
ВОССТАНОВЛЕНИЕ СНИМКОВ фМРТ ПО ПРОСМАТРИВАЕМОМУ ВИДЕОРЯДУ

Киселев Никита
kiselev.ns@phystech.edu

Грабовой Андрей
grabovoy.av@phystech.edu

13 апреля 2023 г.

АННОТАЦИЯ

Исследуется проблема восстановления зависимости между показаниями датчиков фМРТ и восприятием внешнего мира человеком. Проводится анализ зависимости между последовательностью снимков фМРТ и видеорядом, просматриваемым человеком. На основе исследования зависимости предлагается метод аппроксимации показаний фМРТ по просматриваемому видеоряду. Для анализа предложенного метода проводится вычислительный эксперимент на выборке, полученной при томографическом обследовании большого числа испытуемых. (1)

Ключевые слова: нейровизуализация · фМРТ · видеоряд · зависимость между данными

1 Введение

Совокупность методов, визуализирующих структуру и функции человеческого мозга, называется *нейровизуализацией*. Методы нейровизуализации, такие как ЭКГ, КТ, МРТ и фМРТ, используются для изучения мозга, а также для обнаружения заболеваний и психических расстройств.

Функциональная магнитно-резонансная томография или *фМРТ* (англ. *fMRI*) является разновидностью магнитно-резонансной томографии и основана на изменениях в токе крови, вызванных нейронной активностью мозга [4]. Эти изменения происходят не моментально, а с некоторой задержкой. Она возникает из-за того, что сосудистая система достаточно долго реагирует на потребность мозга в глюкозе [5]. Изображения, получаемые с помощью фМРТ, показывают, какие участки мозга активированы при выполнении испытуемым определенных заданий.

Настоящая работа посвящена восстановлению зависимости между снимками фМРТ и видеорядом. Используется предположение, что такая зависимость существует. Кроме

того, предполагается, что между снимком и видеорядом есть постоянная задержка во времени [5]. Проверяется зависимость снимка фМРТ от одного изображения и предыдущего снимка. Время задержки выступает в качестве гиперпараметра модели. На основе анализа зависимости предлагается метод аппроксимации показаний фМРТ по просматриваемому видеоряду.

Метод фМРТ играет большую роль в нейровизуализации, однако имеет ряд важных ограничений. В работах [6, 7] рассматриваются временное и пространственное разрешения фМРТ. Временное разрешение является существенным недостатком данного метода. Другой недостаток фМРТ — неизбежно возникающие шумы, связанные с движением объекта в сканере, сердцебиением и дыханием человека, тепловыми флуктуациями самого прибора и т. д. В работе [8] предлагаются методы подавления вышеперечисленных шумов на основе графов и демонстрируется их эффективность в задаче выявления эпилепсии и депрессии. (2)

Обобщением уже естественных для обработки изображений 2D сверток в CNN являются 3D свертки [9]. Они агрегируют информацию как по времени, так и по пространству. Однако это приводит к сильному увеличению количества используемых параметров. В настоящей работе используется наиболее современная архитектура — Transformer. Впервые она была предложена в статье [10]. Не так давно появилась адаптация архитектуры Transformer для работы с видео [11]. Данная архитектура состоит из кодировщика и декодировщика, (3) каждый из которых в свою очередь состоит из отдельных слоев. Использование механизма Attention [10] позволяет значительно повысить качество работы модели.

Данные, на которых проводятся проверка гипотезы зависимости и демонстрация работы построенного метода, представлены в работе [2]. Этот набор данных был получен при обследовании группы из 63 испытуемых. Тридцать из них проходили обследование фМРТ. Им предлагалось выполнить одно и то же задание — просмотреть короткий аудиовизуальный фильм. Для него в рассматриваемой работе были сгенерированы аннотации, содержащие в том числе информацию о времени появления и исчезновения отдельных слов, объектов и персонажей. Методы аудио- и видеоаннотирования подробно излагаются в [3] и [1].

