroSky MindFlex EEG headset brain waves data (2014). https://ieeexplore.ieee.org/abstract/document/6822382/. Accessed 11 June 2020
11. EEG: The Ultimate Guide. http://neurosky.com/biosensors/eeg-sensor/ultimate-guide-to-eeg/. Accessed 10 June 2023
12. Amran, M.F.M., et al.: A machine learning approach using EEG signals to identify emotions and performance level among university students. In: AIP Conference Proceedings, vol. 2617, no. 1. AIP Publishing (2022)
13. Xu, J., Zhong, B.: Review on portable EEG technology in educational research. Comput. Human Behav. **81**, 340–349 (2018). https://doi.org/10.1016/j.chb.2017.12.037
14. Demos, J.: Getting started with neurofeedback (2005)
15. Making Sense of EEG Bands. http://neurosky.com/2015/05/greek-alphabet-soup-making-sense-of-eeg-bands/. Accessed 10 June 2020
16. Marzbani, H., Marateb, H.R., Mansourian, M.: Neurofeedback: a comprehensive review on system design, methodology and clinical applications (2016) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4892319/. Accessed 10 June 2020

17. Enriquez-Geppert, S., Huster, R.J., Herrmann, C.S.: EEG-neurofeedback as a tool to modulate cognition and behavior: a review tutorial. Front. Hum. Neurosci. **11**(2017). https://doi.org/10.3389/fnhum.2017.00051
18. Li, M., Xu, H., Liu, X., Lu, S.: Emotion recognition from multichannel EEG signals using K-nearest neighbor classification. Technol. Health Care **26**(S1), S509–S519 (2018). https://doi.org/10.3233/THC-174836
19. Zheng, W.-L., Zhu, J.-Y., Lu, B.-L.: Identifying Stable Patterns over time for Emotion Recognition from EEG. IEEE Trans. Affect. Comput. **3045**(c), 1 (2017). https://doi.org/10.1109/taffc.2017.2712143
20. Doma, V., Pirouz, M.: A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals. J. Big Data **7**(1) (2020). https://doi.org/10.1186/s40537-020-00289-7
21. Guo, K., Candra, H., Yu, H., Li, H.: EEG-based emotion classification using innovative features and combined SVM and HMM classifier (2017). https://ieeexplore.ieee.org/abstract/document/8036868/?casa_token=6b3Sd6P6JU0AAAAA:papsunoooZmbTqK50VHHNGpBYNDMwp88D_kCNMMsXMpC4m1iL7e1hUPCFDu4xCtBtPjZ2S3OY1bw. Accessed 11 June 2020
22. Hemanth, D.J.: EEG signal based Modified Kohonen neural networks for classification of human mental emotions. J. Artif. Intell. Syst. **2**(1), 1–13 (2020). https://doi.org/10.33969/ais.2020.21001
23. Keelawat, P., Thammasan, N., Numao, M., Kijsirikul, B.: Spatiotemporal emotion recognition using deep CNN based on EEG during music listening (2019). http://arxiv.org/abs/1910.09719. Accessed 17 June 2020
24. Zhang, T., Zheng, W., Cui, Z., Zong, Y., Li, Y.: Spatial-temporal recurrent neural network for emotion recognition. IEEE Trans. Cybern. **49**(3), 939–947 (2019). https://doi.org/10.1109/TCYB.2017.2788081
25. Anubhav, Nath, D., Singh, M., Sethia, D., Kalra, D., Indu, S.: An efficient approach to EEG-based emotion recognition using LSTM network. In: Proceedings - 2020 16th IEEE International Colloquium on Signal Processing and its Applications, CSPA (2020)



Comparison of Game Development Framework and Model for Parkinson Disease Rehabilitation

Muhammad Fairuz Abd Rauf[1](✉), Saliyah Kahar[2], Mohd Fahmi Mohamad Amran[1], Suziyanti Marjudi[3], Zuraidy Adnan[4], and Rita Wong[1]

[1] Faculty of Defence Science and Technology, National Defence University Malaysia, 57000 Kuala Lumpur, Malaysia
fairuz.rauf@upnm.edu.my
[2] Management and Science University, 40100 Shah Alam, Malaysia
[3] Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Malaysia
[4] Universiti Selangor, 45600 Bestari Jaya, Malaysia

Abstract. Parkinson's disease patients are consulted to frequently exercise and undergo physiotherapy to reduce the progression rate of their disease. However, this can be a problem when the caretakers that bring them to the rehabilitation center are unavailable due to work or other related issues. The cost to go to rehabilitation is also expensive. Therefore, patients tend to exercise at home where there is no one to monitor except for their caretakers. This can lead to problems in evaluating the progress of the patient's rehabilitation. The solution for this issue is to develop a game-based rehabilitation. This paper aims to investigate the element for a game-based Parkinson's disease rehabilitation. Then, the elements are used to develop the preliminary framework which later can be used to develop game-based exercise. A total of eight frameworks and models are studied to identify the core elements. The seven elements which are commonly used are game design, game mechanics, game engine, patient, therapist, results, and database.

Keywords: Game Development Framework · Serious Games · Parkinson

1 Introduction

Past research has used rehabilitation frameworks to serve the underserved community in Malaysia. There are games to rehabilitate the sharpness of sensory in blind and deaf patients; rehabilitate autism, stroke, and dementia patients; and even rehabilitate Alzheimer's patients (Avola et al., 2018; Ayed et al., 2019; Dewald et al., 2016; Ienca et al., 2017; Yap et al., 2019). Research has also found that underserved community in Malaysia prefer to do activities within the community support group than go to the rehabilitation center. One of the research projects mentions that one-third of its community sample prefers to stay at home and do self-rehabilitation, even though 31% of those community members end up not doing the correct rehabilitation activity (Latif et al., 2020).

Despite this, underserved community members trust their community support group more in terms of activity that mimics the rehabilitative effects of rehabilitation because

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 343–356, 2024.
https://doi.org/10.1007/978-3-031-53549-9_35

the community group provides activities that sustain the rehabilitation environment and help its underserved members to improve motor skills; social and language development; vocational skills; self-management skill; creative skill; music therapy; as well as sports (Elor et al., 2018; Latif et al., 2020). This means that rehabilitation framework research can be done through the participation and sampling of the underserved community members such as PD patients within its community support group.

Based on Malaysia's 2020 report on technological trend and usage, more than 80% of the Malaysian population has experience in using computers, phones, and other devices in their daily lives (Department of Statistics Malaysia, 2020a). This means that the general Malaysian population has basic technological usage and engagement and has a high probability to adapt well to a game-based technology for rehabilitation. This is further supported by many other types of research in the computing, social, and medical field (Fincham, 2013; Latif et al., 2020; Muñoz et al., 2019). Studies related to underserved communities have resulted in both acceptance and rejection of technology due to known and informed consent; data privacy, confidentiality, and sharing conventions; unwanted social implication and attention; and also legal implications (Nebeker et al., 2017). Underserved individuals usually defer to the support and recommendations of their community support group. Hence, developers can review both individual and group Technology Acceptance directly through the community support group. For underserved communities with a high rate of legal problems, developers are advised to prepare a full informed consent document with legal implications covering the issue. However, if the underserved community is willing and encourages its members to participate, the development can be proceeded with and without informed consent documents.

However, the success of the game-based rehabilitation depends on its clear development features (Herrlich et al., 2017). There are currently a scanty number of PD rehabilitation framework with clear game development features in Malaysia. Harzhing's Publish or Perish tools is used to extract papers on game development in Malaysia between the year 2016 to 2020. The keywords used are "Malaysia", "Framework", "Game Design", and "Parkinson's Rehabilitation". A total of 922 papers were extracted and filtered according to research that focuses on the game development framework and its features. The filter resulted in only 183 papers highlighting the game development framework and its features. This means that only 19.85% of the research focus on PD rehabilitation frameworks in Malaysia with game development features between 2016 to 2020. Therefore, there is a gap in understanding the framework elements and mapping its features.

2 Existing Framework and Model for Rehabilitation Using Various Technologies

2.1 Introduction

A suitable framework is required to assist in designing exergames for PD patients. Currently, there is a limited framework available that focuses on rehabilitation for PD patients. Some samples of past framework are put up as sample of framework discussion. One of the structures which focuses on using Virtual Reality (VR) technology to rehabilitate patients.

2.2 Framework for Rehabilitation Using Wiimote

The framework in Fig. 1 shows the framework for rehabilitation using Wiimote which uses virtual reality as tools that can help to rehabilitate PD patient. Based on Fig. 1, it is divided into four processes where the human will be connected to a specific remote, the patient will play the exergames, data will be transmitted to a clinician online, and finally, the clinician can view or keep track of progress from the clinician terminal. Basically, four elements have been identified for the framework which includes patient, game design, database and physiotherapist or clinician. Each of these elements are critical in the development and assessment of the patient.

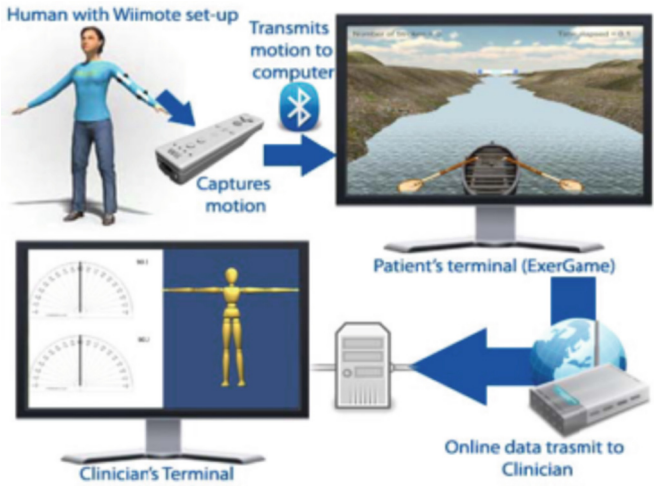


Fig. 1. Framework for rehabilitation using Wiimote (I. Paraskevopoulos & Tsekleves, 2013)

2.3 Framework for Gait-Based Recognition Using Kinect

Another framework available for rehabilitating PD patient is shown in Fig. 2. Figure 2 shows the framework for gait-based recognition using Kinect. It focuses solely on gait recognition and does not cater to other types of rehabilitation.

The framework is divided into four stages which include features extraction, dissimilarity calculation, sparse representation in dissimilarity space and recognition. In the first stage, which features extraction, the gait sequence of the patient is identified, and their walking pattern will be extracted. In the second stage, dissimilarity calculation is conducted to identify differences with a normal walking pattern. In the third stage, a dissimilarity matrix of the training sequence will be generated to fill the gap of the walking pattern of the PD patient. Finally, the last stage, which is the fourth stage, is recognition. Based on framework on Fig. 2, this framework focuses only on gait for PD patient and its main function is to compare gait – through calculations between a PD patient and a healthy person. Hence, there is limited focus on elements for game development.

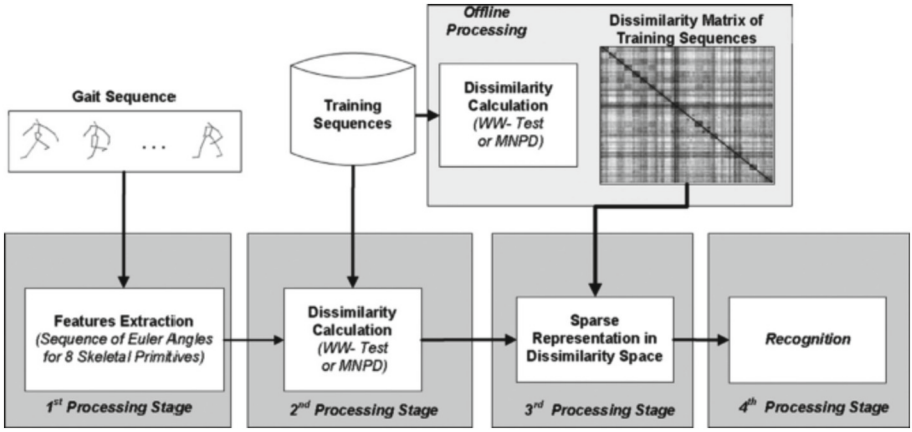


Fig. 2. Framework for gait-based recognition using Kinect (Kastaniotis et al., 2015)

2.4 Framework for Game-Based Cognitive Rehabilitation

Another framework that focuses on rehabilitation through gaming is shown in Fig. 3. It is divided into four components: condition, process, activity, and output. In the first stage, the condition of the patient is being identified. In the next stage, with the help from the therapist, rehabilitation objectives have been set and game characteristics are being adopted to cater to the objectives that are set. The next stage involves the activity of playing the custom games regularly. Three main characteristics under the game cycle will affect the patient at this stage which are system feedback, user judgement and user behavior.

Finally, in the outcome stage, the result is generated, and it can be reflected on the patient itself. The framework in Fig. 3 manages to identify the best elements for game-based rehabilitation. Since it is for cognitive rehabilitation, the focus is on games that involve thinking such as solving puzzles or memorization.

2.5 Framework for Game Design for PD Patient and Patient with Stroke

Figure 4 provides another framework which can be used for designing games for PD patient as well as patient with stroke. It is divided into four major elements which are experimental accuracy study, games for stroke and PD, evaluation and results and guidelines. For the experimental accuracy study, a comparative study between two systems is conducted which are Wiimote MoCap to Vicon System. The next stage involves games for stroke and PD. Few elements involve at this stage includes 3D asset collection and import to game engine, games design and games engine script. Afterwards, the next stage is the pilot testing which involves pilot testing of games with patients. The final stage is the results and game design guidelines which involves validation of the results.

This framework is one of the best in terms of game design for PD patients and patient with stroke. Elements that should be highlighted are the game engine, game design and game engine scripts. Unfortunately, the framework lacks input from patient at the early

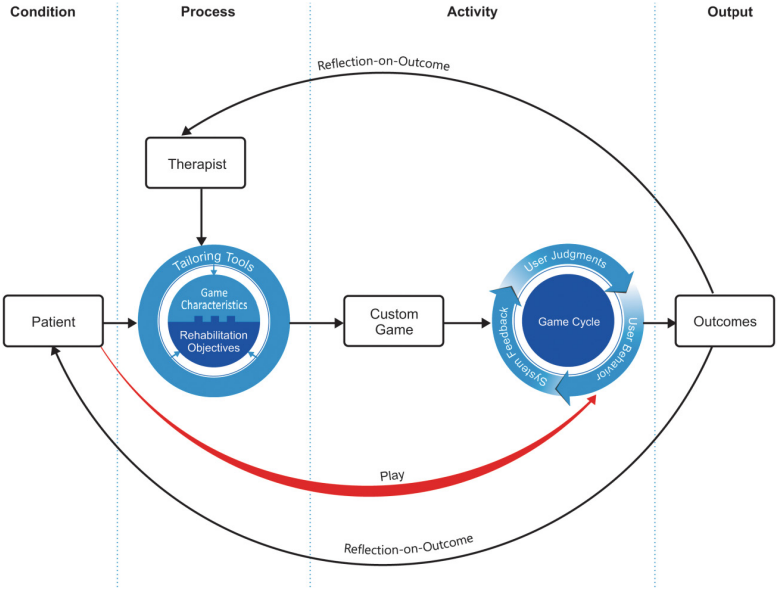


Fig. 3. Framework for game-based cognitive rehabilitation (Elaklouk, 2014)

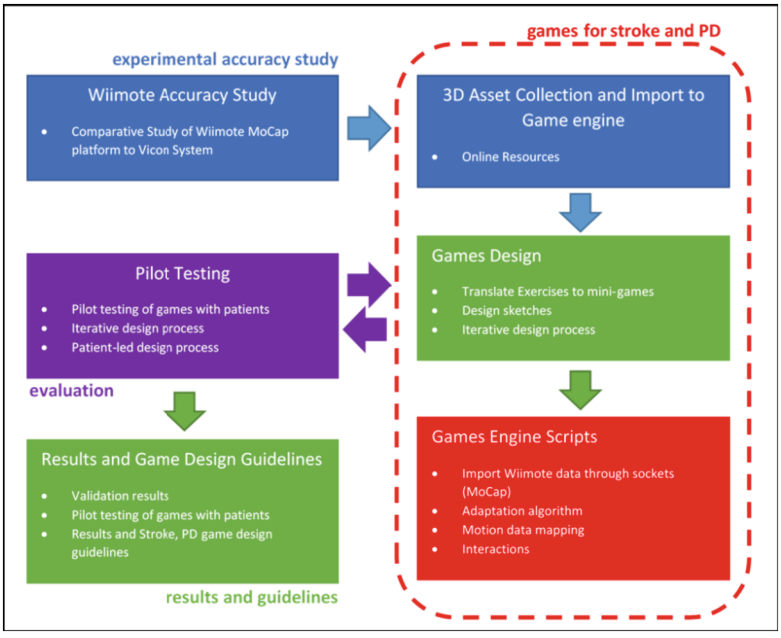


Fig. 4. Framework for game design for PD patient and patient with stroke (I. T. Paraskevopoulos, 2014)

stage and experts from healthcare to ensure that the game that have been developed from this framework can be used to gain maximum efficacy.

2.6 Framework for Development of Serious Games for Motor Skills Rehabilitation

Figure 5 provides another framework for the development of serious games for motor skills rehabilitation. It consists of five layers which are user layer, input/output layer, game engine layer, database layer and web application layer. The user layer is composed of the patient and the therapist. The Input/output layer is responsible to interact with the user. The game engine is responsible for the game logic. The database layer is responsible for storing data such as patient's data and the web application layer consumes data stored on the database to provide insights and report patient's progress. This framework manages to accommodate each element for a development of a serious game but not in terms of game design.

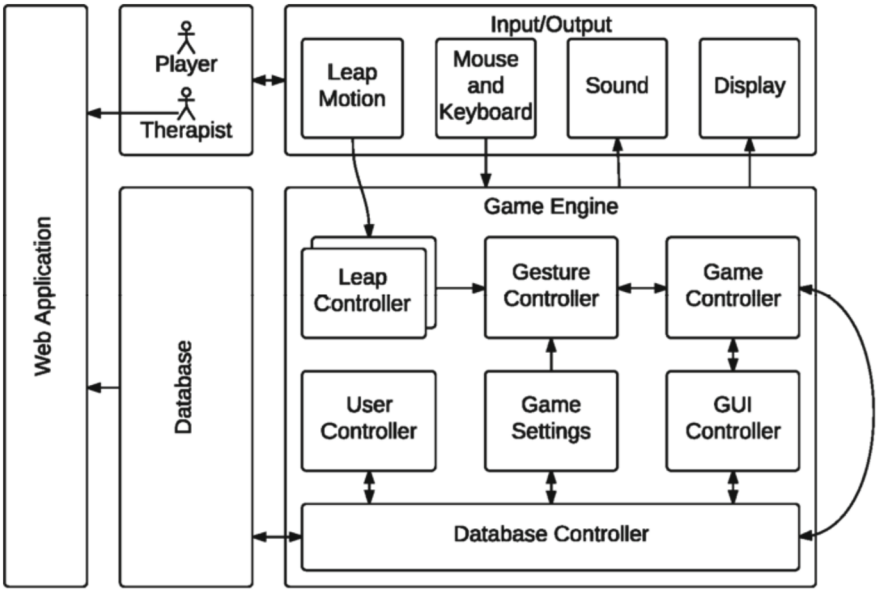


Fig. 5. Framework for development of serious games for motor skills rehabilitation (Foletto et al., 2017)

Basically, the framework focuses on the development process but few elements should be added especially in terms of game design since the objective of the framework is clear – which is to develop serious games (games with specific purpose).

2.7 Architecture and Game Engine Framework for Serious Games in Health Rehabilitation

Figure 6 shows the architecture as well as the game engine framework for serious games in health rehabilitation. The main modules of the game engine framework are game engine, game database, social networking, competition, user management and profiling, and logging and monitoring. Game Engine represents the most generic component of the game logic. Game Database is the repository of all the games. Social networking is a mechanism for a patient to group together in a social network. Competition is responsible for creating the interaction mechanism of competition. User management and profiling are managing user profile. Logging and monitoring are the monitor progress of each patient.

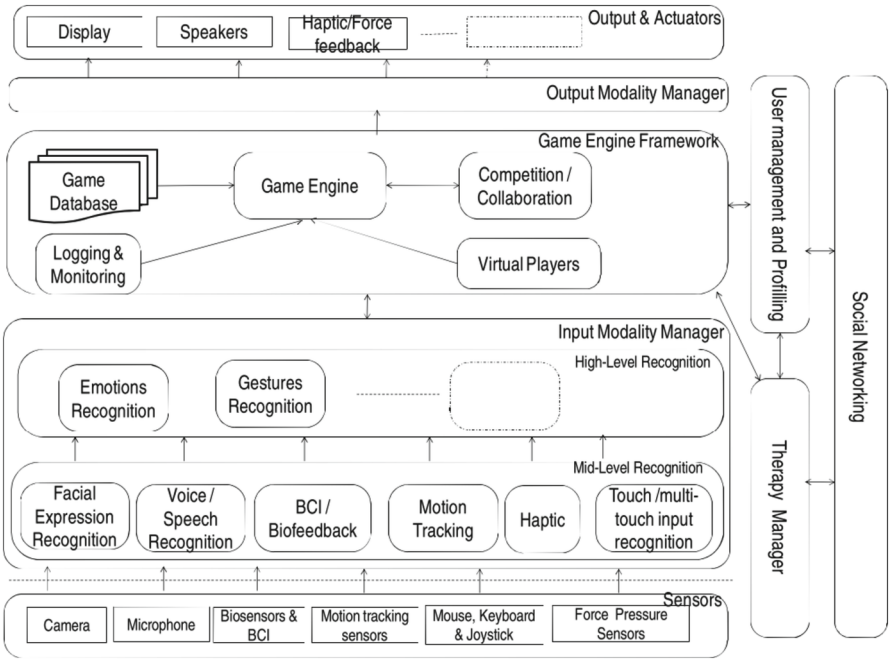


Fig. 6. Architecture and game engine framework for serious games in health rehabilitation (P. A. Rego, Moreira, & Reis, 2014)

This framework focuses more on the game engine while integrating different types of input or output methods which includes sensors and other devices.

2.8 Framework for Serious Games as a Structural Class Diagram for Learning

Figure 7 shows a conceptual framework for serious games as a Structural Class Diagram for learning. It consists of nine elements which include capability, instructional content, intended learning outcome, game attributes, learning activity, game mechanics, games

genre, game achievement and reflection. This framework focuses on the learning outcome. However, for rehabilitation purposes which focuses on teaching patient to carry out simple daily activities, this framework can be applied.

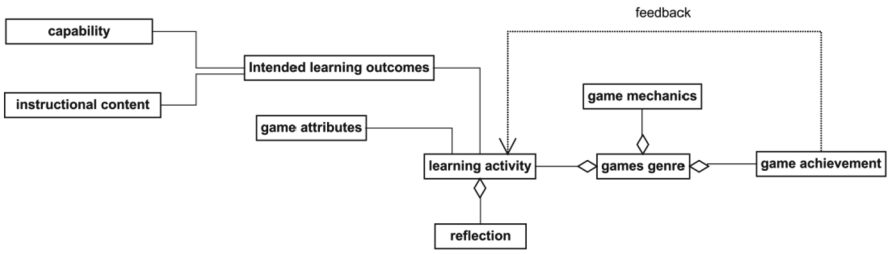


Fig. 7. Framework for serious games as a Structural Class Diagram for learning (Yusoff et al., 2009)

For example, if a pilot spends few hours playing aviation computer games, they can perform better in flights (Yusoff et al., 2009). Although this framework is more towards structural class diagram, it may be considered since some of the elements are important for allowing patient to relearn movement. Some of the important elements that can be considered for future development of framework (later known as MyPard) are game mechanics and game attributes.

2.9 Serious Game Design Assessment Framework

Figure 8 shows a serious game design assessment framework. This framework focuses solely on the purpose of the game and its impact. It is important that the game's purpose acts as the driving force for the player to carry out repetitive actions and achieve the objectives of playing the game such as learning, rehabilitation or even reducing weight.

Serious game design assessment framework consists of six important elements which are fiction or narrative, aesthetic graphics, mechanic, purpose, coherence, framing and content. Description for each component is purpose, content, game mechanics, fiction narrative, aesthetics, graphics, framing, and coherence. Purpose refers to a game's purpose to impact its players. It depends on the objectives of the game. Content refers to information, fact and data offered and used in the game. Game mechanics refer to methods invoked by agents for interacting with the game world. Fiction and narrative refer to creating fictional space and how it relates to the game purpose. Aesthetics and graphics refer to audio-visual language, conceptualized, chosen and developed by the designers for the visualization involved in the game. Framing refers to the playing literacy of the game. Coherence refers to ensuring that all elements are related to each other. For this framework, game mechanics is being given focus since it clearly states that within the game mechanics, learning curve, rules, goals and rewards can encourage player to engage continuously with the game.



Fig. 8. Serious game design assessment framework (Yusoff et al., 2009)

2.10 Findings

These are only a sample of past frameworks to understand the visualization of the framework. Preliminary observation shows that 7 elements stand out. Table 1 summarizes each framework that is available in developing and assesses serious games for rehabilitation and learning:

Based on the framework obtain and compared in previous table, these elements are crucial for a rehabilitation-based exergames:

Player/Patient: Refers to player or patient who plays the game. The objective of the game is to serve the player or patient.

Therapist/Clinician: Refers to another party who are in charge of monitoring the patient's progress and ensure their safety.

Game mechanic: Construct of rules or method designed for interaction of the game.

Game engine: Game engines are tools available for game designers to code and plan out a game quickly and easily without building one from the ground up. Some of the examples include Unity, HTML5, CRYENGINE and Torque.

Game design: Art of applying design and aesthetics to create a game for entertainment or for educational, exercise, or experimental purposes. Each game design might be different to suit the objectives of the game.

Database: Set of data held in a computer and can be viewed later. It can be viewed, accessed, and updated.

Results: Results is the one that keeps the player motivated to play since it will complement the objective of the game. It can be viewed as progress and can be use as benchmark to track achievement or progress.

Table 1. Comparison between each framework

No.	Author	Framework Objective	Number of elements	Elements involve
1	(I. Paraskevopoulos & Tsekleves, 2013)	Schematic diagram of a context using VR	4	Human with Wiimote setup, patients terminal, online data transmit to clinician, clinician terminal
2	(Kastaniotis et al., 2015)	Gait based recognition using Kinect	4	Features extraction, dissimilarity calculation, sparse representation in dissimilarity space, recognition
3	(Elaklouk, 2014)	Game-based cognitive rehabilitation	6	Patient, therapist, tailoring tools, custom game, game cycle, outcome
4	(I. T. Paraskevopoulos, 2014)	Game design for PD patient and patient with stroke	6	Wiimote accuracy study, 3D asset collection and import to game engine, games design, games engine script, pilot testing, results and game design guidelines
5	(Foletto et al., 2017)	Development of serious games for motor skills rehabilitation	5	Web application, player & therapist, input/output, game engine, database
6	(P. A. Rego et al., 2014)	Serious games in health rehabilitation	8	Sensors, input modality manager, game engine framework, output modality manager, output & actuators, user management and profiling, therapy manager, social networking

(*continued*)

Table 1. (*continued*)

No.	Author	Framework Objective	Number of elements	Elements involve
7	(Yusoff et al., 2009)	Serious games as a Structural Class Diagram for learning	9	Capability, instructional content, intended learning outcomes, game attributes, game mechanics, game game achievement, game genre, learning activity, reflection
8	(Ricciardi et al., 2015)	Serious game design assessment	7	Purpose, content, game mechanics, fiction and narrative, aesthetics and graphics, framing, coherence

3 Matrix for Framework

3.1 Comparison

Each framework has its own uniqueness since it caters to different objectives. Table 2 provides mapping for each framework and the elements involved.

3.2 Document Analysis

In order to fully map out and compare the most recent frameworks and their elements between the year 2016 to 2020, 183 documents were analyzed to map the game development framework features and elements. Out of all 183, 141 of those papers highlight clearly and specifically the element and its features. 30 documentations were chosen since they focus on game design and human computer interaction. Analysis was conducted using MAXQDA software. Findings are presented in Table 3.

Table 2. Matrix for each framework

No.	Author	Framework Elements							
		Player/Patient	Therapist/Clinician	Game Mechanic	Game Engine	Game Design	Database	Results	Others
1	(I. Paraskevopoulos & Tsekleves, 2013)	/	/	/				/	/
2	(Kastaniotis et al., 2015)	/		/			/	/	/
3	(Elaklouk, 2014)	/	/	/		/		/	/
4	(I. T. Paraskevopoulos, 2014)	/		/	/	/		/	/
5	(Foletto et al., 2017)	/	/	/	/		/	/	/
6	(P. A. Rego et al., 2014b)	/	/	/	/		/	/	/
7	(Yusoff et al., 2009)	/		/		/		/	/
8	(Ricciardi et al., 2015)	/	/	/		/		/	/

Table 3. Document analysis of each element.

No	Elements	Percentage (%)
1	Patient	83.33
2	Therapist	53.33
3	Game Engine	33.33
4	Game Mechanics	70.00
5	Game Design	73.33
6	Result	83.33
7	Database	70.00

4 Summary

Based on the comparison table and document analysis provided in the earlier section of this paper, it is clear that these elements such as patient, therapist, game engine, game mechanics, game design, result and database are crucial. Patients are required because they are the main stakeholders that need rehabilitation. Therapists are the ones that plan types of rehabilitation and its scoring for the patient. Game related elements such as game engine, game mechanics, game design results and database are based on the feedback from therapist. Game designers need to plan and develop accordingly based on the input provided by therapists. Each of these elements must be included in the next stage of the research which is development of the game development framework.

Acknowledgment. The authors greatly acknowledge Ministry of Higher Education Malaysia and Universiti Pertahanan Nasional Malaysia for the financial support. Special thank you to the reviewers for their valuable comments and suggestions.

References

Avola, D., Cinque, L., Foresti, G.L., Marini, M.R., Pannone, D.: VRheab: a fully immersive motor rehabilitation system based on recurrent neural network. Multimed. Tools Appl. **77**(19), 24955–24982 (2018). https://doi.org/10.1007/s11042-018-5730-1

Ayed, I., Ghazel, A., Jaume-i-Capó, A., Moyà-Alcover, G., Varona, J., Martínez-Bueso, P.: Vision-based serious games and virtual reality systems for motor rehabilitation: a review geared toward a research methodology. Int. J. Med. Informatics **131**(May), 103909 (2019). https://doi.org/10.1016/j.ijmedinf.2019.06.016

Dewald, J.P.A., Ellis, M.D., Acosta, A.M., McPherson, J.G., Stienen, A.H.A.: Implementation of impairment- based neurorehabilitation devices and technologies following brain injury. In: Neurorehabilitation Technology, 2nd edn. (2016). https://doi.org/10.1007/978-3-319-28603-7_18

Department of Statistics Malaysia. Malaysia Pocket Stats Quarter 1 2020 (2020)

Elor, A., Teodorescu, M., Kurniawan, S.: Project star catcher. ACM Trans. Accessible Comput. **11**(4), 1–25 (2018). https://doi.org/10.1145/3265755

Elaklouk, A.M.S.: Framework for game-based cognitive rehabilitation (2014)
Fincham, J.E.: The public health importance of improving health literacy. Am. J. Pharm. Educ. **77**(3), 3–4 (2013)
Foletto, A.A., Cordeiro d'Ornellas, M., Cervi Prado, A.L.: Serious games for parkinson's disease fine motor skills rehabilitation using natural interfaces. Stud. Health Technol. Inform. **245**, 74–78 (2017). https://doi.org/10.3233/978-1-61499-830-3-74
Herrlich, M., Smeddinck, J.D., Soliman, M., Malaka, R.: "Grab-that-there": live direction for motion-based games for health. In: Conference on Human Factors in Computing Systems - Proceedings, Part F1276, pp. 2622–2629 (2017). https://doi.org/10.1145/3027063.3053212
Ienca, M., et al.: Intelligent assistive technology for Alzheimer's disease and other dementias: a systematic review. J. Alzheimer's Dis. **56**(4), 1301–1340 (2017). https://doi.org/10.3233/JAD-161037
Kastaniotis, D., Theodorakopoulos, I., Theoharatos, C., Economou, G., Fotopoulos, S.: A framework for gait-based recognition using Kinect. Pattern Recogn. Lett. **68**, 327–335 (2015). https://doi.org/10.1016/j.patrec.2015.06.020
Latif, R.A., Ismail, R., Mohd Mahidin, E.M.: Game-based rehabilitation program for community-based centers in Malaysia. Int. J. Adv. Sci. Eng. Inf. Technol. **10**(2), 640–646 (2020). https://doi.org/10.18517/ijaseit.10.2.10228
Muñoz, J.E., Gonçalves, A., Rúbio Gouveia, É., Cameirão, M.S., BermúdezBadia, S.: Lessons learned from gamifying functional fitness training through human-centered design methods in older adults. Games Health J. **8**(6), 387–406 (2019). https://doi.org/10.1089/g4h.2018.0028
Nebeker, C., Murray, K., Holub, C., Haughton, J., Arredondo, E.M.: Acceptance of mobile health in communities underrepresented in biomedical research: barriers and ethical considerations for scientists. JMIR Mhealth Uhealth **5**(6), e87 (2017). https://doi.org/10.2196/mhealth.6494
Paraskevopoulos, I., Tsekleves, E.: Use of gaming sensors and customised exergames for parkinson's disease rehabilitation: a proposed virtual reality framework. In: 2013 5th International Conference on Games and Virtual Worlds for Serious Applications, VS-GAMES 2013 (2013). https://doi.org/10.1109/VS-GAMES.2013.6624247
Paraskevopoulos, I.T.: The development and applications of serious games in the public services: defence and health. **13**(2), 197–212 (2014). https://doi.org/10.1177/154411130201300209
Rego, P.A., Moreira, P.M., Reis, L.P.: Architecture for serious games in health rehabilitation paula. **276**(May 2016) (2014). https://doi.org/10.1007/978-3-319-05948-8
Yap, J.K.Y., Pickard, B.S., Chan, E.W.L., Gan, S.Y.: The role of neuronal NLRP1 Inflammasome in Alzheimer's disease: bringing neurons into the neuroinflammation game. Mol. Neurobiol. **56**(11), 7741–7753 (2019). https://doi.org/10.1007/s12035-019-1638-7
Yusoff, A., Crowder, R., Gilbert, L., Wills, G.: A conceptual framework for serious games. In: Proceedings - 2009 9th IEEE International Conference on Advanced Learning Technologies, ICALT 2009, pp. 21–23 (2009). https://doi.org/10.1109/ICALT.2009.19



Information and Communication Skills for Higher Learners Competence Model

Muhammad Hasbi Abd Rahman, Jazurainifariza Jaafar, and Miftachul Huda(✉)

Universiti Pendidikan Sultan Idris, Tanjong Malim, Perak, Malaysia
muhar@fsk.upsi.edu.my, halimelhuda@gmail.com

Abstract. The sole means through which humans may interact with others in their surroundings, either vocally or non-verbally, is through communication. One subset of soft skills is communication ability. The communication abilities of students have been improved through a variety of methods, particularly at Higher education context. Even if the improvement in the standard of academic accomplishment and student learning is regarded as positive, the effort is still thought to not have reached the right aim yet. The purpose of this study is to investigate the communication-related soft skills of university students, examine students' communication abilities through the Ulul Albab professional module, and further develop a model of exceptional university students' communication abilities based on the Ulul Albab professional module. This study used qualitative methodology including document analysis, interviews, and observations. Based on Ulul Albab's professional module, the data collected are analyzed to create a Higher Learners Competent Model (HLCM) that can improve students' soft skills, particularly communication abilities.

Keywords: Soft Skill · information skill · communication skill · Higher Learners Competent Model

1 Introduction

Particularly in the millennial era, when everyone is vying for power and influence, having strong communication skills is a significant asset. Communication abilities are a necessity that cannot be disregarded. It is important to research and gain new skills. Talking a lot or being adept at communicating does not necessarily equate to being smart, and many people start to lose their credibility as soon as they start to speak [37, 38]. Speaking is important, but possessing communication skills, or the art of conversation, is the foundation that every creature called human [1] needs to have if they want to be a great human being, be feared, and rule the world. As a result, teaching the science of communication and its applications must be given top importance when preparing a country's youth. This is so that we may address issues in our social, professional, and personal lives through efficient communication mainly with the trust insurance [39–41]. According to [2], effective communication is a fundamental requirement that complements a prosperous existence, much like our needs for food, housing, and clothing. Every person will interact or communicate with their parents, family, friends, society, coworkers, and other people.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 357–375, 2024.
https://doi.org/10.1007/978-3-031-53549-9_36

Today's graduates need to be marketable beyond just possessing a degree, diploma, or certificate; employers also value soft skills when hiring new hires. Very competitive students have strong leadership qualities, are creative, have excellent communication skills, are disciplined can adapt to any situation [42, 43]. The Covid-19 pandemic, which broke out at the end of 2019, has had a significant influence on the employment prospects for upcoming graduates who will finish their education.

The current COVID-19 virus outbreak has generated alarm on a global scale. The economic and educational sectors, in particular, have been impacted by Covid-19 [3]. According to information from the World Health Organization (WHO) up until August 26, 2022, the COVID-19 virus has caused 6,459,684 fatalities and 596,873,121 positive infections overall [4]. The economic sector in particular has been greatly impacted by the spread of COVID-19, as has the Malaysian economy as a whole, including agriculture, industry, construction, manufacturing, tourism, education, and services [3]. The Malaysian government has implemented the Movement Control Order (MCO) in all sectors affected to prevent the virus from spreading among Malaysians in response to the Covid-19 epidemic.

When all colleges and universities were shut down to stop the COVID-19 outbreak, the education industry was also impacted. In the education industry, online learning is frequently employed to replace in-person instruction, particularly in higher education institutions [5, 44, 45]. Students' soft skills are declining as a result of communication barriers between lecturers and other classmates, which is one of the repercussions of Malaysia's closure of the education system, particularly in higher learning institutes [6]. Even though students' academic performance increases when they study online, there may be communication barriers once they graduate that would limit their capacity to find employment appropriately to the expertise and experience [46, 47]. The COVID-19 epidemic, which has crippled numerous industries and raised the unemployment rate in Malaysia at the moment, is the problem that has emerged today. University graduates must then compete for jobs with workers who have prior work experience referring to the technological advancement and development [48, 49].

