

ОТЗЫВ

на диссертацию Малаховой Екатерины Юрьевны на тему: «Исследование и интерпретация при помощи искусственных нейронных сетей механизмов описания зрительных образов в нижневисочной коре головного мозга приматов» на соискание ученой степени кандидата биологических наук по специальности 1.5.5 – физиология человека и животных

Представленная на соискание диссертация Малаховой Е.Ю. посвящена изучению механизмов переработки зрительной информации в высших отделах коры головного мозга. Работа выполнена в области нейрофизиологии, но автор широко использовала вычислительные методы. Следует отметить в области теории репрезентации формируется новая парадигма интерпретации данных, и работа Е.Ю. Малаховой логично вписывается в этот контекст. В отличие от множества работ, в которых искусственные нейросети применяются постфактум как вспомогательный инструмент анализа, здесь предпринята попытка использовать их и как модель, и как методологическое средство для построения активного эксперимента. Такой подход выходит за рамки традиционной нейрофизиологической парадигмы. Это требует пересмотра самого способа взаимодействия с экспериментальными данными.

Диссертация Малаховой основана на репрезентативных экспериментальных данных, полученные при ее непосредственном участии, при строгом соблюдении этических и методологических требований. Ею показано, что регистрируемая активность нейронов нижневисочной коры макака сопоставляется с архитектурами глубоких сверточных и генеративных нейросетей, причем верификация результатов проводится через несколько взаимодополняющих аналитических подходов. Малахова получила важные результаты из них несколько следует отметить: 1) разработку программного комплекса для целенаправленной генерации стимулов на основе модели зрительной коры, позволяющего варьировать стимульный материал в процессе эксперимента; 2) демонстрацию возможности сопоставления пространств признаков нейросетей и нейронных популяций, выявление фазовых переходов в представлениях и анализ их динамики на основе сопоставления матриц сходства; 3) применение фрагментно-ориентированной интерпретации избирательности нейронов, аппроксимированных через ориентационные и цветовые каналы.

Методология работы построена на сходстве между функционированием сверточных нейронных сетей, оптимизированных для задач распознавания, и ответами нейронов зрительной коры. При этом одним из практически важных результатов является подтверждение возможности адаптивной генерации искусственных стимулов, вызывающие целенаправленную активацию нейронных ансамблей.

Отметим, что полученное доказательство оппонентной системы «записи-обработки-распознавания» зрительной информации весьма близко к принципам голографической модели Карла Прибрама: на первичных низкоуровневых этапах (начальные сверточные слои) доминирует статистика входного сигнала, затем она «рассыпается» на ориентационные и хроматические составляющие, а уже на высокоуровневых сверточных слоях проявляются признаки решения задачи распознавания. В эту же «голографическую» логику вписывается резкая перестройка внутреннего описания зрительного образа, переходящая, как пишет автор, в «продолжительную стадию, сопровождающуюся трансформацией представлений, максимальной сложностью описания и распределенным кодированием». Заключительный процесс интенсивной консолидации признаков и

кардинального переструктурирования описания весьма схож со сверточным механизмом восстановления оптического волнового фронта. На это же указывает обнаруженный автором «фазовый характер обработки зрительной информации»: сначала происходит резкий разрыв связей с входными данными, затем следует этап трансформации представлений и распределенного кодирования, а на заключительном этапе наблюдается восстановление признаков под задачи распознавания.

Еще один существенный результат работы — доказательство того, что комбинирование визуальных признаков, полученных при помощи фрагментов, позволяет объяснить инвариантное к ракурсу распознавание лиц как в нижневисочной коре, так и в искусственных сверточных нейронных сетях, построенных, как показано во многих работах, на функциях взаимной корреляции, аналогичных математическому описанию фурье-голограмм.

Замечания.

1. В работе заявлена аналогия между слоями сверточной сети и уровнями зрительной обработки, но не ясно, какова значимость таких сопоставлений. Используемая архитектура (VGG16) достаточно стабильна, но сегодня имеются более гибкие и физиологически мотивированные модели (например, архитектуры с итеративной переработкой, attention-механизмами). Желательно было бы указать, насколько устойчивы выводы при смене модели.
2. В работе указано, что создание оптимального стимула для кортикальной колонки начинается с генерации случайного вектора в пространстве признаков. Интересно было бы выяснить, насколько сильно на результат распознавания влияет выбор начальной точки в многомерном пространстве параметров. В таких условиях случайный старт вполне может закончиться попаданием в локальный, а не в глобальный оптимум.
3. Хотелось бы понять каким образом влияют на конечный результат аппроксимации функции нейронов (через фрагменты натуральных изображений) по отдельности те самые четыре ориентационных и три хроматических канала, указанных в работе. В автореферате об этом информации нет.

Заключение. Диссертация Малаховой Екатерины Юрьевны представляет собой завершенное научное исследование высокого уровня. Представленная работа содержит оригинальные результаты, обладает значительной теоретической и практической значимостью и соответствует критериям, предъявляемым к кандидатским диссертациям. Считаю, что автор заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.5 – физиология человека и животных.

Стафеев Сергей Константинович,
д.т.н., профессор Физико-технического факультета НИУ ИТМО,
профессор Института когнитивных исследований СПбГУ

02 июня 2025 года

Подпись Стафеева
удостоверяю
Менеджер ОПС
Пономарева О.В.



Унив. ИТМО
Кронверкский, 49 А
СПб, РФ, 197101

7 931 207 48 81

stafeevs@yahoo.com