2 Постановка задачи

Задана частота кадров $\nu \in \mathbb{R}$ и продолжительность $t \in \mathbb{R}$ видеоряда. Задан видеоряд

$$\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_{\nu t}], \quad \mathbf{p}_\ell \in \mathbb{R}^{W \times H \times C}, \quad (4)$$

с шириной, высотой и числом каналов изображения W, H и C соответственно.

Обозначим частоту снимков фМРТ $\mu \in \mathbb{R}$. Задана последовательность снимков

$$\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_{\mu t}], \quad \mathbf{s}_\ell \in \mathbb{R}^{X \times Y \times Z}, \quad (2)$$

где X, Y и Z — размерности воксельного изображения.

1. В аннотации также пишут несколько слов об итоге работы. Этой части не было, когда мы только начинали статью, но можно добавить сейчас. Также в аннотации или во введении можно отметить практическое применение вашей работы.
2. Данные методы используются в Вашей работе?
3. Определения
4. Размерность p_i есть, но нет информации о том, чем являются значения. Для s_i расшифровка значений есть в предполагаемом методе. Возможно стоит перенести сюда, так как здесь впервые встречается данное обозначение.

Задача состоит в построении отображения, которое бы учитывало задержку Δt между снимком фМРТ и видеорядом, а также предыдущие томографические показания. Формально, необходимо найти такое отображение \mathbf{g} , что

$$\mathbf{g}(\mathbf{p}_1, \dots, \mathbf{p}_{k_\ell - \nu \Delta t}; \mathbf{s}_1, \dots, \mathbf{s}_{\ell-1}) = \mathbf{s}_\ell, \quad \ell = 1, \dots, \mu t, \quad (3)$$

где для ℓ -го снимка фМРТ номер соответствующего изображения k_ℓ определяется по формуле

$$k_\ell = \frac{\ell \cdot \nu}{\mu}. \quad (4)$$

3 Предлагаемый метод восстановления снимков фМРТ

Обозначим снимок фМРТ как $\mathbf{s}_\ell = [v_{ijk}^\ell] \in \mathbb{R}^{X \times Y \times Z}$, где $v_{ijk}^\ell \in \mathbb{R}_+$ — значение соответствующего вокселя. Предположим, что каждый снимок зависит только от одного изображения и предыдущего снимка. Тогда соответствующее отображение можно записать в виде

$$\mathbf{g}(\mathbf{p}_{k_\ell - \nu \Delta t}; \mathbf{s}_{\ell-1}) = \mathbf{s}_\ell - \mathbf{s}_{\ell-1} = \boldsymbol{\delta}_\ell, \quad \ell = 2, \dots, \mu t. \quad (5)$$

где $\boldsymbol{\delta}_\ell = [v_{ijk}^\ell - v_{ijk}^{\ell-1}] = [\delta_{ijk}^\ell] \in \mathbb{R}^{X \times Y \times Z}$ — разность между двумя последовательными снимками.

Для каждого изображения из видеоряда имеем вектор признаков описания размерности d :

$$\mathbf{x}_\ell = [x_1^\ell, \dots, x_d^\ell]^\top \in \mathbb{R}^d, \quad \ell = 1, \dots, \nu t.$$

Используется архитектура нейронной сети ResNet152 *ссылка или описание* без последнего линейного слоя.