Every person has to be taught fundamental soft skills beginning at home and in the primary grades [7, 8]. But in higher education, students learn and practice the soft skills they will need in the profession. The significance of pupils' understanding of soft skills in upholding positive working potentials which could link into the connections in the future [50, 51]. For this reason, students should emphasize developing their soft skills to place themselves in a company where they would subsequently be able to contribute more effectively [7, 9].

Today's university graduates are believed to believe that earning a degree will enable them to land a decent job with a bright future [52, 53]. But on the other hand, from the perspective of the employer, they will look for applicants that not only possess expertise in the industry relevant to the business but also possess specific soft skills that enable these job candidates to be able to address job challenges [9, 54]. The value of an employee possessing these soft skills is predicated on the idea that they will enable them to accomplish a variety of duties assigned by the employer referring to their expertise [10, 55].

Many different methods have been employed, particularly at university, to enhance students' soft skills [56, 57]. Even while the progress in the standard of academic accomplishment and student learning is seen as good, the efforts are still thought to not have yet reached the the objective goal as their main destination journey [58, 59]. The capacity to master fundamental soft skills like communication, language, creative thinking, problem-solving, and leadership must be emphasized if graduates are to be considered marketable in today's society as the response to the contemporary demand and need amidst the technological advancement [60, 61].

Beginning in 2018, Pusat Ulul Albab Universiti Pendidikan Sultan Idris (UPSI) will start to develop a QEI Module based on the Quran, Encyclopedic, and Ijtihadik. University students who have demonstrated remarkable leadership, language, creative thinking, and problem-solving abilities have been generated by a comprehensive module that focuses on the development and improvement of soft skills and training for students.

The goal of this research is to examine the communication skills of students at the Higher Education Institute (HEI) through the Ulul Albab professional module and to build a model of excellent students in communication skills at the HEI based on the Ulul Albab professional module to meet current requirements and apply the Model Excellent students as a program that is the best example of developing a model of excellent students in communication skills.

2 Literature Review

2.1 Soft Skill Definition

Individuals' soft skills, which are generic abilities, can enhance performance, interaction, and career advancement. Individual traits from the perspective of association, personality, language proficiency, and personality are examples of soft skills [8, 62]. Soft skills are necessary for a career, especially if it involves interacting with people frequently [9]. There are many various definitions of generic skills because the emphasis of the skills learned relies on the professional field that the graduate will enter [10, 63]. According to [11], the term "generic talents" refers to abilities that enable learning in any subject and may be applied in a variety of contexts, including higher education or the job. [12] The main point refers to assert that a person's marketability in the workplace can be impacted by their traits and abilities mainly amidst the technological development [64, 65]. Few students even recognize the value of soft skills, let alone make an effort to develop them while in college.

Soft skills, according to [13], may be continuously mastered by anyone in any work and can be extensively practiced. Graduates may be able to develop soft skills in their universities, but they will encounter difficult situations that will test and hone their strategic abilities [66, 67]. Thinking, communication, creativity, management, and leadership skills, problem-solving, lifelong learning, responsibility, and social and teamwork skills are among the abilities that are frequently practiced [12].

A generic term for skills used in other nations, such as Australia, is "meta-skills", which describes specific talents or traits graduates must have to meet industry and economic expectations [14]. In the United Kingdom, the terms "score skills", "transferable

skills", and "competence skills" are used to describe the skills that must be included in every learning process at the higher education level.

Employers' desired "employability abilities" and "technical skills" are referred to as "generic skills" by [17]. How these skills are utilized or developed affects how widespread they are, according to [18]. First, "generic skills" refers to a person's unique abilities and requirements for a profession and are a confluence of "personal skills" and "workplace skills" that are adaptable to changing demands. The majority of them list "communication skills", "social and interpersonal", "planning and organization", "problem-solving", "critical thinking", "information", and "reading and writing skills" (literacy) regardless of the terms used by various countries, such as "key skills", "core skills", "transferable skills", "personal transferable skills", and "employability skills" [10, 15, 17].

The development of these two abilities will assist a person deal with a variety of situations that test the mind, particularly at work [8]. Soft skills can also be referred to as interpersonal or personal skills. Individual relationships with other people are referred to as interpersonal relationships, which involve the ability to listen, communicate, negotiate, solve problems, and make decisions [7, 13]. Personal skills, on the other hand, include the capacity for processing knowledge, the desire for lifelong learning, the capacity for critical thought and decision-making, and the ability to make plans to accomplish objectives [8]. If training soft skills will be more preferred by employers, an individual's skills will aid and be able to increase work performance.

2.2 Communication Skills

The Latin word "communis", which means oneness or fostering togetherness between two or more people, is where the word "communication" originates [16, 19]. Based on this definition, it is clear that communication encompasses all actions that include the dissemination of information to one or more parties and the subsequent exchange of feedback between the parties and the information source [68, 69]. If there is no comparable meaning, then communication will not occur since understanding the interpretation of information between the sender and recipient is a necessary component of communication. Stuart's viewpoint is further supported by [20], who claimed that management messages on communication attempt to produce a common understanding. As with [21], who describes communication as contact between the speaker and receiver that aims to establish a common understanding.

Another meaning of communication is the human practice of exchanging ideas, knowledge, or emotions. According to [22], communication is the act of interacting with another person through speech, the exchange of ideas, gestures, or symbols. This interaction takes place to accomplish objectives including information exchange, behavior modification, persuasion, and alteration of the physical and mental environments. [23] definition of communication as the process of exchanging information, views, and feelings from one person with another strengthens this term even further.

The Higher Education Institute (HEI) is a particularly social institution that places a strong emphasis on the value of excellent classroom communication in fostering universal social development [70, 71]. An interactive process and effective learning allow for student interaction, peer assistance, responsibility, self-confidence building, and other

activities. It is conducted in an engaging and enjoyable environment, which may encourage pupils to express ideas and opinions clearly [72, 73]. Effective communication also emphasizes social skills so that students may respond to their surroundings and difficulties with self-control and achieve and develop their greatness [24].

[25, 26] asserts that effective communication depends on the alignment of specific talents. Speaking skills, questioning skills, empathy skills, listening skills, and other abilities must be mastered. Students require these skills to establish effective communication with others. Students who possess these abilities can develop their capabilities because they will become more courageous, self-assured, and able to exercise more self-control [74, 75]. According to [27], communication skills are the capacity of a person to transmit or send a message to an audience. It's crucial to communicate to function in daily life. Since human beings cannot exist in one location without interacting with one another, communication activities at the same time are both a requirement and a component of society's organization [76, 77]. There are many different ways to categorize different forms of communication, but in general, depending on how it is delivered, it can be categorized as either verbal or non-verbal, including written communication.

2.3 Verbal Communication

Verbal communication is any form of communication that employs spoken or written symbols [28]. One or more syllables are used in symbols and vocal messages. Most of the words we say fall under the category of deliberate speech, which is done with the intent of orally communicating with others. In addition to being delivered orally or in writing, explicit communication also tends to be two-way, and non-verbal cues are frequently used to gauge the effectiveness of verbal communication [78, 79].

There are two main categories of verbal communication: speaking and writing and listening and reading. Written communication is verbal-nonvocal, whereas spoken communication is verbal-vocal. Speeches or discussions are examples of vocal verbal communication, whereas correspondence is an example of verbal-nonvocal communication [1]. Hearing and listening are two distinct processes; while listening merely entails picking up sound vibrations, hearing also involves deriving meaning from what is heard clearly from the information source [80, 81]. The four components of listening are listening, paying attention, understanding, and remembering. Reading is a method for learning from written material.

If the presenter uses words to convey information in this oral communication. The way that words are utilized, including their intonation, is crucial. The presenter should speak in a voice that is appropriate for the message to be delivered for the audience to hear and comprehend what is being said by the presenter more clearly [82, 83]. For instance, use a little greater volume when speaking about crucial topics to get the other person's attention.

Verbal communication is complemented by nonverbal communication. This is because verbal communication is ineffective and won't be effective if the presenter does not effectively communicate information via non-verbal cues [84, 85]. For instance, the speaker is currently communicating or conversing, whether by hand gestures, body bending, or eye contact. This demonstrates how tightly related and interdependent verbal and non-verbal communication is.

Words, thoughts, and decisions are easier to transmit verbally than nonverbally, therefore verbal communication is equally crucial [86]. The information is given more clearly and is understood by listeners or readers. Additionally, the presenter needs to confirm that the information has been understood by the audience or information recipient [87]. As a result, it's critical to take note of feedback from receivers to make sure that communication follows expectations [88, 89].

2.4 Non-verbal Communication

A communication process called non-verbal communication involves expressing ideas without using words. Nonverbal communication is carried out through body language, expressions, and eye or face contact [29]. Non-verbal communication also includes the use of physical cues like clothing, hairstyles, and other physical features, as well as vocal cues like emphasis, intonation, voice, and speaking style. Non-verbal communication is typically defined by non-verbal and non-verbal communication experts as "not utilizing words" and "not equalizing communication" [1].

Non-verbal communication refers to forms of signaling that use symbols rather than words. Here are some examples of nonverbal communication, including body language and facial emotions. Kinesics refers to the study of posture, movement, and facial expressions in the human body. Kinesics is derived from the Greek word kinesis, which means motion [30]. Body language and facial emotions are examples of motion signals in communication.

This nonverbal communication encompasses body language, gesture, facial emotion, and body shape completely (posture) [30]. In other words, the presenter can already convey the meaning he wants to convey using body language, specific facial expressions, and gestures rather than utilizing words. Sometimes facial emotions, body language, and gestures are unintentional, as when someone blushes in embarrassment or begins to perspire out of anxiety. Another illustration is how, when we have a headache, we place our palms on our eyebrows to show that we are feeling lightheaded or sick to our stomach. One example of nonverbal communication is this.

2.5 Writing Communication

Humans may communicate in writing in addition to verbally speaking and reading. Written language is actually the most advanced and sophisticated method of communication. Writing is both a skill and a means of expressing one's thoughts and feelings. The capacity to write well is usually combined with enhancing the essential platform from visual, motor, and conceptual abilities [90, 91]. Students primarily demonstrate their understanding of advanced academic subjects through writing. A person must first master mechanical writing abilities before using writing as a means of communication. Only after mastering these skills can a person use writing as a means of communication.

Information is conveyed through written communication using a variety of signs, symbols, illustrations, and graphics. The information or things that need to be communicated can be written down or printed. Delivering complex information, such as statistics and other significant data, that is difficult to express through voice or discussion, requires the use of written communication [1]. The results of this written communication can also

be continuously addressed since it allows information to be documented so that it can be used as a reference or reference at a later time [92, 93]. Some people refer to this written communication as visual communication (visual communication).

This written communication must be succinct and understandable to be effective. A good and accurate written report uses proper grammar and avoids using excessive or extraneous words to effectively deliver the necessary information [94, 95]. Letters, meeting minutes, memos, and many other written communications presented in the organization's written form are just a few instances of written communication.

2.6 One Way Communication

Self-efficacy, or the individual's conviction in their capacity to perform a task successfully, is one technique to assess the impact of training [31]. According to this theory, people can alter their behavior depending on their cognitive processes and the connections they make with their surroundings. A critical evaluation of changes in performance and behavior, which can be fostered by intervention and quality improvement programs, is encouraged by the appraisal of their skills and competencies [32]. The recipient of the information is not required to respond to the information being transmitted. More talkative individuals frequently instruct, describe, explain, and explain to listeners referring to the information properly in line with the topic being discussed [96, 97]. Additionally, it is a speech made to a sizable crowd about something. Consider lectures. Talk is a type of explanation-based communication in which the presenter covers a lot of ground in a short amount of time. A speaker must prepare by reading and selecting key materials to present a case for a dilemma and a conclusion.

2.7 Two Ways Communication

The material being presented by the presenter calls for a response or notice from the audience member. The transmitter and the receiver interact during two-way communication (receiver). When the presenter requests a response or feedback, both speakers and listeners can speak at the same time and express thoughts, ideas, or other information [33]. One of the obvious instances is the argument. Research, analysis, and discussion are all parts of the debate, which is communication. A debate is a formal discussion format in which supporters and opponents converse back and forth. When there is opposition, contesting the arguments of the proponents, a two-way dialogue will take place. Meetings, discussions, conversations, and any other kind of communication involving two or more people that asks for feedback are also counted as two communication ways.

3 Methods

A qualitative study with interviews, observations, and document analysis was employed as the research methodology. Ulul Albab's QEI Module, which was implemented in numerous IPTA/S, will be evaluated for effectiveness, and this will be done mostly through the analytical technique. There are several of them, including Universiti Pendidikan Sultan Idris (UPSI), International Islamic University Malaysia (UIAM), Universiti Malaysia Terengganu (UMT), Universiti Malaysia Sabah (UMS), Universiti

Malaysia Sarawak (UNIMAS), Universiti Teknologi Mara (UiTM) Perak Campus, Universiti Teknologi Mara (UiTM) Kedah Campus, Universiti Teknologi Mara (UiTM) Sabah Campus, Universiti Sains Islam Malaysia (USIM), Institut Perguruan Kuala Lumpur, Institut Perguruan Sabah, Selangor International Islamic University College (KUIS), SEGI College and Asia E-University.

To get pertinent information from the use of the QEI Professional Module, analysis was also done on related documents such as student performance before and after. The respondents in this survey are primarily bachelor's degree candidates, including freshmen, students in the middle of semesters, and students nearing the conclusion of semesters. In addition to speaking with nearly 20 representatives of IPTA/S and other government and non-government organizations in the Peninsula and Sabah Sarawak who were represented by the participants, the researcher also conducted data triangulation and respondents-based interviews to assess the effectiveness of the development of this Outstanding Student Model. This is consistent with [35]'s assertion that the triangulation method can be viewed as a sculpture if seen from several angles to acquire a clear picture of the sculpture's shape. Therefore, by examining the sculpture from all sides, the researchers will arrive at a persuasive conclusion [98, 99].

The results of the study on the suitability of the QEI Ulul Albab Professional Module to be used as a Higher Learners Competent Model (HLCM) that meets the current requirements of the market and employers will then be strengthened by the researcher's direct observation of respondents in the implementation and effectiveness of the Ulul Albab Professional Module. After that, professionals in the fields of soft skills and legal education will study and assess all of this data. All of this data is subsequently examined to produce research findings and conclusions that concentrate on the student communication development process and modules, with the panel or trainer placing attention on talent or participation in the modules used [100]. Students or participants will be evaluated before, during, and after the QEI Module training based on Ulul Albab and several additional modules on the aspects of public speaking, impromptu lecturing, and interview session. IAtlas was used as a tool to examine all forces to aid in the qualitative data analysis process from data organization to coding to coding to data description [101]. The examined data is then presented to specialists for confirmation of the development of the Higher Learners Competent Model at HEI.

This study is a design and development study, as indicated at the outset. [36] defines design and development research as "the systematic study of design, development, and evaluation processes to establish an empirical foundation for the creation of instructional and non-instructional products and tools, as well as new or improved modules that govern their development". It is a thorough examination of the design, development, and assessment processes to establish an empirical foundation for the creation of tools, new modules, or product alterations that affect the development of instructional or non-instructional goods [102]. Consequently, the goal of this study is to develop a Higher Learners Competent Model (HLCM). The Model Development offers broad findings and doesn't focus on the product explicitly, but rather on the practices, environments, and circumstances that support its use.

4 Analysis and Discussion

4.1 Ulul Albab QEI Profesional Module

An Ijtihadik, Encyclopedic, and Quranic-based QEI Module was developed by UPSI's Ulul Albab Center commencing in 2018. Since 2018, a complete module that emphasizes the development and improvement of soft skills and training for students has created IPT graduates that exhibit the qualities of students who thrive in areas like leadership, language, creative thinking, and exceptional problem-solving. The Quranic, Encyclopedic, and Ijtihadik components of this QEI module total three [34]. The Basic Module, Intermediate Module, and Advanced Module are the three (3) modules that make up the Ulul Albab Professional Program. The Quranic, Encyclopedic, and Ijtihadik components are the same in each of the Ulul Albab Professional Program's three modules.

The main element of Ulul Albab's three parts is the Quranic. Along with its root word, which is derived from the words of the Al-Quran itself, "Quranic" refers to imparting knowledge and understanding of the essence of the teachings from the Al-Quran itself to people [34]. This knowledge and understanding include aspects of reading, *tajwid*, memorization, *tadabbur*, *I'jaz* (miracles/science), *ibrah*, as well as applications in daily life. The Al-Quran generation's building based on Quranic verses is following the Sunnah's prescriptions. This center, the sole Ulul Albab Center in Malaysia and the entire world takes the initiative and duty of creating Quranic-based rules that are intended to develop the next generation of Al-Quran. There are three levels to this module (Basic, Intermediate, and Advanced) The Al-Quranic component of ULUL Albab's Professional Module has been demonstrated to be a comprehensive and useful module to be implemented for the Malaysian community, especially among those who yearn for the knowledge and abilities of Al-Quran translated in the form of structured and organized modules. The Quranic component is controlled by expert figures who have expertise across the fields of religion and science.

The Ulul Albab model's Encyclopedic concept emphasizes and develops communication skills in addition to a variety of information and other abilities, including fluency in foreign languages. Encyclopedias play a significant role in the development of characters. Students can apply their Encyclopedic intelligence through the Encyclopedic Component Module as a preparation to become skilled future leaders and become community referral sources. *Ijtihadik* is implemented based on a method of problem-solving that is directed by wise thinking and action to produce something with originality and invention, in keeping with the overall meaning of the term "Ulul Albab" itself, which refers to people who are constantly thinking (Fig. 1).

4.2 Higher Learners Competent Models (HLCM)

According to the results of the study done in 2020–2022, which used the three modules of the Ulul Albab Professional Program, a considerable improvement may be shown following the use of this module. The majority of respondents, particularly new students, as well as students in the middle and after the semester, lack the courage to speak up and engage in conversation, where this is significant contribution in the competent model [103, 104]. The self-confidence of students with poor communication skills was



Fig. 1. Ulul Albab Profesional Program [34]

observed to increase when the Ulul Albab Professional Program Module was implemented at many higher education institutions. Similar to this, students whose communication skills were previously ordinary are improving as a result of their training in the Ulul Albab Professional Program. Students who already have strong communication skills can benefit from this Ulul Albab Professional Program to "Sharpen" their leadership abilities. Before, during, and following the program, the students or participants are assessed based on their public speaking, impromptu lecturing, and interviewing skills.

The outcomes of interviews with various responders and third parties to get feedback on the efficacy of this QEI Module to confirm this conclusion. This can be demonstrated by the fact that those respondents who followed the QEI Module's implementation were chosen to serve on the Student Representative Council at their respective universities. Some respondents serve as hosts for university-sponsored activities [105, 106]. As a result, the Higher Learners Competent Model (HLCM) was created to help students strengthen their communication abilities. Innovative and creative thinking, communication strategies, and information distribution are all components of HLCM. Students that thrive in soft skills can be created using the three HLCM components. In the communication component, the emphasis is placed on the student's capacity to communicate ideas, cultivate personal communication skills, practice active listening, and offer feedback.

The Higher Learners Competence Model is depicted in Fig. 2. The component of communication skills is emphasized as the most crucial soft skill by the Higher Learners Competent Model (HLCM). The four components of the HLCM approach are also more focused on encouraging students to engage in one-on-one or group communication [107, 108]. Students that can think critically and solve problems in a variety of contexts have communication skills that are a part of their creative and inventive thinking. Future innovators can be produced by students who are creative and innovative in HLCM.

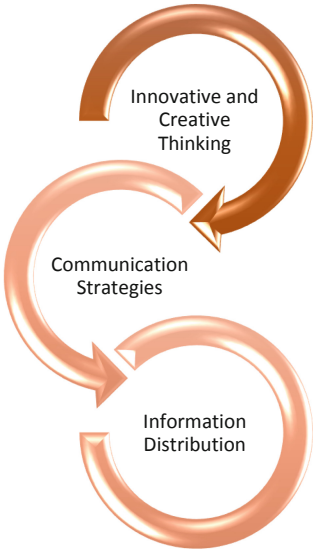


Fig. 2. Higher Learners Competence Model (HLCM)

This paradigm can be applied to both verbal and non-verbal communication techniques. Techniques that improve verbal abilities have a significant impact on students. Verbal communication is the act of communicating through speech [109, 110]. Verbal communication can take place in person or face-to-face situations, such as over the phone or in person. The students have the opportunity to hone their listening, speaking, and language abilities during this oral communication [109, 110]. To receive meaningful feedback or ideas, students might practice their listening abilities. Students are capable of knowing the ins and outs of speaking and speaking well in terms of speaking and language abilities. Students can assure efficient communication in this speaking and language skills course by using simple language, a regulated tone of voice, reasoned argumentation, and repeated repetition of crucial points [111]. When it comes to non-verbal communication, students can use good role models, take care of their appearance, attitude, and body language, as well as a psychological approach to persuade listeners to pay attention to what they are attempting to say.

To guarantee that the communication is efficient and has an impact on the listener, it is crucial to how the information is delivered [112, 113]. The benefit of sharing this knowledge with kids is that it can promote leadership qualities and boost communication confidence. When this information is conveyed effectively, the audience will recall the presenter as having clear talents, being confident and outgoing, active and knowledgeable in the subject matter, and delivering information that is both fascinating and pertinent [114, 115]. Depending on the appropriateness of the event being held, a student's ability to control the audience when presenting information can help them build rapport with them, draw attention to themselves and their message, and increase attention and focus [116]. It can also help make the presentation more interactive.

5 Conclusion and Recommendation

Communication is crucial since it allows people to engage and connect. The process of communication starts with the explosion of abstract concepts that are born in the mind of someone who wants to receive or communicate information that is sent directly or indirectly by using written, spoken, or visual codes.

We are born into a world of communication that is full of occurrences that may be observed by human senses and processed by the mind to produce specific meanings, which is why communication is so crucial to humans. Humans are compelled to engage in communication activities including talking, phoning, discussing, lecturing, offering advice, providing briefings, and so forth.

One of the specific purposes of communication is to help two people build relationships with one another. When two people engage, they exchange information, and this exchange creates meaning, which serves as a bond between the two parties. The more information that is communicated, the more meaning is created, and the greater the improvement in understanding between people. They then get to know each other's personalities and actions, which ultimately fosters a sense of camaraderie between them. Additionally, communication has the power to affect other people's attitudes, values, and beliefs. A person's capacity to affect how other people think is a crucial element in obtaining success in any management organization, job, or personal life. Additionally, communication links various community groups disseminate communal information and warnings and spreads social culture.

If a communication effort satisfies three criteria—understanding, maintaining the target's confidence, and follow-up action. It is considered to be effective. Understanding is the act of receiving a message in a way that is both comprehensible to the listener and consistent with the sender's intended meaning. If the message sent is comprehended but not the same as what the message's sender intended, communication is said to be ineffective in this context. Due to differences in age, experience, education, beliefs, culture, and other factors, this occurs. Additionally, this occurs because listeners, in particular, do not exercise active listening, which entails being prepared to listen, question something if not understood, focus on the primary point, and if required, take down what is said.

Communication is effective if the listener of the message follows it up with action that is consistent with the intended goal of the message communicated. Follow-up behaviors can take many different forms, including actions, verbal responses, body language, and more. Communication impediments are more accurately and methodically identified with the aid of effective communication. Channeling instructions, fostering cooperation, and building a sense of teamwork all depend on effective communication. Effective communication can also enhance productivity and job satisfaction while resolving the majority of organizational issues.

In addition to helping to resolve issues that arise, effective communication can also strengthen interpersonal bonds. According to communication experts, most issues are caused by ineffective communication, and most problems may be solved by effective communication. Understanding the tenets and theories of effective communication and putting them into practice can foster peace between professors and students at the higher education level. In a classroom setting, communicating refers to the process by which

lecturers share experiences with their students, and conversely, it also refers to the process by which students share their experiences with the lecturer. As professors and students exchange knowledge and experience, they genuinely develop a deeper understanding of one another. Students with effective communication skills are those who are liked, popular, and successful in the circle class. Effective communication requires a person to possess a variety of abilities, such as those for negotiation, persuasion, and public speaking.

Students need to take their communication conduct seriously and with concern. By taking this action, uncertainty in communication that leads to a message being misunderstood or misread can be avoided. The foundation of harmonious human interactions is effective communication. This hormone connection promotes cooperative problem solving, precise direction mobilization, and teamwork-based morale building within the workplace.

The construction of HLCM based on the Ulul Albab Professional Program Module serves as the foundation for teaching and enhancing students' soft skills, particularly communication abilities at each level with a formative assessment that may determine a student's advancement to a higher level. Students who excel in soft skills—which are the primary needs and assessments of businesses and society toward contemporary students—are produced by HLCM based on Ulul ALbab's Professional Program Module.

References

1. Ismail, M.Z., Yaacob, R., Abd Rahman, H., Awang Da, N.F., Zakaria, S.: Asas Komunikasi Untuk Pemimpin Pendidikan. Genting Highlans: Institut Aminuddin Baki (2020)
2. Wok, S., Ismail, N., Hussain, M.Y.: Teori-teori komunikasi. PTS Professional Publishing Sdn. Bhd, Kuala Lumpur (2006)
3. Mohamed, A.K., Salaeh, A., Halim, A.A., Omar, S.A.T.S.: Kesan Covid-19 Terhadap Kebolehpasaran Graduan IPT: Pelan Tindakan Kepada Pelajar Fakulti Pengajian Quran & Sunnah. USIM, pp. 788–799 (2020)
4. World Health Organization. https://covid19.who.int/
5. Shukri, S.H.D.B.M., Yunus, M.M.: Kesan pandemik COVID-19 terhadap pencapaian tahap akademik murid. J. Melayu Sedunia **4**(1), 313–343 (2021)
6. Othman, M.Z.: Bagaimana COVID-19 Mempengaruhi Kemahiran Berkomunikasi Graduan Pada Masa Depan? DagangNews.com (2021). https://dagangnews.com/bagaimana-covid-19-mempengaruhi-kemahiran-berkomunikasi-graduan-pada-masa-depan-9568
7. Cimatti, B.: Definition, development, assessment of soft skills and their role for the quality of organizations and enterprises. Int. J. Qual. Res. **10**(1), 97 (2016)
8. Noah, J.B., Abdul Azizm, A.: A Systematic review on soft skills development among university graduates. EDUCATUM J. Soc. Sci. **6**(1), 53–68 (2020). https://doi.org/10.37134/ejoss.vol6.1.6.2020
9. Pachauri, D., Yadav, A.: Importance of soft skills in teacher education programme. Int. J. Educ. Res. Technol. **5**(1), 22–25 (2014)
10. Badusah, J. (ed.): Model Pembangunan Kemahiran Insaniah (Soft Skills) untuk Institusi Pengajian Tinggi Malaysia. Penerbit Universiti Putra Malaysia, Serdang (2006)
11. Jelas, Z., Azman, N.: Generic skills provision in higher education. Int. J. Learn. **12**(5) (2005)

12. Fahimirad, M., Nair, P.K., Kotamjani, S.S., Mahdinezhad, M., Feng, J.B.: Integration and development of employability skills into malaysian higher education context: review of the literature. Int. J. High. Educ. **8**(6), 26–35 (2019)
13. Tsaoussi, A.I.: Using soft skills courses to inspire law teachers: a new methodology for a more humanistic legal education. Law Teach. **54**(1), 1–30 (2020)
14. Pitman, T., Broomhall, S.: Australian universities, generic skills and lifelong learning. Int. J. Lifelong Educ. **28**(4), 439–458 (2009)
15. Brown, J.: Training needs assessment: a must for developing an effective training program. Public Pers. Manage. **31**, 569–578 (2002). https://doi.org/10.1177/009102600203100 41
16. Brown, B.L.: Generic skills in career and technical education myths and realities. ED472363 (2002). http://www.ericfacility.net/extra/search2/view.html
17. Mahamod, Z., Alias, A., Said, N.E.M.: Perbandingan kemahiran generik pelajar opsyen bahasa Melayu di Sebuah IPTA dan IPGM. Asean J. Teach. Learn. High. Educ. **4**(1), 82–95 (2012)
18. Grapragasem, S., Krishnan, A., Mansor, A.N.: Current trends in malaysian higher education and the effect on education policy and practice: an overview. Int. J. High. Educ. **3**(1), 85–93 (2014)
19. Hall, S.: Cultural studies 1983. In: Cultural Studies 1983. Duke University Press (2016)
20. Deacon, D., Pickering, M., Golding, P., Murdock, G.: Researching Communications: A Practical Guide to Methods in Media and Cultural Analysis. Bloomsbury Publishing, USA (2021)
21. Yoder, D.D., Wallace, S.P., Hugenberg, L.W.: Creating Competent Communication. Kendall/Hunt Publishing Company (1996)
22. Abdul Mua'ti @ Zamri Ahmad. Santun komunikasi. Kuala Lumpur: Dewan Bahasa dan Pustaka (2001)
23. Bolarinwa, J.A., Olorunfemi, D.Y.: Organizational communication for organizational climate and quality service in academic libraries. Libr. Philos. Pract. **305**, 1–5 (2009)
24. Dazali, N.S.M., Awang, M.I.: Tahap kemahiran komunikasi dalam kalangan pelajar sarjana muda pendidikan IPTA di utara semenanjung Malaysia. J. Pendidikan Bahasa Melayu **4**(2), 44–56 (2016)
25. Ibrahim, N.A.N., Mahbob, M.H.: Keterampilan Peribadi Pelajar Sebagai Satu Bentuk Kemahiran Komunikasi dan Kepentingannya dalam Pasaran Dunia Pekerjaan. J. Komunikasi: Malays. J. Commun. **37**(1), 209–226 (2021). https://doi.org/10.17576/JKMJC-2021-3701-12
26. Aw, S.: Komunikasi Interpersonal, PT. Graha Ilmu, Yogyakarta (2011)
27. Hafied, C.: Pengantar Ilmu Komunikasi. Raja Grafindo Persada, Jakarta (2011)
28. Megawati, W., Hartono, R.: The impact of teachers' verbal and non-verbal communication on students' motivation in learning English. Engl. Educ. J. **10**(4), 436–448 (2020). https://doi.org/10.15294/eej.v10i4.39157
29. Ananda, J., Solihat, D., Suryana, Y.: Nonverbal communication performed by foreign English teacher. Indon. EFL J. **6**(2), 175–188 (2020). https://doi.org/10.25134/ieflj.v6vi2.3424
30. Hans, A., Hans, E.: Kinesics, haptics and proxemics: aspects of non-verbal communication. IOSR J. Human. Soc. Sci. (IOSR-JHSS) **20**(2), 47–52 (2015)
31. Mata, Á.N.D.S., et al.: Training in communication skills for self-efficacy of health professionals: a systematic review. Hum. Resour. Health **19**(1), 1–9 (2021). https://doi.org/10.1186/s12960-021-00574-3
32. Bandura, A.: Self-efficacy: toward a unifying theory of behavioral change. Psychol. Rev. **84**(2), 191–215 (1977). https://doi.org/10.1037/0033-295X.84.2.191
33. Barida, M., Muarifah, A.: The role of teachers' communication skills on children' well-being in school. In: The 2nd International Conference on Child-Friendly Education (ICCE) (2018)

34. Mohamed, F., et al.: Program Profesional Ulul Albab. Pusat Ulul Albab UPSI (2022)
35. Chua, Y.P.: Kaedah dan statistik penyelidikan: Kaedah penyelidikan. Buku 1. McGraw Hill Education, Kuala Lumpur (2006)
36. Richey, R., Klein, J.: Design and Development Research: Methods, Strategies and Issues. Lawrence Erlbaum Associates, Mahwah (2007)
37. Huda, M.: Empowering application strategy in the technology adoption: insights from professional and ethical engagement. J. Sci. Technol. Policy Manag. **10**(1), 172–192 (2019)
38. Yousefi, S., Tosarkani, B.M.: Exploring the role of blockchain technology in improving sustainable supply chain performance: a system-analysis-based approach. IEEE Trans. Eng. Manag. (2023)
39. Huda, M.: Trust as a key element for quality communication and information management: insights into developing safe cyber-organisational sustainability. Int. J. Organ. Anal. (2023). https://doi.org/10.1108/IJOA-12-2022-3532
40. Syofiarti, S., Saputra, R., Lahmi, A., Rahmi, R.: The use of audiovisual media in learning and its impact on learning outcomes of Islamic cultural history at Madrasah Tsanawiyah Negeri 4 Pasaman. Firdaus J. **1**(1), 36–44 (2021). https://doi.org/10.37134/firdaus.vol1.1.4.2021
41. Huda, M.: Towards digital access during pandemic age: better learning service or adaptation struggling? Foresight **25**(1), 82–107 (2023). https://doi.org/10.1108/FS-09-2021-0184
42. Huda, M.: Digital marketplace for tourism resilience in the pandemic age: voices from budget hotel customers. Int. J. Organ. Anal. **31**(1), 149–167 (2023). https://doi.org/10.1108/IJOA-10-2021-2987
43. Zainul, Z., Rasyid, A., Nasrul, W.: Distance learning model for Islamic religious education subjects in non Internet Server Provider (ISP) areas. Firdaus J. **1**(1), 45–53 (2021). https://doi.org/10.37134/firdaus.vol1.1.5.2021
44. Huda, M.: Between accessibility and adaptability of digital platform: investigating learners' perspectives on digital learning infrastructure. High. Educ. Skills Work-Based Learn. (2023, ahead-of-print). https://doi.org/10.1108/HESWBL-03-2022-0069
45. Huda, M., et al.: Strategic role of trust in digital communication: critical insights into building organizational sustainability. In: Proceedings of the Future Technologies Conference, pp. 387–403. Springer Nature Switzerland, Cham, October 2023
46. Abdul Aziz, N.A., Mohd Razali, F., Saari, C.Z.: Penggunaaan Media Sosial dari Perspektif Psiko Spiritual Islam. Firdaus J. **2**(1), 65–75. https://doi.org/10.37134/firdaus.vol2.1.6.2022
47. Hanafi, H.F., et al.: A review of learner's model for programming in teaching and learning. J. Adv. Res. Appl. Sci. Eng. Technol. **33**(3), 169–184 (2024)
48. Huda, M., Bakar, A.: Culturally responsive and communicative teaching for multicultural integration: qualitative analysis from public secondary school. Qual. Res. J. (2024, ahead-of-print). https://doi.org/10.1108/QRJ-07-2023-0123
49. Syafri, N., Ali, A.H., Ramli, S.: Penggunaan kaedah Inovasi Sambung dan Baca Bahasa Arab (SaBBAr) dalam meningkatkan kemahiran membaca perkataan Bahasa Arab murid di Sekolah Rendah. Firdaus J. **2**(2), 62–71. https://doi.org/10.37134/firdaus.vol2.2.6.2022
50. Huda, M., et al.: Enhancing digital leadership direction: insight into empowering gender violence prevention. In: Mishra, D., et al. (eds.) Communication Technology and Gender Violence. Springer (2024)
51. Ab Rahim, N.M.Z., Saari, Z., Mohamad, A.M., Rashid, M.H., Mohamad Norzilan, N.I.: Konsep Ulul Albab dalam Al-Quran dan hubungannya dengan pembelajaran kursus Sains, Teknologi dan Manusia di UTM Kuala Lumpur. Firdaus J. **2**(2), 72–78 (2022). https://doi.org/10.37134/firdaus.vol2.2.7.2022
52. Huda, M., et al.: Understanding of digital ethics for information trust: a critical insight into gender violence anticipation. In: Mishra, D., et al. (eds.) Communication Technology and Gender Violence. Springer (2024)

53. Hanafi, H.F., Wahab, M.H.A., Selamat, A.Z., Masnan, A.H., Huda, M.: A systematic review of augmented reality in multimedia learning outcomes in education. In: Singh, M., Kang, DK., Lee, J.H., Tiwary, U.S., Singh, D., Chung, W.Y. (eds.) IHCI 2020. LNCS, vol. 12616, pp. 63–72. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68452-5_7
54. Khalili, Huda, M., Rosman, A.S., Mohamed, A.K., Marni, N.: Digital learning enhancement through social network site (SNS). In: Software Engineering Application in Informatics: Proceedings of 5th Computational Methods in Systems and Software 2021, vol. 1, pp. 421–431. Springer (2021)
55. Leh, F.C., Anduroh, A., Huda, M.: Level of knowledge, skills and attitude of trainee teachers on Web 2.0 applications in teaching geography in Malaysia schools. Heliyon **7**(12) (2021)
56. Sukadari, S., Huda, M.: Culture sustainability through co-curricular learning program: learning Batik cross review. Educ. Sci. **11**(11), 736 (2021)
57. Huda, M., Gusmian, I., Mulyo, M.T.: Towards eco-friendly responsibilities: Indonesia field school model cross review. J. Comp. Asian Dev. (JCAD) **18**(2), 1–12 (2021)
58. Rachman, A., Oktoviani, I., Manurung, P.: Digitization of Zakat and Charity BAZNAS Tangerang City through crowdfunding platform tangerangsedekah. id. Firdaus J. **3**(1), 96–106 (2023). https://doi.org/10.37134/firdaus.vol3.1.9.2023
59. Huda, M., Hashim, A.: Towards professional and ethical balance: insights into application strategy on media literacy education. Kybernetes **51**(3), 1280–1300 (2022)
60. Mohamad Shokri, S.S., Salihan, S.: Modul pembangunan Program Huffaz Profesional Universiti Tenaga Nasional: Satu pemerhatian kepada konstruk pembangunan Al-Quran. Firdaus J. **3**(2), 12–23 (2023). https://doi.org/10.37134/firdaus.vol3.2.2.2023
61. Berlian, Z., Huda, M.: Reflecting culturally responsive and communicative teaching (CRCT) through partnership commitment. Educ. Sci. **12**(5), 295 (2022)
62. Huda, M.: Towards an adaptive ethics on social networking sites (SNS): a critical reflection. J. Inf. Commun. Ethics Soc. **20**(2), 273–290 (2022)
63. Omar, M.N.: Inovasi pengajaran & pemudahcaraan menggunakan Aplikasi Ezi-Maq (MAHARAT AL-QURAN) untuk menarik minat pelajar menguasai Ilmu Tajwid. Firdaus J. **2**(2), 79–87 (2022). https://doi.org/10.37134/firdaus.vol2.2.8.2022
64. Huda, M., Sutopo, L., Liberty, Febrianto, Mustafa, M.C.: Digital information transparency for cyber security: critical points in social media trends. In: Advances in Information and Communication: Proceedings of the 2022 Future of Information and Communication Conference (FICC), vol. 2, pp. 814–831. Springer, Cham, March 2022
65. Borham, A.H., et al.: Information and communication ethics in social media for Indigenous people's religious understanding: a critical review. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer (2024)
66. Ali, A.H.: Pelajar Khalifah Profesional Tempaan Ulul Albab sorotan penerapannya berasaskan komponen QEI di UPSI. Firdaus J. **3**(2), 51–63 (2023). https://doi.org/10.37134/firdaus.vol3.2.5.2023
67. Jusoh, A., Huda, M., Abdullah, R., Lee, N.: Development of digital heritage for archaeo-visit tourism resilience: evidences from e-lenggong web portal. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer (2024)
68. Rahim, M.M.A., Huda, M., Borham, A.H., Kasim, A.Y., Othman, M.S.: Managing information leadership for learning performance: an empirical study among public school educators. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer (2024)
69. Hehsan, A., et al.: Digital Muhadathah: framework model development for digital arabic language learning. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer Nature Switzerland, Cham (2024)

