

СВЁРТОЧНАЯ НЕЙРОСЕТЕВАЯ МОДЕЛЬ ПРОСТРАНСТВЕННОГО ВОЗНИКНОВЕНИЯ ИЛЛЮЗИИ МЮЛЛЕРА-ЛАЙЕРА

А.Н. Мамаев, И.А. Горбунов

Иллюзия Мюллера-Лайера — классическая оптическая иллюзия, исследуемая в психологии ещё с 1889 года [Muller-Lyer, 1889]. Её суть заключается в том, что длина двух одинаковых по размеру линии с разным оперением на концах — стрелками, направленными либо наружу, либо внутрь — воспринимается людьми отличающейся: линия с оперением внутрь всегда выглядит длиннее, чем линия с оперением наружу. С момента описания иллюзии было предложено несколько теорий, объясняющих её существование [Bermond, Van Heerden, 1996]. Разные теории предоставляют разные возможности для того, чтобы воссоздать феномен в рамках вычислительной модели. В случае с объяснениями иллюзии через связь с пространственным восприятием, начатыми Грегори [Gregory, 1963], возможно моделировать не только процессы восприятия у субъекта, но и воспринимаемое пространство и наблюдаемые в нём объекты. Для реализации этой идеи мы прибегли к использованию свёрточных алгоритмов машинного обучения и формированию набора данных для обучения при помощи создания виртуального трёхмерного пространства.

Архитектура свёрточных нейронных сетей вдохновлена структурой зрительного анализатора, благодаря этому они предоставляют возможность для моделирования психофизиологических процессов восприятия. Свёрточные нейронные сети — категория нейронных сетей с прямой связью, в которых применяется математическая операция свёртки. Она происходит при помощи фильтров — двумерных числовых массивов, которые «накладываются» на входящие данные и, в зависимости от своей конфигурации, выражаемой в коэффициентах, применяемых к значениям матрицы данных, выделяют те или иные паттерны, как рецептивные поля зрительного анализатора определяют границы контраста, различать прямые линии и углы, исходя из уровня активации

рецепторов сетчатки. Каждый фильтр проходит по массиву данных, и его суммарная «активация» для каждого фрагмента переходит на следующий слой в виде карты признаков — матрицы, в которой закодирована локализация выделенных паттернов. К картам признаков тоже применяется свёртка, до тех пор, пока модель не выделит основные характеристики данных, необходимые для решения задачи. Свёрточные слои позволяют использовать для обучения структурированные данные, представленные в виде матриц, такие как изображения, с наибольшей эффективностью. Из-за этого свёрточные нейронные сети широко используются для решения задач компьютерного зрения.

До сих пор исследования действия оптических иллюзий на функционирование нейронных моделей были немногочисленными. Предыдущие попытки исследовать иллюзию Мюллера-Лайера с использованием свёрточных нейронных сетей были ограничены бинарной классификацией и двухмерными изображениями [Zeman и др., 2013]. Мы предлагаем подход, приближающий условия исследования к реальным и позволяющим получить количественные результаты для более точного сравнения с результатами у людей. Мы предполагаем, что свёрточная нейронная сеть, обученная для измерения размера объекта в пространстве, будет давать искажённые оценки, если ей будет предъявлена иллюзия Мюллера Лайера, как это происходит у людей.

Чтобы воссоздать процесс измерения объектов в трёхмерном пространстве, мы обучили свёрточную регрессионную модель для определения высоты изображённого на картинке куба. Модель получала на входной слой изображения в оттенках серого с разрешением 200×200 пикселей, сигнал проходил через 8 внутренних слоёв (свёрточные, субдискретизации, потом полносвязные) и итоговый результат выводился на один нейрон выходного слоя, активируемый линейно. Предсказанное значение сравнивалось с целевым по формуле среднеквадратичной ошибки, обучение происходило при помощи оптимизатора градиентного спуска ADAM.

Для того, чтобы разнообразить стимульный материал и приблизить условия эксперимента к реальным, мы создали виртуальное пространство, в котором возможно сгенерировать набор данных для обучения. Трёхмерная сцена состояла из объекта и камеры, которая сохраняла изображения из определённой точки в пространстве. Высота и поворот объекта, а также углы наклона камеры и её отдаление от объекта рандомизировались для каждого нового изображения. Мы обучили модель давать предсказания о том, какой размер, выраженный в единицах длины виртуальной среды, был задан для объекта на каждом из снимков. Когда модель достигла высокой точности, мы предъявили рисунки, вызывающие иллюзию Мюллера-Лайера и выяснили, что предсказания модели искажены таким же образом, каким искажаются оценки людей.