Учитывая (4), суммарное число пар (изображение, снимок) равно $N = \mu(t - \Delta t)$. Таким образом, для каждого вокселя задана выборка

$$\mathfrak{D}_{ijk} = \{(\mathbf{x}_\ell, \delta_{ijk}^\ell) \mid \ell = 2, \dots, N\}.$$

Поставлена задача восстановления регрессии

$$y_{ijk} : \mathbb{R}^d \rightarrow \mathbb{R}. \quad (6)$$

Используется линейная модель с вектором параметров

$$\mathbf{w}_{ijk} = [w_1^{ijk}, \dots, w_d^{ijk}]^\top \in \mathbb{R}^d : \quad (7)$$

$$f_{ijk}(\mathbf{x}, \mathbf{w}_{ijk}) = \langle \mathbf{x}, \mathbf{w}_{ijk} \rangle.$$

Для модели f_{ijk} с соответствующим ей вектором параметров $\mathbf{w}_{ijk} \in \mathbb{R}^d$ определим квадратичную функцию потерь с L_2 регуляризацией:

$$\mathcal{L}_{ijk}(\mathbf{w}_{ijk}, \Delta t) = \sum_{\ell=2}^N (f_{ijk}(\mathbf{x}_\ell, \mathbf{w}_{ijk}) - \delta_{ijk}^\ell)^2 + \alpha \|\mathbf{w}_{ijk}\|_2^2, \quad (8)$$

где $\alpha \in \mathbb{R}$ — коэффициент регуляризации.

Требуется найти параметры, доставляющие минимум функционалу потерь $\mathcal{L}_{ijk}(\mathbf{w}_{ijk}, \Delta t)$ при заданном гиперпараметре Δt :

$$\hat{\mathbf{w}}_{ijk} = \arg \min_{\mathbf{w}_{ijk}} \mathcal{L}_{ijk}(\mathbf{w}_{ijk}, \Delta t). \quad (9)$$

Минимум функции потерь находится методом наименьших квадратов. Определим матрицу объектов-признаков

$$\mathbf{X} = [\mathbf{x}_2^T, \dots, \mathbf{x}_N^T]^T = [x_j^i] \in \mathbb{R}^{(N-1) \times d} \quad (10)$$

и вектор, компонентами которого являются разности значений одного и того же вокселя в разных снимках,

$$\Delta \mathbf{v}_{ijk} = [\delta_{ijk}^2, \dots, \delta_{ijk}^N]^T \in \mathbb{R}^{N-1}. \quad (11)$$

Решение можно записать в виде

$$\hat{\mathbf{w}}_{ijk} = (\mathbf{X}^T \mathbf{X} + \alpha \mathbf{I})^{-1} \mathbf{X}^T \Delta \mathbf{v}_{ijk}. \quad (12)$$

Получим формулу для восстановленных снимков фМРТ. Введем матрицу весов

$$\hat{\mathbf{W}} = [\hat{\mathbf{w}}_1^T, \dots, \hat{\mathbf{w}}_{XYZ}^T]^T = [\hat{w}_j^i] \in \mathbb{R}^{XYZ \times d}. \quad (13)$$

Введем для тензоров $\mathbf{s}_\ell, \boldsymbol{\delta}_\ell \in \mathbb{R}^{X \times Y \times Z}$ векторы

$$\mathbf{s}_\ell^R = [v_1^\ell, \dots, v_{XYZ}^\ell]^T, \boldsymbol{\delta}_\ell^R = [\delta_1^\ell, \dots, \delta_{XYZ}^\ell]^T \in \mathbb{R}^{XYZ}.$$

Тогда вектор восстановленного снимка находится по формуле

$$\hat{\mathbf{s}}_\ell^R = \mathbf{s}_{\ell-1}^R + \hat{\boldsymbol{\delta}}_\ell^R = \mathbf{s}_{\ell-1}^R + \hat{\mathbf{W}} \mathbf{x}_\ell. \quad (14)$$

4 Вычислительный эксперимент

Для анализа работоспособности предложенного метода, а также проверки гипотез проведен вычислительный эксперимент.

В качестве данных использовалась выборка, представленная в работе [2]. Набор данных содержит результаты обследования 63 испытуемых. Для тридцати из них известны показания фМРТ. Среди них 16 мужчин и 14 женщин в возрасте от 7 до 47 лет. Средний возраст испытуемого — 22 года.

Размерности изображений и снимков, а также другие характеристики выборки представлены в Таблице 1.