70. Muharom, F., Farhan, M., Athoillah, S., Rozihan, R., Muflihin, A., Huda, M.: Digital technology skills for professional development: insights into quality instruction performance. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer Nature Switzerland, Cham (2024)
71. Muhamad, N., Huda, M., Hashim, A., Tabrani Z.A., Maarif, M.A.: Managing technology integration for teaching strategy: public school educators' beliefs and practices. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer Nature Switzerland, Cham (2024)
72. Wahid, A., et al.: Digital technology for indigenous people's knowledge acquisition process: insights from empirical literature analysis. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer Nature Switzerland, Cham (2024)
73. Huda, M., Taisin, J.N., Muhamad, M., Kiting, R., Yusuf, R.: Digital technology adoption for instruction aids: insight into teaching material content. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer Nature Switzerland, Cham (2024)
74. Alwi, A.S.Q., Ibrahim, R.: Isu terhadap Penggunaan Teknologi Media Digital dalam kalangan guru pelatih jurusan Pendidikan Khas. Firdaus J. **2**(2), 88–93 (2022). https://doi.org/10.37134/firdaus.vol2.2.9.2022
75. Zamri, F.A., Muhamad, N., Huda, M., Hashim, A.: Social media adoption for digital learning innovation: insights into building learning support. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer Nature Switzerland, Cham (2024)
76. Huda, M., et al.: Trust in electronic record management system: insights from Islamic-based professional and moral engagement-based digital archive. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)
77. Yahya, S.F., Othman, M.A.: Penggunaan video dalam Pengajaran dan Pembelajaran Pendidikan Moral Tingkatan 2. Firdaus J. **2**(2), 94–105 (2022). https://doi.org/10.37134/firdaus.vol2.2.10.2022
78. Huda, M., et al.: Empowering Islamic-based digital competence and skills: how to drive it into reconstructing safety strategy from gender violence. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)
79. Cita Sari, D., et al.: Transformation of artificial intelligence in Islamic Edu with Ulul Albab Value (global challenge perspective). Firdaus J. **3**(1), 1–9 (2023). https://doi.org/10.37134/firdaus.vol3.1.1.2023
80. Huda, M., et al.: From digital ethics to digital community: an Islamic principle on strengthening safety strategy on information. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)
81. Zamri, F.A., Muhamad, N., Huda, M.: Information and communication technology skills for instruction performance: beliefs and experiences from public school educators. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)
82. Masud, A., Borham, A.H., Huda, M., Rahim, M.M.A., Husain, H.: Managing information quality for learning instruction: insights from public administration officers' experiences and practices. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)
83. Huda, M., et al.: Digital record management in Islamic education institution: current trends on enhancing process and effectiveness through learning technology. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer Nature Switzerland, Cham (2024)

84. Mohd Nawi, M.Z.: Media variations in education in Malaysia: a 21st century paradigm. Firdaus J. **3**(1), 77–95 (2023). https://doi.org/10.37134/firdaus.vol3.1.8.2023
85. Huda, M., et al.: From technology adaptation to technology adoption: an insight into public school administrative management. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
86. Tan, A.A., Huda, M., Rohim, M.A. Hassan, T.R.R., Ismail, A.: Chat GPT in supporting education instruction sector: an empirical literature review. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
87. Wahid, A., et al.: Augmented reality model in supporting instruction process: a critical review. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
88. Musolin, M.H., Serour, R.O.H., Hamid, S.A., Ismail, A., Huda, M., Rohim, M.A.: Developing personalized Islamic learning in digital age: pedagogical and technological integration for open learning resources (OLR). In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
89. Musolin, M.H., Ismail, M.H., Huda, M., Hassan, T.R.R., Ismail, A.: Towards an Islamic education administration system: a critical contribution from technology adoption. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
90. Musolin, M.H., Ismail, M.H., Farhan, M., Rois, N., Huda, M., Rohim, M.A.: Understanding of artificial intelligence for Islamic education support and service: insights from empirical literature review. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
91. Susilowati, T., et al.: Getting parents involved in child's school: using attendance application system based on SMS gateway. Int. J. Eng. Technol. **7**(2.27), 167–174 (2018)
92. Huda, M., et al.: From digital ethics to digital partnership skills: driving a safety strategy to expand the digital community? In: Digital Transformation for Business and Society: Contemporary Issues and Applications in Asia. Taylor & Francis, UK (2023a)
93. Susilowati, T., et al.: Learning application of Lampung language based on multimedia software. Int. J. Eng. Technol. **7**(2.27), 175–181 (2018)
94. Abadi, S., et al: Application model of k-means clustering: insights into promotion strategy of vocational high school. Int. J. Eng. Technol. **7**(2.27), 182–187 (2018)
95. Huda, M., et al.: Towards digital servant leadership for organisational stability: driving processes in the pandemic age? In: Digital Transformation for Business and Society: Contemporary Issues and Applications in Asia. Taylor & Francis, UK (2023b)
96. Aminudin, N., et al.: The family hope program using AHP method. Int. J. Eng. Technol. **7**(2.27), 188–193 (2018)
97. Wulandari, et al: Design of library application system. Int. J. Eng. Technol. **7**(2.27), 199–204 (2018)
98. Aminudin, N., et al.: Higher education selection using simple additive weighting. Int. J. Eng. Technol. **7**(2.27), 211–217 (2018)
99. Zainuri, A., Huda, M.: Empowering cooperative teamwork for community service sustainability: insights from service learning. Sustainability **15**(5), 4551 (2023)
100. Maseleno, A., et al.: Hau-Kashyap approach for student's level of expertise. Egypt. Inform. J. **20**(1), 27–32 (2019)

101. Huda, M., Teh, K.S.M.: Empowering professional and ethical competence on reflective teaching practice in digital era. In: Dikilitas, K., Mede, E., Atay, D. (eds.) Mentorship Strategies in Teacher Education, pp. 136–152. IGI Global, Hershey (2018). https://doi.org/10.4018/978-1-5225-4050-2.ch007
102. Huda, M., Teh, K.S.M., Nor, N.H.M., Nor, M.B.M.: Transmitting leadership based civic responsibility: insights from service learning. Int. J. Ethics Syst. **34**(1), 20–31 (2018a). https://doi.org/10.1108/IJOES-05-2017-0079
103. Huda, M., Mulyadi, D., Hananto, A.L., Nor Muhamad, N.H., Mat Teh, K.S., Don, A.G.: Empowering corporate social responsibility (CSR): insights from service learning. Soc. Responsib. J. **14**(4), 875–894 (2018b)
104. Huda, M., Yusuf, J.B., Jasmi, K.A., Nasir, G.A.: Understanding comprehensive learning requirements in the light of al-Zarnūjī's Ta'līm al-Muta'allim. Sage Open **6**(4), 1–14 (2016a)
105. Huda, M., Yusuf, J.B., Jasmi, K.A., Zakaria, G.N.: Al-Zarnūjī's concept of knowledge ('ilm). SAGE Open **6**(3), 1–13 (2016b)
106. Huda, M., Jasmi, K.A., Mohamed, A.K., Wan Embong, W.H., Safar, J.: Philosophical investigation of Al-Zarnuji's Ta'lim al-Muta'allim: strengthening ethical engagement into teaching and learning. Soc. Sci. **11**(22), 5516–5551 (2016c)
107. Huda, M., et al.: Innovative teaching in higher education: the big data approach. Turk. Online J. Educ. Technol. **15**(Spec. Issue), 1210–1216 (2016)
108. Othman, R., Shahrill, M., Mundia, L., Tan, A., Huda, M.: Investigating the relationship between the student's ability and learning preferences: evidence from year 7 mathematics students. New Educ. Rev. **44**(2), 125–138 (2016)
109. Anshari, M., Almunawar, M.N., Shahrill, M., Wicaksono, D.K., Huda, M.: Smartphones usage in the classrooms: learning aid or interference? Educ. Inf. Technol. **22**(6), 3063–3079 (2017)
110. Huda, M., Sabani, N., Shahrill, M., Jasmi, K.A., Basiron, B., Mustari, M.I.: Empowering learning culture as student identity construction in higher education. In: Shahriar, A., Syed, G. (eds.) Student Culture and Identity in Higher Education, pp. 160–179. IGI Global, Hershey (2017a). https://doi.org/10.4018/978-1-5225-2551-6.ch010
111. Huda, M., Jasmi, K.A., Alas, Y., Qodriah, S.L., Dacholfany, M.I., Jamsari, E.A.: Empowering civic responsibility: insights from service learning. In: Burton, S. (ed.) Engaged Scholarship and Civic Responsibility in Higher Education, pp. 144–165. IGI Global, Hershey (2017f). https://doi.org/10.4018/978-1-5225-3649-9.ch007
112. Huda, M., et al.: Innovative e-therapy service in higher education: mobile application design. Int. J. Interact. Mob. Technol. **11**(4), 83–94 (2017g)
113. Huda, M., et al.: From live interaction to virtual interaction: an exposure on the moral engagement in the digital era. J. Theor. Appl. Inf. Technol. **95**(19), 4964–4972 (2017j)
114. Huda, M., Maseleno, A., Jasmi, K.A., Mustari, I., Basiron, B.: Strengthening interaction from direct to virtual basis: insights from ethical and professional empowerment. Int. J. Appl. Eng. Res. **12**(17), 6901–6909 (2017k)
115. Huda, M., Haron, Z., Ripin, M.N., Hehsan, A., Yaacob, A.B.C.: Exploring innovative learning environment (ILE): big data era. Int. J. Appl. Eng. Res. **12**(17), 6678–6685 (2017l)
116. Maseleno, A., et al.: Combining the previous measure of evidence to educational entrance examination. J. Artif. Intell. **10**(3), 85–90 (2017)

Twitter Sentiment Analysis with Machine Learning for Political Approval Rating

Rodrigo Loayza Abal(✉), Juan J. Soria, and Lidia Segura Peña

Universidad Tecnológica del Perú, Lima, Peru
u18101212@utp.edu.pe

Abstract. This study provides an insightful analysis of Peruvian political approval rating sentiment, using Twitter data and applying various ranking algorithms. Despite the challenging context of political instability, the research achieved a high degree of accuracy, with Linear SVC classification leading the way. The detailed breakdown of the results and the evident correlation between tweet sentiment and real-world events provide compelling validation of the chosen methods. The authors' exploration of various machine learning techniques amplifies the relevance of the study. The research collected 8274 tweets from the @presidenciaperu account, employing API v2 during the month of April 2023, regarding the government's political approval rating, with the objective of identifying the accuracy of Machine Learning algorithms NB multinomial, NB Bernoulli, Support Vector linear classifier, Logistic regression classifier and KNeighbors classifier from the sentiment analysis of Tweets in Spanish language. Tweets were processed using PLN, words were vectorized with the bag of words algorithm, allowing to build a vocabulary of 5773 tweets with negative (0) and positive (1) polarity in tweets with the support of Python and BETO. Five machine learning sentiment analysis techniques were compared, resulting in an accuracy of 96.3636% (F1Score = 0.98) for the linear SVC, an accuracy of 95.3246% (F1Score = 0.98) for the KNeighbors classifier, an accuracy of 95.2380% (F1Score = 0.98) for the logistic regression classifier, and a tie in accuracy of 94.8051% for NB multinomial and NB Bernoulli. The results indicate that the optimal algorithm was the Support Vector linear classifier with 96.3636% accuracy applied in a Peruvian political approval index environment.

Keywords: Machine Learning · Twitter · Linear SVC · Bernoulli NB · KNN classifier · logistic regression classifier · sentiment analysis

1 Introduction

The political approval index has become a very important information tool in the governance of a country, since it allows politicians who hold a government position to know and respect the public opinion of the citizens who elected them in relation to their commitments and promises made during the campaign. Knowing the public opinion of citizens is important for understanding and shaping the type of government in operation in a country [1]. Currently, the approval index is an information tool widely used by

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 377–397, 2024.
https://doi.org/10.1007/978-3-031-53549-9_37

governments to know the degree of acceptance of citizens in relation to their political actions [2]. Also defined as a metric, the political approval index is elaborated by several pollsters in Peru, and issued on a monthly basis, with the purpose of providing information regarding the level of support that a politician or political party receives from the population over a period of time to obtain this information, a sample of the population is surveyed using statistical methods [3–5]. In a recent opinion poll conducted by the Institute of Peruvian Studies (IEP), it is mentioned that the current administration of the President of the Republic reached its highest level of disapproval in her term of office, reaching 80% disapproval in June 2023 [5, 6]. This reflects an overall negative outlook for the month of June 2023 in relation to his administration. Although the information issued by these studies is reliable, it should be taken into account that they cannot capture the most recent opinions of people due to the speed with which they change, so in some cases they reflect an outdated image [2]. As a result of the coup d'état and attempted dissolution of Congress by former Peruvian President Pedro Castillo, the current Peruvian President Dina Boluarte assumed the presidential mandate in December 2022, reaching 71% of citizen disapproval according to IEP, and at the end of January she reached 76% of disapproval, which stands out as a critical month in her administration, due to a significant increase of 5% of citizen disapproval compared to the following months, which increased in a 1% [7, 8]. During this critical month for the current government there was a progressive impact on presidential disapproval related to each of the stages of the protests, i.e., each action taken by the government to control the situation had a negative impact on the level of presidential approval and on the intensity of the demonstrations, which generated several tragedies in Peru, and likewise, the abuse of police power manifested by the citizens [7, 9, 10]. Due to their monthly issuance, the studies issued by the polling entities do not show a degree of specificity in short periods of time for better governmental decision making. Due to this problem, there are currently several investigations that exploit and use alternative data sources for the analysis of the approval rate in real time. In this sense, social networks have become a great alternative source of information for this type of process, since they present a simple dynamic in which people can share information, their opinions and express their emotions in relation to a specific topic [11]. In Peru, between 2021 and 2022, an increase in the use of social networks was reported, reaching 83.8% of the total population [12]. The increase in the use of social networks has generated a large amount of data, providing an opportunity to capture and analyze public opinion. Among the social networks most used by politicians is Twitter, a platform that presents a simple experience to share information with thousands of people [2]. Today, politicians and political parties actively use these platforms as strategic tools to share messages and interact with citizens, which has generated a great change in the way politics is done [13]. Lu *et al.* [14] mentions that, during the 2016 U.S. presidential election, politicians and political parties intensively leveraged social networks to interact with voters, with the social network Twitter being the most used by both political parties. This phenomenon evidenced a significant change in the political sphere of a technologically advanced country, where networks had a great impact on political communication. This new way of doing politics was also adapted in other countries, such as El Salvador. In the 2019 presidential elections in El Salvador, social networks established themselves as the main means of political communication for the

political party that won the majority of votes in the country for the first time in its history [13]. This event represented a change in the campaign strategies of the candidates in the country's future elections.

The issuance of these large volumes of data has manifested itself in a significant increase in research related to sentiment analysis (SA) and the use of data generated in social networks by applying Natural Language Processing (NLP) and Machine Learning techniques (ML) [1, 11, 15, 16]. Much of these studies have shown high accuracy in predicting pass rate using AS, supervised learning (SL) algorithms and social network data, among the best performing supervised learning algorithms in the researches are mentioned Support Vector Machine (SVM), K-nearest neighbor (KNN) and Naive Bayes (NB) algorithm [1, 2, 17, 18].

The present research aims to evaluate the precision, F1 score and accuracy of the most widely used supervised learning algorithms in the field of sentiment analysis and data classification, such as Bernoulli Naive Ba-yes (BNB), Multinomial Naive Bayes (MNB), Support Vector Classification (SVC), Logistic Regression Classifier (LRC) and k-Neighbors Classifier (KNF). These algorithms were used to analyze a dataset concerning the Peruvian presidential account in Spanish on the social network Twitter, to propose a supervised Machine Learning algorithm to optimally determine the accuracy of the political approval index in Peru.

2 Data Collection and Processing

The research firstly had a previous registration on the Twitter Developer Platform website, secondly, access to the academic version of the Twitter API was requested, which offers greater access to information and a larger monthly limit of tweets. The request for the academic version of the API began on April 8 and ended on April 22, 2023. Thirdly, through the use of the R programming language and its editor R Studio, a total of 8,724 tweets were extracted in Spanish related to the @presidenciaperu account during the period from April 29 to May 8, 2023. It should be noted that we did not extract the tweets related to the Twitter account of the current president Peru that had as user @DinaErcilla, because it was deactivated on January 31, 2023 [19]. This strategic approach allowed us to obtain a representative sample of the opinions of Twitter users in relation to the presidential account, thus generating an updated database to carry out sentiment analysis in the political context of Peru. The data collected from the API were also stored in an Excel file for further statistical processing.

Data cleaning is a fundamental step in sentiment analysis, because the quality and integrity of the data directly influence the results obtained. In the context of sentiment analysis, unprocessed data often have noise and non-relevant elements that can negatively affect the accuracy of the results [2]. Data cleaning focuses on removing unwanted elements that are not relevant to the analysis. By using Python programming language version 3 and importing the necessary libraries such as Pandas, RE and TLTK, we proceeded to remove special characters, punctuation marks, numeric characters, usernames that continued from an at (@), URLs and the text of tweets were converted to lowercase, which simplifies the text and allows a focus on the keywords that provide the meaning to the text. In addition, lemmatization and stemming were applied, which allows words

to be reduced to their root form and helps to group similar terms and to obtain a more general view of the opinions expressed. Likewise, tokenization and normalization were applied to ensure that the data are structured in a uniform way, which facilitates their processing. In this way, cleaning helps to improve the quality and validity of the results obtained, by eliminating noise and making the data fit correctly.

To obtain the polarity, either positive, negative, or neutral, of each of the tweets in our dataset to evaluate the supervised learning classification algorithms, we used the BETO model, a pre-trained model designed for natural language processing in Spanish. BETO is the Spanish version of BERT, which is a bidirectional transformer model developed by Google researchers [20]. To use this model, the transformers library was installed, and its pipeline class was imported to run the sentiment analysis model called "finiteautomata/beto-sentiment-analysis" from the Hugging Face platform [21–23]. The result of this analysis was downloaded in an XLSX format file for further evaluation with the classification algorithms.

3 Methodology

The text classification problem is given an information $d \in$ X of a document, where X is the high-dimensional document space and a fixed set of classes or categories $\mathbb{C} = \{c_1, c_2, c_3, \cdots, c_J\}$, where the classes are defined by humans for the needs of an application. A training set D of labeled documents <d, c> is given, where $\langle d, c\rangle \in X \times \mathbb{C}$.

$$\langle d, c\rangle = \langle government, ministers, peru, dina, justice, presidency\rangle$$

Using a machine learning algorithm, it is desired to learn a classifier or classification function γ:X $\rightarrow$ C that assigns documents to classes with five classification algorithms [24].

3.1 Machine Learning Algorithms

Machine Learning allows to automatically analyze information in a supervised way, using techniques to catalog a set of data [25]. Classification algorithms find patterns in the input data and classify them into various sections, and then compare those same data and place them into a section, predicting which topic it refers to. The classification algorithms selected in this study are multinomial NB, Bernoulli NB, linear Support Vector classifier, Logistic regression classifier and KNeighborsclassifier, because they are much more flexible and simple algorithms [26].

Linear Support Vector Classifier

The principle of support vector machines was invented by Vapnik in 1982 at AT&T laboratories, as one of the results of computational learning theory [27]. This algorithm can be used for both regression and classification. SVM searches for a hyperplane, called a support vector, of separation with the largest possible distance. In Fig. 1, a line is shown that linearly separates a group of data, classifying them into two classes correctly. Thus, when a new element is entered, the side of the hyperplane will decide to which class to

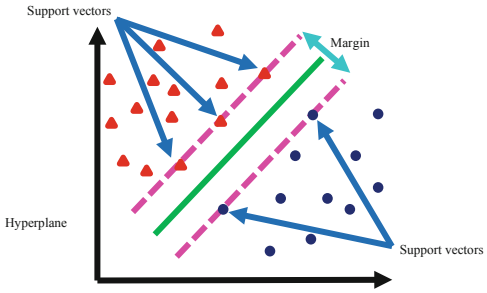


Fig. 1. Gráfica de Support *Vector Machine*

assign it. It is important to note that the larger the margin, the more likely it is that the classifier will work with future data [28].

The hyperplane separated by potentially better classes (among an infinite number of candidates) is identified using support vector machine (SVM) technology. It attempts to maximize the distance between the closest positive and negative instances of the hyperplane. Therefore, the task of machine learning is to determine the support vector that maximizes the margin. The specific mechanisms for finding the optimal set of support vectors exceeding the ambition of an introductory text [29].

The Machine Learning Support Vector Machine (SVM) classification model aims to find a hyperplane h of dimension (n − 1) that separates the Tweets labeled −1 from those labeled +1 with a maximum margin ρ by finding closer points between the classes forming the support vectors, declaring the best line that divides the two classes and at the same time maximizes the distance of the hyperplane to the support vectors (r) as shown in Fig. 2.

The Support Vector Machine model finds w and b such that $\rho = \frac{2}{\|w\|}$ is maximized and satisfies Eq. (1) [30]

$$\begin{cases} w^T x_i + b \geq 1, si\ y_i = 1 \\ w^T x_i + b \leq -1, si\ y_i = -1 \end{cases}; \forall (x_i; y_i), i = 1, 2, 3, .., n \tag{1}$$

Require: $\boldsymbol{X}$ and $\boldsymbol{y}$ loaded with training labeled data, $\boldsymbol{\alpha}$ <= 0 or $\boldsymbol{\alpha}$ <= partially trained SVM

1. C <= some value (10 for example)
2. **repeat**
3. **for all** $\{x_i, y_i\}, \{x_j, y_j\}$ do
4. Optimize α_i and α_j
5. **end for**
6. **until** no changes in α or other resource constraint criteria met

Ensure: Retain only the support vectors ($\alpha_i > 0$)

Fig. 2. Pseudocode Support Vector Machine

Naive Bayes Classifier

It is a multinomial probabilistic learning method, where the probability that a document d belongs to the class c is calculated by (2)

$$P(c|d) \propto P(c) \prod_{1 \leq k \leq n_d} P(t_k|c) \tag{2}$$

where $P(t_k|c)$ is the conditional probability that the term t_k correctly appears in class c in the document, where the objective is to find the best a posteriori most probable class c for the document, represented by Eq. (3)

$$C_{map} = \underset{c \in \mathbb{C}}{\operatorname{argmax}} \hat{P}(c|d) = \underset{c \in \mathbb{C}}{\operatorname{argmax}} \hat{P}(c) \prod_{1 \leq k \leq n_d} \hat{P}(t_k|c) \tag{3}$$

where $P(c)$ y $P(t_k|c)$ are estimated from the training set, with the pseudocode shown in Fig. 3 [24].

TRAINMULTINOMIALNB (C, D)
1. V ← EXTRACTVOCABULARY (D)
2. N ← COUNTDOCS (D)
3. for each $c \in C$
4. do N_c ← COUNTDOCSINCLASS (D, c)
5. prior[c] ← $\frac{N_c}{N}$
6. $text_c$← CONCATENATETEXTOFALLDOCSINCLASS (D, c)
7. for each $t \in V$
8. do T_{ct} ←COUNTTOKENSOFTERM ($text_c$, t)
9. for each $t \in V$
10. do condprob [t][c] ← $\frac{T_{ct}+1}{\sum_{t'} (T_{tc'}+1)}$
11. return V, prior condprob

APPLYMULTINOMIALNB (C, V, prior, condprob, d)
1. w ← EXTRACTTOKENSFROMDOC (V, d)
2. for each $c \in C$
3. do score [c] ← logprior[c]
4. for each $t \in W$
5. do score [c] += logcondprob[t][c]
6. return arg $max_{c \in C}$ $score$ [c]

Fig. 3. Naive Bayes algorithm (multinomial): Training and testing

Multinomial Model NB

To classify by levels a query q as used in Eq. (4)

$$P(d|q) \propto P(d) \prod_{t \in q} (1 - \lambda) P(t|M_c) + \lambda P(t|M_d) \tag{4}$$

In the particular case we have the Eq. (4).

The multinomial NB model in the particular case, when $\lambda = 1$ Eq. (5) is obtained with the pseudocode shown in Fig. 4. [24]

$$P(d|q) \propto P(d) \prod_{t \in q} P(t|M_d) \tag{5}$$

TRAINBERNOULLINB (C, D)
1. V ← EXTRACTVOCABULARY (D)
2. N ← COUNTDOCS (D)
3. for each $c \in C$
4. do N_c ← COUNTDOCSINCLASS (D, c)
5. prior[c] ← $\frac{N_c}{N}$
6. for each $t \in V$
7. do N_{ct} ←COUNTDOCSINCLASSCONTAININGTERM(D, c, t)
8. condprob [t][c] ← $\frac{N_{ct}+1}{N_c+2}$
9. return V, prior, condprob
APPLYBERNOULLINB (C, V, prior, condprob, d)
1. V_d ←EXTRACTTERMSFROMDOC (V, d)
2. for each $c \in C$
3. do score [c] ← logprior[c]
4. for each $t \in V$
5. do if $t \in V_d$
6. then score[c] +=logcondprob[t][c]
7. else score[c] +=log (1- condprob[t][c])
8. return arg $max_{c \in C}$ $score$ [c]

Fig. 4. NB algorithm (bernoulli model): Training and testing

Logistic Regression Classification Model

The objective of binary logistic regression is to train a classifier that makes a binary decision about the class of a new input observation, which is defined by Eq. (6) [30]

$$decision(x) = \begin{cases} 1; \ if \ P(y = 1|x) > 0.5 \\ 0; \ otherwise \end{cases} \tag{6}$$

Uses the sigmoid classifier for the only input $x = [x_1, x_2, x_3, x_4, \cdots, x_n]$ which is defined by $\sigma(z) = \frac{1}{1+e^{-z}}$ [31],

where the classifier can be 1 which are observations that belong to the class or 0 which are the observations that do not belong to the class defined in Eq. (7), (8) which we want to know the probability $P(y = 1|x)$ that this observation belongs to the class (Fig. 5) [24], where the decision is defined by:

$$decision(x) = \begin{cases} 1; \ if \ P(y = 1|x) > 0.5 \\ 0; \ otherwise \end{cases} \tag{7}$$

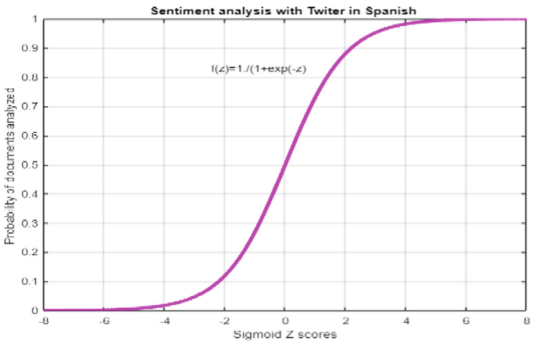


Fig. 5. Sigmoid logistic regression function

In addition

$$\begin{cases} \mathrm{P(y=1)} = \frac{1}{1+\mathrm{e}^{-(\mathrm{wx+b})}} \\ \mathrm{P(y=0)} = \frac{\mathrm{e}^{-(\mathrm{wx+b})}}{1+\mathrm{e}^{-(\mathrm{wx+b})}} \end{cases} \tag{8}$$

K-nearest Neighbors

The k nearest neighbors determines the decision boundary locally. For 1NN we assign each document to the class of its nearest neighbor. For kNN we assign each document to the majority class of its k nearest neighbors where k is a parameter. According to the contiguity hypothesis, we expect that a test document d has the same label as the training documents located in the local region surrounding a ***d*** [32]. The decision limits on 1NN are concatenated segments of the Voronoi tessellation that decomposes each document object into cells, in which the plane is divided into convex polygons, each containing its corresponding document. The KNN algorithm can categorize its dataset using a test query similarity function by considering the k training data examples closest to the query being tested. A graphical representation of the K-nearest neighbor is visualized in Fig. 6 [33].

Figure 7 shows three classes presidency, government, and ministers in a two-dimensional space, where documents are shown as ellipses, diamonds and Xs. The decision limits were chosen to separate the classes, where classifying a new document determines the region in which it is located and we assign it the class presidency.

Figure 6 shows the pseudocode of the KNN nearest neighbor algorithm KNN, in its training and testing [24, 34].

The centroid of a class c is computed as the vector average or center of mass of its members with (9)

$$\vec{\mu}(c) = \frac{1}{|D_c|} \sum_{d \in D_c} \vec{v}(d) \tag{9}$$

where D_c is the set of documents in whose class it is found [35] (Fig. 8).

TRAIN-**kNN** (C, D)
1. $D' \leftarrow$ PREPROCESS (D)
2. k $\leftarrow$ SELECT-k (C, D')
3. Return D', k
APPLY-kNN (C, D', k, d)
1. $S_k \leftarrow$ COMPUTENEARESTNEIGHBORS (D', k, d)
2. for each $c_j \in C$
3. do $p_j \leftarrow \frac{|S_k \cap c_j|}{k}$
4. return arg $max_j\, p_j$

Fig. 6. KNN algorithm: Training and testing

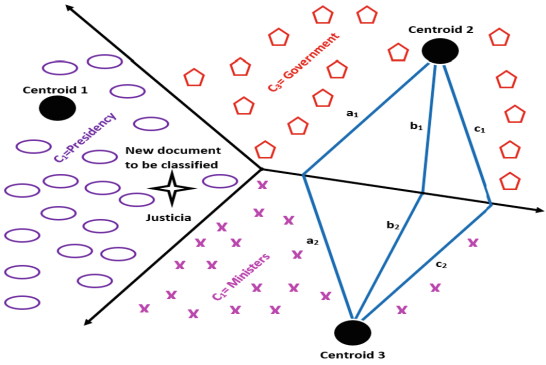


Fig. 7. KNN classification scheme

TRAINROCCHIO ($\mathbb{C}, D$)
1. for each $c_j \epsilon D$
2. do $D_i \leftarrow \{d: < d, c_j > \in D\}$
3. $\overrightarrow{\mu_j} \leftarrow \frac{1}{|D_j|} \sum_{d \in D_j} \vec{v}(d)$
4. return $\{\overrightarrow{\mu_1}, \overrightarrow{\mu_2}, \ldots, \overrightarrow{\mu_j}\}$
APPLYROCCHIO $(\{\overrightarrow{\mu_1}, \overrightarrow{\mu_2}, \ldots, \overrightarrow{\mu_j}\}, d)$
1. return arg $min_i\, |\overrightarrow{\mu_i} - \vec{v}(d)|$

Fig. 8. Rocchio classification training and testing

Cosine Similarity

To measure the distance between two documents, the cosine similarity was applied, which means that the closer two word vectors are, the smaller the angle between them and, therefore, the greater the cosine of the angle between the vectors represented in Eq. (2).

The vector derived from document d is denoted by $\vec{v}(d)$ with a component in the vector for each dictionary term [36]. The similarity quantization between two documents d_1 y d_2 in this vector space, is determined by the cosine similarity of their vector representations $\vec{V}(d_1)$ and $\vec{V}(d_2)$ respectively, represented in Eq. (10) [37]

$$sim(d_1, d_2) = \frac{\vec{V}(d_1).\vec{V}(d_2)}{\left\|\vec{V}(d_1)\right\|.\left\|\vec{V}(d_2)\right\|} \tag{10}$$

If the equation is normalized, it can be rewritten by the Eq. (11) (Fig. 9).

$$sim(d_1, d_2) = \vec{v}(d_1).\vec{v}(d_2) \tag{11}$$

$$\text{where} \begin{cases} \vec{v}(d_1) = \frac{\vec{V}(d_1)}{\left|\vec{V}(d_1)\right|} \\ \vec{v}(d_2) = \frac{\vec{V}(d_2)}{\left|\vec{V}(d_2)\right|} \end{cases} \tag{12}$$

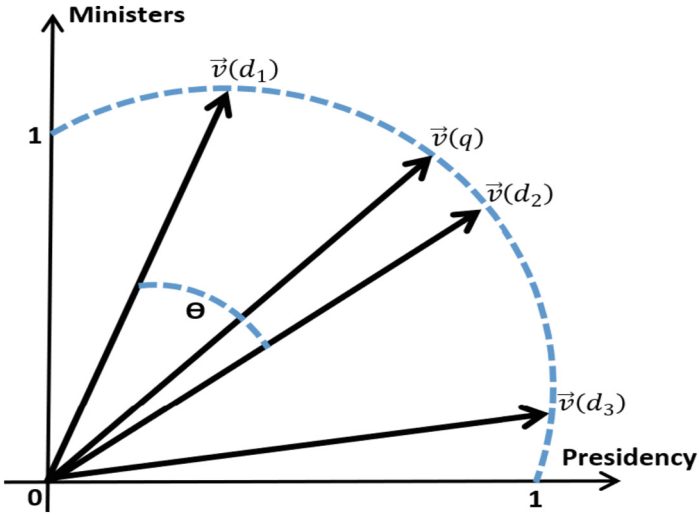


Fig. 9. Cosine similarity

3.2 Method

The method of sentiment analysis with Twitter user API data was developed in 10 stages [38] point out that it starts with data collection, which seeks to extract information from the Twitter API Application Programming Interface. To perform the AS, machine learning methods were applied, which consists of classification algorithms for large volumes of data. Then, data preprocessing was performed, through various techniques that performed the cleaning and normalization of the texts. The techniques used in this study were text filtering, tokenization, transformed cases and empty word elimination, as shown in the flow chart in Fig. 10.

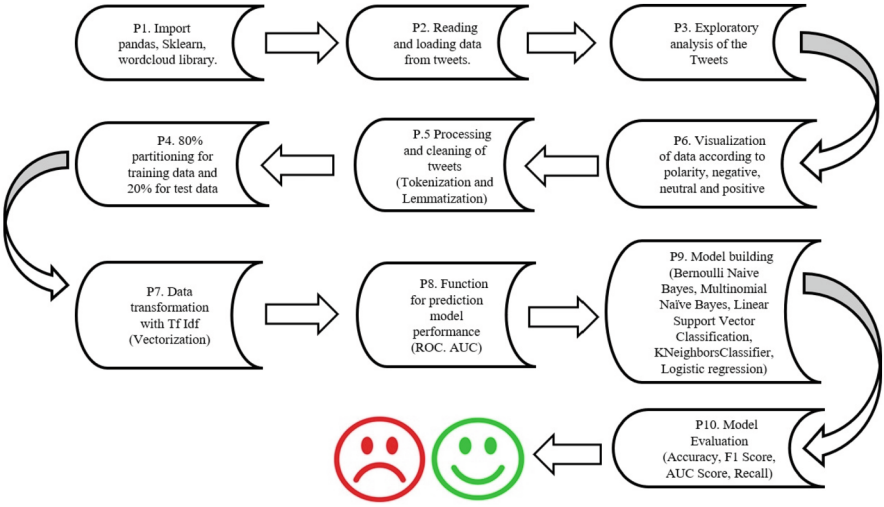


Fig. 10. Sentiment analysis process of Tweet information

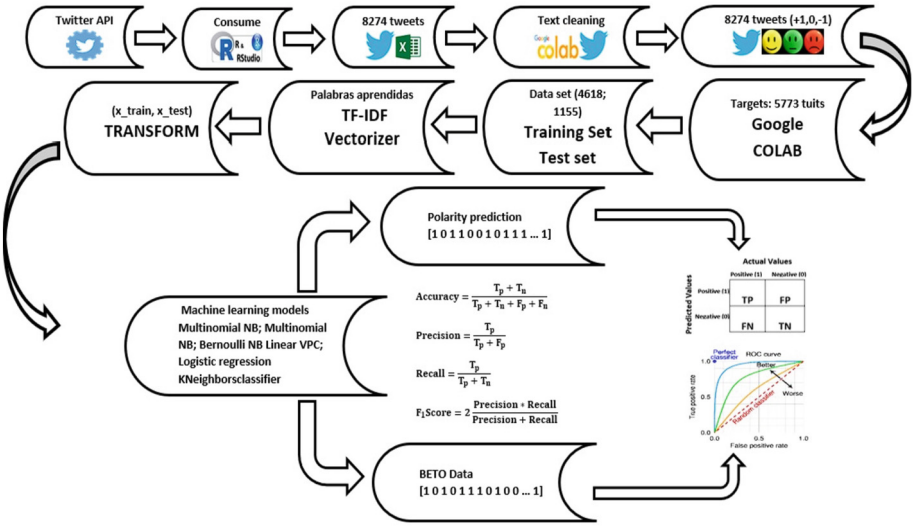


Fig. 11. Research work method

The natural language processing (NLP) computational linguistics and text analysis allows identifying and extracting subjective information from resources [39]. We evaluated word similarity of sentiment lexicons with a precision metric of natural language processing (NLP) computational linguistics and text analysis allows identifying and extracting subjective information from resources [39]. We assessed word similarity of sentiment lexicons with a polarity consistency accuracy metric of polarity between each sentiment word and its N nearest N words with Eq. (13) shown [40] (Fig. 11).

$$Accuray = \frac{\sum_{i=1}^{\#Lex} \sum_{j=1}^{N} \beta\left(w_i; c_{ij}\right)}{\#Lex \times N} \tag{13}$$

where #Lex is the number of words of the sentiment lexicon, w_i is the i-ésima word word closer to w_i in the lexicon with cosine similarity and $\beta\left(w_i; c_{ij}\right)$ is a function that indicates an image 1 si w_i y c_{ij} have the same polarity and 0 otherwise (Fig. 12).

$$\begin{aligned} \textbf{Accuracy} &= \frac{\mathbf{T_p}+\mathbf{T_n}}{\mathbf{T_p}+\mathbf{T_n}+\mathbf{F_p}+\mathbf{F_n}} & \textbf{Precision} &= \frac{\mathbf{T_p}}{\mathbf{T_p}+\mathbf{F_p}} \\ \textbf{Recall} &= \frac{\mathbf{T_p}}{\mathbf{T_p}+\mathbf{T_n}} & \mathbf{F_1}\textbf{Score} &= 2\frac{\textbf{Precision}*\textbf{Recall}}{\textbf{recision}+\textbf{Recall}} \end{aligned} \tag{14}$$

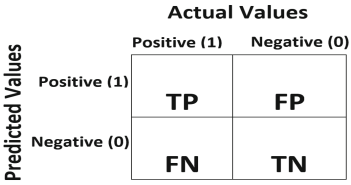


Fig. 12. Indicadores de la matriz de confusión

4 Results

The results of the analysis of users' tweets against the Peruvian presidential account @presidenciaperu, were performed in sub-processes of statistical and econometric analysis under the Machine Learning algorithms under study, analyzing the structure and interpretation of the mini texts.

