

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. Проанализирована зависимость между показаниями датчиков и восприятием внешнего мира человеком. Эффективность предложенного подхода демонстрируется на наборе данных, собранных у большой группы людей, когда они смотрели короткий аудиовизуальный фильм.

Ключевые слова: фМРТ, видеоряд, Трансформер модель.

1 Введение

Человеческий мозг один из самых интересных объектов исследования [7]. Внутречерепные записи человека являются редким и ценным источником информации о мозге. Поэтому исследование методов получения данных о функциональной активности коры головного мозга актуально в наши дни.

Одним из методов исследования активности головного мозга является функциональная магнитно-резонансная томография [8]. 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 связи мозгового кровотока и активности нейронов. 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 Алёны Белинской [1] собраны современные возможности фМРТ в нейровизуализации — общее название нескольких методов, позволяющих визуализировать структуру, функции и биохимические характеристики мозга.

В работе Юлии Березункиной [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 Шарасва [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

0. Я думаю, до объединения работ, лучше Киселева лучше не указывать. Это все-таки Ваша работа

1. Моделей типа Трансформер

2. Кривовато. Может "собранных в процессе просмотра фильма"?

3. Здесь нужно длинное тире

4. Грамматически предложение получилось не очень корректным. Давайте разобьем. И напомним "Под нейровизуализацией понимается ..."

5. Лучше с цифры не начинать предложение. Перефразируйте. И пере "в отличие" нужна запятая

6. см. пункт 1

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. При получении метода использовались два основных предположения. Первая гипотеза о зависимости между результатами фМРТ и просматриваемым фильмом. Второе предположение заключается в том, что реакция мозга, фиксируемая фМРТ, на информацию, поступающую от органов зрения, происходит не мгновенно, а с некоторой задержкой [6]. Полученная в ходе экспериментов корреляционная картина между данными в выборке подтверждает зависимость между показаниями фМРТ и восприятием внешнего мира человеком.

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

Пусть Ω — видеоряд, ν — частота кадров, t — продолжительность видеоряда:

$$\Omega = (\omega_1, \dots, \omega_{\nu t}), \quad (1)$$

где $\omega \in \mathbb{R}^{W_\omega \times H_\omega \times C_\omega}$ — изображение, W_ω — ширина изображения, H_ω — высота изображения и C_ω — число каналов.

Введем также \mathcal{S} — последовательность фМРТ снимков, μ — частота снимков:

$$\mathcal{S} = (s_1, \dots, s_{\mu t}), \quad (2)$$

где $s \in \mathbb{R}^{X_s \times Y_s \times Z_s}$ — фМРТ снимок, X_s , Y_s , Z_s — размерность одного измерения.

Также считаем, что известно несколько дополнительных измерений фМРТ \mathcal{S}_0 того же испытуемого. Необходимо построить отображение f :

$$f(\omega_1, \dots, \omega_{i-\nu\Delta t}, \mathcal{S}_0) = s_i, \quad (3)$$

которое учитывает задержку Δt , между фМРТ картиной и моментом получения информации зрительными органами.

3 Вычислительный эксперимент

4 Анализ ошибки

5 Заключение

6 *

Список литературы

7. Ссылка или пояснение понятия энкодера и декодера
8. Я бы написал "механизм внимания" (термин существует)
9. Не хватает сказуемого
10. Последнее предложение хорошо бы подкрепить формулой, которая иллюстрирует Ваше требование

-
- [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.
 - [6] Вячеслав Евгеньевич Демидов. *Как мы видим то, что видим*. Знание, 1979.
 - [7] ТА Жумакова, ШО Рыспекова, ДД Жунистаев, НМ Чурукова, АМ Исаева, and ИО Алимкул. Тайны человеческого мозга. *Международный журнал прикладных и фундаментальных исследований*, (6-2):230–232, 2017.
 - [8] ВЛ Ушаков, ВМ Верхлютов, ПА Соколов, МВ Ублинский, ВБ Стрелец, АЮ Аграфонов, АВ Петрайкин, and ТА Ахадов. Активация структур мозга по данным фМРТ при просмотре видеосюжетов и припоминании показанных действий. *Журнал высшей нервной деятельности им. ИП Павлова*, 61(5):553–564, 2011.