Произведено разделение выборки на тренировочную и тестовую в соотношении 70% и 30% соответственно. Критерием качества восстановления снимка фМРТ служит

Таблица 1: Описание выборки

Название	Обозначение	Значение
Продолжительность обследования	t	390 с
Частота кадров видео	ν	25 с ⁻¹
Частота снимков фМРТ	μ	1,64 с ⁻¹
Размерности изображения	W, H, C	640, 480, 3
Размерности снимка	X, Y, Z	40, 64, 64

пикселей?

MSE — сумма квадратов отклонений между истинным и восстановленным снимками, усредненная по всем вокселям каждого снимка из тестовой выборки.

Для сокращения времени работы алгоритма производится предварительное сжатие снимка фМРТ с помощью сверточного слоя MaxPool3D. Рассматриваются коэффициенты сжатия 1, 2, 4 и 8. Значения вокселей нормализуются на $[0; 1]$ процедурой MinMaxScale.

возможно описание / ссылка?

Проанализирована зависимость качества восстановления снимков фМРТ от гиперпараметра Δt . Использовалось предварительное 8-кратное сжатие снимка. Зависимость метрики MSE от гиперпараметра Δt представлена на Рис. 1. Для построения графика производилось усреднение по испытуемым. Обозначены границы среднеквадратичного отклонения. Наблюдается характерный минимум при $\Delta t \approx 10$ с.

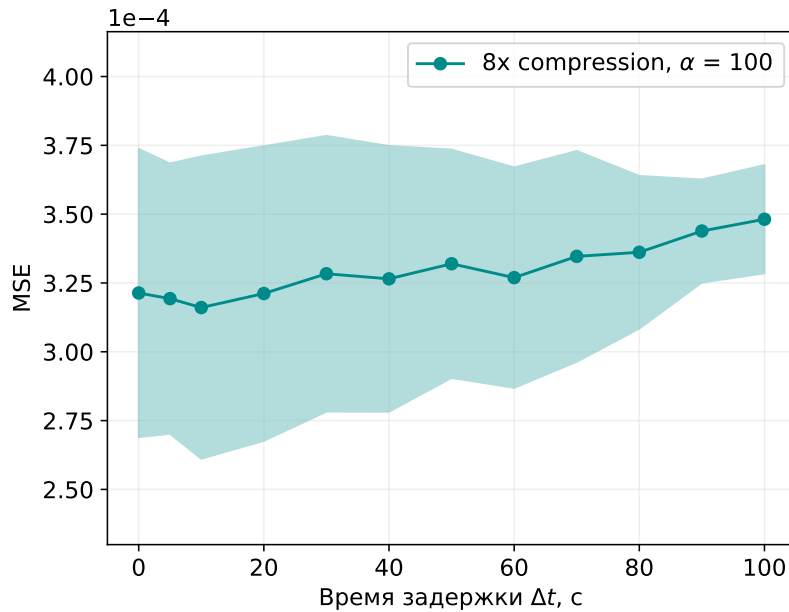


Рис. 1: Зависимость метрики MSE от гиперпараметра Δt на снимках из тестовой выборки

На Рис. 2 представлены срезы истинного и восстановленного снимков из тестовой выборки. На Рис. 2.(с) можно наблюдать разность между ними. Был выбран 4-ый испытуемый, $\Delta t = 5$ с, коэффициент сжатия 1, коэффициент регуляризации $\alpha = 1000$. Рассмотрен 20-ый срез по первому измерению 100-го снимка в последовательности. Значе-

ния вокселей лежат в отрезке $[0; 1]$, поэтому ошибка порядка $10^{-3} \div 10^{-2}$ свидетельствует о достаточно точном предсказании.