4.1 Partial Results of Document Analysis with Logistic Regression

The text shown in Fig. 13 was analyzed.

Defining the text variables in Table 1 we found the following probabilities from the sentiment analysis of Twitter users.

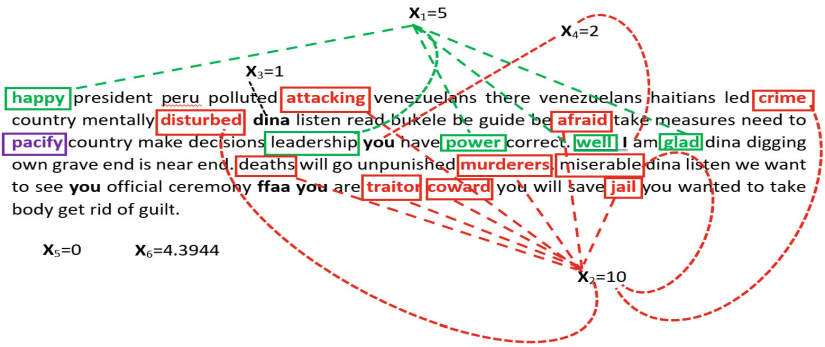


Fig. 13. Mini document Tweets with six analyzed characteristics of vector x

Table 1. Texto Text with six variables analyzed with Logistic Regression

Var	Definition	Value in Fig. 13
x_1	Count (positive lexicon words $\in$ doc)	5
x_2	Count (negative lexicon words $\in$ doc)	10
x_3	$\begin{cases} 1;\ if\ \text{dina}\ \in doc \\ 0;\ otherwise \end{cases}$	1
x_4	Count (1st and 2nd pronouns $\in$ doc)	2
x_5	$\begin{cases} 1;\ if\ !\in doc \\ 0;\ otherwise \end{cases}$	0
x_6	Ln (Word count of doc)	$Ln(81) = 4.3944$

Six features were taken for text analysis from Fig. 13 with a vector of weights $\overline{x} = \left[\frac{5}{2}, -1, \frac{-6}{5}, \frac{-1}{2}, 3, \frac{1}{10}\right]$ and with bía of b = 0.1, in which the logistic regression model obtained a probability for the positive tokens of 69.83472604234761%

$$
\begin{aligned}
&P(+|x) = P(y = 1|x) = \sigma(w.x + b) \\
&= \sigma\left(\left[\tfrac{5}{2}, -1, \tfrac{-6}{5}, \tfrac{-1}{2}, 3, \tfrac{1}{10}\right] \cdot [5, 10, 1, 2, 0, 4.3944] + 0.1\right) \\
&= \sigma\left(\left[\tfrac{5}{2}(5) - 1(10) - \tfrac{6}{5}(1) + \tfrac{1}{2}(2) + (3)(0) + \tfrac{3}{10}(4.3944)\right] + 0.1\right) \\
&= \sigma([0.73944] + 0.1) \\
&= \sigma(0.83944)
\end{aligned}
$$

$$P(+|x) = P(y = 1|x) = \sigma(0.83944) = \frac{1}{1 + e^{-0.83944}} = 0.6983472604234761$$

For negative tokens, the probability obtained was 69.83472604234761% cuyo calcula se muestra en el siguiente procedimiento

$$P(-|x) = P(y = 0|x) = 1 - \sigma(0.83944) = 0.3016527395765239$$

4.2 Partial Result of Analysis with Multinomial Naive Bayes Algorithms

In the application of the Naive Bayes multinomial classification algorithm, the probabilities of the tokens analyzed were found and are shown in Table 2.

Table 2. Analysis of Mini Tweets with Naive Bayes

	docID	Word in document	in c = president
Training Test	1	government, ministers, government	Yes
	2	government, government, peru	Yes
	3	government, dyne	Yes
	4	boluarte, justice, government	No
Test set	5	government, government, government, boluarte, justice	

Four terms were extracted $n_d = \{t_1, t_2, t_3, t_4, t_5\}$, where $t_1 = government$, $t_2 = government$, $t_3 = government$, $t_4 = boluarte$, $t_5 = justice$ and the probabilities for each category were obtained using the Eq. (15) [41].

$$\hat{P}(t|c) = \frac{T_{tc} + 1}{\left(\sum_{t' \in V} T_{ct'}\right) + B'} \tag{15}$$

The following probabilities were obtained:

$$\hat{P}(t_1|c) = \hat{P}(gobierno|c) = \frac{5+1}{8+6} = \frac{3}{14}; \hat{P}(t_2|c) = \hat{P}(boluarte|c) = \frac{0+1}{8+6} = \frac{1}{14}$$

$$\hat{P}(t_3|c) = \hat{P}(justicia|c) = \frac{0+1}{8+6} = \frac{1}{14}; \hat{P}(t_1|\overline{c}) = \hat{P}(gobierno|\overline{c}) = \frac{1+1}{3+6} = \frac{2}{9}$$

$$\hat{P}(t_2|\overline{c}) = \hat{P}(boluarte|\overline{c}) = \frac{1+1}{3+6} = \frac{2}{9}; \hat{P}(t_3|\overline{c}) = \hat{P}(justicia|\overline{c}) = \frac{1+1}{3+6} = \frac{2}{9}$$

Luego se aplicó la ecuación (16),

$$P(c|d) \propto P(c).\prod_{1 \le k \le n_d} P(t_k|c) \tag{16}$$

obtained the ranked probabilities

$$P(c|d_5) \propto P(c).[P(t_1|c).P(t_2|c).P(t_3|c).P(t_4|c).P(t_5|c)]$$
$$P(c|d_5) \propto \tfrac{3}{4}.\left[\left(\tfrac{3}{7}\right)\left(\tfrac{3}{7}\right)\left(\tfrac{3}{7}\right)\left(\tfrac{1}{14}\right)\left(\tfrac{1}{14}\right)\right] = \tfrac{81}{268912} = 0.000301237799$$
$$P(\overline{c}|d_5) \propto P(c).[P(t_1|c).P(t_2|c).P(t_3|c).P(t_4|c).P(t_5|c)]$$
$$P(\overline{c}|d_5) \propto \tfrac{1}{4}.\left[\left(\tfrac{2}{9}\right)\left(\tfrac{2}{9}\right)\left(\tfrac{2}{9}\right)\left(\tfrac{2}{9}\right)\left(\tfrac{2}{9}\right)\right] = \tfrac{32}{236196} = 0.0001354807024$$

The classifier assigns the test document to c = president. The classification decision is that the three occurrences of the positive indicator government in d5 outweigh the occurrences of the two negative indicators boluarte and justice.

4.3 Partial Result of the Analysis with Bernoulli Algorithms

Applying Bernoulli's model for table no. Xxx, four terms were extracted $n_d = \{t_1, t_2, t_3, t_4, t_5\}$, where $t_1 = government$, $t_2 = government$, $t_3 = government$, $t_4 = boluarte$, $t_5 = justice$ and the probabilities for each category were obtained using the Eq. (17) [36].

$$\hat{P}(t|c) = \frac{N_{tc}+1}{N_c+2}; \text{in addition} \quad P(c|d) \propto \prod_{1\leq i\leq M} P(U_i = e_i|c) \tag{17}$$

The following probabilities were obtained: $\hat{P}(gobernment|c) = \frac{3+1}{3+2} = \frac{4}{5}$; $\hat{P}(boluarte|c) = \frac{0+1}{3+2} = \frac{1}{5}$

$$\hat{P}(justice|c) = \frac{0+1}{3+2} = \frac{1}{5}; \hat{P}(minister|c) = \hat{P}(dina|c) = \hat{P}(peru|c) = \frac{2}{5}$$
$$\hat{P}(government|\overline{c}) = \frac{1+1}{1+2} = \frac{2}{3}; \hat{P}(boluarte|\overline{c}) = \frac{1+1}{1+2} = \frac{2}{3}$$
$$\hat{P}(justice|\overline{c}) = \frac{1+1}{1+2} = \frac{2}{3}; \hat{P}(minister|c) = \hat{P}(dina|c) = \hat{P}(peru|c) = \frac{1}{3}$$

The following equation was then applied (18) [24].

$$P(c|d) \propto P(c).\prod_{1\leq k\leq n_d} P(t_k|c) \tag{18}$$

obtaining the ranked probabilities for the positive tokens

$$\hat{P}(c|d_5) \propto \hat{P}(c).\hat{P}(gobgovernment|c).\hat{P}(justice|c).\hat{P}(boluarte|c)$$
$$\left[1-\hat{P}(minister|c)\right]\left[1-\hat{P}(dina|c)\right]\left[1-\hat{P}(peru|c)\right]$$

$$\hat{P}(c|d_5) \approx \frac{3}{4}.\frac{4}{5}.\frac{1}{5}.\frac{1}{5}.\left(1-\frac{2}{5}\right).\left(1-\frac{2}{5}\right)\left(1-\frac{2}{5}\right) \approx \frac{3}{125}.\left(\frac{3}{5}\right)\left(\frac{3}{5}\right)\left(\frac{3}{5}\right) \approx 0.005184$$

Likewise, negative tokens were obtained.

$$P(\overline{c}|d_5) \propto \frac{1}{4}.\left[\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{3}\right)\right] =$$
$$\frac{1}{4}.\left[\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\right] = \left(\frac{16}{729}\right) = 0.0219478737997$$

4.4 Analysis of Similarity in Tweets with Cosine Similarity

Table 3. Analysis of mini words with cosine similarity

docID	Word in document	Words	Vector frequency
1	government, government, government, boluarte, justice	Government, boluarte, justice	(3,1,1)
2	boluarte, justice, government	boluarte, justice, government	(1,1,1)

Taking the documents from the Table 3, $\vec{V}(d_1) = (3, 1, 1)$; $\vec{V}(d_2) = (1; 1, 1)$ a cosine similarity is obtained with the Eq. (19) [42].

$$sim(d_1, d_2) = \frac{\vec{V}(d_1).\vec{V}(d_2)}{\left\|\vec{V}(d_1)\right\|.\left\|\vec{V}(d_2)\right\|} \tag{19}$$

$$\left\|\vec{V}(d_1)\right\| = \sqrt{3^2 + 1^2 + 1^2} = \sqrt{11}; \quad \left\|\vec{V}(d_2)\right\| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$$

$$sim(d_1, d_2) = \frac{(3,1,1).(1;1;1)}{\sqrt{11}.\sqrt{3}} = \frac{5}{\sqrt{15}} = 1.2909944487358056$$

The radial angle is 0.276165236973, which means that the similarity between documents 1 and 2 have a slight similarity in their wording as it is close to zero.

4.5 Analysis of Tweets with the Five Machine Learning Algorithms

The Python tool was used to process the 5773 clean tweets under the Machine Learning classification algorithms, where the database and codes are hosted on the GitHub account https://github.com/RodLA/sa_twitter_president_peru to corroborate the results shown in Table 4.

Table 4. Results of the classification algorithms

Models	Value	Precision	Recall	F1-Score	accuracy
Multinomial NB	0	0.95	1.00	0.97	0.948051
	1	1.00	0.05	0.09	
Bernoulli NB	0	0.95	1.00	0.97	0.948051
	1	0.00	0.00	0.00	
Linear SVC	0	0.96	1.00	0.98	0.963636
	1	0.96	0.35	0.51	
Logistic Regression	0	0.95	1.00	0.98	0.952380
	1	1.00	0.13	0.23	
KNeighbors classifier	0	0.97	0.99	0.98	0.953246
	1	0.62	0.38	0.47	

Table 5 shows the different Tweets collected cleanly criticizing the Peruvian presidential account, confirming that the polarity of the texts is negatively inclined (Fig. 14, 15 and 16).

Table 5. Mini results of tweets from the @presidenciaperu account.

N°	Tweets
1	happy president of peru urged to attack venezuelans there venezuelans haitians led crime country mentally deranged
2	dina listen read read bukele be guide be afraid take measures need to pacify country make decisions leadership you have power correct
3	well i'm glad dina digging own grave end about end about end about
4	respects cashier cerron endorsing congratulating ffaa police kill civilians
5	murderer jail awaits you soon people will have justice
6	killings will go unpunished murderers
7	chabelita was going to lose merit medal give them pleasure rocío veronika sigrid mirtha company adjust
8	be careful, get your hands in caviar dina boluarte be careful, we are watching you
9	well dina first woman president peru glad elections paraguay carried inconveniences give room left generates hunger division violence ideology popular humanist capitalism facilitates poor entrepreneurship
10	government fight corruption front remain doubts coherence appoint minister justice friend chosen white collars suitability

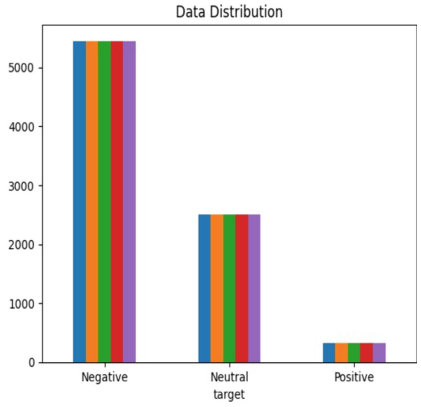


Fig. 14. Polarity of the 8274 Tweets with Beto

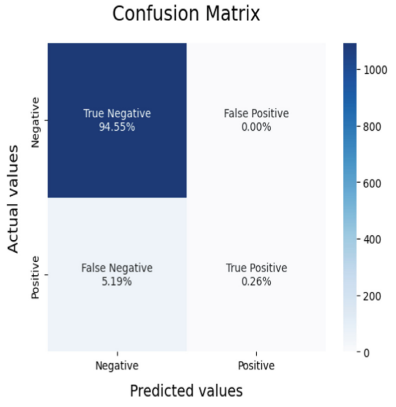


Fig. 15. Confusion matrix of the Linear Support Vector Classification Model

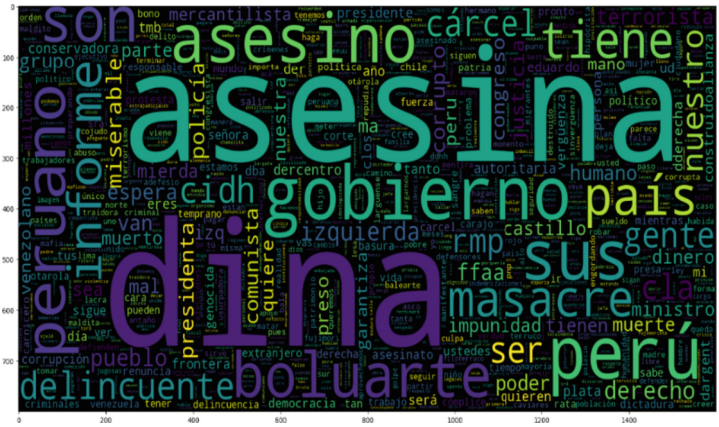


Fig. 16. N- gram of the analyzed Tweets

5 Conclusion

The research set out to identify the accuracy of the five Peruvian political approval rating sentiment analysis ranking algorithms that demonstrate the best level of accuracy of the 5773 clean Spanish-language tweets analyzed from the Twitter social network. The results of the classification from the Tweets reflected that the predominant sentiment in the micro texts collected on Twitter was the negative polarity, which could be associated to the period of collection of the Tweets which was planned during a period of political instability in the Peruvian context. The data collection was from June 2023, a very representative date for Twitter users for the analysis of the presidential crisis.

The partial result of a mini-text of the 8274 tweets downloaded from the Twitter API with Peruvian presidential account regarding the political approval index, analyzed with six vector features with the logistic regression algorithm found 69.83% positive polarity and 30. 17% negative polarity of the 5773 clean tweets analyzed.

Likewise, a high similarity was obtained between the two documents with a radial angle of 0.2762. In addition, a mini-text of four documents was taken for training and 1 document for testing with the Naive Bayes algorithm, finding a probability for the category presidency of 0.000301237799 and its complement of 0.0001354807024, demonstrating that the positive indicator of the word government surpasses the appearances of the two negative words boluarte and justice. With the Bernoulli classifier a probability of 0.005184 was obtained and for the negative tokens 0.0219478737997. Regarding the Machine Learning techniques applied, it is concluded that the Linear SVC classification algorithm obtained the first place with an accuracy of 0.963636, in second place, the KNeighbords classifier with 0.953246, in third place, the logistic regression with 0.952380 and Multinomial NB with Bernoulli NB with a tie of 0.948051.

References

1. Navarro, L.F.S.: Peruvian president's approval rating based on sentiment analysis on tweet data. TECHNO Rev. Int. Technol. Sci. Soc. Rev. Revista Internacional de Tecnología, Ciencia y Sociedad **11** (2022). https://doi.org/10.37467/REVTECHNO.V11.4396
2. Yauri, J., Solis, L., Porras, E., Lagos, M., Tinoco, E.: Approval rating of peruvian politicians and policies using sentiment analysis on Twitter. Int. J. Adv. Comput. Sci. Appl. **13**(6), 812–818 (2022). https://doi.org/10.14569/IJACSA.2022.0130696
3. Ipsos | Investigación de Mercados. https://www.ipsos.com/es-pe. Accessed 16 June 2023
4. CPI - compañia peruana de estudios de mercado y opinión pública sac. https://cpi.pe/index.html. Accessed 16 June 2023
5. IEP | Instituto de Estudios PeruanosInstituto de Estudios Peruanos. https://iep.org.pe/. Accessed 16 June 2023
6. Desaprobación del Congreso: el 91% de los peruanos reprueban gestión de los parlamentarios | Actualidad | La República. https://larepublica.pe/politica/actualidad/2023/06/25/iep-el-91-de-los-peruanos-desaprueba-al-congreso-dina-boluarte-poder-ejecutivo-2276450. Accessed 26 June 2023
7. IEP: Base enero II 2023: total de entrevistados-Nacional (1214) Base enero II 2023: total de entrevistados-Nacional, 1214. https://iep.org.pe/wp-content/uploads/2023/01/Informe-IEP-OP-Enero-II-2023-completo-v2.pdf. Accessed 15 June 2023
8. El Peruano - Resolución del Congreso que declara la permanente incapacidad moral del Presidente de la República y la vacancia de la Presidencia de la República - RESOLUCION – 001-2022-2023-CR - PODER LEGISLATIVO - CONGRESO DE LA REPUBLICA. https://busquedas.elperuano.pe/normaslegales/resolucion-del-congreso-que-declara-la-permanente-incapacida-resolucion-001-2022-2023-cr-2132939-1/. Accessed 6 July 2023
9. Paro nacional: un recuento de lo que se vivió desde las manifestaciones en la Toma de Lima del 19 de enero. https://data.larepublica.pe/paro-nacional-un-recuento-de-lo-que-se-vivio-en-las-manifestaciones-de-la-denominada-toma-de-lima-del-19-de-enero/. Accessed 6 July 2023
10. IPSOS: Encuesta Nacional Urbana-Rural. https://www.ipsos.com/sites/default/files/ct/news/documents/2023-01/Informe%20Encuesta%20Nacional%20Urbano%20Rural%20-%20Perú%2021%20al%2013%20de%20enero%202023_0.pdf. Accessed 8 July 2023
11. Gómez-Torres, E., et al.: Influencia de redes sociales en el análisis de sentimiento aplicado a la situación política en Ecuador. Enfoque UTE **9**(1), 67–78 (2018). https://doi.org/10.29019/ENFOQUEUTE.V9N1.235
12. Digital 2022: Peru — DataReportal – Global Digital Insights. https://datareportal.com/reports/digital-2022-peru. Accessed 28 June 2023
13. Orantes Rivera, E.T.: El dispositivo de la propaganda en las redes sociales de la campaña presidencial de El Salvador (2018–2019). Comunicación y Medios **30**(43), 62 (2021). https://doi.org/10.5354/0719-1529.2021.58774
14. Lu, Y., et al.: Pattern recognition and artificial intelligence (2020). https://link.springer.com/book/10.1007/978-3-030-59830-3. Accessed 30 June 2023
15. Martínez-Cámara, E., Martín-Valdivia, M.T., Ureña-López, L.A., Montejo-Ráez, A.R.: Sentiment analysis in Twitter. Nat. Lang. Eng. **20**(1), 1–28 (2014). https://doi.org/10.1017/S1351324912000332
16. Linares, R., Herrera, J., Cuadros, A., Alfaro, L.: Prediction of tourist traffic to Peru by using sentiment analysis in Twitter social network. In: Proceedings - 2015 41st Latin American Computing Conference, CLEI 2015 (2015). https://doi.org/10.1109/CLEI.2015.7360051
17. AminiMotlagh, M., Shahhoseini, H.S., Fatehi, N.: A reliable sentiment analysis for classification of tweets in social networks. Soc. Netw. Anal. Min. **13**(1) (2023). https://doi.org/10.1007/s13278-022-00998-2

18. Soria, J.J., De la Cruz, G., Molina, T., Ramos-Sandoval, R.: Comparative approach of sentiment analysis algorithms to classify social media information gathering in the Spanish language. In Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Data Science and Algorithms in Systems. CoMeSySo 2022. LNNS, vol. 597, pp. 762–773. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-21438-7_64
19. Dina Boluarte: Presidenta desactiva su cuenta oficial de Twitter, pero mantiene su perfil en Facebook adelanto de elecciones protestas en Lima protestas en el Perú | POLITICA | EL COMERCIO PERÚ. https://elcomercio.pe/politica/actualidad/dina-boluarte-presidenta-desactiva-su-cuenta-oficial-de-twitter-pero-mantiene-su-perfil-en-facebook-adelanto-de-elecciones-protestas-en-lima-protestas-en-el-peru-noticia/. Accessed 9 July 2023
20. Limaylla-Lunarejo, M.I., Condori-Fernandez, N., Luaces, M.R.: Requirements classification using FastText and BETO in Spanish documents. In: Ferrari, A., Penzenstadler, B. (eds.) Requirements Engineering: Foundation for Software Quality. REFSQ 2023. LNCS, vol. 13975, pp. 159–176. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-29786-1_11/COVER
21. finiteautomata/beto-sentiment-analysis · Hugging Face. https://huggingface.co/finiteautomata/beto-sentiment-analysis. Accessed 23 June 2023
22. Pérez, J.M., Giudici, J.C., Luque, F.: pysentimiento: a Python toolkit for sentiment analysis and SocialNLP tasks, June 2021. https://arxiv.org/abs/2106.09462v1. Accessed 9 July 2023
23. Cañete, J., Chaperon, G., Fuentes, R., Jorge, P.: Spanish pre-trained BERT model and evaluation data. In: PML4DC at ICLR, vol. 2020, pp. 1–10 (2020)
24. Christopher, M.: An Introduction to Information Retrieval, vol. 38, no. c (2009). https://doi.org/10.1210/endo-38-3-156
25. Montesinos, L.: Análisis de sentimientos y predicción de eventos en Twitter. Santiago De Chile, 12–16 (2014)
26. Chen, W., Xu, H., Jia, L., Gao, Y.: Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. Int. J. Forecast. **37**(1), 28–43 (2021). https://doi.org/10.1016/j.ijforecast.2020.02.008
27. Pauli, P.: Análisis de sentimiento: comparación de algoritmos predictivos y métodos utilizando un lexicon español. Instituto tecnológico de Buenos Aires - ITBA, Buenos Aires (2019). http://ri.itba.edu.ar/handle/123456789/1782. Accessed 4 June 2023
28. Tutor, S.: Autora: Desirée García Soriano Tutor: Rubén Martín Clemente (2021)
29. Pedersen, R., Schoeberl, M.: An embedded support vector machine. In: Proceedings of the Fourth Workshop on Intelligent Solutions in Embedded Systems, WISES 2006, pp. 79–89 (2006). https://doi.org/10.1109/WISES.2006.237155
30. Daniel, J., Martin, J.H.: Chapter 5 - Speech and Language Processing (2023)
31. Ananthakumar, U., Sarkar, R.: Application of logistic regression in assessing stock performances. In: Proceedings - 2017 IEEE 15th International Conference on Dependable, Autonomic and Secure Computing, 2017 IEEE 15th International Conference on Pervasive Intelligence and Computing, 2017 IEEE 3rd International Conference on Big Data Intelligence and Compu, vol. 2018, pp. 1242–1247 (2018). https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.199
32. Lin, S.Y., Kung, Y.C., Leu, F.Y.: Predictive intelligence in harmful news identification by BERT-based ensemble learning model with text sentiment analysis. Inf. Process Manag. **59**(2) (2022). https://doi.org/10.1016/j.ipm.2022.102872
33. Uddin, S., Haque, I., Lu, H., Moni, M.A., Gide, E.: Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. Sci. Rep. **12**(1) (2022). https://doi.org/10.1038/s41598-022-10358-x
34. Rodríguez-Ibánez, M., Casánez-Ventura, A., Castejón-Mateos, F., Cuenca-Jiménez, P.M.: A review on sentiment analysis from social media platforms. Exp. Syst. Appl. **223**, 119862 (2023). https://doi.org/10.1016/j.eswa.2023.119862

35. Charpentier, A.: Computational Actuaril Science with R. Taylor & Francis Group, Canada (2015)
36. Liu, B.: Sentiment analysis: mining opinions, sentiments, and emotions (2015). https://doi.org/10.1017/CBO9781139084789
37. Russell, M.A.: Mining the Social Web: Analyzing Data from Facebook (2019)
38. Arango Pastrana, C.A., Osorio Andrade, C.F.: Aislamiento social obligatorio: un análisis de sentimientos mediante machine learning. Suma de Negocios **12**(26), 1–13 (2021). https://doi.org/10.14349/sumneg/2021.v12.n26.a1
39. Liu, B.: Document sentiment classification. In: Sentiment Analysis and Opinion Mining. Lectures on Human Languaje Technologies. Springer, Cham (2012). https://doi.org/10.1007/978-3-031-02145-9_3
40. Chaibi, M., Benghoulam, E.L.M., Tarik, L., Berrada, M., El Hmaidi, A.: An interpretable machine learning model for daily global solar radiation prediction. Energies (Basel) **14**(21) (2021). https://doi.org/10.3390/en14217367
41. Jurafsky, D., Martin, J.: Naive Bayes and sentiment classification. In: Speech and Language Processing, p. 1024 (2019)
42. Liu, B.: Sentiment Analysis Mining Opinions, Sentiments, and Emotions, vol. 2 (2011)



Evaluation of Multiplatfom Component for Biometric Authentication in Low-Code Programming Platform – Case Study

Zdzisław Sroczyński(✉)

Faculty of Applied Mathematics, Silesian University of Technology,
23 Kaszubska Str., 44-100 Gliwice, Poland
zdzislaw.sroczynski@polsl.pl

Abstract. Biometric authentication has emerged as a prominent method for secure user identity verification. Integrating biometric authentication into software applications, particularly within the 'low-code/no-code' paradigm, causes unique challenges. This paper presents a case study demonstrating the seamless implementation of a pre-built biometric authentication component within a selected low-code programming platform. This approach describes and elaborates an existing component that empowers developers to incorporate biometric authentication without delving into complex technological details. This platform-agnostic component enhances security and convenience, helping with effective development of iOS and Android applications sharing the whole or the majority of the source code. The study underscores the significance of the 'low-code/no-code' paradigm in simplifying the rapid development of cross-platform applications while advancing user security.

1 Introduction

Biometric authentication has become increasingly popular as a secure and efficient method for verifying user identity [1]. However, integrating biometric authentication into software applications can be a challenging task, especially for low-code programming platforms. In this paper, we present a case study of the implementation of a pre-built component for biometric authentication that is available on a selected low-code programming platform. Our approach involves the use of an existing component that allows developers to implement biometric authentication without extensive knowledge of the underlying technologies. The pre-built component is designed to ensure security and be platform-agnostic, making it an ideal solution for low-code programming platforms. Through our case study, we showcase the effectiveness of the component by demonstrating its integration in test mobile applications running on iOS and Android platforms. Our research contributes to the software engineering and more efficient development of secure authentication mechanisms for software applications, particularly in low-code programming platforms.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 398–409, 2024.
https://doi.org/10.1007/978-3-031-53549-9_38

Low-code programming has emerged as a popular approach for developing software applications, especially in the context of Rapid Application Development (RAD) [2]. In this section, we will explore the basics of low-code programming and its connection to RAD.

Low-code programming is an approach to software development that allows developers to create applications using a visual interface and pre-built components, rather than traditional coding. This approach is based on the idea of abstraction, which involves hiding the complexity of the underlying technologies and providing developers with a simplified interface for creating software applications. The low-code programming model provides a rapid development cycle, reducing the amount of time and resources needed to create applications. This approach allows for the rapid creation of highly adaptable multi-platform software applications.

Low-code programming is closely related to RAD, which is a methodology that emphasizes the rapid development of software applications by using pre-built components and a visual interface [3]. RAD is based on the idea of prototyping, which involves quickly building and testing a working version of an application to gather feedback from users and stakeholders, allowing developers to create software applications in a short period of time and to iterate on them quickly. This approach has become popular in recent years, as it helps organizations to respond quickly to changing business needs and market conditions.

Low-code programming and RAD share many similarities, as both approaches emphasize the rapid development of software applications using pre-built components and a visual interface. Both approaches allow developers to create applications quickly, reducing the time and resources needed to bring products to market. Low-code programming extends the RAD approach by providing a more abstracted interface and a more streamlined development cycle, enabling even non-technical users to create software applications quickly and easily.

In recent years, the concept of 'low-code/no-code' has received significant attention from researchers in the field of software engineering. This heightened interest has led to many publications exploring various aspects of this paradigm. Researchers have delved into the topic through both comprehensive review papers, such as [4–11], which provide an overview of the field, and more in-depth studies, like [12–18], which examine detailed solutions and specific aspects of 'low-code/no-code' development. This growing body of research reflects the increasing recognition of the importance and potential of 'low-code/no-code' approaches in modern software development practices.

Low-code programming is a powerful approach to software development that is closely related to the RAD methodology. This approach allows developers to create software applications quickly and efficiently, using pre-built components and a visual interface. Low-code programming extends the RAD approach by providing a more abstracted interface and a more streamlined development cycle, enabling even non-technical users to create software applications quickly and

easily. This approach is becoming increasingly popular as organizations search for immediate response to changing business needs and market conditions.

Low-code programming is a modern approach to software development that aims to simplify the application development process. It is usually a visual, drag-and-drop method of building software applications. This approach follows the Rapid Application Development (RAD) paradigm, which focuses on building software applications in a fast and iterative manner.

One example of a low-code programming platform is RAD Studio, an Integrated Development Environment (IDE) that enables developers to create software applications for various platforms, including Windows, macOS, iOS, and Android. RAD Studio uses the FireMonkey Cross-Platform Framework – FMX [19,20], a multi-platform library that provides a rich set of UI controls, graphic effects, and animations. FMX allows developers to create visually appealing applications with responsive layouts that adapt to different screen sizes and resolutions [21].

FMX is an example of a low-code programming library that provides developers with pre-built components that can be easily dragged and dropped into their applications. The library includes a range of UI controls, such as buttons, labels, edit boxes, list boxes, and other essential components. Developers can also use FMX to add multimedia elements, such as audio and video playback, to their applications [22–26].

In summary, low-code programming is a powerful approach to software development that enables developers to create applications quickly and easily. RAD Studio and FMX are examples of a low-code programming IDE and library, respectively, that provide developers with the tools they need to create visually appealing, responsive applications for multiple platforms, including iOS and Android.

2 Native Implementations of the Biometric Authentication for Mobile Platforms

In the following sections, we will delve into the native solutions and APIs associated with biometric security in the Android and iOS operating systems. Our objective is to illustrate the substantial differences between these two platforms and highlight the potential time and effort savings that can be achieved through the introduction of a cross-platform environment equipped with a component that accomplishes these tasks uniformly.

2.1 Implementing Biometric Authentication for Android

Biometric authentication in Android provides a secure and convenient method for users to authenticate their identity using their biometric data, such as fingerprints, face recognition, or iris scanning. Android provides an API for integrating biometric authentication into applications, which is supported by most modern Android devices.

To implement biometric authentication in an Android app, developers can use the BiometricPrompt API, which provides a standardized, user-friendly interface for biometric authentication. The BiometricPrompt API handles all the interactions with the biometric hardware, such as fingerprint scanners or facial recognition sensors, and ensures that the authentication process is secure and reliable.

There is a sketch example of how to implement fingerprint authentication using the BiometricPrompt API in Kotlin in Listing 1.1 below.

Listing 1.1. Biometric Authentication for Android in Kotlin

```
val promptInfo = BiometricPrompt.PromptInfo.Builder()
    .setTitle("Authentication required")
    .setSubtitle("Please use your fingerprint to authenticate")
    .setDescription("This app requires fingerprint authentication")
    .setNegativeButtonText("Cancel")
    .build()

val biometricPrompt = BiometricPrompt(this,
    ContextCompat.getMainExecutor(this),
    object : BiometricPrompt.AuthenticationCallback() {
        override fun onAuthenticationError(errorCode: Int, errString: CharSequence)↩
            {
            // handle authentication error
        }
        override fun onAuthenticationSucceeded(result: BiometricPrompt.↩
            AuthenticationResult) {
            // handle authentication success
        }
        override fun onAuthenticationFailed() {
            // handle authentication failure
        }
    })
biometricPrompt.authenticate(promptInfo)
```

This code creates a BiometricPrompt.PromptInfo object that specifies the details of the authentication request, such as the title, subtitle, and description. It then creates a BiometricPrompt object that takes an AuthenticationCallback object, which provides callbacks for handling authentication success and failure.

Finally, the biometricPrompt.authenticate(promptInfo) method is called to start the authentication process. The user will see a prompt asking to authenticate using a fingerprint, and the AuthenticationCallback methods will be called based on the result of the authentication process.

Overall, the BiometricPrompt API provides a solid and secure way to implement biometric authentication in Android apps, and Kotlin provides a concise and expressive language for coding the implementation.

2.2 Implementing Biometric Authentication for iOS

Biometric authentication in iOS, also known as Touch ID or Face ID, allows users to authenticate themselves using their biometric data, such as fingerprints or facial recognition. iOS provides an API helping developers to integrate biometric authentication into their apps, which is supported by most modern iOS devices.

To implement biometric authentication in an iOS app, developers can use the LocalAuthentication framework, which provides a high-level API for interacting with the biometric hardware and handling the authentication process.

There is a sketch example of how to implement Touch/Face ID authentication using the LocalAuthentication framework in Swift in Listing 1.2 below.

Listing 1.2. Biometric Authentication for iOS in Swift

```
import LocalAuthentication

let context = LAContext()

var error: NSError?
if context.canEvaluatePolicy(.deviceOwnerAuthenticationWithBiometrics, error: ↩
    &error) {
    let reason = "Authenticate with Touch/Face ID"
    context.evaluatePolicy(.deviceOwnerAuthenticationWithBiometrics, ↩
        localizedReason: reason) {
        (success, error) in
        if success {
            // Handle authentication success
        } else {
            // Handle authentication failure
        }
    }
} else {
    // Touch ID not available, handle error
}
```

This code creates an instance of the LAContext class, which represents the local authentication context. It then checks if the device supports biometric authentication by calling the canEvaluatePolicy method with the .deviceOwnerAuthenticationWithBiometrics policy.

If biometric authentication is supported, the app can request authentication by calling the evaluatePolicy method with the same policy and providing a reason string for the user. The evaluatePolicy method takes a closure that is called with the result of the authentication process, either success or failure.

Overall, the LocalAuthentication framework provides an efficient way to implement biometric authentication in iOS apps, and Swift provides a modern and expressive language for coding the implementation.

3 Multiplatform Implementation of the Biometric Authentication in FMX/RAD Studio

The FMX.BiometricAuth.TBiometricAuth class provides a simple and convenient means of authorization via biometrics, such as Face ID/Touch ID or Fingerprint recognition, on both Android and iOS devices.

It is important to note that for Android, TBiometricAuth is currently supported only on Android 10 or higher. To utilize TBiometricAuth, developers should paste it from the component palette (e.g., on a form or datamodule) and configure its properties according to their requirements.

For Android compatibility, it is crucial to ensure that the PromptDescription and PromptTitle properties have appropriate values. Developers should also specify the desired strengths for BiometricStrengths. Additionally, they should make sure to check the UseBiometric permission in the project properties (Fig. 1).

On iOS, developers must provide a meaningful description for NSFaceIDUsageDescription in project's options. This text will be presented to the user during the initial use of biometric verification.

Eventually, developers should create event handlers for the OnAuthenticateSuccess or OnAuthenticateFail events to define the actions to be taken in those scenarios. The code in the OnAuthenticateSuccess event handler should

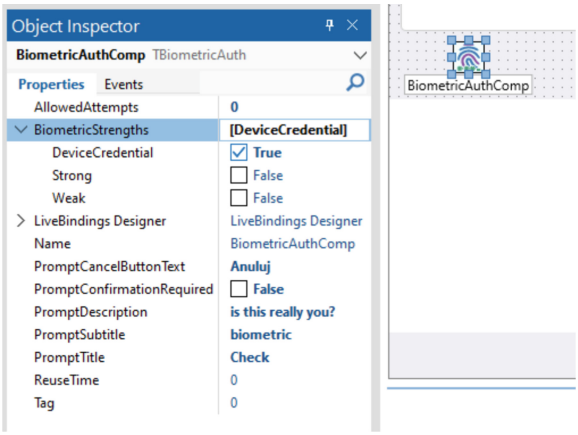


Fig. 1. TBiometricAuth component and its properties during development in the RAD Studio IDE.

be synchronized with the main UI thread with the use of TThread.Synchronize method. To initiate the biometric authorization process, developers should call the Authenticate method, which will present the system dialog for biometric authorization to the user.