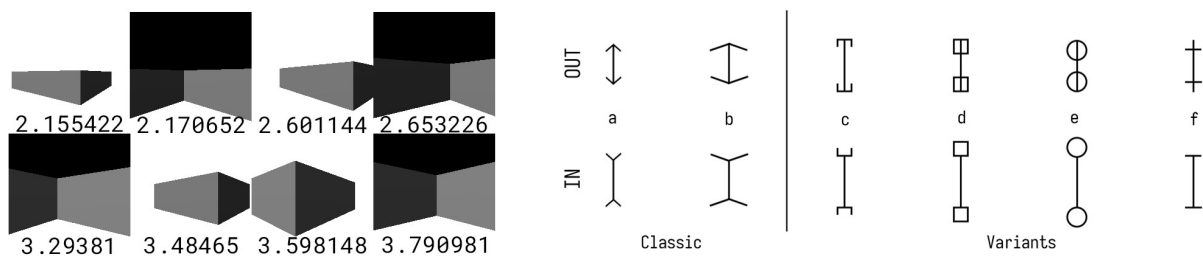


Рисунок 1. Примеры стимульного материала: изображения для обучения и валидации модели (слева), тестовые изображения, вариации стимулов иллюзии Мюллера-Лайера (справа). А, В: классические пары стимулов, С–Е: иные варианты из литературы [Hatwell, 1960; Parker, Newbigging, 1963], F: новая пара стимулов.

Для того, чтобы подтвердить результаты, мы повторили процедуру с 30 независимыми моделями и проанализировали результаты как с помощью фреквентистских (дисперсионный анализ с повторными измерениями), так и байесовских методов: Bayes Factor, Leave One Out cross-validation (LOO), Watanabe-Akaike Information Criterion (WAIC). Приоритетными были выбраны байесовские методы, поскольку они лучше подходят для оценки нелинейных моделей. Исследовалось влияние факторов направленности оперения стрелки,

его ширины и угла наклона. Мы установили влияние как отдельных факторов, так и их взаимодействия на предсказания модели.

С помощью алгоритма Markov Chain Monte Carlo (MCMC) были созданы четыре байесовских модели: нулевая (оперение не влияет на предсказания нейронной сети), in-out (фактор направления), narrow-wide (фактор ширины), полная модель (два фактора).

При сравнении моделей при помощи Bayes Factor мы определили, что полная двухфакторная наиболее вероятна (10^{37}), следом за ней модели, учитывающие только фактор ширины и угла оперения (10^{33}) и фактор направления стрелок (10^{11}). «Нулевая» модель, предполагающая отсутствие влияние исследуемых факторов, была определена как наименее вероятная.

Сравнение моделей с применением LOO и WAIC подтвердило вероятность полной модели, однако изменило положение однофакторных моделей в рейтинге. Ниже приведены показатели критериев и стандартные ошибки для каждой из моделей:

1. Полная модель: $LOO=270,7\pm 10,5$; $WAIC=270,73\pm 10,48$.
2. Модель фактора направления: $LOO=245,21\pm 9,33$; $WAIC=245,22\pm 9,32$.
3. Модель ширины/угла наклона оперения: $LOO=194,34\pm 6,86$; $WAIC=194,35\pm 6,86$.
4. Модель нулевой гипотезы: $LOO=186,66\pm 7,69$; $WAIC=186,66\pm 7,69$.

В работе мы также учитываем критику обоснования иллюзии через связь с пространственным восприятием, заключающуюся в проявлении феномена не только при использовании «стрелок», но и нестандартных версий стимульного материала, не имеющих очевидного сходства с элементами объёмных предметов. Мы выяснили, что модели свойственно стойкое проявление иллюзии Мюллера-Лайера как в случае со стандартными, так и нестандартными стимульными парами. Дополнительно мы провели количественное сравнение эффекта различий между вариациями стимулов, и выяснили, что различие между

стимулами в парах всегда положительное, то есть одна из стрелок (с оперением, направленным наружу) постоянно оценивается как бóльшая.