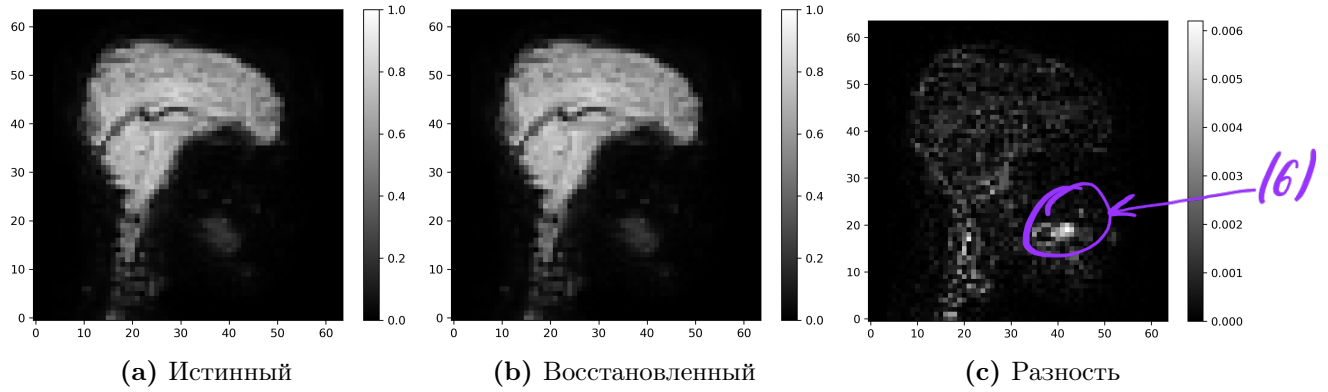


Рис. 2: Срез снимка фМРТ из тестовой выборки

Проведен анализ зависимости MSE от коэффициента регуляризации α . Рассматривались коэффициенты сжатия 1, 2, 4 и 8. Соответствующие графики приведены на Рис. 3. Для построения графика производилось усреднение по испытуемым. Обозначены границы среднеквадратичного отклонения. Из графиков видно, что оптимальное значение коэффициента $\alpha \approx 100$.

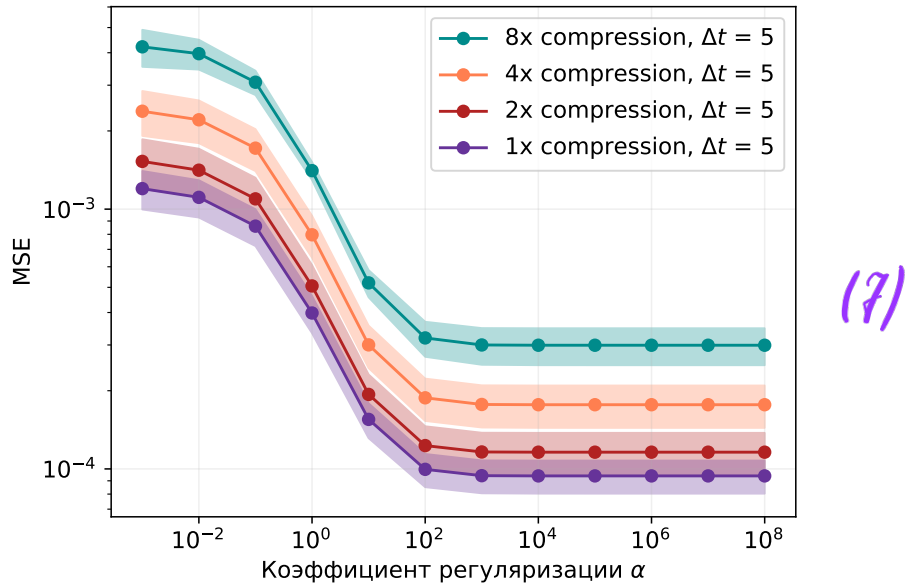


Рис. 3: Зависимость метрики MSE от коэффициента регуляризации α на снимках из тестовой выборки

Построен график распределения значений компонент вектора весов модели. Для построения производилось усреднение по всем вокселям для 4-го испытуемого. Результат представлен на Рис. 4. Веса модели не лежат в окрестности какого-то конкретного

значения, а распределены нормально. Это может говорить о том, что построенная модель, вероятно, имеет определенную статистическую значимость.