The TBiometricAuth component streamlines the implementation of biometric authentication in Android and iOS applications, enhancing user security and convenience. The utilization of this component offers a remarkable advantage: the ability to design a cross-platform application without delving into the intricate, platform-specific implementation details, which can vary significantly for each platform. This exemplifies a quintessential scenario where RAD (Rapid Application Development) or Low-Code programming truly shines, allowing developers to focus their efforts on the core objectives of their application, while abstracting away the complexities of platform-specific biometric authentication. This approach not only enhances development efficiency but also underscores the power of RAD and Low-Code methodologies in simplifying the creation of robust, multiplatform applications.

3.1 Hiding the Secret According to the Application Life-Cycle

Enhancing the security of a mobile application is essential in contemporary digital landscape. To safeguard sensitive information, such as secrets or authentication tokens, when the application is sent to the background, it is necessary to obscure this data. Additionally, mitigating the risk of unauthorized access through screen capture or peering at the app in the list of recently used applications is equally crucial. In the FMX environment, these security measures can be implemented by attaching the appropriate event handling function to the IFMXApplicationEventService interface.

Two key events, TApplicationEvent.WillBecomeInactive and TApplicationEvent.WillBecomeForeground, play pivotal roles in this endeavor. When the application is about to become inactive, it is the ideal moment to conceal sensitive information. Subsequently, as the application is about to return to the foreground, the authentication process should be revalidated to ensure that only authorized users regain access.

For iOS, an additional step is necessary to force a refresh of the application's view. This can be achieved by invoking:
WindowHandleToPlatform(Handle).View.setNeedsDisplay.
These solutions collectively enhance the application's security, making it more resilient to potential vulnerabilities and ensuring that user data remains confidential even in dynamic mobile environments.

On the other hand, for the Android platform, another additional step is necessary to fulfil security measures, involving a modification to one of the source files within the FMX library, namely FMX.Platform.Android.pas. In this file, adjustments are made within the method called:
TPlatformAndroid.HandleApplicationCommandEvent
to handle two specific commands.

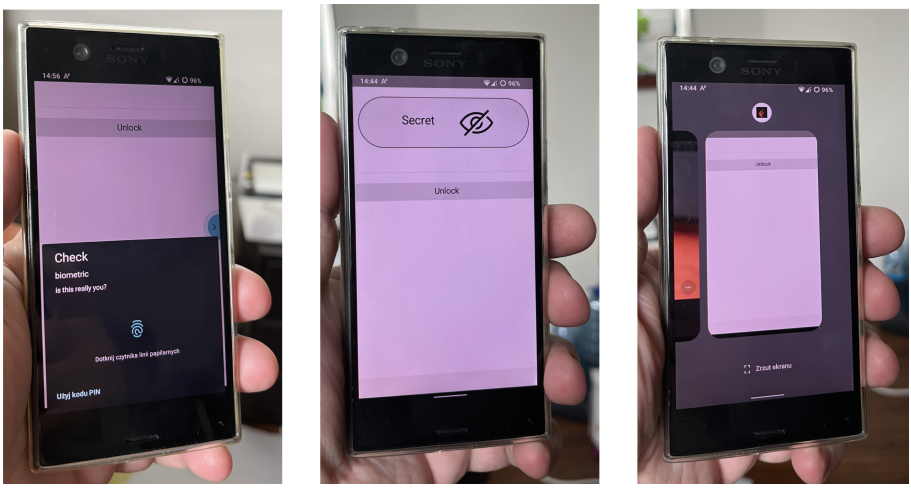
Firstly, when the command is TAndroidApplicationCommand.LostFocus, signifying that the application is losing focus as it transitions into the background, certain actions must be taken. To ensure the confidentiality of sensitive information, the application's main window is marked with parameter:
MainActivity.getWindow.setFlags(8192, 8192),
which is equivalent to FLAG_SECURE. This effectively treats the content of the window as secure, preventing it from being captured in screenshots or displayed on non-secure screens. Additionally, the function hiding the secret is invoked to conceal any confidential data.

On the other hand, when the ACommand is:
TAndroidApplicationCommand.GainedFocus,
indicating that the application is returning to the foreground, different actions are necessary. In this scenario, MainActivity.getWindow.clearFlags(8192) is used to remove the FLAG_SECURE setting previously applied during the LostFocus event. This action allows the application to be interacted with again without the heightened security restrictions. Furthermore, the function initializing the biometric authentication process is called, ensuring that security is promptly reestablished as the application becomes active once more.

These modifications within FMX.Platform.Android.pas play a pivotal role in enhancing the application's security on the Android platform. By utilizing FLAG_SECURE (8192), the content within the application's window is treated as secure, preventing unauthorized access via screenshots or non-secure displays. This comprehensive approach safeguards sensitive data and user privacy effectively.

3.2 Tests in Real-Life Environment

The test application utilizing the FMX library's TBiometricAuth component was developed with the use of RAD Studio 11 Alexandria environment and subsequently deployed on both Android 12 and iOS 16 devices. In Fig. 2, we showcase images of the application actively engaged in biometric authentication, highlighting two critical aspects: one with the confidential data visibly obscured and the other with the data securely hidden while browsing recently opened applications. Notably, the pictures provided here represent actual photographs of the application in action on mobile devices due to implemented blockade of screenshots.

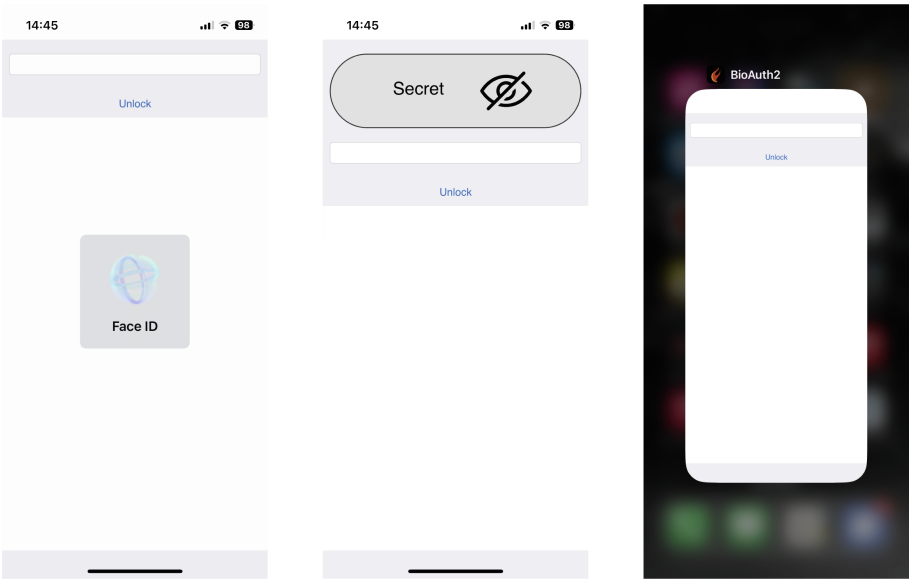


(a) Unlock dialog (b) Authentication success (c) App in background

Fig. 2. The test application developed with the use of FMX framework in RAD Studio running at Android 12.

It is worth noting that the very same application, compiled, installed, and executed on an iPhone running iOS 16, flawlessly executed biometric authentication tasks without any modifications to the source code. The screens of the application in action on iOS 16 are shown in Fig. 3.

This seamless cross-platform compatibility, where the application effortlessly transitioned between Android 12 and iOS 16, speaks volumes about the effectiveness of the development approach. With minimal code adjustments, if any, the application properly leveraged the native biometric authentication capabilities of iOS, further underscoring the power of RAD Studio and the FMX library in delivering consistent and reliable user experience across different mobile platforms.



(a) Unlock dialog (b) Authentication success (c) App in background

Fig. 3. The test application developed with the use of FMX framework in RAD Studio running at iOS 16.

Experiments with this multi-platform application, built using the FMX library, have demonstrated seamless integration with the platform-specific biometric recognition dialogues. What's particularly impressive is that developers required minimal in-depth knowledge regarding the details of biometric authentication implementation. Actually, the process primarily involved configuring a few essential parameters within the TBiometricAuth component. This successful test underscores the power of RAD Studio and the FMX library in streamlining the development of robust, cross-platform applications in the low-code manner, while simultaneously enhancing user security and experience.

4 Summary

Biometric authentication has gained popularity as a secure and efficient means of verifying user identity. Integrating biometric authentication into software applications can be a challenging software engineering task. This paper presents a case study of implementing a pre-built biometric authentication component within a selected low-code programming platform. The approach involves utilizing an existing component that allows developers to implement biometric authentication without in-depth knowledge of underlying technologies. The pre-built component is designed to be easy to use, secure, and platform-agnostic, making it

an ideal solution for low-code programming platforms. Through a case study, the paper demonstrates the component's effectiveness by integrating it into test mobile applications running on iOS and Android platforms.

Low-code programming is a contemporary approach to software development that simplifies the application development process. It leverages a visual, drag-and-drop method for building software applications, allowing developers to create applications quickly and easily. This approach is rooted in the Rapid Application Development (RAD) paradigm, emphasizing the swift and iterative creation of software applications. The 'low-code/no-code' paradigm plays significant role in revolutionizing the software engineering landscape.

The paper explores the implementation of biometric authentication in Android and iOS. For Android, the BiometricPrompt API is introduced, providing a standardized interface for biometric authentication. The API streamlines interactions with biometric hardware, ensuring a secure and reliable authentication process. A code snippet in Kotlin showcases fingerprint authentication using the BiometricPrompt API.

Similarly, for iOS, the paper discusses implementing biometric authentication using the LocalAuthentication framework. This framework offers a high-level API for interacting with biometric hardware and handling authentication. A Swift code example illustrates Touch/Face ID authentication using the LocalAuthentication framework.

Finally, the FMX.BiometricAuth.TBiometricAuth class is introduced as a convenient means of biometric authorization on both Android and iOS devices. The paper highlights configuration requirements for Android and iOS, emphasizing the simplicity of integrating the TBiometricAuth component. The TBiometricAuth component allows developers to streamline biometric authentication implementation across different platforms, abstracting away platform-specific complexities. This approach exemplifies the power of RAD Studio and the FMX library in simplifying the development of cross-platform applications within the 'low-code/no-code' paradigm.

Additionally, the paper emphasizes the importance of obscuring sensitive data when the application is sent to the background. Some extra measures are also taken to prevent unauthorized access through screen capture or peering at the app in the list of recently used applications. These security measures are implemented by attaching event handlers to the IFMXApplicationEventService interface. The paper also highlights the additional steps required for iOS and Android platforms to further enhance security.

The paper presents real-life tests of the application developed using the FMX library's TBiometricAuth component. The application was successfully deployed on Android 12 and iOS 16 devices, showcasing its functionality during biometric authentication. The tests emphasize the seamless cross-platform compatibility of the application, requiring minimal or no code adjustments between Android and iOS. This case study demonstrates the effectiveness of RAD Studio and the FMX library in delivering consistent user experiences across diverse mobile platforms, while enhancing user security within the 'low-code/no-code' paradigm.

References

1. Sarkar, A., Singh, B.K.: A review on performance, security and various biometric template protection schemes for biometric authentication systems. Multimedia Tools Appl. **79**, 27721–27776 (2020)
2. Kulkarni, M.: Deciphering low-code/no-code hype-study of trends overview of platforms and rapid application development suitability. Int. J. Sci. Res. Publ. **11**(7) (2021)
3. Kralev, V., Kraleva, R.: Methods and tools for rapid application development. In: International Scientific and Practical Conference World Science, vol. 1/4, pp. 21–24. ROST (2017)
4. Sahay, A., Indamutsa, A., Di Ruscio, D., Pierantonio, A.: Supporting the understanding and comparison of low-code development platforms. In: 2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 171–178. IEEE (2020)
5. Alsaadi, H.A., Radain, D.T., Alzahrani, M.M., Alshammari, W.F., Alahmadi, D., Fakieh, B.: Factors that affect the utilization of low-code development platforms: survey study. Rom. J. Inf. Technol. Autom. Control/Revista Română de Informatică şi Automatică **31**(3) (2021)
6. Yan, Z.: The impacts of low/no-code development on digital transformation and software development. arXiv preprint arXiv:2112.14073 (2021)
7. Benac, R., Mohd, T.K.: Recent trends in software development: low-code solutions. In: Arai, K. (ed.) FTC 2021. LNNS, vol. 360, pp. 525–533. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-89912-7_41
8. Bucaioni, A., Cicchetti, A., Ciccozzi, F.: Modelling in low-code development: a multi-vocal systematic review. Softw. Syst. Model. **21**(5), 1959–1981 (2022)
9. Bock, A.C., Frank, U.: Low-code platform. Bus. Inf. Syst. Eng. **63**, 733–740 (2021)
10. Hirzel, M.: Low-code programming models. arXiv preprint arXiv:2205.02282 (2022)
11. Gomes, P.M., Brito, M.A.: Low-code development platforms: a descriptive study. In: 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–4. IEEE (2022)
12. Patkar, N., Chiş, A., Stulova, N., Nierstrasz, O.: Interactive behavior-driven development: a low-code perspective. In: 2021 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C), pp. 128–137. IEEE (2021)
13. Lichtenthäler, R., Böhm, S., Manner, J., Winzinger, S.: A use case-based investigation of low-code development platforms. In: ZEUS, pp. 76–83 (2022)
14. Galhardo, P., Silva, A.R.D.: Combining rigorous requirements specifications with low-code platforms to rapid development software business applications. Appl. Sci. **12**(19), 9556 (2022)
15. Torres, A., Kapralos, B., Da Silva, C., Peisachovich, E., Dubrowski, A.: Moirai: a no-code virtual serious game authoring platform. Virtual Worlds **1**, 147–171 (2022)
16. Gerasimov, A., Michael, J., Netz, L., Rumpe, B.: Agile generator-based GUI modeling for information systems. In: Dahanayake, A., Pastor, O., Thalheim, B. (eds.) M2P 2020. CCIS, vol. 1401, pp. 113–126. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72696-6_5
17. Martins, R., Caldeira, F., Sa, F., Abbasi, M., Martins, P.: An overview on how to develop a low-code application using OutSystems. In: 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), pp. 395–401. IEEE (2020)

18. Da Cruz, M.A., de Paula, H.T., Caputo, B.P., Mafra, S.B., Lorenz, P., Rodrigues, J.J.: OLP-a restful open low-code platform. Future Internet **13**(10), 249 (2021)
19. Kovačević, Ž.: RAD studio: developing mobile applications. In: CARNetova korisnička konferencija-CUC 2014 (2014)
20. Sroczyński, Z.: Designing human-computer interaction for mobile devices with the FMX application platform. Theor. Appl. Inf. **26**(1–2), 87–104 (2014)
21. Bier, A., Sroczyński, Z.: Development of the multi-platform human-computer interaction for mobile and wearable devices. In: Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T., Kacprzyk, J. (eds.) DepCoS-RELCOMEX 2019. AISC, vol. 987, pp. 57–68. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-19501-4_6
22. Teti, D.: Delphi Cookbook. Packt Publishing Ltd. (2016)
23. Chandler, G.: FireMonkey development for iOS and OS X with Delphi XE2. Coogara Consulting (2012)
24. Glowacki, P.: Expert Delphi. Packt Publishing Ltd. (2017)
25. Magni, A.: Delphi GUI Programming with FireMonkey: Unleash the Full Potential of the FMX Framework to Build Exciting Cross-Platform Apps with Embarcadero Delphi. Packt Publishing, Limited (2020)
26. Cornelius, D.: Fearless Cross-Platform Development with Delphi: Expand Your Delphi Skills to Build a New Generation of Windows, Web, Mobile, and IoT Applications. Packt Publishing Ltd. (2021)

Leveraging Machine Learning and Raspberry Pi for Enhanced Wildlife Remote Monitoring and Localization

Fabrice Manzi[1], Emmanuel Tuyishime[2(✉)], Antoine Hitayezu[1], Gedeon Muhawenayo[3], Philibert Nsengiyumva[1], and Kayalvizhi Jayavel[1]

[1] African Center of Excellence in Internet of Things, University of Rwanda, Kigali, Rwanda
k.jayavel@arts.ac.uk
[2] Electronics and Computers Department, Transilvania University of Brasov, Brasov, Romania
emmmanuel.tuyishime@unitbv.ro
[3] Rwanda Space Agency (RSA) Geospatial Team, Kigali, Rwanda

Abstract. The tourism sector faces challenges, particularly concerning the expectations of wildlife enthusiasts visiting the national parks. The occasional difficulty in spotting favored wildlife species can lead to a dissonance between anticipated and actual experiences, potentially impacting visitors' overall park experience. To address this issue, we propose a novel system integrating the You Only Look Once (YOLOv5) machine learning model with a Raspberry Pi and an intuitive client application to enrich tourists' experience. This study presents the outcomes of our proposed system, showcasing its potential to assist tourists in wildlife localization and effective trip planning. Beyond enhancing tourist encounters, this system holds promise in the domains of wildlife conservation and animal behavior research, furnishing an advanced tool for monitoring and comprehending animal populations in their natural habitats, and contributing significantly to conservation initiatives. While this study presents a single system within a specific location, the scalability of this approach across entire national parks is evident.

Keywords: Embedded Computing · Machine Learning · YOLOv5 · Raspberry Pi · Object Detection

1 Introduction

Tourism, as a global phenomenon, has emerged as a substantial driver of economic growth for countries across the world. The ability of this industry to stimulate economic development, generate foreign exchange, and create employment opportunities has rendered it an indispensable component of many nations' economies. Rwanda stands as an example of a country where tourism's transformative impact has been particularly pronounced. According to the report [1], tourism plays a significant role in Rwanda's economy by contributing substantially to its foreign exchange earnings. Additionally, this industry has a distinct tendency to create a larger share of formal sector employment in comparison to other sectors.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 410–423, 2024.
https://doi.org/10.1007/978-3-031-53549-9_39

Nevertheless, the tourism sector in Rwanda does face certain challenges. A common challenge among tourists exploring the nation's beloved national parks is the occasional difficulty of spotting their favorite wildlife species. This situation revolves around the gap between tourists 'earnest desires to observe specific fauna and the inherent unpredictability associated with such encounters. While seeing animals in their natural habitat is exciting, it's not always guaranteed. The mismatch between reality and expectations can affect the overall perception of the park visit.

Presently, the predominant method employed to ascertain the whereabouts of the animal relies heavily on the conjecture of park guides. Relying on guides' predictions is practical but not foolproof. Guides are knowledgeable about the park and animal behavior, but pinpointing the exact location of animals is challenging. Unfortunately, the accuracy of their estimations is often lacking, which can lead to subpar experiences for both tourists and park staff. While guides' estimations are rooted in insight and expertise, there's room for a more precision-guided approach.

This study sets out to explore a brighter perspective on this matter. We aim to examine the potential of using modern technology to enhance tourists' encounters with wildlife. Our approach involves leveraging machine learning (ML), coupled with a Raspberry Pi, and a user-friendly client application to enrich the tourist experience. We're working with the cameras in the park and using an ML model based on YoloV5 for object detection, alongside a dataset of animal images. YOLOv5 is an object detection algorithm and architecture that builds upon the YOLO (You Only Look Once) series of real-time object detection models [2]. The model is optimized for IoT devices like the Raspberry Pi, ensuring its functionality while minimizing any significant loss in performance. Upon analyzing the image captured by the camera, the model identifies the animal's species. Simultaneously, the system captures crucial details including location, date, and time, which are then transmitted to the cloud database through an API. To make this data accessible, a user-friendly dashboard is designed to visualize and map the collected information for public utilization.

Through this research, we hope to contribute to the comprehension of the potential synergy between the leveraged technologies in elevating the animal-viewing encounter for tourists visiting Rwanda's national parks. Moreover, the insights we gather have the potential to extend beyond the immediate context, offering valuable perspectives to both park management and policymakers dedicated to elevating the comprehensive tourism experience within the country.

As we delve into the details of our study, the rest of the paper is organized as follows: Sect. 2 provides a review of the related works in the field, highlighting the existing literature and establishing the groundwork for our ongoing research endeavors. Section 3 discusses the theory related to the employed explores the steps we've taken to create this innovative solution. Section 4 discusses the results obtained from the deployed system, demonstrating the effectiveness of the proposed system. Finally, in Sect. 5, we provide the concluding remarks and recommendations.

2 Related Work

An approach to animal detection that has garnered considerable interest involves employing computer vision methodologies to identify animals nearing roads and issuing warnings. For instance, Sharma et al. [3] presented a study that utilizes computer vision techniques for automatic animal detection on highways to prevent animal-vehicle collisions. The study involves the combination of a Histogram of Oriented Gradients (HOG) and a boosted cascade classifier for feature extraction, along with image processing techniques implemented in OpenCV software. The system also incorporates distance estimation and driver alert mechanisms based on the detected animal's proximity to the vehicle. In [4], Santhanam et al. propose a model that can detect animals and alert the driver. They utilize a deep learning algorithm, specifically convolutional neural networks, for animal detection and classification in real-time from live camera images. The model is trained using a vast open-source dataset, which allows for the segregation and identification of animals.

Singh et al. [5] provide insights into the limitations of current animal detection models in generalizing from training images of animals in their natural habitats to deployment scenarios in man-made environments. The study proposes a solution using synthetic data generation, which involves extracting animals from images of natural habitats and inserting them into target scenes, to improve the performance of detectors trained on synthetic images. In their study, the use of transfer learning is highlighted as a practical approach to adapt models trained on benchmark datasets for real-world deployment. The study demonstrates the competitiveness of lightweight models like RETINA and YOLO for multi-camera mobile deployment, providing a practical solution for animal detection in real-time video sequences. The practical implication of their study is to accurately detect animals in man-made environments to prevent accidents and ensure public safety.

An alternate strategy for animal detection and tracking encompasses the utilization of wireless sensor networks. In [6], Bhatt et al. present a GPS-based animal tracking system that incorporates a GSM modem, microcontroller, and various sensors to track the location of animals, sense temperature, and motion, and alert the user with SMS. The developed system is intended to be attached to the neck of animals for real-world wildlife tracking. In [7], Antonio et al. propose a methodology for animal detection in images captured by road cameras, which involves extracting features from image regions and using Machine Learning (ML) techniques to classify areas as either animal or non-animal. The goal of the study is to reduce accidents and promote environmental conservation. In their endeavor, Two ML techniques, KNN, and Random Forest, were compared using synthetic images to identify animals on roads. Authors claim that the KNN learning model is more reliable than Random Forest.

An additional strategy for animal tracking involves the application of automated radio telemetry, as outlined in [8]. In this study, automated radio-telemetry receivers are strategically positioned on towers to continuously monitor the movements and whereabouts of animals fitted with radio collars. The collected data is then wirelessly transmitted to a laboratory server, offering researchers convenient access. This technique presents the advantage of round-the-clock animal tracking capability; however, its implementation necessitates the deployment of radio collars, which might not be feasible or ethically sound for all species.

In relation to our case study, the current system employed to ascertain animals' locations for visitors relies on verbal communication and a manual reporting structure. The guides report animal sightings to the reception, and this data is documented on a blackboard. The recorded information encompasses the location of the sighting, the animal's species, the sighting time, and the animals' activities during that specific moment. An alternative approach involves the recognition that various animals exhibit distinct movement patterns at different times. Thus, the guides possess insights on where to direct their attention based on the time of day. For instance, around 1 p.m., many animals tend to gather by the lake to drink water. The guides might advise tourists to visit the lake during this timeframe to increase their chances of observing a variety of animals. This approach presents several drawbacks and limitations: I) The information displayed on the blackboard might not be current or precise, as it hinges on the timely reporting of guides and manual data entry. II) The process of manual reporting consumes time and is susceptible to errors, contributing to a less efficient approach. III) The blackboard approach has limited visibility and lacks interactivity, preventing visitors from engaging with the information, such as the ability to filter or search for specific animals.

Considering the shortcomings of the current approach, there is an evident necessity for a more efficient, precise, and user-friendly mechanism for tracking animal locations in national parks. As our contribution, we introduce a novel system designed to overcome these limitations. The proposed system harnesses the potential of Raspberry Pi, combined with Machine Learning, specifically YOLOv5, to elevate the tourist experience. That involves the creation of a user-friendly application that enables tourists to access real-time information about the locations of their favorite animal species, facilitating more informed trip planning. To our knowledge, this study is the first to employ the presented approach in enhancing the animal-viewing experience for tourists at Rwanda's national parks.

3 Material and Method

In this section, we will delve into the various techniques utilized in implementing our system. This entails an exploration of the essential hardware and software components, the communication technology employed within the system, and the types of data utilized. Furthermore, we will elaborate on the relevant theoretical framework, which we consider pivotal for grasping the rationale behind the materials employed in this study.

During our on-site visit, we discovered that cameras have been strategically positioned in various vital zones throughout the national park. These cameras serve multiple purposes, which include ensuring the safety of the animals, detecting any instances of mortality, and ensuring compliance with park regulations by tourists and guides.

Our methodology revolves around the utilization of these pre-existing cameras, which are strategically deployed across the park. We plan to establish connections between these cameras and Raspberry Pi devices to access their feeds. However, as a pilot project within this study, we initially employed a webcam to demonstrate our prototype. This approach allows us to showcase the viability of our system before extending it to integrate with the park's existing camera infrastructure.

Currently, within Akagera National Park, there exist two primary roads for tourist access: Lakeshore Road and Hills Road. Lakeshore Road is equipped with approximately

60 to 80 cameras, while Hills Road boasts about 40 cameras. Despite this coverage, it's important to note that not all areas frequented by tourists are currently under surveillance. To address this gap, we will collaborate with park management to identify these underserved regions and install additional cameras as needed.

When the camera is actively recording video, we will capture individual frames as images and pass them to the model for animal identification. Upon successful identification, key information such as the animal's species, date, time, and precise location will be transmitted to a cloud-based database. We have developed a dashboard interconnected with the cloud database, which actively monitors changes in the database and subsequently refreshes its display to reflect the most recent information accordingly. This real-time feedback loop ensures that pertinent data is readily accessible to tourists and park management alike, facilitating informed decision-making.

3.1 Raspberry Pi

The Raspberry Pi, a compact and powerful minicomputer roughly the size of a standard credit or debit card [9–11], offers exceptional versatility. It can seamlessly connect to a computer monitor or TV and is compatible with a standard keyboard and mouse [9]. Additionally, it can accommodate multiple sensors and actuators simultaneously, making it ideal for innovative embedded system projects [11]. Raspberry Pi boards find applications across a wide spectrum of purposes, from home automation and zero-powered smartphone development to serving as universal remote controls for smart home appliances [9]. However, it has its drawbacks, such as the absence of internal storage (eMMC), reliance on slower SD cards, and the lack of a dedicated graphics processor. Overheating can be an issue with extended usage due to the absence of integrated cooling, and it cannot run the Windows OS which is the most widely used operating system. Figure 1 showcases the Raspberry Pi Board.

In the context of this research, we utilized the Raspberry Pi 2 as the hosting platform for the Yolov5 model and camera code. Furthermore, we employed it to transmit the captured data to the cloud. The Raspberry Pi 2 is equipped with a 900 MHz quad-core ARM Cortex-A7 CPU. It features [10] a 100 Base Ethernet, 4 USB ports, 40 GPIO pins, a Full HDMI port, a combined 3.5mm audio jack and composite video, a Camera interface (CSI), Display interface (DSI), Micro SD card slot, and VideoCore IV 3D graphics core.

3.2 Image Dataset Acquisition

As highlighted in the preceding section, the camera assumes a pivotal role in this research, serving as the primary means for continuous animal recording. Our strategy involves leveraging the pre-existing cameras strategically situated in key areas of the national park, which are originally intended for various purposes. However, for this pilot project, we have opted to employ a webcam as our initial proof of concept. A webcam is essentially a basic video camera that captures both still images and videos when connected to a computer [12]. In our study, we have interfaced it with the Raspberry Pi via a USB connection to capture field images. A Python script facilitates seamless communication between the webcam and the Raspberry Pi.

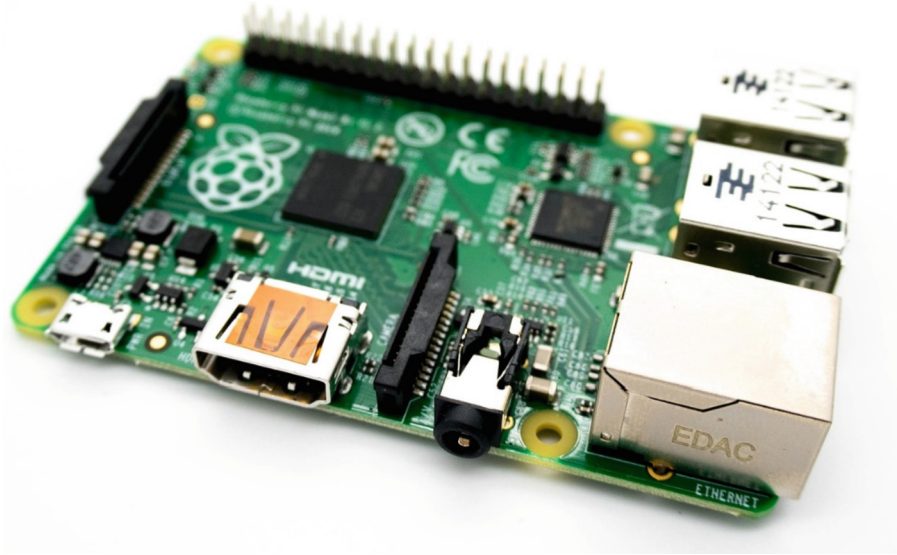


Fig. 1. Raspberry Pi

As previously mentioned, it's worth noting that the Raspberry Pi lacks internal storage capacity. To address this challenge, we have implemented a system where all the images captured are transmitted to a cloud-based database. These captured images encompass a wide variety of species, including animals such as gorillas, elephants, buffalo, giraffes, zebras, leopards, hyenas, lions, impalas, chimpanzees, monkeys, crocodiles, and even individuals. To augment our dataset, we sourced additional images from publicly available datasets.

This curated dataset underwent preprocessing and annotation procedures. Subsequently, it was partitioned into three distinct subsets: training, validation, and test sets, as depicted in Fig. 2. The training set served as the foundation for training the YOLOv5 model, while the validation set was instrumental in assessing the model's performance during training. The test set is reserved to evaluate the model's final performance. This partition process ensures that the model generalizes effectively, and that the accuracy achieved on the test set accurately represents the model's performance on unseen data.

3.3 Object Detection Model (YOLOv5)

YOLOv5 is a model within the You Only Look Once (YOLO) family. The YOLO models are commonly used for object detection tasks and are a popular choice for various computer vision applications [13]. YOLOv5 [14, 15] builds upon the YOLO series'(YOLOv1, YOLOv2, YOLOv3, YOLOv4) real-time object detection capabilities and is designed to detect objects within images and videos efficiently. It is known for its speed and accuracy in detecting objects in a single pass through a neural network. Compared to its predecessor, YOLOv5 introduced improvements in architecture, training techniques, and performance [16]. The model offers different sizes (small, medium,

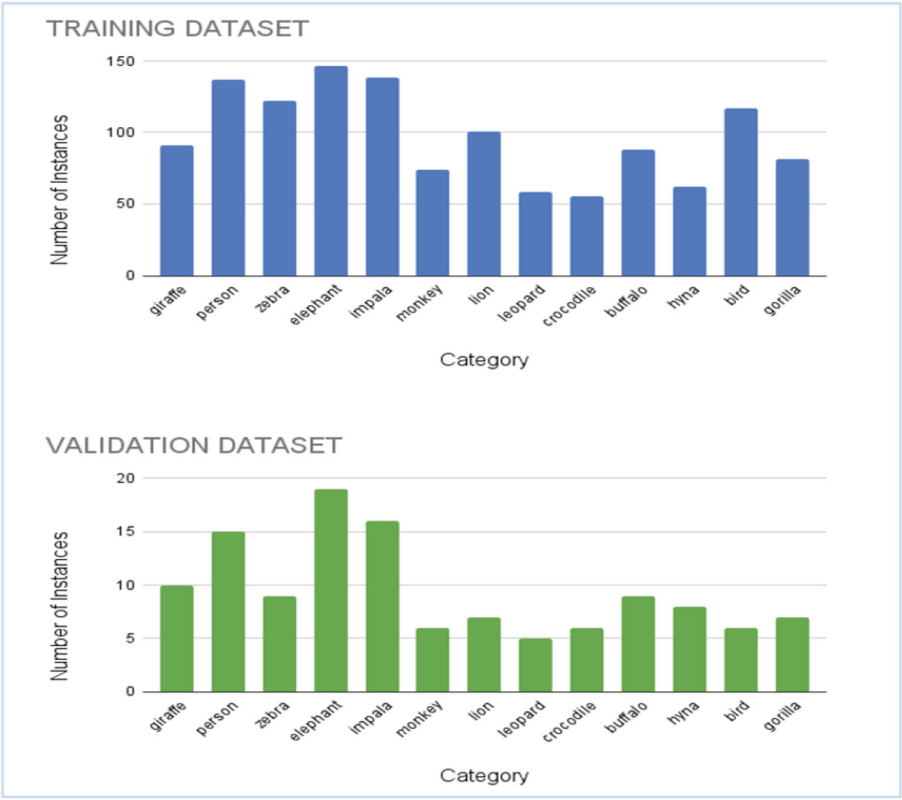


Fig. 2. Image dataset curation.

large, extra-large) [14]to accommodate varying hardware and performance requirements. YOLOv5 has been implemented using the Python programming language, departing from the previous versions that were primarily written in C. This shift simplifies the process of installation and integration, particularly on IoT devices [2]. The YOLO model processes an image through a single pass of a Fully Convolutional Neural Network (FCNN), hence "You Only Look Once". This streamlined approach results in faster object detection compared to the traditional Region-based CNN model, which involves a multi-step detection process [14].

The YOLO model architecture, as depicted in Fig. 3, consists of three main components [16]. First, there is the "Backbone," a convolutional neural network responsible for all the learning processes. Its role is to extract essential features from the input image, and in the case of YOLOv5, it utilizes Cross Stage Partial (CSP) Networks as its foundation. The "Neck" component follows, which is tasked with creating feature pyramids. It achieves this by employing a set of layers that blend and combine image characteristics before they are passed for prediction. Finally, the "Head" component receives the output from the neck, and it is from here that box and class predictions are made. The head can be configured as either one-stage for dense prediction or two-stage for sparse prediction.

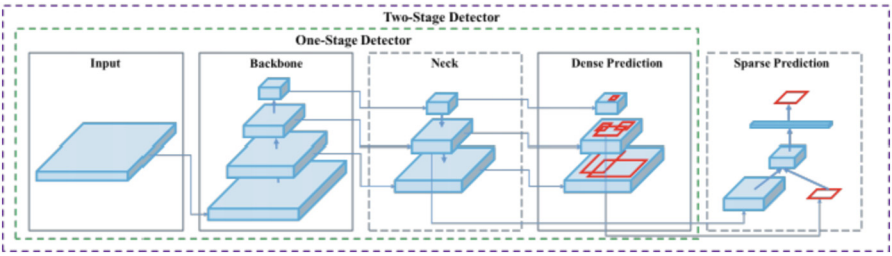


Fig. 3. Architecture of YOLO model [16].

In this study, we employed YOLOv5s, a versatile model within the YOLOv5 family. YOLOv5s offers several significant advantages for object detection tasks. Firstly, it stands out for its remarkable speed and lightweight nature, making it exceptionally suitable for real-time applications. Additionally, YOLOv5s is memory-efficient, demanding less memory than its counterparts. This characteristic facilitates its deployment on resource-constrained devices like Raspberry Pi, mobile phones, and other embedded systems. While we may not claim the title of the most accurate model in the family, YOLOv5s consistently delivers strong accuracy, making it a reliable performer across a broad spectrum of object detection tasks. Moreover, it is capable of handling diverse object detection challenges, ranging from detecting small objects in cluttered environments to identifying large objects in high-resolution images.

3.4 React JS Dashboard

A client application developed with React JS is used to display animal locations on maps. React JS is chosen for its ability to create reusable components and build user interfaces. The application integrates with the YOLOv5 object detection model running on a Raspberry Pi. The detection data is sent to the cloud, and animal location information is retrieved and displayed on maps using the ArcGIS API for JavaScript. The React JS app also provides features like filtering animals by class and displaying additional detection information (such as time and date). This application plays a crucial role in enhancing the tourists' experience by enabling them to visualize animal locations on maps.

3.5 Design and Implementation

Figure 4 provides a comprehensive overview of the proposed system architecture, encompassing data collection, model training, and deployment. The training phase, illustrated in Fig. 5, involved exposing the model to a dataset of annotated animal images and their corresponding labels, facilitating its learning and improved accuracy in animal detection. The integration of hardware and software is facilitated through APIs. When the system identifies an animal, it utilizes an API to store its data in a PostgreSQL cloud database. This continuous process operates at a frame rate ranging from 15 to 30 frames per second (fps) once the device is powered on, ensuring uninterrupted functionality.

Figure 6 illustrates the hardware components, including the Raspberry Pi, webcam, and the case composing the configuration for this system.

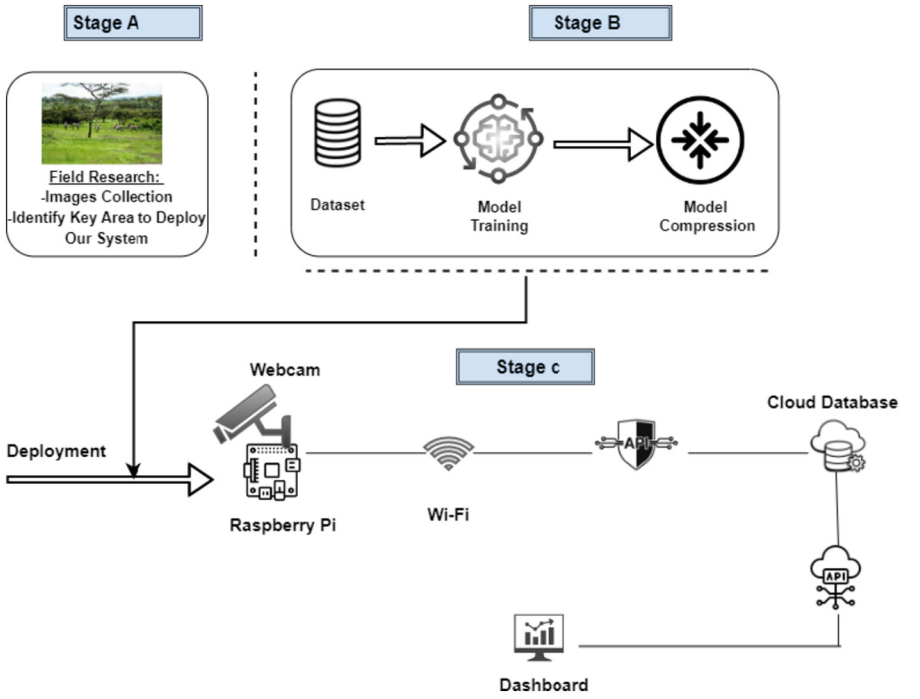


Fig. 4. Proposed system architecture.

