

Моделирование показаний фМРТ по видео, показанному человеку.

Дорин Д. Д.^{1,2}, Киселев Н. С.^{1,2}, Грабовой А. В.²

dorin.dd@phystech.edu

¹Организация; ²Организация

В данной работе исследуется задача прогнозирования показаний датчиков фМРТ по видеоряду, показанному человеку. Предложен метод аппроксимации показаний фМРТ по видеоряду на основе трансформер моделей. Подтверждена зависимость между показаниями датчиков и восприятием внешнего мира человеком. Эффективность нашего подхода демонстрируется на наборе данных, собранных у большой группы людей, когда они смотрели короткий аудиовизуальный фильм.

Ключевые слова: *фМРТ, видеоряд, трансформер модель.*

1 Введение

Несмотря на множество достижений современной науки, человеческий мозг остается одним из самых загадочных объектов. Внутречерепные записи человека являются редким и ценным источником информации о мозге. Поэтому исследование методов получения данных о функциональной активности коры головного мозга и зависимостей между результатами методов и восприятием окружающего мира человеком актуально в наши дни.

Одним из методов исследования активности головного мозга является функциональная магнитно-резонансная томография, далее – фМРТ. В работе исследуется зависимость между показанием фМРТ и видеорядом, просмотренным человеком. фМРТ – разновидность магнитно-резонансной томографии, которая проводится с целью измерения гемодинамических реакций – изменений в потоке крови, вызванных нейронной активностью головного или спинного мозга. Этот метод основывается на том, что мозговой кровоток и активность нейронов связаны между собой. Когда область мозга активна, приток крови к этой области также увеличивается. фМРТ позволяет определить активацию определенной области головного мозга во время нормального функционирования под влиянием различных заданий, например, зрительных, когнитивных, моторных, речевых. В работе Алены Беляевской [1] собраны современные возможности фМРТ в нейровизуализации. Рассмотрим статьи, посвященные работе с данными фМРТ. В работе Юлии Березуцкой [2] собран обширный набор данных, состоящий из видеорядов, просмотренных человеком, и соответствующих снимков фМРТ. Одна из проблем при работе с данными нейровизуализации – шум, вызванный движением головы, биением сердца, тепловыми эффектами и др. В работе Максима Шараева [3] рассмотрены подходы к подготовке, предварительной обработке, шумоподавлению, направленные на устранение артефактов, вредных для распознавания образов, а также методы классификации данных нейровизуализации.

Наиболее известные методы обработки видео основаны на 3D свертках. 3D в отличие от 2D свертки одновременно работают с пространственной и временной частью информации. Существенный недостаток данных методов – сильное увеличение числа параметров модели и большие вычислительные затраты. В работе используется более современная архитектура – Трансформер модель. Впервые модель Трансформер была предложена в статье «Attention Is All You Need» Ashish Vaswani [4]. Архитектура активно применяется в области машинного перевода. А в 2022 году появилась работа Shen Yan [5] на тему адаптации архитектуры Трансформер для работы с видеорядами. Данная архитектура

учитывает пространственно-временные зависимости и повышает скорость обучения за счет attention слоев. Сама модель состоит из кодирующего компонента, декодирующего компонента и связи между ними. Каждый компонент состоит из стека энкодеров и декодеров соответственно. Входящая последовательность, поступающая в энкодер, сначала проходит через attention слой, помогающий энкодеру посмотреть на другие слова во входящем объеме во время кодирования конкретного элемента. Выход attention слоя отправляется в нейронную сеть прямого распространения. Аналогично устроен декодер, за исключением наличия еще одного слоя внимания, помогающего фокусироваться на релевантных элементах.

В данной работе предлагается метод аппроксимации показаний датчиков фМРТ по видеоряду. Полученная в ходе экспериментов корреляционная картина между фМРТ снимками и видеорядами подтверждает зависимость между показаниями фМРТ и восприятием внешнего мира человеком. По нашим сведениям, на данный момент нет статей, которые бы исследовали возможность предсказания снимков фМРТ по видео или подтверждали зависимость между данными.

Проверка метода проводится на выборке, представленной в работе [2]. Набор данных включает в себя в себя записи фМРТ 30 участников в возрасте от 7 до 47 лет во время выполнения одинаковой задачи и записи внутричерепной электроэнцефалографии 51 участника в возрасте от 5 до 55 лет.

2 Постановка задачи

3 Вычислительный эксперимент

4 Анализ ошибки

5 Заключение

end

6 *

Список литературы

- [1] A. A. Belyaevskaya, N. V. Meladze, M. A. Sharia, D. V. Ustyuzhanin, and M. H. Zashezova. MODERN POSSIBILITIES OF FUNCTIONAL MAGNETIC RESONANCE IMAGING IN NEUROIMAGING. *Medical Visualization*, (1):7–16, February 2018.
- [2] Julia Berezutskaya, Mariska J. Vansteensel, Erik J. Aarnoutse, Zachary V. Freudenburg, Giovanni Piantoni, Mariana P. Branco, and Nick F. Ramsey. Open multimodal iEEG-fMRI dataset from naturalistic stimulation with a short audiovisual film. *Scientific Data*, 9(1), March 2022.
- [3] Maxim Sharaev, Alexander Andreev, Alexey Artemov, Alexander Bernstein, Evgeny Burnaev, Ekaterina Kondratyeva, Svetlana Sushchinskaya, and Renat Akzhigitov. fmri: preprocessing, classification and pattern recognition, 2018.
- [4] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017.
- [5] Shen Yan, Xuehan Xiong, Anurag Arnab, Zhichao Lu, Mi Zhang, Chen Sun, and Cordelia Schmid. Multiview transformers for video recognition, 2022.