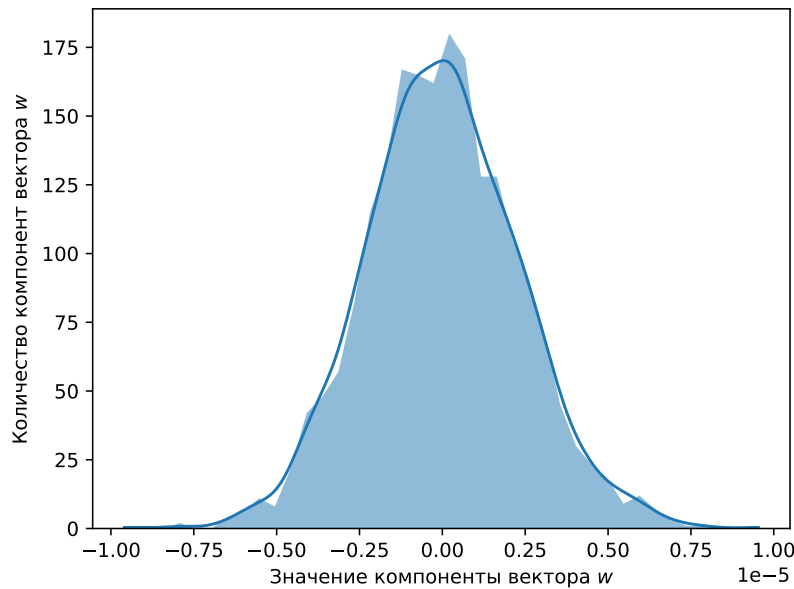


Рис. 4: Распределение значений компонент вектора весов

Проведена проверка гипотезы инвариантности весов модели относительно человека: можно ли восстановление снимка фМРТ одного испытуемого, используя матрицу весов другого. Использовалась метрика MSE на тестовой выборке. Результаты представлены в Таблице 2. Рассмотрены 4-ый и 7-ый испытуемые. Матрица весов 4-го использовалась для восстановления снимков 7-го. Значения MSE практически совпадают. Это свидетельствует о справедливости рассматриваемой гипотезы. (8)

Таблица 2: Проверка гипотезы инвариантности весов модели относительно человека

Матрица весов	Истинная	Подмешанная
MSE	$1.02 \cdot 10^{-4}$	$1.05 \cdot 10^{-4}$

Рассмотрено качество работы метода на случайном шуме. В качестве матрицы \mathbf{X} из (10) взята матрица случайных чисел из равномерного распределения на $[0; 1)$. Произведено сравнение с результатами на настоящей матрице объектов-признаков. К первому снимку 35-го испытуемого последовательно прибавляются все восстановленные изменения значений вокселей. В результате имеем последний снимок последовательности. На Рис. 5 представлены срезы последнего истинного и восстановленного снимков из тестовой выборки. На Рис. 5.(с) можно наблюдать разность между ними. Результаты на случайном шуме продемонстрированы на Рис. 6. Можно видеть, что разность между истинным и восстановленным снимками при работе со случайным шумом значительно выше. Численные результаты приведены в Таблице 3.

5. Нет описания основной модели в разделе вычислительный эксперимент. Выше написано про использование архитектуры нейронной сети ResNet152. Для читателей, которых интересует именно вычислительный эксперимент, стоит добавить описание части от предобработки до представления результатов в вычислительном эксперименте.
6. Описание ситуаций на снимках, для которых модель работает намного хуже, чем в других частях.
7. Чем плохи результаты при больших значениях α ?
8. Результаты есть только для сравнения двух человек? Есть ли результаты для других пар людей? Зависит ли разница от сравнения томографий мужчины и женщины, взрослого и ребёнка? Возможно, стоит сделать усреднение по разным парам сравниваемых томографий.