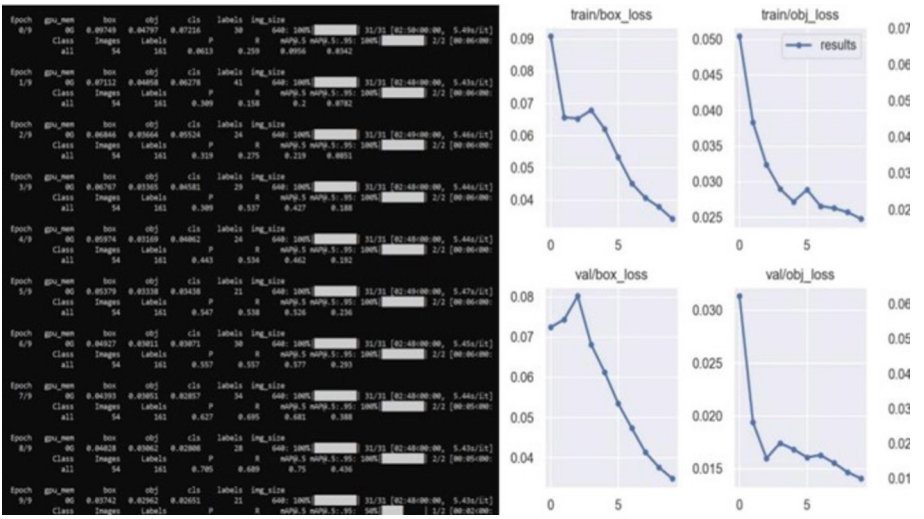


Fig. 5. Model training.



Fig. 6. Hardware setup.

4 Results Discussions

The model demonstrated its capability by accurately detecting all specified animal categories, achieving an impressive accuracy rate of 94.4%. This accuracy underscores the effectiveness of the employed algorithm. A separate test dataset reaffirmed the algorithm's efficacy. Figure 7 provides insights into the model's performance, displaying mean average precision and recall metrics with sample predictions. The right side of the figure highlights the predicted animals and their respective categories.

After deploying the model on the Raspberry Pi, details regarding the identified animal are transmitted to a cloud database for subsequent retrieval and display on the dashboard. Figure 8 illustrates that this information encompasses the captured image, stored in a base64 string format, along with the animal's name, location, date, and time. This pop-up feature simplifies the process for users to gather crucial information about each animal sighting, eliminating the need to navigate through multiple pages and significantly improving the user experience.

The developed dashboard includes an insights page, broadening the scope of information accessible to end-users. Among the featured insights, Fig. 9 highlights the 'most searched' animal, while Fig. 10 presents the 'Big Five' category, indicating the peak time and month when these animals are most frequently spotted. The Big Five category comprises animals that visitors most desire to see, namely Lion, Buffalo, Elephant, Rhino, and Leopard, known for their distinctive character in Akagera National Park.

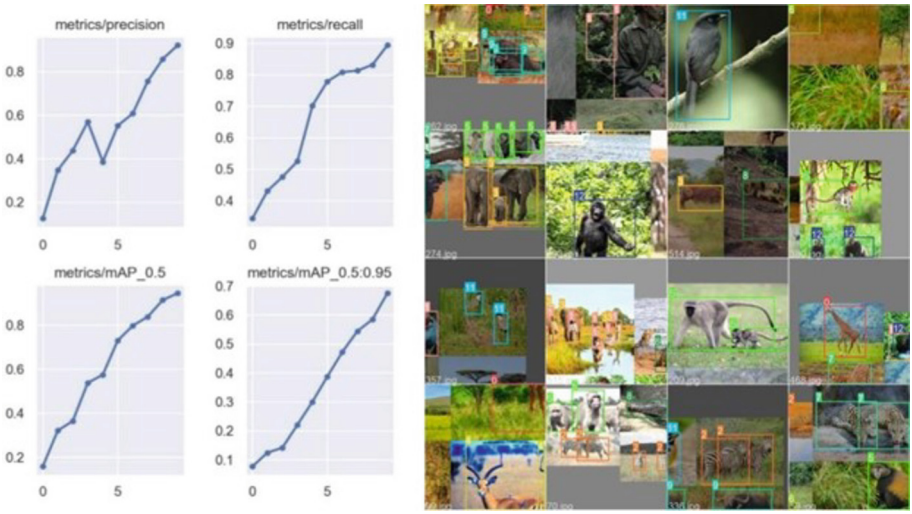


Fig. 7. Model metrics.

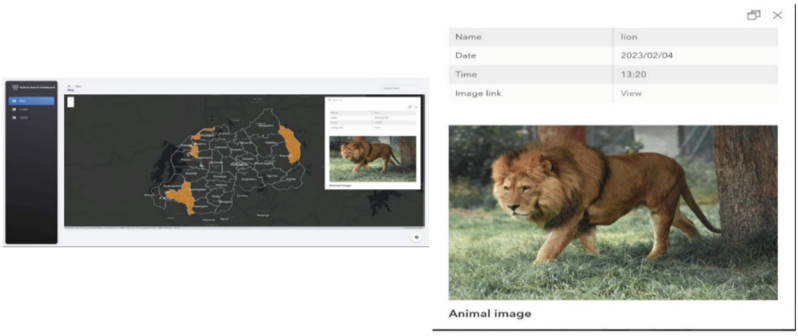


Fig. 8. Pop-up with information about the detected animal.

These insights extend the dashboard's utility beyond real-time animal detection, offering valuable data for park management. The data reveals which animals garner the most interest from visitors, enabling targeted marketing campaigns to promote lesser-known but equally intriguing species. Moreover, these insights can empower park management to customize conservation and educational programs to align with visitor interests, enhancing engagement and educational value. Additionally, the data provides insights into animals that may require increased conservation efforts or attention due to their popularity.

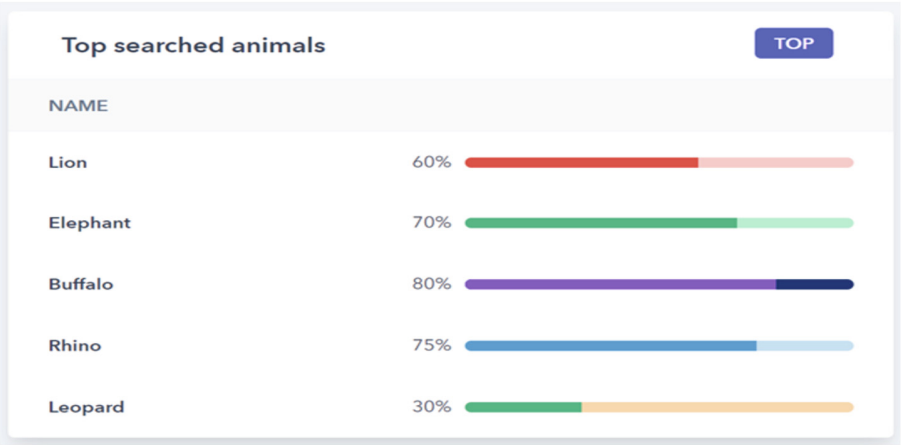


Fig. 9. The most searched animal insight.

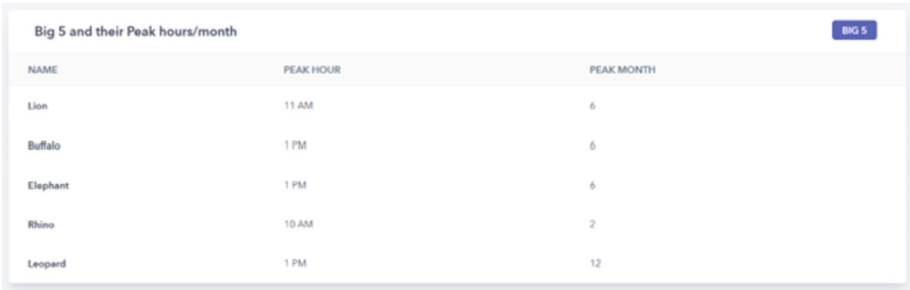


Fig. 10. Peak time and month when Big 5 are most frequently spotted.

5 Conclusion and Recommendation

This study has successfully demonstrated the creation of a cost-effective and efficient system that combines embedded system skills with machine learning to assist tourists in locating animals and planning their visits effectively. The system, as demonstrated by the dashboard, provides not only the location of animals but also other useful insights. The utilization of the YOLOv5 model proves its effectiveness in accurately identifying various animal types. While this study discusses a single system in one location, its potential is significant when deployed across the entire national park. This expansion will require collaboration with park management to ensure reliability, effectiveness, and ethical considerations regarding the animals and their natural habitats. The proposed system has the potential to contribute significantly to wildlife conservation and animal behavior research by offering an improved means of monitoring and studying animal populations in their natural environments, enhancing our understanding of their behavior, and supporting conservation efforts.

As for recommendations, the project's success hinges on deploying the system extensively throughout the park's key areas. Further research must explore the handover process between cameras in areas where the view of animals is interrupted due to camera placement or terrain. This exploration could benefit from machine learning algorithms and predictive models to enhance the effectiveness of camera handovers, ultimately improving animal monitoring and visitor experiences in the park. Additionally, to enhance location tracking accuracy, investigating factors influencing animal locations, such as food availability, terrain, and vegetation cover, can lead to more accurate location tracking systems and better-informed conservation strategies.

Acknowledgements. This work was jointly supported by the African Center of Excellence in Internet of Things (ACEIoT), College of Science and Technology, University of Rwanda.

References

1. Rwanda economic update: nature-based tourism holds tremendous economic Potential, https://www.worldbank.org/en/news/press-release/2023/02/21/rwanda-afe-economic-update-nature-based-tourism-holds-tremendous-economic-potential
2. Thuan, D.: Evolution of yolo algorithm and Yolov5: the state-of-the-art object detention algorithm (2021). https://www.theseus.fi/handle/10024/452552
3. Sharma, S.U., Shah, D.J.: A practical animal detection and collision avoidance system using computer vision technique. IEEE Access. **5**, 347–358 (2017). https://doi.org/10.1109/access.2016.2642981
4. Santhanam, S., Panigrahi, S.S., Kashyap, S.K., Duriseti, B.K.: Animal detection for road safety using deep learning. In: 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA) (2021)
5. Singh, A., Pietrasik, M., Natha, G., Ghouaiel, N., Brizel, K., Ray, N.: Animal detection in man-made environments. In: IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 1438–1449. IEEE Xplore (2020)
6. Datt Bhatt, N., Kumar, S., Rai, S., Bastakoti, J.: GPS based animal tracking with SMS Alert. In: KEC Conference (2019)
7. Antônio, W.H.S., Da Silva, M., Miani, R.S., Souza, J.R.: A proposal of an animal detection system using machine learning. Appl. Artif. Intell. **33**, 1093–1106 (2019). https://doi.org/10.1080/08839514.2019.1673993
8. Kays, R., et al.: Tracking animal location and activity with an automated radio telemetry system in a tropical rainforest. Comput. J. **54**, 1931–1948 (2011). https://doi.org/10.1093/comjnl/bxr072
9. Ghael, H., Solanki, D., Sahu, G.: A review paper on raspberry pi and its applications. Int. J. Adv. Eng. Manage. (IJAEM). **2**, 225 (2021). https://doi.org/10.35629/5252-0212225227
10. Alex David, S., Ravikumar, S., Rizwana Parveen, A.: Raspberry Pi in computer science and engineering education. In: Thalmann, D., Subhashini, N., Mohanaprasad, K., Murugan, M. (eds.) Intelligent Embedded Systems. LNEE, vol. 492, pp. 11–16 (2018). Springer, Singapore. https://doi.org/10.1007/978-981-10-8575-8_2
11. Kurniawan, A.: Introduction to Raspberry Pi. In: Raspbian OS Programming with the Raspberry Pi, pp. 1–25. Apress, Berkeley, CA (2018)
12. Sugumaran, N., Vijay, G.V., Annadevi, E.: Smart surveillance monitoring system using raspberry pi and pir sensor. Int. J. Innov. Res. Adv. Eng. (IJIRAE). **4**, 11–15 (2017)

13. Jocher, G.: Ultralytics YOLOv5 - Ultralytics YOLOv8 Docs. https://docs.ultralytics.com/yolov5/
14. Marko, H., Jelecevic, L., Gledec, G.: A comparative study of YOLOv5 models performance for image localization and classification. In: Central European Conference on Information and Intelligent Systems, pp. 349–356. Faculty of Organization and Informatics Varazdin 2022 (2022)
15. Wen, H., Dai, F., Yuan, Y.: A study of YOLO algorithm for target detection. In: Proceedings of International Conference on Artificial Life and Robotics, vol. 26, pp. 622–625 (2021). https://doi.org/10.5954/icarob.2021.os13-9
16. Solawetz, J.: YOLOv5 new version - improvements and evaluation. https://blog.roboflow.com/yolov5-improvements-and-evaluation/



A Software Library for Managing Groups of Collector Motors in Robotics

Rinat Galin(✉), Daniiar Volf, Saniya Galina, and Mark Mamchenko

V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow 117997, Russia
rinat.r.galin@yandex.ru

Abstract. The article considers group management and control of collector motors with different types of drivers. The paper deals with a scheme where the microcontroller manages a cyberphysical system, which is a combination of heterogeneous motors and drivers with different control principles. Motors perform a single task in each group, and the composition of the group can be changed by reassigning each motor (or multiple motors) to other groups. We set the problem of controlling groups of motors under given conditions and propose the corresponding software solution. The presented software library allows to simplify the implementation of algorithms to control robotic devices by using a high-level programming interface to control the groups of collector motors, and create an abstraction layer to specify operating modes for single motors and their groups without taking into account specific electric control schemes and features of each model of a motor or its driver. This greatly simplifies the work with the motors for the end user, and allows the use of heterogeneous motors and drivers in a single robotic project, as well as the replacement of motor drivers with their compatible analogues. The proposed software library can be used both in Arduino projects, and other frameworks and environments for programming robotics.

Keywords: DC Motor · Motor Driver · Microcontroller · Driver Control Principle · Group Management · Group Control · User Programming Interface

1 Introduction

At present, the market for electronic control modules (drivers) for DC motors used in robotic systems is very extensive. Among the most common motor drivers for wheeled and tracked robots are L298N, TB6612 (SunFounder), and MX1508. These drivers are auxiliary power modules for microcontrollers (usually AVR/STM32) to control DC motors. To simplify the work with the motor drivers, different software solutions (libraries) can be used. An example is the STM32F10x Standard Peripherals Library [1], created by STMicroelectronics, which has required functionality and macros to facilitate work with the STM32F10x family of microcontrollers. Another similar library is Stepper [2], which is designed to work with unipolar or bipolar stepper motors for Arduino projects. Other known Arduino-compatible libraries are AFMotor and Freeduino Motor Shield [3, 4], that enable a programming interface to operate ordinary DC, servo, and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 424–432, 2024.
https://doi.org/10.1007/978-3-031-53549-9_40

stepper motors. For Raspberry microcomputers Adafruit Motorkit [5] is most commonly used as an auxiliary library for programming collector and stepper motors.

However, none of the considered libraries support management and control of groups of motors. To prove the relevance of this problem, consider two scenarios.

In the first scenario, we use a simple cyberphysical system (CPS) to manage heterogeneous DC motors – three drivers connected to the microcontroller (the electrical wiring diagram is shown in Fig. 1) control four motors (Motor 1, Motor 2, Motor 3, and Motor 4). During the operation of the CPS, it becomes necessary to group multiple motors (for example, in order to perform a single task) and manage them at the group level. Each group has a specific task, and the motors can be moved from one group to another (if necessary). In case of single-type drivers, the problem of controlling a group of motors is trivial and can be solved quite easily. However, the use of drivers of different manufacturers implies that they may differ in the implemented principles and logic of electrical control [6]. For example, the L298N driver uses three electrical pins (Table 1), the TB6612 driver multiplexes four separate control signals (Table 2), while the MX1508 driver requires only two input signals to control two motors (Table 3). Define the following control scheme:

1. For the MX1508 driver, the A1 and A2 pins are used both for control of the direction and speed of Motor 1 – via the pulse width modulation (PWM);

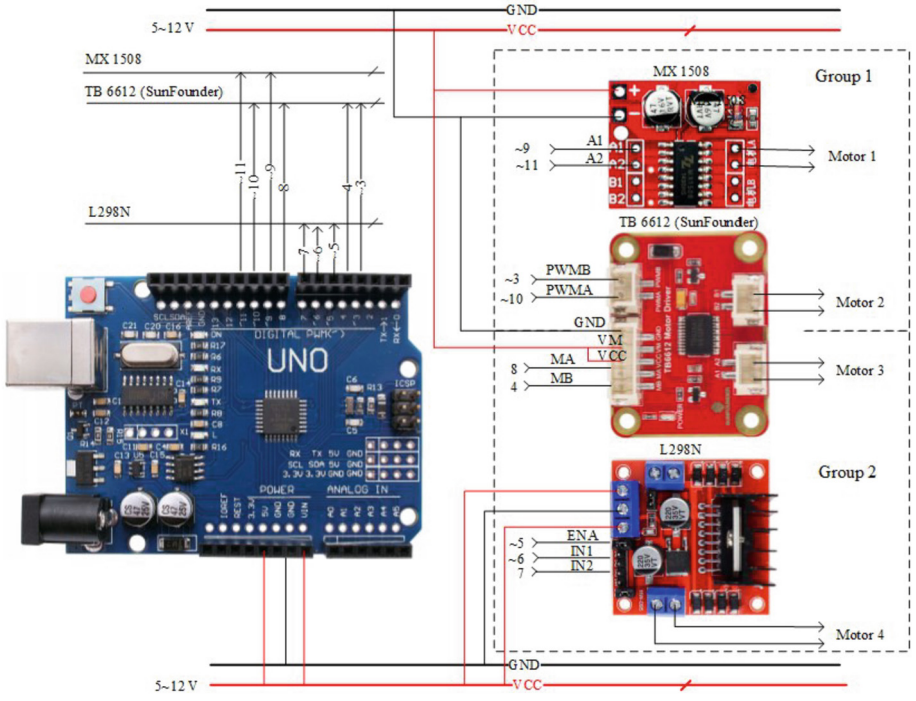


Fig. 1. Electrical wiring diagram for the L298N, TB6612, and MX1508 drivers connected to the microcontroller.

2. For the TB6612 driver, the rotation directions for Motor 2 and Motor 3 are set via MA and MB pins, and the rotation speed is set via PWMA and PWMB pins;
3. For the L298N driver, the direction of rotation of Motor 4 is controlled via the IN1 and IN2 pins, and the ENA pin is used to control the motor speed using PWM.

Table 1. L298N driver control logic.

IN1 pin	IN2 pin	ENA pin
HIGH	LOW	PWM
LOW	HIGH	PWM
LOW	LOW	LOW
HIGH	HIGH	LOW

Table 2. TB6612 (SunFounder) driver control logic.

MA pin	PWMA pin	Result
HIGH	PWM	Forward
LOW	PWM	Backward
HIGH	LOW	Idle
LOW	LOW	Halt

Table 3. MX1508 driver control logic.

A1 pin	A2 pin	Result
PWM	LOW	Forward
LOW	PWM	Backward
LOW	LOW	Idle
HIGH	HIGH	Halt

Consider another scenario. Motor 2 and Motor 3 are controlled according to the specifications of the driver they are physically connected to (e.g., TB6612), but during the operation it becomes necessary to combine Motor 1 and Motor 2 into one group (Group 1), and Motor 3 with Motor 4 – into Group 2. However, according to the wiring diagram in Fig. 1, Motor 1 and Motor 4 are controlled by different types of drivers.

As a result, it is clear that during the stages of development, operation, and further maintenance of software for the robotic systems, it is necessary to take into account the management and control logic for each specific type of driver [7, 8]. Thus, in case of using of a large number of drivers of different types the program code for motors control may become cumbersome and difficult to maintain.

2 Problem Statement and the Proposed Solution

To solve this problem, we developed a specialized software library in C++ that provides a level of abstraction and an interface for group control of collector-type motors [9].

The first basic idea of the software is to represent a bundle of different types of motors and drivers as a single uniform software abstraction – a set of containers:

$$D_1 = \{m_1, m_2, \ldots, m_P\}, D_2 = \{m_{P+1}, m_{P+2}, \ldots, m_K\}, \ldots, D_M = \{m_{K+1}, m_{K+2}, \ldots, m_N\}, \quad (1)$$

where m_i is an object-oriented model of i motor.

Management of software models (individual motors in the form of objects, and groups of motors – as containers) is carried out by a special module (manager) via a high-level interface.

Figure 2 shows the link between the physical objects and software entities that are created in the library to control and manage a group (groups) of motors with different types of drivers. Each abstract motor is "mapped" as an independent program control object, that can be added to a group, regardless of the driver connected to it (e.g., Group 1 and Group 2).

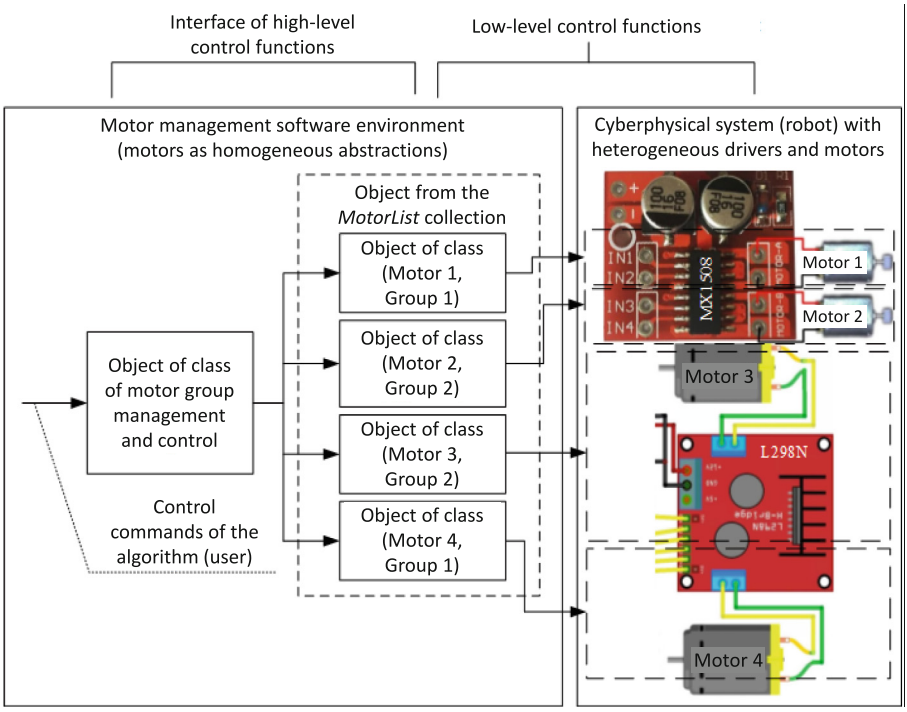


Fig. 2. Links between the physical and software objects for management and control of a group of four motors connected to different types of drivers.

The core of the library is two software classes – *Motor* and *MotoDriver*. *Motor* is responsible for creating software abstractions of physical motors and implements low-level motor driver management functions. The *MotoDriver* class controls *Motor* class objects via high-level programming methods for motor group management [10, 11]. The user interface of mapping of i motor (creating software abstraction entity) is as follows:

$$m_i = Motor(arg1, arg2, arg3, arg4), i = 0, \ldots N, \tag{2}$$

where $arg1$ is the PWM output on the microcontroller; $arg2$ – is the first control output; $arg3$ is the second control output; $arg4$ is the motor type.

Depending on the type of driver, *arg1* and *arg3* variables may not necessarily be declared as the input to the function (for example, when $arg2 = arg3$), and there is no constructor and method overloading in classes at the current stage of software implementation.

Then each container with motors is presented as a number of subsets – "motor unique number – motor model" pairs for each motor in the container. Any container can be represented by at least one "motor number – motor model" pair, i.e. each group should have at least one motor in it. Thus, depending on the hardware capabilities of the microcontroller, software objects of the class and a set of pairs are created:

$$S_j = \{\{1, m_1\}, \{2, m_2\}, \ldots, \{P, m_P\}\}, j = 0, \ldots, M, \tag{3}$$

where P is the unique identifier of the motor (integer); M – is the number of containers.

After that, the object of the control class of j group of motors is created using the *MotoDriver* function with S_j set as an input:

$$obj_j = MotoDriver(S_j), j = 0, \ldots, M. \tag{4}$$

To manage individual motors (excluding group features), three methods have been developed: obj_j.*Clockwise_Rotation*([0:255], i), obj_j.*Counterclockwise_Rotation* ([0:255], i), and obj_j.*Halt*(i). The *Clockwise_Rotation* method causes the rotor to rotate clockwise, and *Counterclock_Rotation* is responsible for counterclockwise rotation. Both methods are called with two identical parameters. The first input parameter ([0:255]) is the PWM cycle value. The second parameter i is the unique identifier of the motor. The third method – *Halt* – stops the rotation of the rotor.

The second basic idea of the library is the group to which i motor belongs. After the stage of initialization, all motors are assigned to group 0 (by default). Any single motor or a number of motors can be reassigned to another group [12, 13] using the obj_j.*MotorToGroup*(i, q) method, where $q \neq j$ is a unique group identifier. This method can be called as soon as *MotoDriver* class objects are initialized, as well as during the execution of user's control algorithms.

The direction of rotation of all the motors of q group, as well as their simultaneous stop is specified using obj_j.*Clockwise_Group_Rotation(q)*, obj_j.*Counterclockwise_Group_Rotation(q)*, and obj_j.*Halt_Group*(q) methods.

Overall, the developed library includes the following software components (Fig. 3):

1. *motodriver.h* header file, which declares the main classes and methods, links the external *ArduinoSTL* library to use *map* and *string* classes. This file has only a list of fields and functions, bodies of the functions are not included;

2. *motordriver.cpp* implementation file with classes and methods;
3. *type-declaration.h* header file with a list of supported motor drivers.

Library class methods are implemented according to C++ 17 standard. The *Motor* class contains *motorID*, *groupId* and *shield_driver_type* public fields, corresponding to the unique identifier of the motor, the group to which it belongs, and the type of driver connected. Private fields of the class (*pwm_pin*, *direction*, and *dcycle*) can be accessed using the *get* and *set* methods (getters and setters). All *MotoDriver* class methods are public and implement group management over *Motor* class collections of objects.

The source code of the developed library, as well as examples of using it are available on the GitHub developer web page [14].

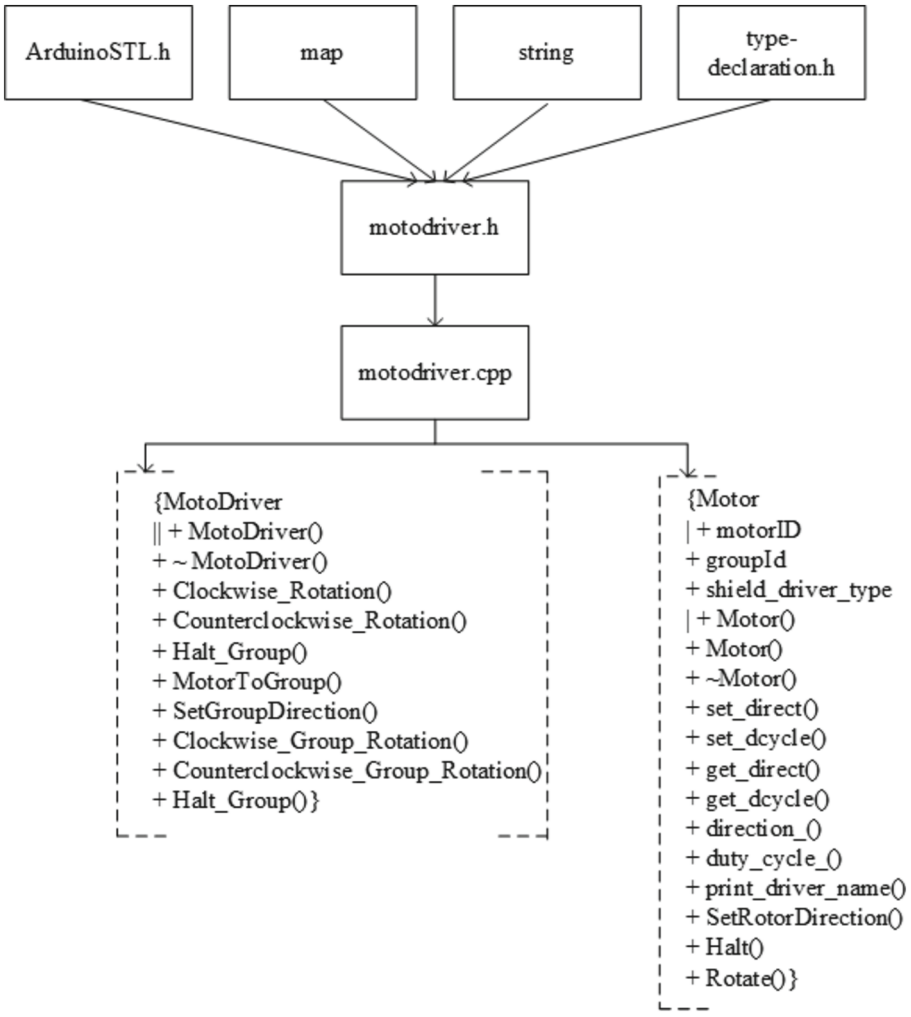


Fig. 3. Structure of the library for group management and control of collector motors.

2.1 Practical Experiment

For testing and validation of the developed library an experimental stand was built (see Fig. 4) in accordance with the scheme given in Fig. 1: L298N, TB6612, and MX1508

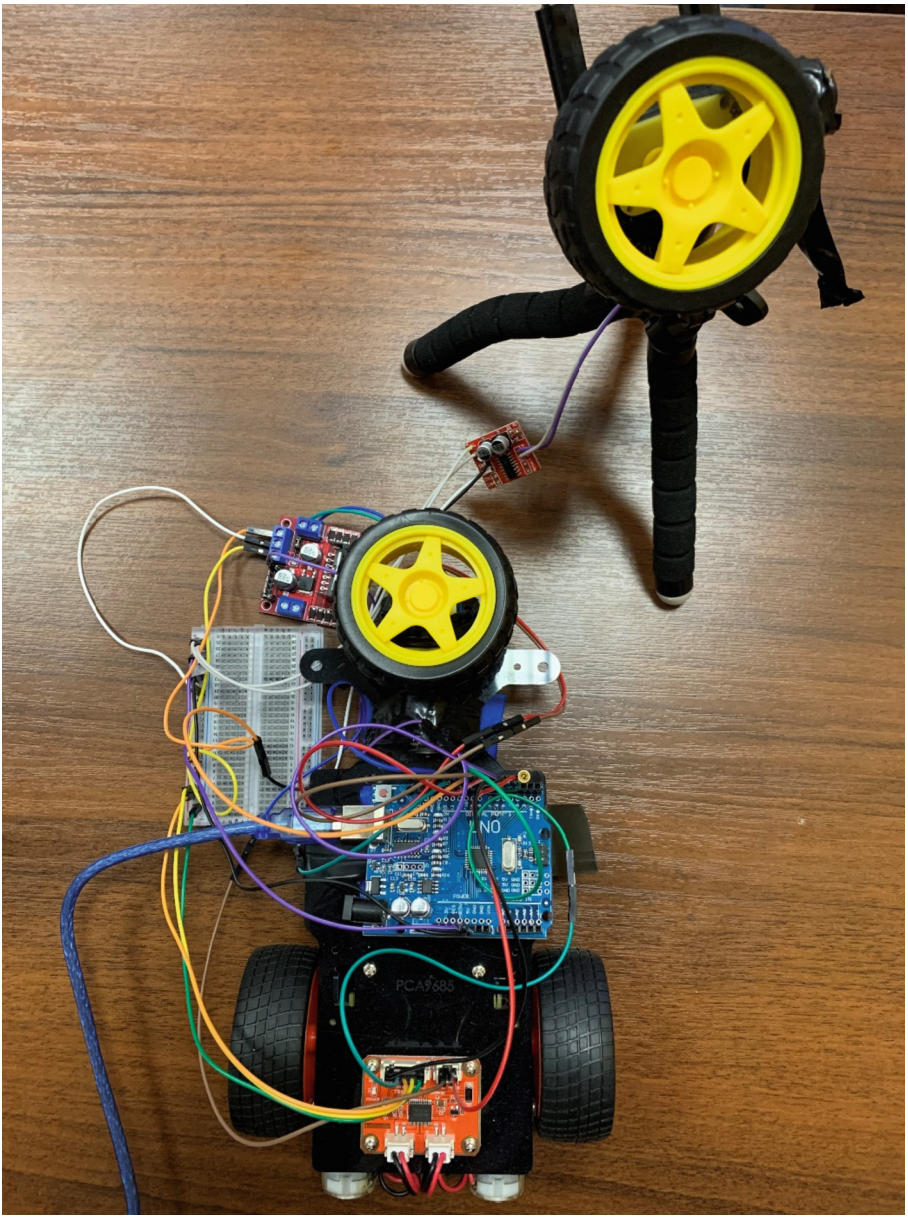


Fig. 4. The experimental stand.

motor drivers were connected to the ATmega328p microcontroller (Arduino UNO). During the experiment, drivers and motors were combined into different groups. Each group was assigned a task of rotating the wheels clockwise or counterclockwise, while the composition of the groups changed instantly ("on the fly"), and the tasks of the motor groups also changed multiple times (clockwise rotation/counterclockwise rotation/halt). In general, the experiment showed that the library obtains all the above-mentioned functionality. Combining motors with heterogeneous drivers into groups makes the control and management of the robot much easier in comparison to controlling individual independent motors and their drivers.

3 Conclusion

The article proposes a solution to control and manage heterogeneous motors in cyber-physical systems, its architecture, as well as program implementation (a software library). The presented library is intended to simplify the development of motorized robotic devices, allowing the user to abstract from the hardware specifications and focus on the implementation of algorithms to control the robot. Throughout the practical experiment, it was found that combining motors with heterogeneous drivers into groups using the proposed library makes it much easier to operate the robot compared to the control of individual independent motors and their drivers. In the future, it is planned to expand the list of supported motor drivers, as well as to adapt the library for other software development frameworks.

References

1. STM32F10x standard peripherals library. https://github.com/wajatimur/stm32f10x-stdperiph-lib. Accessed 20 Sep 2022
2. Stepper library for Arduino. https://github.com/arduino-libraries/Stepper. Accessed 20 Sep 2022
3. AFMotor library. https://github.com/adafruit/Adafruit-Motor-Shield-library. Accessed 20 Sep 2022
4. Freeduino motor shield library. https://github.com/adafruit/Adafruit-Motor-Shield-library. Accessed 20 Sep 2022
5. Adafruit Motorkit Library. https://docs.circuitpython.org/projects/motorkit/en/latest/. Accessed 20 Sep 2022
6. Tiwari, D., Miscandlon, J., Tiwari, A., Jewell, G.W.: A review of circular economy research for electric motors and the role of industry 4.0 technologies. Sustainability **13**(17):9668, 1–19 (2021)
7. Shagin, A.V., Naung, Y., Khaing, Z.M.: Control method of DC motor with closed loop system of precision positioning system. Izvestiya vuzov. Elektronika (Proc. Univ. Electron.) **23**(3), 304–308 (2018) (In Russ.)
8. Shestakov, A.V.: Modeling and experimental analysis of dynamic characteristics of asynchronous motor. In: 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), pp. 1–4. IEEE, Piscataway (2019)
9. Al-Mahturi, F.S., Samokhvalov, D.V., Bida, V.M.: Parameters identification of a brushless DC motor by specification. In: 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), pp. 558–561. IEEE, Piscataway (2018)

10. Galin, R., Shiroky, A., Magid, E., Meshcheryakov, R., Mamchenko, M.: Effective functioning of a mixed heterogeneous team in a collaborative robotic system. Robot. Autom. Control Syst. **6**(20), 1224–1253 (2021). (In Russ.)
11. Galin, R., Meshcheryakov, R., Kamesheva, S.: Distributing tasks in multi-agent robotic system for human-robot interaction applications. In: Ronzhin, A., Rigoll, G., Meshcheryakov, R. (eds.) Interactive Collaborative Robotics. ICR 2020. LNCS, vol. 12336, pp. 99–106. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60337-3_10
12. Podvesovskii, A., Zakharova, A., Korostelyov, D., Kuzin, A.: DecisionMaster: a multi-criteria decision support system with ability to combine different decision rules. SoftwareX **16**(100814), 1–8 (2021)
13. Galin, R., Mamchenko, M., Meshcheryakov, R.: Analysis of the allocation and implementation of tasks in the heterogeneous team of the collaborative robotic system. In: Ronzhin, A., Shishlakov, V. (eds.) Electromechanics and Robotics. Smart Innovation, Systems and Technologies, vol. 232, pp. 109–119. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-2814-6_10
14. Motodriver universal Arduino library. https://github.com/Runsolar/motodriver. Accessed 20 Sep 2022

EdApp as a Tool to Intensify Foreign Language Professional Training in the Digitalization of the Educational Environment

Dmitrii Burylin, Damir Ibraimov, Nadezhda Chernova, Natalia Katakhova, Irina Osliakova(✉), Tatiana Kudinova, and Svetlana Katahova

MIREA-Russian Technology University, Moscow, Russia
oslyakova@mirea.ru

Abstract. The paper considers the problem of intensification of foreign language professional training under the conditions of digitalization of educational environment in a technological university. Theoretical and methodological and comparative analysis in the course of experimental work showed that the use of EdApp digital educational platform contributes to the development and improvement of professional foreign language competences of students. #COMESYSO1120.

Keywords: Intensification · Foreign Language Training · Digitalization · Digital Educational Platform · Moodle · EdApp · Foreign Language Competence

1 Introduction

After the education industry moved from traditional education to online learning, it was further united with big data and created a data warehouse to create a new innovative online education that greatly simplified student learning and teacher teaching [11]. As information and communication technology has developed, online education has become more feasible from technological, economic and operational perspectives [12]. The introduction of information technology tools in the educational process makes it possible to prepare highly qualified competent specialists to solve production issues in the professional field of their activities. The effectiveness of the educational process is enhanced by its intensification.