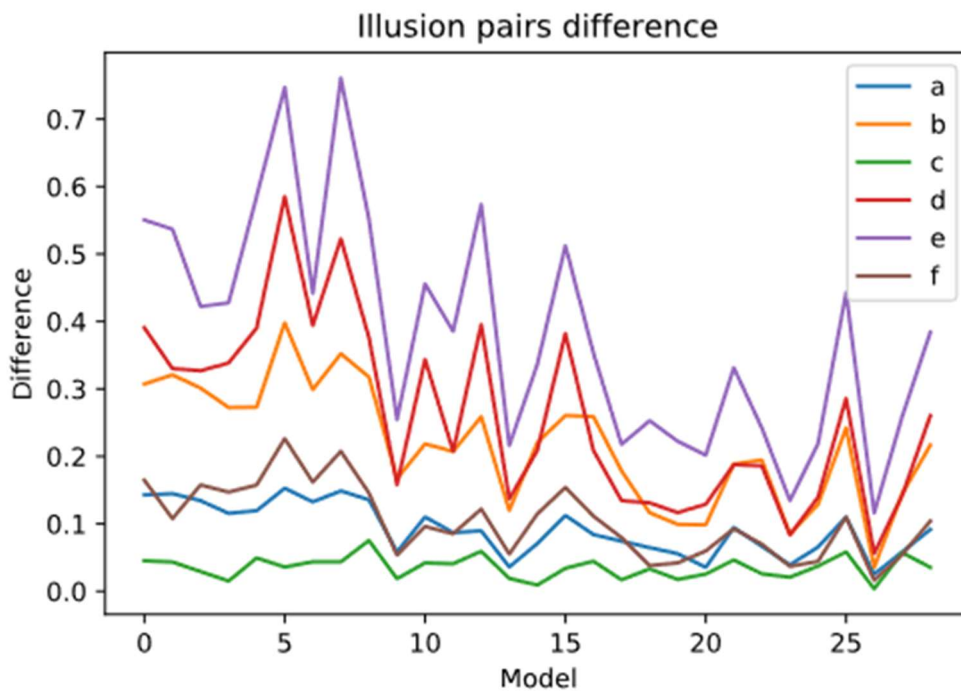


Рисунок 2. Разница между предсказанным размером линий в парах, данным 30 независимыми моделями, сравнение между вариантами стимулов.

Результаты работы поддерживают теоретическую связь иллюзии Мюллера-Лайера с развитием пространственного восприятия человека и открывают перспективы для изучения процессов зрительного восприятия с использованием моделей машинного обучения и пространственной симуляции. Обучение каждой новой модели происходит «с чистого листа», поэтому свойства, которые она приобретает, неминуемо зависят от используемых данных. Обучая модель решать пространственные задачи в изолированных условиях мы вынуждаем её выделять признаки, релевантные выполняемой функции. Совершая перенос этих признаков на уже нерелевантную задачу (измерение линии на плоскости), мы наблюдаем эффект иллюзии, и предполагаем, что эти же признаки вызывают иллюзии в зрительном восприятии человека. Однако, использование нестандартных стимулов приводит к тому же эффекту, хотя они и лишены признаков, которые могли бы напоминать о перспективе. Ранее этот факт рассматривался как опровержение идеи о роли пространственного восприятия в

появлении иллюзии, однако, поскольку даже непосредственное обучение решению исключительно пространственных задач вызывает такой же эффект, он должен быть пересмотрен. Мы предполагаем, что существование иллюзии может обуславливаться двумя факторами. С одной стороны, пространственная суммация на нейронах рецептивных полей и активация свёрточных фильтров может объяснять, почему иллюзию вызывают как стимулы, напоминающие углы объёмных фигур, так и совершенно не относящиеся к пространственному восприятию. С другой стороны, исследования, ограничивающиеся обучением модели только классификации плоских линий, всё же не смогли в точности повторить результаты, полученные на выборке людей [Zeman и др., 2013], что говорит о необходимости рассматривать и другие возможные варианты объяснения эффекта. В этом смысле, текущее исследование предлагает один из таких вариантов, не только основанный на классических теориях, но и подкреплённый симуляцией с использованием машинного обучения. Для того, чтобы глубже изучить восприятие иллюзии обученной моделью, в дальнейшем будет возможно проследить изменение её внутренних параметров: конфигурацию фильтров, карты признаков, сформированные для отдельных стимулов, активацию нейронов на конечных слоях сети, отвечающих за наиболее абстрактное представление значимых признаков изображения.

1. Bermond B., Van Heerden J. The Muller-Lyer illusion explained and its theoretical importance reconsidered // *Biol Philos.* 1996. Т. 11. № 3. С. 321–338.
2. Gregory R. L. Distortion of Visual Space as Inappropriate Constancy Scaling // *Nature.* 1963. Т. 199. № 4894. С. 678–680.
3. Hatwell Y. Etude de quelques illusions géométriques tactiles chez les aveugles // *L'Année psychologique.* 1960. Т. 60. № 1. С. 11–27.
4. Muller-Lyer F. Optische urteilstauschungen // *Archiv fur Anatomie und Physiologie, Physiologische Abteilung.* 1889. Т. 2. С. 263–270.
5. Parker N. I., Newbigging P. L. Magnitude and decrement of the Müller-Lyer illusion as a function of pre-training. // *Canadian Journal of Psychology/Revue canadienne de psychologie.* 1963. Т. 17. № 1. С. 134–140.
6. Zeman A. и др. The Müller-Lyer Illusion in a Computational Model of Biological Object Recognition // *PLoS ONE.* 2013. Т. 8. № 2. С. e56126.