Theoretical and methodological aspects of this issue have been dealt with by a number of researchers I.A. Zimnaya, G.A. Kitaygorodskaya, G.K. Lozanov and others [1]. In the works of A.N. Leontev intensification is considered as a controlled and controlled learning curve, forming professional foreign language competence by comparing speech components of activity with foreign language skills and abilities. V.A. Slastenin, N.A. Polkovnikov, V.V. Petrunsky in their works are aimed at preparing a socially adaptive personality taking into account the psychophysical capabilities of the learner through the use of digital didactic and methodological support.

The problem of intensification of foreign language professional training of technological university students in the conditions of digitalization of the educational environment is relevant. The availability of foreign language communicative skills is a huge

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 433–438, 2024.
https://doi.org/10.1007/978-3-031-53549-9_41

advantage for a competent specialist. One of the key tasks of a teacher of a foreign language discipline is the implementation of variable differentiated methods of formation of speech skills, taking into account individual characteristics of students and psychological and pedagogical conditions of the learning process [2]. At MIREA - Russian technological university foreign language training presupposes only 2 academic hours a week during 2 years (where the 1st year of training - GeneralEnglish, the 2nd year - SpecialEnglish), which is extremely insufficient for full mastering of specifics of professional language. On this basis, the foreign language teacher focuses on extracurricular and independent work, aimed at consolidation and rehearsal of the material learned in the practical classroom lessons with the help of special digital technology.

2 Methods

In recent years, digital forms of learning, in particular distance learning, forced because of the epidemiological situation in the world, have predetermined unique opportunities to obtain absolutely any specialty, thereby strengthening the issue of intensifying professional training in higher education. In this article, we will introduce 10 on-demand learning tools that one can use to make sure that students can access relevant learning materials at the right time.

EdApp [8] is an award-winning LMS for microlearning [9] that is perfect for on-demand learning needs. It is a mobile-centric platform, which means students can access learning materials anytime, anywhere. Using a microlearning model, information is turned into easily digestible snippets combined with gamification elements. As a result, this increases the effectiveness of learning, student engagement, and retention of knowledge. The compact nature of microlearning also makes it easier to create, update, and distribute than traditional formats. This ensures that students have up-to-date information at their fingertips as soon as they need it.

EdApp also features a proprietary drag-and-drop tool that makes it easy to create your own microlessons using more than 80 intuitive templates. There is no need to start from scratch, as access to a content library of 100+ courses that are fully editable and customizable is provided. The course library covers a wide range of topics such as retail, construction, leadership, self-improvement, and many more. EdApp provides essential course material on a variety of subjects.

Key Features:

- Mobile-first - Lessons run easily on both iOS and Android devices. They are also fully responsive and perfectly formatted for any web browser of your choice.
- Editable Microlearning Content Library - EdApp has partnered with the best in the business to create a growing library of courses filled with expertly designed lessons. It is possible to personalize them and rebrand them to your liking.
- Drag-and-drop lesson creation tool - there is the ability to view the lessons you create on different devices and see changes on the fly so you know exactly what students will see. Canva integration feature allows to add visually appealing images directly from the authoring tool.

Advantages:

- Students can access lessons on any mobile device and download them for offline learning
- Short, "bite-sized" lessons help students understand information better
- Easy to create lessons without any coding or instructional design knowledge
- Choose from more than 100 microlearning courses that can be deployed right out of the box
- Create SCORM-compliant courses directly from the authoring tool without the need for developers.

An example of one of today's most popular free LMSs is Moodle. The system provides users with tools for both creating courses and monitoring the learning process. The key advantages are simplicity and ease of use, extensive opportunities for organizing the educational process and monitoring the knowledge of students and relative friendliness to third-party developments and the process of their implementation in the system [10].

3 Results

Moodle is the unified educational digital platform, both during distance learning and during full-time education, as an auxiliary tool in RTU MIREA. In the course of our research on intensification of professional foreign language training we conducted a survey among the first-year students of the Institute of Radioelectronics and Informatics "Is Moodle digital educational platform effectively used in the course of professional foreign language training?" [3]. As a result, only 40% of the respondents under study emphasize the effectiveness of the Moodle digital learning platform in learning a foreign language, citing the availability of information resources in a single digital space, the ability to track progress on the studied material. 12% of the students found it difficult to answer. Thus, almost half of the surveyed students (48%) doubted the effectiveness of the Moodle digital learning platform for foreign language training, citing system interruptions, lack of timely feedback from the teacher, the identity of the tasks performed. That is why we decided to consider other online platforms allowing to diversify and differentiate the learning process in the logic of professional foreign language training. All because for all its versatility Moodle has a number of significant disadvantages:

- In order to organize training and create full-fledged courses, the system must be installed. This will require a server or hosting, a domain name and so on. As a rule, for this organization has to hire an IT-specialist. In the end, setting up the service becomes an expensive task.
- Moodle is demanding on the server. Free hosting allows to install only older versions of the system. In addition, the service consumes a lot of resources, which increases the financial cost.
- Despite frequent updates, Moodle has a rather outdated, not user-friendly interface. If required to make course management more convenient, the entire system will have to be configured separately.
- Moodle - a large system with lots of features, some of which are not used. They can be removed only with the help of programmers. This removal is not possible for all functions.

- If you don't have minimal experience managing an LMS, Moodle will require a long and serious study. It is impossible to immediately understand the abundance of sections, attachments and forms here. Such a method as "let's run it, and then we'll figure it out" will not work.
- The system is geared more toward universities, colleges, and other educational institutions. Most plugins are created specifically for education. Moodle lacks flexibility and scalability for the commercial sector. To integrate the system with the necessary business solutions, it will have to make a lot of effort and attract experienced professionals.
- The analytics tools are there but they do not provide much depth and coverage for business issues.
- Moodle does not issue certificates of completion.

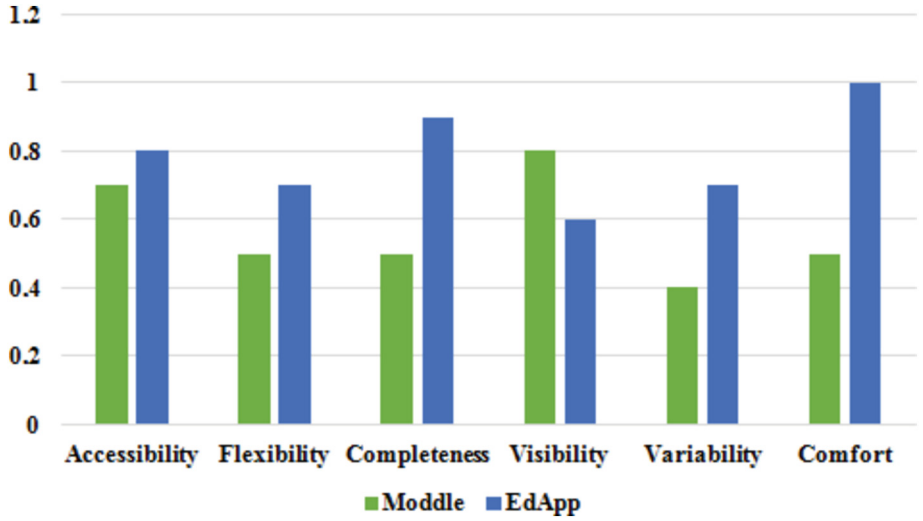


Fig. 1. Comparative analysis of Moodle and EdApp digital education platforms, %.

That is why, when considering these two platforms, we chose EdApp. This educational platform considers a wide range of educational courses developed by internationally recognized experts, offering both template and customized solutions. The developed software allows learners to install the application not only on personal computers, but also on smartphones, smart watches, etc. Thus, emphasizing the accessibility and flexibility of the educational platform EdApp. The developers of this digital educational platform paid special attention to microlearning in which any user can create their own authoring tool, improving professional and technical skills and abilities to organize an effective learning process [4, 5].

Regarding the intensification of professional foreign language training the digital educational platform EdApp is a multifaceted interactive system of training courses in English, which is an effective solution for extracurricular and independent work of students. Using EdApp will allow to form professional foreign language competence

according to the individualized trajectories built by the foreign language teacher to consolidate the lexical aspects and expand the vocabulary within the studied topics. The application of this digital educational platform was assessed by empirical research method - questionnaire survey (Fig. 1).

4 Discussion

First-year technical students of the Institute of Radioelectronics and Informatics of MIREA answered a number of questions, which resulted in the following data of comparative analysis of two educational platforms Moodle and EdApp as an additional integrated tool for professional and language training. The highest indicators of EdApp are "comfort", which implies the use of training courses without time frames, and "completeness", which reflects the content of the topic presented. However, the indicator "visibility" is surprisingly significantly lower in EdApp compared to the language in which the educational platforms are navigated.

5 Conclusion

Thus, the digital educational platform EdApp in intensifying professional foreign language training in the conditions of digitalization of the educational environment, as an additional tool, can be used as an effective integrated component of foreign language training. We can conclude that the intensified approach to foreign language training on the basis of digital educational platforms forms professional foreign language competence for intercultural interaction and building interpersonal relationships, effectively influences the formation of a socially adaptive personality and develops communicative skills and abilities.

References

1. Babansky, Y.K.: Intensification of the learning process. Znanie, Moscow, USSR (1987)
2. Druzhinina, A.I., Tsyplakova, S.A.: Intensification: education development trend and factor in improving the quality of education. In: Proceedings of the II All-Russian student scientific and practical conference "Modern trends in the modernization of Russian education: problems and prospects", pp. 46–48. Branch of Voronezh State University, Voronezh, Russia (2016)
3. Zagvyazinsky, V.I., Zakirova, A.F., Atakhanov R., et al.: Qualitative and quantitative methods of psychological and pedagogical research. Publishing center "Academy", Moscow, Russia (2013)
4. Tsarapkina, Y.M., Shkarupina, V.V.: Application of interactive technologies in the educational process as a basis for self-development of students. Human. Educ. **3**(19), 87–92 (2014)
5. Polat, E.S., Bukharkina, M.Y.: Modern pedagogical and information technologies in the education system. Publishing center "Academy", Moscow, Russia (2008)
6. Hurhish, A., Sterlikova, M.: Sustaining connections: in-class learning and game-based learning in ELT. Int. Sci. J. «Grail Sci.» **12–13**, 476–480 (2022). https://doi.org/10.36074/grail-of-science.29.04.2022.083
7. GetApp. Moodle vs EdApp Comparison. https://www.getapp.com/education-childcare-software/a/moodle/compare/edapp/. Accessed 21 Nov 2016

8. EdApp. There's a better way to train. https://www.edapp.com/. Accessed 21 Nov 2016
9. EdApp. Microlearning Blog. https://www.edapp.com/blog/microlearning-guide/. Accessed 21 Nov 2016
10. Starichkova, J.V., Rogov, I.E., Tomashevskaya, V.S.: Developing the data management component of an academic discipline program for an educational management information system. Russ. Technol. J. **11**(1), 7−17. https://doi.org/10.32362/2500-316X-2023-11-1-7-17
11. Cui, Y., et al.: A survey on big data-enabled innovative online education systems during the COVID-19 pandemic. J. Innov. Knowl. **8**(1), 100295 (2023). https://doi.org/10.1016/j.jik.2022.100295
12. Palvia, S., et al.: Online education: worldwide status, challenges, trends, and implications. J. Glob. Inf. Technol. Manag. **21**(4), 233–241 (2018). https://doi.org/10.1080/1097198X.2018.1542262

Virtual Reality as a Toolkit in the Professional Training of Students

Tatiana A. Shchuchka(✉), Natalia A. Gnezdilova, Nataliya V. Chernousova, and Pavel V. Pankin

Federal State Budgetary Educational Institution of Higher Education, «Bunin Yelets State University», 28, Kommunarov Street, Yelets 399770, Russia
tasiaelez@mail.ru

Abstract. Modernization of the educational system is closely related to the virtual reality, which in this context act as a necessary tool for digital transformation and educational content. The aim of the study is to discover the potential of virtual reality as a tool for digital transformation of education and to assess its application in teaching practice. The following scientific approaches and methods were used in the work: discourse approach, which determines the consideration of the problem within the boundaries of discourses by methods of approach: analysis, synthesis, matching, comparison, analogy; information approach, which operates with the processes of perception, processing of information flows by selective selection; systemic approach, which gives the rationale for the relationship of theories in educational research by system and structural and functional methods. The study reveals the potential of virtual reality as a tool for digital transformation of education in the professional training of students. Based on the analysis of the questionnaire data of teachers and students, the high effectiveness of the virtual reality potential is identified based on the experience of its methodological implementation, and the satisfaction with this process is observed among the majority of students and teachers. Identification and purposeful implementation of virtual reality potential as a tool for digital transformation of education in the practice of professional training of students in Bachelor's, Specialist's and Master's degree programs at Bunin Yelets State University allowed to optimize the process of training specialists based on qualitatively new conclusions, transforming traditional didactic techniques under the influence of new conditions of digitalization of education, which can increase the involvement of students in the process of mastering the discipline and help the teacher when working in this format. #COMESYSO1120.

Keywords: Virtual Reality · Education · Digitalization Tools

1 Introduction

The primary task of modern education is its modernization in accordance with the requirements of the time and the emergence of fundamentally new technologies and forms of learning organization. At the same time, proposing the introduction of innovative means and methods of educational work, it should be understood that the solution to

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 439–450, 2024.
https://doi.org/10.1007/978-3-031-53549-9_42

this difficult task lies in a systemic multidimensional paradigm and can only be implemented on an integrated platform. Modernization of the educational system as a whole, which we have in mind, is closely connected with virtual reality tools, which in this context act as a necessary tool for digital transformation of education.

However, on this way there are objective difficulties associated with terminological nomenclature, differences in definitions, ways of application, which is due to the high rate of development and improvement of the proposed technologies, including virtual reality (VR), augmented reality (AR) and mixed reality MR, which determines the relevance of this study.

The problem is stated, an attempt to solve which is undertaken in the present article: what is the potential of virtual reality as a tool for digital transformation of education and the results of its application in professional training in the evaluation of students and teachers?

Analysis of psychological and pedagogical publications [1–8] showed their insufficiency in the direction of identifying the potential of virtual reality as a tool for digital transformation of education in the professional training of students, which indicated:

- the purpose of the study, consisting in the resolution of the problem;
- research objectives, consisting in identifying the potential of virtual reality as a tool for digital transformation of education and assessing the result of its implementation in the practice of professional training by questioning teachers and students in Bachelor's, Specialist's and Master's degree programs of the Federal State Budgetary Educational Institution of Higher Education "Bunin Yelets State University".

2 Related Works

Russian and foreign literature states that computer introduction into human life has caused cardinal changes in all spheres of human activity and this process will only develop in the future [9–16]. Not only in the foreseeable future, but already in the present, innovative computer technologies are widely used in all sectors of human economic, social, scientific and educational activities. Only recently many companies whose activities are related to the development of products based on high technology have announced themselves on the domestic market.

So, for example, Holo Group company is the leading Russian company on mixed reality products and solutions development, the main direction of which is virtual and augmented reality technologies and solutions development, realizes interesting diversified projects in medicine and construction, topography and excursions sphere. As an example, let us take three of the company's products, which, in our opinion, have practically unlimited application.

1) MR Builder helps construction and architectural companies, customers and clients "see" buildings before they appear, in their actual size and volume on the building site or in any landscape, as well as "see" interiors, engineering equipment of buildings and much more. The main clients of the company today are represented by state corporations and major enterprises, both domestic and foreign.

2) MR Guide allows to create exhibition stands on the digital platform, conduct museum tours, make entertaining and informative virtual journeys. Excursions with educational content in any field of knowledge, business presentations to enhance the status and prestige of an organization, booth presentations at conferences, etc. are possible. With MR Guide, the user can effortlessly create virtual content for any space by interacting with the software interface.
3) Holo Study is a series of educational applications. The main goal of the project is to create a toolkit for spatial understanding of complex phenomena. Actually, it is an engaging series of science classes where teachers and students get to interact with halograms, resulting in learning in a fun teamwork process.

Considering the offers of companies to develop products and solutions for virtual reality, scientists in psychological and pedagogical studies note that, in accordance with the Strategy for Digital Transformation of Science and Higher Education in Russia, "the transition to personalized organization of the educational process is one of the conditions for successful use of the pedagogical potential of virtual reality-based educational tools" [17]; present the results of research on the influence of virtual reality tools on the cognitive abilities of students [18, 19]; consider immersive virtual reality environments as educational tools [18, 20–22].

3 Methodology. Materials and Methods

Scientific approaches and methods were applied in solving the problem:

- discourse approach, defining the consideration of the problem in the boundaries of discourses by methods of approach: analysis, synthesis, matching, comparison, analogy
- information approach, operating the processes of perception, processing of information flows by methods of selective selection, questioning;
- systematic approach, which gives the rationale for the interconnection of theories in pedagogical research by systemic and structural-functional methods.

4 Results

Each of the considered virtual reality technologies has its purpose and possibilities, that is, it has potential in the education environment.

Purposeful implementation of the potential of virtual reality as a tool of digital transformation of education in the practice of professional training of students in Bachelor's, Specialist's and Master's degree programs of Bunin Yelets State University allowed to optimize the process of training specialists based on qualitatively new findings, transforming traditional didactic techniques under the influence of new conditions of digitalization of learning.

The potential of VR technology is widely used in solving educational problems. Thus, VR helmets and headsets, immersing the user in an unreal world simulated by the computer, allow him to operate in a virtual reality that has nothing to do with reality, which determines both the pros and cons of the analyzed technology. The advantages

are the ability to recreate the virtual world as intended by its developer in order to manipulate the user within this world. The disadvantages are the disconnection of the user from reality, which is fraught with health risks.

With the help of AR technology, the user is able to link computer-processed information with a picture of the real world, combining them layer by layer. In this way, interesting Internet projects in education can be put into practice. AR technology allows to animate drawn images, add multimedia elements to traditional book illustrations, and much more. AR technology is constantly evolving. One of the promising areas of this development is the creation of techniques for presenting information and phenomena in a user-friendly visual form, which expands the traditional search capabilities associated with the computer interface, and brings them to a new level of interaction with virtual objects.

Finally, the widely used mixed reality MR technology is fundamentally different from VR and AR, as it allows the user to interact with the real world through the virtual world, opening up new possibilities in science and education [9]. By scanning specified objects, the MR headset creates their volumetric models, superimposing the obtained information on real parameters. Scientific and production video conferences organized on the basis of MR are characterized by a high level of realism, creating the effect of personal presence.

MR technology is used in the study of humanities disciplines. Studying cultures of different countries and different nations, visiting museums anywhere in the world, memorable places and places of historical events, studying sights and archeological artifacts, rare and lost crafts, reconstruction of historical events, practicing linguistic skills of articulation and listening, reading rare books and archive documents, work in major scientific libraries, virtual travels and much more is possible with the help of this technology.

In the field of science disciplines, MR technology offers the possibility of "transferring" to science laboratories, as well as the opportunity to build up one's own experience base and interact with objects of different nature and characteristics. MR technology provides students with an immersion in a problematic environment where work in the real world involves risks to life and health, and practicing skills of interaction with this environment without fear of possible consequences in case of a mistake made. In addition, besides possible risks, the use of this technology significantly reduces training costs by providing access to high-tech work and expensive equipment in the virtual environment.

A promising application of innovative technology is the creation of virtual excursions for tourism training areas in the university, which allow, without leaving the classroom to visit any country, museums and attractions, to visit hard-to-reach places, to touch rare artifacts from any museums and storage facilities of the planet, to see the life and traditions of peoples [10], assess the environment, geological features of mineral deposits and production processes associated with them, as well as many other things. A particular direction of this technology is the development of panoramic video in three-dimensional format, which has also shown its effectiveness in the learning process.

The most widespread application of MR technology at the moment is in the courses of teaching medicine at the medical faculty of the university. Human anatomical atlas with its volumetric visualization in real dimensions, virtual journey inside human body,

pinpoint and detailed study of any biological and physiological process, journey through blood vessels, through cell at any stages of disease and remission - all this is available to today's students already.

In addition, educational complexes are actively used in the practice of physician training:

- Virtual training complex "Virtual Surgery" (see. Fig. 1);
- Virtual training complex "Dentistry";
- Virtual procedural simulator "ECG technique with a 15-channel electrocardiograph";

Simulator-expert system for training and testing knowledge of medical professionals "Virtual therapist".



Fig. 1. Working with the virtual training complex "Virtual Surgery".

Obviously, this kind of work requires a high degree of professionalism. It cannot be implemented by a single specialist, as it requires teamwork, where some of its participants are responsible for the computer software, while others are responsible for filling it with specific educational content. Therefore, it is logical that it was the university, which accumulated a huge pedagogical and theoretical potential, that focused on this work, having understood and appreciated the preferences it offered.

Innovative technologies in the field of game platforms and methods of information exchange have largely contributed to the creation of qualitatively new conditions for the implementation of games in the educational environment of higher education, significantly expanding the range of their use.

In the field of professional education, educational games, which are denoted by the English term “serious games”, are multifunctional. Their main role comes down to the fact that they not only involve the student in the learning process, but reveal to him the creative magnetism of this process, its creativity, thereby increasing the motivation to learn and master a profession.

Other advantages of educational video games include the ways of learning the material they offer, ranging from independent work or teamwork to the mentoring of a virtual tutor. It is known that for the assimilation of the topic to the learner requires different time depending on the level of intellectual development, from the existing knowledge base. In the course of educational computer game any user to study the topic can spend personally the required amount of time required to achieve the required level.

Speaking of the advantages of VR technology, it is impossible to remain silent about the disadvantages. Certainly, there are many of them, and they are very serious. Is the new technology safe for human life? It is difficult to answer this question in the affirmative. The long-term consequences of their use have not been studied, due to the historically short period of use of digital tools. However, it is already quite obvious that the use of new technologies is associated with negative effects on human health [23–26]. First of all, this and the impact on the spine, eyes and nervous system, because VR headsets, in the process of creating different kinds of images for each eye of the user, rely on the functions of the nervous system. Therefore, it is quite understandable that products of these technologies are not recommended for use by children due to the low level of research on their long-term effect on human vision and nervous system, and especially on a child [19].

Working with a computer, including VR screen, causes general fatigue, dry eyes. The feeling of being immersed in an illusory reality creates other problems. People with visual impairments cannot work with headsets at all or partially, as virtual immersion in an illusory world is associated with eye and headaches, increased blood pressure, etc. In addition, when working with VR headsets, a person experiences illusions similar to the sensations in real reality. So, simulating the sea rocking can cause a real attack of seasickness in people prone to it, which is a serious violation of the vestibular apparatus. Nor should we forget that while working with VR headsets, the human brain experiences overloads resembling hallucinations, which causes nausea as a response to the lost connection with reality. And, of course, we should not forget that working with VR headsets is addictive, that is, a desired and conscious substitution of real life with fictional life. All of this must be taken into account when organizing the work.

Medical research in recent years [14, 19, 26] has shown that VR technologies can have not only a negative impact on human mental health, but also a positive one. Thus, they can be used in the correction of many mental deviations, both acquired (mental states) and congenital (mental diseases). A wide field of activity for VR opens up in the field of psychotherapy. However, the VR content presented today in this field does not fill the field of even the routine activity of the psychotherapist, as it is developed only for individual mental disorders - mainly stress of different kinds, unwarranted anxiety and worry, phobias, etc.

It is quite understandable that medical research of VR possibilities is only at the initial stage. We believe that medical faculties of universities, which have specialists

and necessary equipment for such work, can help to carry out necessary comprehensive and systematic research, the result of which are age-specific methods and regulations for medically correct use of VR in various fields of human activity, including education, developed by scientists.

Proceeding from the aim of the research and based on the analysis of the potential of virtual reality technologies in the educational process, we carried out a survey of students with the following questions of the questionnaire:

1) What features of virtual reality technology would you consider to be its undoubted advantages?
2) Indicate the most significant disadvantages of virtual reality technology as a tool for implementing an educational course.
3) Assess the degree of your involvement in mastering the educational course with the use of virtual reality tools on a five-point scale, where "1" - very low degree of involvement, "2" - low degree of involvement, "3" - medium degree of involvement, "4" - high degree of involvement, "5" - very high degree of involvement.

Teachers of the educational organization were asked to answer the questions:

1) What is the effectiveness of applying the potential of virtual reality in the experience of its methodological implementation? Is there satisfaction with this process?
2) What qualities, in your opinion, should teachers have when applying virtual reality technologies, and what is necessary to establish or develop these qualities?
3) Has the degree of learner involvement in the learning process changed, and has there been a qualitative change in the learning outcomes? Assess the degree of students' involvement in mastering the educational course with the use of virtual reality tools on a five-grade scale, where "1" - very low degree of involvement, "2" - low degree of involvement, "3" - medium degree of involvement, "4" - high degree of involvement, "5" - very high degree of involvement.

The study involved 11 teachers and 185 students enrolled in the programs:

- Specialist's degree 31.05.01 Medicine - 22 people, 31.05.02 Pediatrics - 1 person (disciplines "Anatomy", "Medical informatics. Basics of Artificial Intelligence", "Medical Information Systems");
- Bachelor's degree 44.03.04 Professional training (by branches),profile Health care and public health - 29 people, 01.03.02 Applied mathematics and computer science (profile) Applied programming and information systems - 28 people, 09.03.01 Informatics and computer engineering (profile) Informatics and computer engineering (profile) and Information technologies and artificial intelligence - 32 people, 10.03.01 Information security (profile) Organization and technologies of information protection (by branch or in the field of professional activity) - 19 people, (disciplines "Information Technologies. Fundamentals of Artificial Intelligence", "Medical Information Systems", "Intelligent Information Systems", "Networks and Information Transmission Systems", "Administration of Information Systems", "Pattern Recognition Theory");
- Master's degree 09.04.01 Informatics and Computer Engineering, profile Management of digital transformation of medical organizations - 8 people, 35.04.04 Agronomy (profiles) Agrobusiness and Adaptive Farming Systems - 28 people (disciplines

"Medical Information Systems", "Artificial Intelligence in Healthcare", "Innovative Technologies in Agronomy");

- Secondary vocational education 33.02.01 Pharmacy - 10 people, 34.02.01 Nursing - 8 people (disciplines "Informatics", "Anatomy").

Yandex forms were used as a tool in the survey process (Figs. 2, 3).

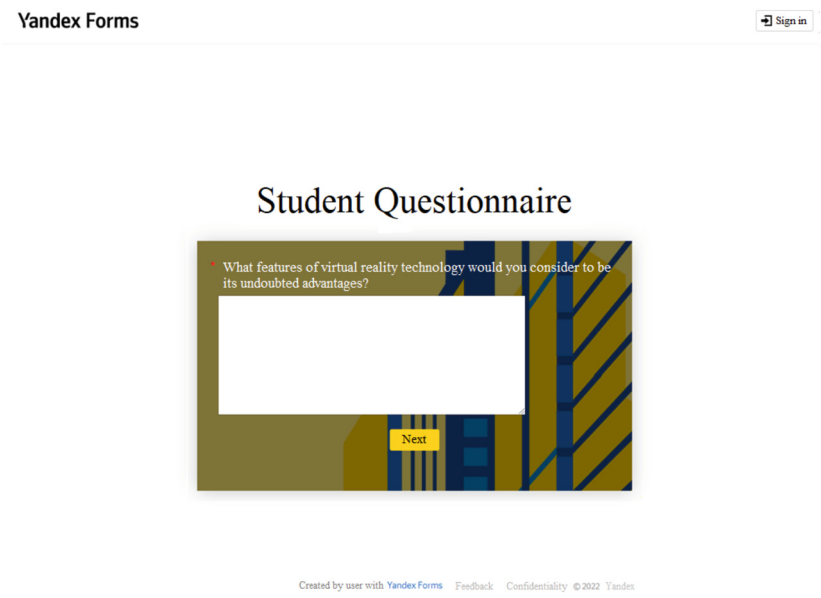


Fig. 2. Student questionnaire.

The results of the questionnaire are automatically sent to the email of the faculty member who coordinates the process.

The analysis of the data (see Fig. 4) revealed that 87% of students in assessing the degree of involvement in learning the course using the tools of virtual reality reported a high degree of involvement and a very high degree of involvement in the process, while emphasizing the undoubted advantages of virtual reality technology, including the use of a virtual machine, simulation systems in the study of IT - technology and technological processes in agriculture, 3-D atlas of human anatomy in the study of medical sciences. There are also disadvantages such as work with VR screen causes general fatigue, dry eyes in students with visual impairment.

The survey of teachers showed the following results. The high effectiveness of the virtual reality potential, based on the experience of its methodological implementation, was noted, and the satisfaction with this process was observed among the majority of teachers.

Teachers with insufficient proficiency in information technology are an exception.

When assessing the degree of involvement in mastering the training course using virtual reality tools, 92% of the teachers indicated a high degree of involvement and a very high degree of involvement in the process (Fig. 5).

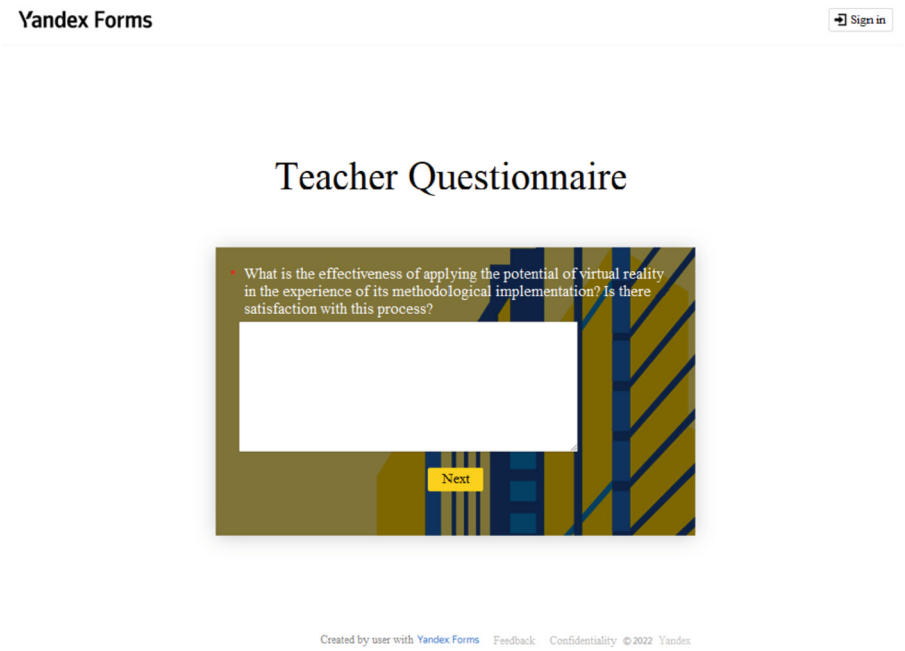


Fig. 3. Teacher questionnaire.

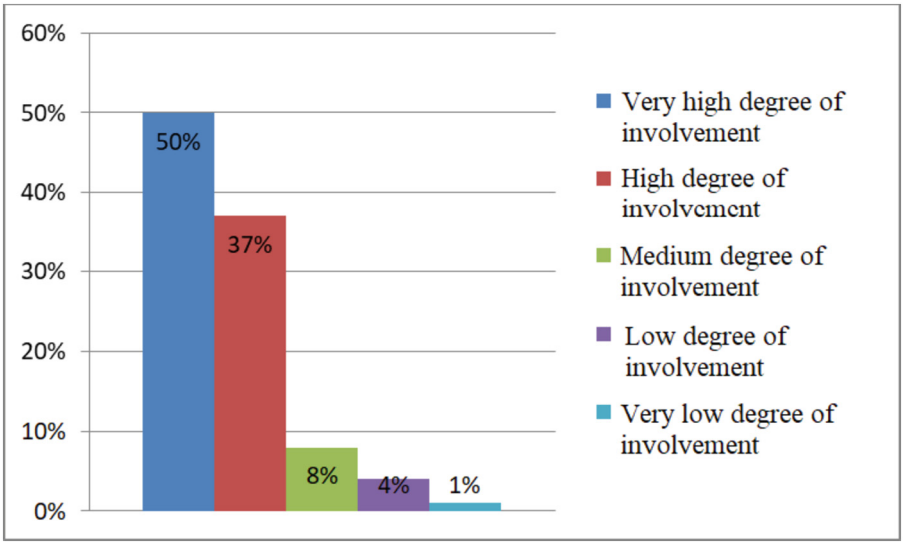


Fig. 4. Students' assessment of the degree of involvement in mastering the training course with the use of virtual reality tools.

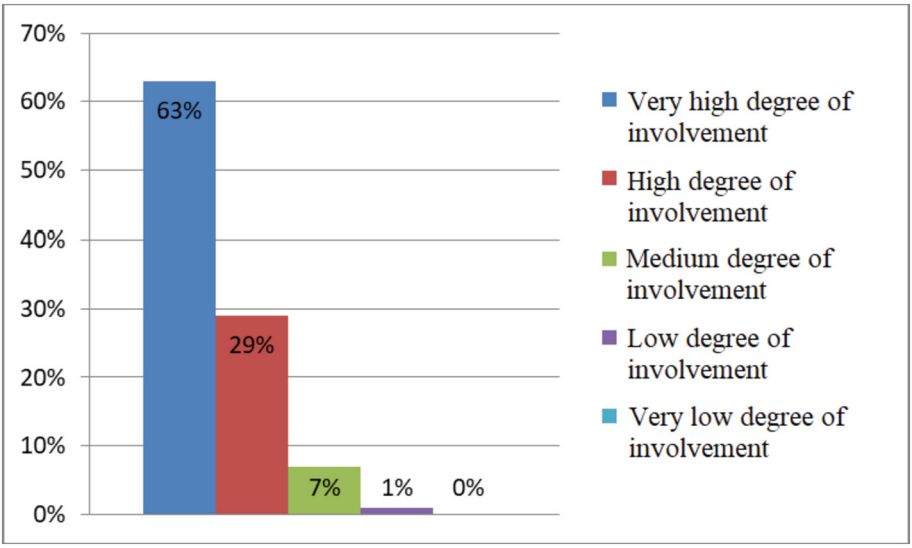


Fig. 5. Teachers' assessment of the degree of students' involvement in the educational course with the use of virtual reality tools.

5 Discussion

Based on the results of the study, it should be stated that video games, game simulators and other virtual reality programs used in the training process have proven to be a highly effective tool to provide low-cost professional training, reduce health risks in mastering dangerous operations and in general reduce the time spent on mastering training material as compared to traditional tools.

In this regard, today, in light of the new opportunities, it is time to review the methods and methodologies of the educational system that run counter to the current requirements. And here universities, as educational centers, come to the aid of educators and psychologists by providing them with the necessary methods of using innovative tools with broad educational potential.

6 Conclusion

The study revealed the potential of virtual reality as a tool for digital transformation of education and evaluated the results of its implementation in the practice of professional training by questioning teachers and students in Bachelor's, Specialist's and Master's degree programs of the Federal State Budgetary Educational Institution of Higher Education "Bunin Yelets State University". The advantages and disadvantages of this process were revealed.

According to the results of the study, we believe that the development of fundamentally new pedagogical methods, which have not yet been tested by the educational system, is to be continued. In our opinion, the prerequisite for new developments should

be a strong link to traditional education, which has deep historical and cultural roots, and at the same time, new methods should ensure the transition of traditional education to a new educational culture that meets the needs of the time. This work will be done in the foreseeable future. Its main results will be increased motivation for learning and, consequently, the level of education, increased interest of participants in the educational process in increasing the volume and quality of their own knowledge, and in general the formation of a new type of professional, possessing modern technology and distinguished by creative thinking, rapid response and the desire for self-development.

The present study, which was conducted within the framework of one university, does not claim to be universal, but reveals the trends that arouse the interest of professional educators in the development of new courses.

References

1. Velichkovsky, B.B., et al.: Cognitive control and sense of presence in virtual environments. Exp. Psychol. **9**(1), 520 (2016). https://doi.org/10.17759/exppsy.2016090102
2. Kochkin, D.A., Shevchenko, G.I.: Review of virtual reality tools used in the educational process. Student science for information society development. In: IV All-Russian Scientific and Technical Conference, vol. 1, p. 238. Publishing house of NCFU, Stavropol (2016)
3. Uvarov, A.U.: Informatization as digital transformation of education. Informatization of education and e-learning methodology. In: M.V. Noskov. II International Scientific Conference. Krasnoyarsk 25–28 September 2018, vol. 1, pp. 67–72. Siberian Federal University, Krasnoyarsk (2018)
4. Uvarov, A.Y.: Technologies of artificial intelligence in education. Inf. Educ. **4**(293), 7–15 (2018)
5. Aczél, P., Benedek, A., Veszelszki, Á.: Beyond persuasion – rhetoric in a virtual world. virtual reality - real visuality. Virtual, visual, veridical. Vis. Learn. **7**, 29–40 (2017)
6. Aczél, P.: Virtual reality and education – world of teachcraft? Perpect. Innov. Econ. Bus. **17**(1), 6–22 (2017). https://doi.org/10.15208/pieb.2017.02
7. Burdea, G., Coiffet, P.: Virtual Reality Technology, 2nd edn. Wiley, Hoboken (2003)
8. Liu, D., Bhagat, K.K., Gao, Y., Chang, TW., Huang, R.: The potentials and trends of virtual reality in education. In: Liu, D., Dede, C., Huang, R., Richards, J. (eds.) Virtual, Augmented, and Mixed Realities in Education. Smart Computing and Intelligence, pp. 105–130. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-5490-7_7
9. Barabanshchikov, V.A., Selivanov, V.V.: Interaction of the subject and virtual reality: psychological development and personal determination. Monograph. Universum, Moscow 479 (2019). http://www.psychlib.ru/inc/absid.php?absid=392124. Accessed 14 Feb 2022
10. Voyskunsky, A.E.: Psychology and the Internet. Acropolis, Moscow, p. 439 (2010)
11. Ermakov, S.S.: Modern e-learning technologies: analysis of the influence of gamification methods on students' involvement in the educational process. Mod. Foreign Psychol. **9**(3), 4758 (2020). https://doi.org/10.17759/jmfp.2020090304
12. Bell, M.: Toward a definition of 'virtual worlds'. J. Virtual Worlds Res. **1**(1), 1–5 (2008). https://doi.org/10.4101/jvwr.v1i1.283
13. Dawley, L., Dede, C.: Situated learning in virtual worlds and immersive simulations. In: Spector, J., Merrill, M., Elen, J., Bishop, M. (eds.) Handbook of Research on Educational Communications and Technology, pp. 723–734. Springer, New York (2014). https://doi.org/10.1007/978-1-4614-3185-5_58
14. Study into the use of virtual reality. KALLIDUS (2017). https://www.kallidus.com/vr-study-pr/, Accessed 14 Feb 2022

15. Swift, R., Allatt, D.: Virtual reality in education: Oculearning: Our path to reality. Amazon, Kindle Edition, Seattle p. 35 (2016)
16. The state of virtual reality for education. Unimersiv (2016). https://unimersiv.com/the-stateof-virtual-reality-for-education/. Accessed 14 Feb 2022
17. Uvarov, A.Y.: Technologies of virtual reality in education. Sci. School **4**, 108–118 (2018)
18. Kochkin, D.A., Shevchenko, G.I.: Technological levels of virtual reality representation in the context of the educational process. Student science for information society development. In: Proceedings of the III All-Russian Scientific and Technical Conference, part. 1, p. 313. Publishing house of NCFU, Stavropol (2015)
19. Selivanov, V.V., Selivanova, L.N.: Effectiveness of virtual reality application for youth and adult education. Continuous Educ. XXI century **1**(9), 120 (2015)
20. Obraztsov, I.V., Belov, V.V.: Virtual simulators in the practice of technical education. In: Proceedings of the All-Russian competition of students' and post-graduate students' research works "Innovative technologies in the educational process", vol. 2. Belgorod (2011)
21. Suvorov, K.A.: Virtual reality systems and their application. T-Comm Telecommun. Transp. **9**, 140–143 (2013)
22. Hew, K.F., Cheung, W.S.: Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: a review of the research. Br. J. Edu. Technol. **41**(1), 33–55 (2010). https://doi.org/10.1111/j.1467-8535.2008.00900.x
23. Selivanov, V.V.: Psychology of virtual reality: tutorial. Smolensk State University Press, Smolensk p. 152 (2015). http://www.psychlib.ru/inc/absid.php?absid=392127. Accessed 14 Feb 2022
24. Dede, C.J., Jacobson, J., Richards, J.: Introduction: virtual, augmented, and mixed realities in education. In: Liu, D., Dede, C., Huang, R., Richards, J. (eds.) Virtual, Augmented, and Mixed Realities in Education. Smart Computing and Intelligence, pp 1–16. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-5490-7_1
25. Kuksa, I., Childs, M.: Making sense of space. The design and experience of virtual spaces as a tool for communication. Chandos – Elsevier, Oxford p. 206 (2014)
26. Riva, G., Dakanalis, A., Mantovani, F., Sunder, S.S.: Leveraging psychology of virtual body for health and wellness. The handbook of the psychology of communication technology. Wiley-Blackwell, Chichester, pp. 528–547 (2015). https://doi.org/10.1002/9781118426456.ch24

Methods for the Formation of an Automated Distribution of Pursuers in Group Pursuit

Alexander Dubanov(✉)

Banzarov Buryat State University, Ulan-Ude, Russia
alandubanov@mail.ru

Abstract. At present, it is relevant to equip moving robotic systems with autonomous control elements. Optimization issues in pursuit problems were reduced to constructing optimal trajectories, while the aspects of automated distribution over targets in group pursuit were not raised. The purpose of this article is to build a model for the automated distribution of pursuers by targets in group pursuit. A model of group pursuit of multiple goals is proposed for consideration. The methods of pursuit by individual objects are various modifications of the methods of parallel rendezvous, chase and proportional rendezvous. First of all, a matrix is constructed, which reflects the number of pursuers and the number of targets. CoMeSySo #1120.

Keywords: Pursuer · Target · Trajectory · Sample · Matrix

1 Introduction

Differential games, pursuit problems are distinguished by the variety of goal setting for process objects. In the group pursuit model described in the article, a process is considered where the targets move along predetermined trajectories. But the predetermination of the movement of goals is not of fundamental importance. The distribution of pursuers by goals is automated based on the solution of the objective function. After finding the automated distribution, the control parameters of the pursuers' movement are modified. In the article example, this is the parameter of the minimum curvature of the trajectory, which makes it possible to ensure the simultaneous achievement of goals. The task of optimizing the group pursuit of a set of goals is a promising direction in the development of such a discipline as optimal motion control in tasks related to automated decision making and autonomous control.

The purpose of this article is to form the principles of the automated distribution of pursuers by targets based on the chosen objective function. In the model of the article, algorithms for modifying the trajectories of pursuers are proposed to achieve goals simultaneously or according to a set schedule. The issues of forming a library of pursuit methods are also considered.

In works [1–4], the issues of coordinated behavior of a group of pursuers and targets were considered. For general theoretical and practical issues in pursuit problems, the

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 451–459, 2024.
https://doi.org/10.1007/978-3-031-53549-9_43

works [5–9] were considered. The sources [10–13] were used to analyze the aiming of the pursuer at the target.

In this article, the formation of a matrix for achieving goals by pursuers is carried out. The assignment of targets to pursuers occurs according to the following principle. There is an enumeration of all possible combinations of achieving the goals by the pursuers, from which the combination of the minimum value of the criterion from the generated set with the maximum value is selected.

Consider group pursuit of a set of targets: N pursuers are chasing M targets. Let's form a matrix of distribution of pursuers by targets:

$$\Psi_{ij}, i = 1..N, j = 1..M.$$

Each cell Ψ_{ij} contains information about the phase coordinates of the i- th pursuer and the j - th target. The matrix Ψ_{ij} contains information about the method by which the i -th pursuer pursues the j - th goal.

Based on the data stored in the cells of the matrix, it accesses the library for calculating the pursuer's control vectors.

In each cell of the matrix Ψ_{ij} can be calculated, as an example, the predicted time for the i-th pursuer to reach the j-th goal: t_{ij}.

In each received sample $A_k = \left\{ \Psi_{i_1 k j_1 k} \ldots \Psi_{i_n k j_n k} \right\}$, one should find the maximum value of the arrival times $t_k = Max\left\{t_{ij}\right\}$. Suppose, in the sample indicated in Table 1, one should find the maximum value from the times: $\{t_{21}, t_{23}, t_{32}, t_{41}\}$.

The maximum value of t_k in the sample A_k is dictated by the fact that all values of t_{ij}, depending on the velocity vectors of the pursuers and targets, as well as their allowable angular velocities, can be changed upwards to the value of the parameter t_k.

In the sample A_k, it is possible to increase the remaining values of the times for the pursuers to reach the goals t_{ij} to the value t_k, due to the velocity vectors of the pursuer and the target, the values of the angular velocities.

When an array of samples $\{A_k\}$ with corresponding values of times $\left\{t_{ij}\right\}$ is obtained, then the minimum time $t_{min} = Min\{t_k\}$ should be found. Thus, the optimal time for group achievement of multiple goals at the same time will be obtained.

2 Theory

2.1 Algorithms for Calculating the Next Step of the Pursuer and Estimating the Time the Pursuer Reaches the Goal

On Fig. 1 shows a block diagram of the algorithm for calculating the next step and the speed vector of the pursuer.

On Fig. 2 shows a block diagram of the function for calculating the time and distance of reaching the target by the pursuer. The variable ε is the threshold value of the distance from the pursuer to the target, upon reaching which the target is considered to be achieved.

In the case when the target moves along a predetermined trajectory, the block diagram shown in Fig. 2 can give an estimate of the time t_{ij} for the i-th pursuer to achieve the j-th goal. As an output parameter of the function shown in Fig. 2 may be the number of iterations of the pursuit process N_{it}. The value of N_{it}, the number of iterations, is the

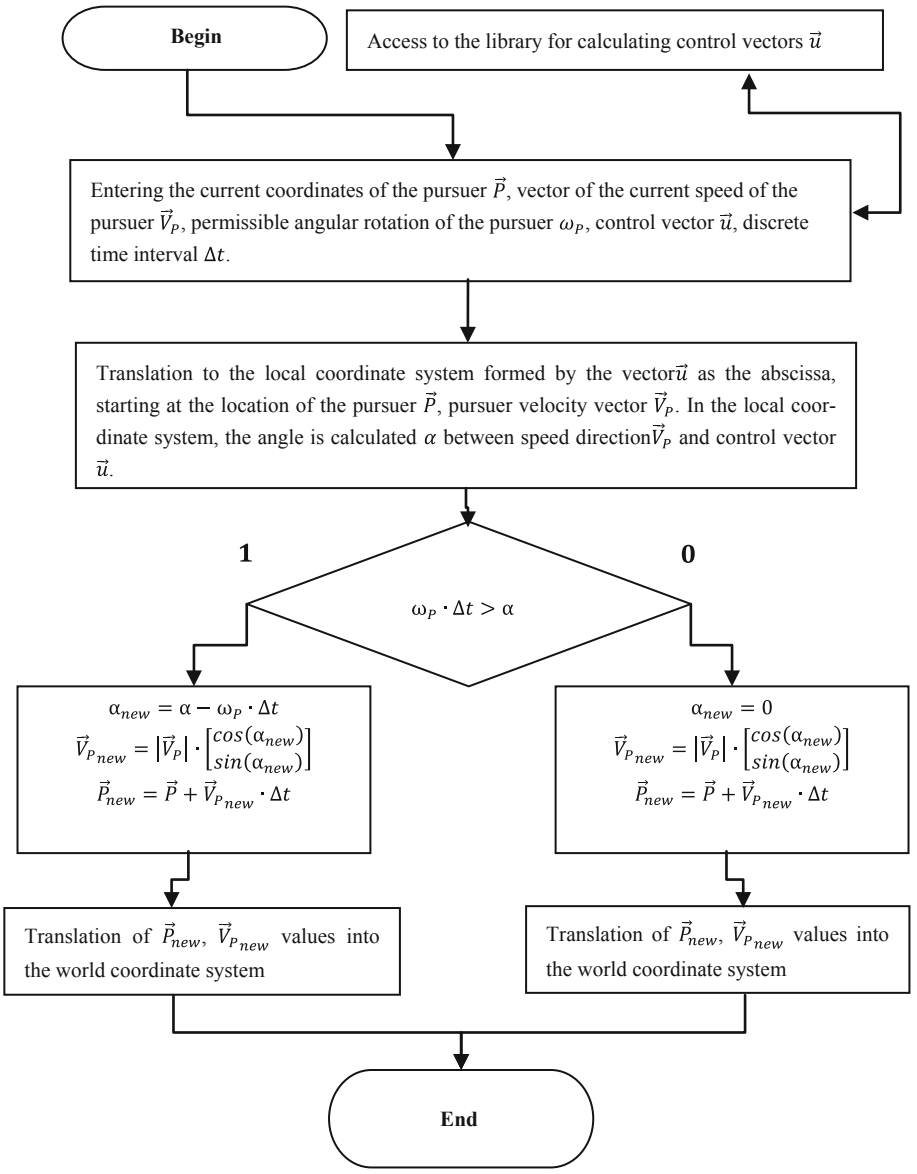


Fig. 1. Block diagram for calculating the phase coordinates of the pursuer at the next step

output parameter of the function for calculating the time and distance to reach the goal by the pursuer.

In the case when the goal takes reciprocal steps to avoid achievement, it is proposed to estimate the time differently. To do this, the predicted trajectories need to be built as composite segments of straight lines, circular arcs, square and cubic parabolas and other known lines, so as not to solve boundary value problems in the calculation cycle.

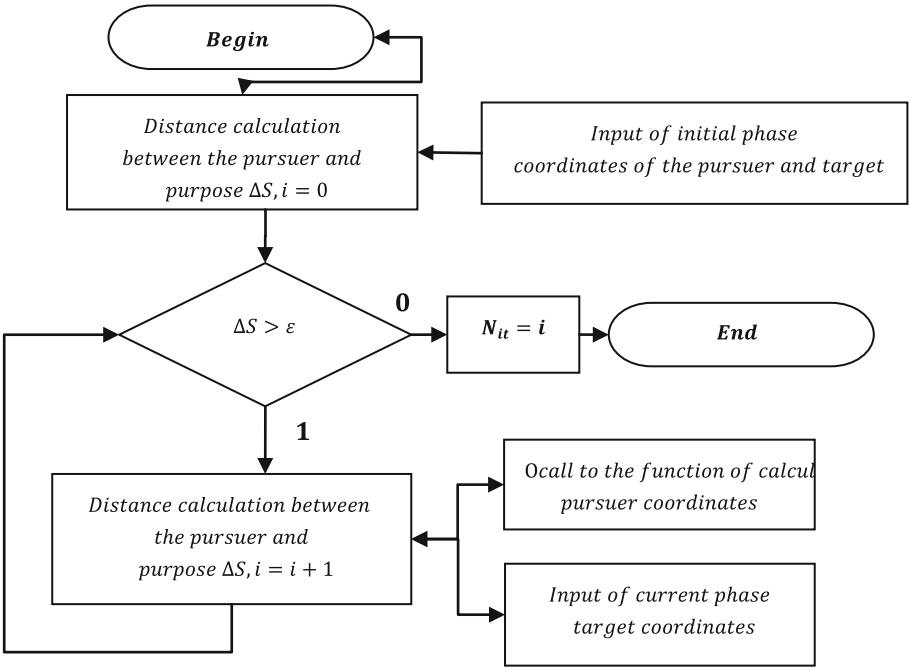


Fig. 2. Block diagram of the function for calculating the time and distance of reaching the target by the pursuer

2.2 Formation of a Library of Control Vector Calculations

It is assumed that the distribution matrix Ψ_{ij}, где $i = 1..N, j = 1..M$of target pursuers, will be built on each discrete time interval. Each cell of the matrix Ψ_{ij} will store information about the method of pursuit. Based on this information, the library of functions for calculating the control vectors $\overrightarrow{u}$ will be accessed (Table 1).

This library of control vector calculations contains the methods of the chase in the plane, in space and on the surface. Parallel approach methods are calculated on the plane, in space and on the surface. Proportional approach methods are calculated on the plane and space. Three-point methods are calculated on the plane and in space. Modified pursuit methods are calculated on the plane and in space, when the pursuer can be controlled by changing the allowable curvature of the trajectories. Modified methods of parallel approach are calculated on the plane and in space.

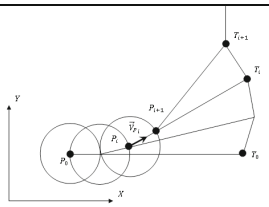
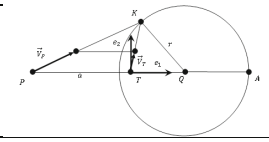
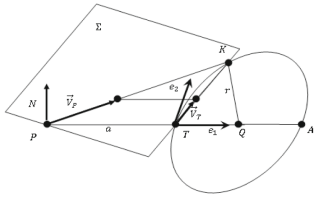
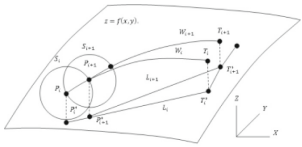
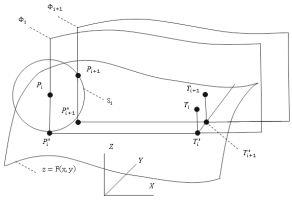
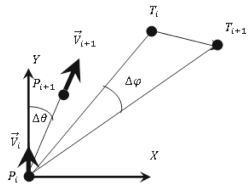
A modification of the parallel approach and chase methods allows one to build a network of predictable trajectories that allow different boundary conditions.

Table 1 does not reflect all methods for calculating control vectors. This is meant to be an open, upgradable library of functions.

2.3 An Example of Applying Matrix Modeling to Group Pursuit

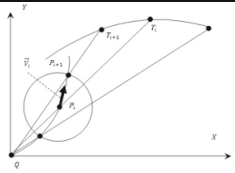
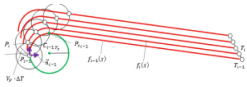
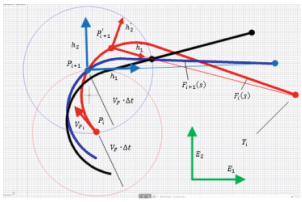
Consider an example of group pursuit (Fig. 3).

Table 1. Methods of pursuing a target moving along a certain trajectory

1	**Chase method on plane and space:** $\vec{u}_i = \frac{\vec{T}_i - \vec{P}_i}{\left\|\vec{T}_i - \vec{P}_i\right\|}$, $\vec{T}_i$ **- target position point,** $\vec{P}_i$ **- pursuer's position [14].**	
2	**Method of parallel approach on a plane:** $\vec{u} = \frac{K-P}{\|K-P\|}$, T **- target position point,** P **- pursuer's position,** K **- point on Apollonius' circle, uniquely defined by points** P **,** T **and the target's velocity tor** $\vec{V}_T$ **[15].**	
3	**Method of parallel approach in space:** $\vec{u} = \frac{K-P}{\|K-P\|}$, T **- target position point,** P **- pursuer's position,** K **- point on Apollonius' circle, the circle of Apollonius lies in the plane Σ, dotted** P**,** T **and the target's velocity vector** $\vec{V}_T$ **[16]. The above example shows the case when the pursuer's velocity vector is arbitrarily directed. After some time, the speed of the pursuer is directed to a point on the circle of Apollonius.**	
4	**Surface chase method:** $\vec{u}_i = \frac{P_{i+1} - P_i}{\|P_{i+1} - P_i\|}$, **where** P_{i+1} **- is the result of the intersection of the surface** $z = f(x, y)$**, plane** $P_i P_i^* T_i$ **and sphere** S_i **centered on a point** P_i **and radius** $\left|\vec{V}_T\right| \cdot \Delta t$**. Point** P_i^* **is point orthogonal projection** P_i **on the plane** XY**. It is necessary to express the control vector in order to make a unified call to the library [17].**	
5	**Method of parallel approach on the surface:** $\vec{u}_i = \frac{P_{i+1} - P_i}{\|P_{i+1} - P_i\|}$, **where** P_{i+1} **- is the result of the intersection of the surface** $z = f(x, y)$**, plane** $P_{i+1} P_{i+1}^* T_{i+1}$ **and sphere** S_i **centered on a point** P_i **and radius** $\left|\vec{V}_T\right| \cdot \Delta t$**. Point** P_{i+1}^* **- orthogonal projection of the point** P_{i+1} **on the plane** XY**. It is necessary to express the control vector in order to make a unified call to the library. One-parameter family of planes** Φ_i **parallel to each other [18].**	
6	**Proportional convergence method:** $\frac{d\theta}{dt} = k \cdot \frac{d\varphi}{dt}$, $\Delta\varphi = arccos\left(\frac{\|T_i\|^2 + \|T_{i+1}\|^2 - \|T_i - T_{i+1}\|^2}{2 \cdot \|T_i\| \cdot \|T_{i+1}\|}\right)$, $\Delta\theta = k \cdot arccos\left(\frac{\|T_i\|^2 + \|T_{i+1}\|^2 - \|T_i - T_{i+1}\|^2}{2 \cdot \|T_i\| \cdot \|T_{i+1}\|}\right)$, $P_{i+1} = \begin{bmatrix} V_P \cdot \Delta t \cdot cos\left(k \cdot arccos\left(\frac{\|T_i\|^2 + \|T_{i+1}\|^2 - \|T_i - T_{i+1}\|^2}{2 \cdot \|T_i\| \cdot \|T_{i+1}\|}\right)\right) \\ V_P \cdot \Delta t \cdot sin\left(k \cdot arccos\left(\frac{\|T_i\|^2 + \|T_{i+1}\|^2 - \|T_i - T_{i+1}\|^2}{2 \cdot \|T_i\| \cdot \|T_{i+1}\|}\right)\right) \end{bmatrix}$, $\vec{u}_i = \frac{P_{i+1} - P_i}{\|P_{i+1} - P_i\|}$	

(*continued*)

Table 1. (*continued*)

7	**Three-point approach method:** $(\boldsymbol{P}_{i+1}-\boldsymbol{P}_i)^2=\left(\vec{\boldsymbol{V}}_i\cdot\Delta\boldsymbol{t}\right)^2$,$\vec{\boldsymbol{u}}_i=\frac{\boldsymbol{P}_{i+1}-\boldsymbol{P}_i}{\lvert\boldsymbol{P}_{i+1}-\boldsymbol{P}_i\rvert}$. $\boldsymbol{P}_{i+1}=(\boldsymbol{1}-\boldsymbol{\tau})\cdot\boldsymbol{Q}+\boldsymbol{\tau}\cdot\boldsymbol{T}_{i+1}$ **The application of the method is convenient when the target moves along a ballistic trajectory.**	
8	**Modification of the method of parallel approach on a plane: a network of parallel lines is being built** $\boldsymbol{f}_{i+1}(\boldsymbol{s})=\boldsymbol{f}_i(\boldsymbol{s})+\boldsymbol{T}_{i+1}-\boldsymbol{T}_i$**, where** $\boldsymbol{s}$ **- line arc length,** $\boldsymbol{T}_i$ **- array of reference points of the target trajectory. Equation solution** $(\boldsymbol{f}_{i+1}(\boldsymbol{s})-\boldsymbol{P}_i)^2=(\boldsymbol{V}_P\cdot\Delta\boldsymbol{t})^2$ **relative to parameter** $\boldsymbol{s}$ **will find the value** $\boldsymbol{s}^*$**, which will match** $\boldsymbol{P}_{i+1}=\boldsymbol{f}_{i+1}(\boldsymbol{s}^*)$. $\vec{\boldsymbol{u}}_i=\frac{\boldsymbol{P}_{i+1}-\boldsymbol{P}_i}{\lvert\boldsymbol{P}_{i+1}-\boldsymbol{P}_i\rvert}$. **As a line family** $\boldsymbol{f}_i(\boldsymbol{s})$ **can serve as lines of arbitrary configuration [19].**	
9	**Modification of the method of parallel approach on a plane: network is being built** $\boldsymbol{f}_i(\boldsymbol{s})$**, where** $\boldsymbol{s}$ **- line arc length,** $\boldsymbol{T}_i$ **- array of reference points of the target trajectory. The condition is met that the end of the line** $\boldsymbol{f}_i(\boldsymbol{s})$ **passes through a point** $\boldsymbol{T}_i$**, a dot** $\boldsymbol{P}_i$ **incident line** $\boldsymbol{f}_i(\boldsymbol{s})$**. That is, in other words, it is used as a pattern. Equation solution** $(\boldsymbol{f}_{i+1}(\boldsymbol{s})-\boldsymbol{P}_i)^2=(\boldsymbol{V}_P\cdot\Delta\boldsymbol{t})^2$ **relative to the parameter** $\boldsymbol{s}$ **will allow you to find the value** $\boldsymbol{s}^*$**, which will match** $\boldsymbol{P}_{i+1}=\boldsymbol{f}_{i+1}(\boldsymbol{s}^*)$. $\vec{\boldsymbol{u}}_i=\frac{\boldsymbol{P}_{i+1}-\boldsymbol{P}_i}{\lvert\boldsymbol{P}_{i+1}-\boldsymbol{P}_i\rvert}$. **As a line family** $\boldsymbol{f}_i(\boldsymbol{s})$ **can serve as lines of arbitrary configuration [20].**	

It is necessary to form matrices Ψ_{ij}, где $i=1..3, j=1..2$, corresponding to possible samples $A_k, k=1..6$ (Table 2). This is followed by a call to find the maximum value $t_k=Max\{t_{ij}\}$. After calculating the times of reaching, it is found that the pursuer P_1 has the longest time to reach the target T_1 from the sample A_2.

In the situation shown in Fig. 3, all pursuers reach the target using the modified parallel approach method, which corresponds to line 8 of Table 1.

In the example in Fig. 3 shows the case when there was a calculation of the simultaneous achievement of goals. In the pursuit model indicated in line 8 of Table 1, the curvature of the trajectory should not exceed a certain value. Therefore, in the model, the initial radius of curvature of the trajectory increases for the pursuers P_2 and P_3, which is shown in Fig. 3.

A sample A_k is formed, in which the pursuer P_i catches up with T_j. Further, there is a primary estimate of the time to reach t_{ij}. The estimation of the time t_{ij}, in this example, occurs in the calculation of the length of the straight section to the target and in the calculation of the length of the arc of the mating circle of the allowable radius. Then, the choice of the maximum value $t_k=Max\{t_{ij}\}$ is made. The increase in time t_{ij} to the value t_k, in this model, occurs due to the increase in the pursuer P_i of the radius of the mating circle from the value r_i to the value $r_i+\delta r_i$.

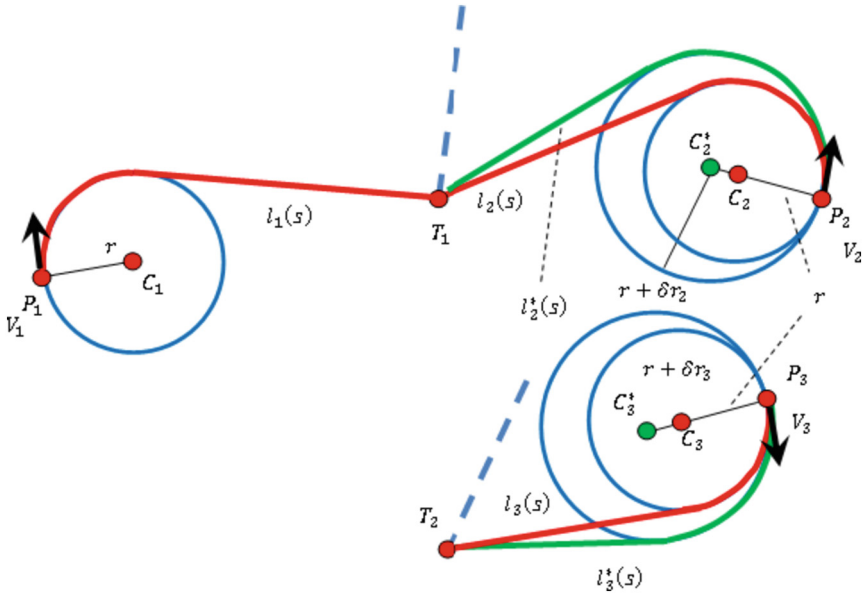


Fig. 3. Scheme of group pursuit of multiple goals

Table 2. Samples corresponding to the distribution of pursuers by goals

Persuers		Targets		Targets		Targets		Targets		Targets		Targets	
		1	**2**	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**2**
	1	x		x		x			x		x		x
	2		x	x			x	x		x			x
	3		x		x	x		x			x	x	
Selections		A_1		A_2		A_3		A_4		A_5		A_6	

Figure 4 is supplemented with an animated image showing the process of group pursuit of multiple targets [21].

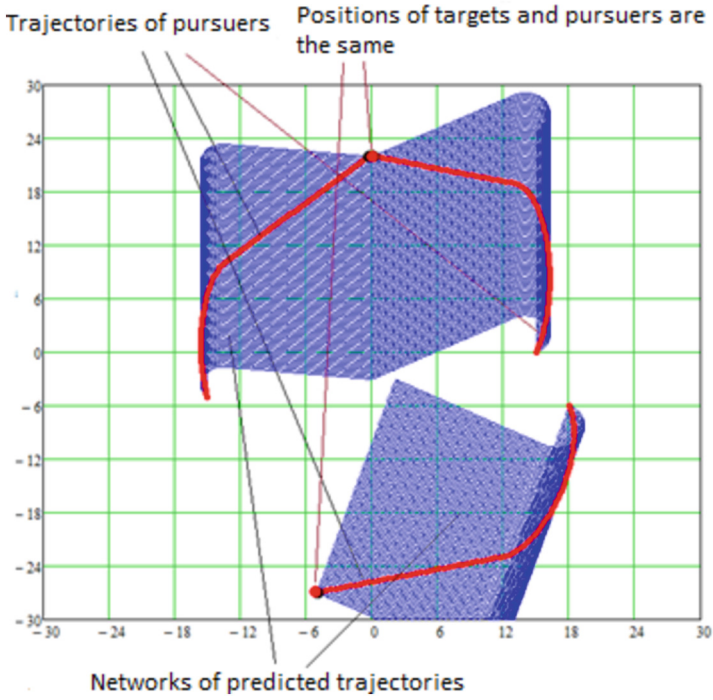


Fig. 4. The final phase of group pursuit

3 Conclusion

Modeling the matrix of the distribution of pursuers by goals can also be used by the party representing the goals, which sets itself the task of not being achieved.

It is assumed that the matrix of the distribution of pursuers by goals will be generated at each moment of time. Targets and pursuers may disappear, new ones may appear.

The method of forming a matrix for the distribution of pursuers by goals can be in demand when designing virtual reality systems for game tasks, in which the process of group pursuit, escaping, and evasion will be simulated.

References

1. Rappoport, I.S.: Group approach strategies in the method of resolving functions for quasilinear conflict-controlled processes. Cybern. Syst. Anal. **55**(1), 149–163 (2019). https://docs.yandex.ru/docs/view?url=ya-mail%3A%2F%2F1818328349551 19499%2F1.2&name=1.pdf&uid=198750017
2. Bannikov, A.C.: Somenon-stationary problems of group pursuit. Proc. Soft Inst. Math. Comput. Sci. UdSU **1**(41), 3–46 (2013)
3. Хачумов, М.В.: Решение задачи следования за целью автономным летательным аппаратом / М. В. Хачумов// Искусственный интеллект и принятие решений — 2015. — №2. —С. 45–52

4. Хачумов, М.В.: Задачи группового преследования цели в условиях возмущений. Искусственный интеллект и принятие решений — 2016 —№2.—С. 46–54
5. Abramyants, T.G., Maslov, E.P., Yakhno, V.P.: Evasion of multiple target in three-dimensional space. Autom. Remote Control **69**, 737–747 (2008). https://doi.org/10.1134/S0005117908050019
6. Samatov, B.T.: On problems of group pursuit under integral constraints on controls. Cybern. Syst. Anal. **49**(5), 132–145 (2013). https://docs.yandex.ru/docs/view?url=ya-mail%3A%2F%2F1818328349551194 99%2F1.3&name=6.pdf&uid=198750017&nosw=1
7. Chikrii, A.A.: Game dynamic problems for systems with fractional derivatives. In: Chinchuluun, A., Pardalos, P.M., Migdalas, A., Pitsoulis, L. (eds.) Pareto Optimality, Game Theory and Equilibria. Springer Optimization and Its Applications, vol. 17, pp. 349–386. Springer, New York (2008). https://doi.org/10.1007/978-0-387-77247-9_13
8. Borie, R.: Algorithms and complexity results for pursuit–evasion problems. In: Borie, R., Tovey, C., Koenig, S. Proceedings of the International Joint Conference on Artificial Intelligence (2009). Accessed 11 Mar 2010
9. Sozinov, P.A.: Kinematic analysis of proportional navigation methods as applied to guidance of an anti-aircraft guided missile at a ballistic target. In: Sozinov, P.A., Gorevich, B.N. (eds.) Bulletin of the Almaz-Antey Concern (2), 74–92 (2022). https://doi.org/10.38013/2542-0542-2022-2-74-92
10. Zarchan, P.: Tactical and strategic missile guidance. Am. Inst. Aeronaut. Astronaut. **1095** (2012). https://doi.org/10.1017/S000192400008698X
11. Chikrii A.A.: Conflict controlled processes. Springer, London, 424 p. (2013). https://doi.org/10.1016/S0898-1221(97)90249-2
12. Chikrii A.A., Chikrii G.T. Matrix resolving functions in game problems of dynamics. Proc. Steklov Inst. Math. **291**, 56–65 (2015). https://doi.org/10.1134/S0081543815090047
13. Chung and Furukawa.: A reachability-based strategy for the time-optimal control of autonomous pursuers. Eng. Optim. **40**(1), 67–93 (2008). doi:https://doi.org/10.1080/03052150701593133
14. Video, a model of the chase method on a plane and in space. https://youtu.be/PAu9Qg1dySM (дата обращения: 16.01.2023), Author Alexander Dubanov, @dubanov5999
15. Video, parallel rendezvous model on a plane. https://youtu.be/hGieKXNiuz8 (дата обращения: 16.01.2023) , Author Alexander Dubanov, @dubanov5999
16. Video, model of parallel approach in space. URL:https://youtu.be/8nDUSi3ENB4 (дата обращения: 16.01.2023) , Author Alexander Dubanov, @dubanov5999
17. Video, surface chase method model. https://youtu.be/sU724Db_VMk(дата обращения: 16.01.2023) , Author Alexander Dubanov, @dubanov5999
18. Video, method of parallel approach on the surface. https://youtu.be/06qgINE4j8U(дата обращения: 16.01.2023) , Author Alexander Dubanov, @dubanov5999
19. Video, modification of the parallel approach method. https://www.youtube.com/watch?v=qNXdykK21Z8(дата обращения: 16.01.2023) , Author Alexander Dubanov, @dubanov5999
20. Video, modification of the chase method. https://www.youtube.com/watch?v=UQ5bVKjVqZ4(дата обращения: 16.01.2023) , Author Alexander Dubanov, @dubanov5999
21. Video, simulation results of the pursuit problem. https://www.youtube.com/watch?v=NNJDJOJT34I (дата обращения: 9.07.2022) , Author Alexander Dubanov, @dubanov5999

Author Index

A
Abal, Rodrigo Loayza 377
Abd Rauf, Muhammad Fairuz 334, 343
Adnan, Zuraidy 343
Ahmadova, Sevil 82
Ali, Fazilatulaili 334
Aliyev, Veyis 75
Aly, Shady 51
Asenov, Oleg 130
Azman, Mohd 303, 316

B
Bagirzade, Kerim 82
Bakar, Abu 303, 316
Benda, Tomáš 51
Berdnova, E. V. 245
Bhavani, A. Durga 172
Bobrovs, Vjaceslavs 271, 278, 283, 288
Bobrovs, Vjaceslavs 263
Bogomolov, A. S. 245
Böhm, Josef 35
Bolshakov, Alexey 271
Borjkhani, Mehdi 288
Burylin, Dmitrii 433

C
Chen, Taotao 35
Chernousova, Nataliya V. 439
Chernova, Nadezhda 433

D
Dankov, Yavor 113, 121
Dankova, Andjela 113
Degtyareva, K. V. 14
Dnekeshev, A. A. 245
Dranko, O. I. 245
Dubanov, Alexander 451

E
Emebu, Samuel 58

F
Firdose, Salma 183
Firdus, Elnare 75

G
Galin, Rinat 424
Galina, Saniya 424
Gantimurov, A. P. 14
Gnezdilova, Natalia A. 439
Gospodinov, Nikolay 93
Gurbanov, Abdulaga 82

H
Hamplová, Adéla 51
Hasanah, Uswatun 303, 316
Hitayezu, Antoine 410
Huang, Jiawei 221
Huda, Miftachul 303, 316, 357
Hwang, Hyoseong 208, 221

I
Ibraimov, Damir 433
Islam Khan, Burhan Ul 183
Ivanova, Boyana 103
Ivanova, Ekaterina 1
Ivashchenko, V. A. 245

J
Jaafar, Jazurainifariza 357
Jayavel, Kayalvizhi 410
Joe, Inwhee 208, 221

K
Kahar, Saliyah 343
Katahova, Svetlana 433
Katakhova, Natalia 433
Kislov, Denis 263
Kohout, Jan 35
Kositzyn, A. A. 245
Koval, Olga 271

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 909, pp. 461–463, 2024.
https://doi.org/10.1007/978-3-031-53549-9

Krastev, Georgi 93
Kudinova, Tatiana 433
Kukartseva, O. I. 14
Kulmagambetova, Zhumazhan 254
Kulzhagarova, Bazargul 254
Kurinnaya, Alina 271
Kushnikov, V. A. 245
Kuznetsov, Alexey V. 278, 283

L
Lim, Dongkyun 233
Lim, Ducsun 233

M
Mahmudova, Zarifa 82
Mamchenko, Mark 424
Mammadova, Mehriban 82
Mangla, Neha 172
Manzi, Fabrice 410
Mareš, Jan 35
Marjudi, Suziyanti 343
Mary Pushpa, T. Anita Jones 160
Mat Razali, Noor Afiza 334
Mathew, K. Deepa 160
Matušů, Radek 58
Mayilov, Rauf 75, 82
Medina, Sebastian 26
Mehdiyeva, Almaz 75
Mohamad Amran, Mohd Fahmi 334, 343
Mohan, Venothanee Sundra 334
Mohd Yusop, Nurhafizah Moziyana 334
Muhawenayo, Gedeon 410
Mujahidin 303, 316
Murzagulov, Damir 254
Musolin, Mukhamad Hadi 303, 316

N
Natividad, Reyna 292
Nelyub, V. A. 14
Nsengiyumva, Philibert 410

O
Olivares, Rodrigo 26
Osliakova, Irina 433

P
Pakhomova, T. V. 245
Pankin, Pavel V. 439
Peña, Lidia Segura 377
Peralta-Reyes, Ever 292
Pérez-Pastenes, Hugo 292
Puttaswamy, Chaya 196

R
Rahman, Muhammad Hasbi Abd 357
Ravelo, Camilo 26
Regalado-Méndez, Alejandro 292
Rezchikov, A. F. 245
Robles-Gómez, Edson E. 292
Romero, Rubi 292
Rusev, Rumen 103
Rustamova, Durdana 75

S
Sardarova, Ijabika 82
Seiner, Jakub 67
Selyutin, A. D. 245
Serour, Reda Owis Hassan 303, 316
Serov, Alex 271
Shaikh, Ibrahim 58
Shangytbayeva, Gulmira 254
Sharif, Morteza A. 288
Sharifli, Sabina 75
Shchuchka, Tatiana A. 439
Shipkovenski, Georgi 130
Shivaprasad, Nandini Prasad Kanakapura 196
Shoilekova, Kamelia 103
Shtrom, Igor 271
Smailova, Ulmeken 254
Soria, Juan J. 377
Sroczyński, Zdzisław 398
Steinbach, Jakub 67, 141
Stepanovskaya, I. A. 245
Štícha, Karel 35
Syed Ali, Sharifah Aishah 334

T
Triantafyllou, Serafeim A. 150
Tuyishime, Emmanuel 410
Tynchenko, V. S. 14
Tyrychtr, Jan 51

U
Urbániová, Zuzana 141

V
Valdez-Dominguez, Iris C. 292
Volf, Daniiar 424
Vrba, Jan 67, 141

W
Wong, Rita 343

Y
Yahaya, Yuhanim Hani 334
Yauri, Andi Muhammad 303, 316
Yusifov, Salahaddin 75, 82

Z
Zlatarov, Pavel 1
Zuhri, Muhammad 303, 316