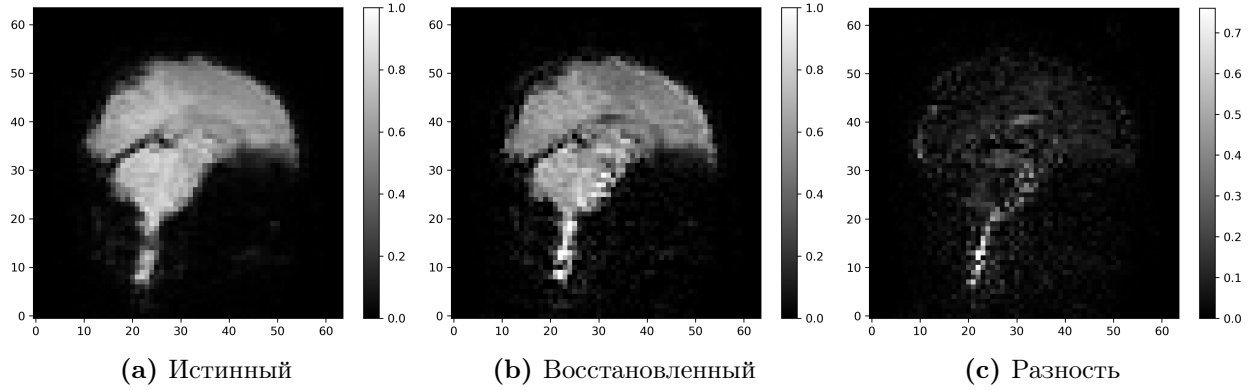


Рис. 5: Срез снимка фМРТ из тестовой выборки

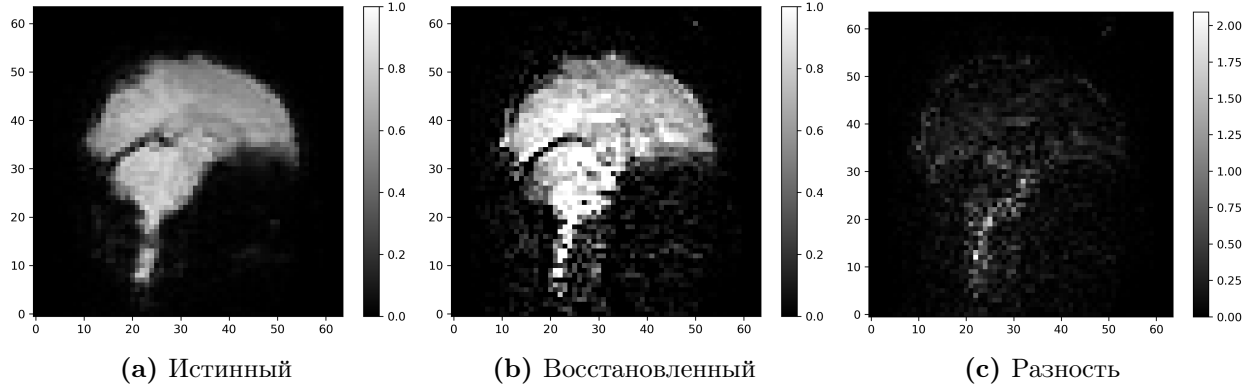


Рис. 6: Срез снимка фМРТ из тестовой выборки (на случайном шуме)

Таблица 3: Качество работы метода на случайном шуме

Выборка	Истинная	Случайный шум
MSE	$2 \cdot 10^{-3}$	10^{-1}

5 Заключение

Построен метод аппроксимации последовательности снимков фМРТ по видеоряду, просматриваемому человеком. Справедлива гипотеза о линейной зависимости между данными. Подтверждена гипотеза о взаимосвязи снимков в последовательности. Проверена гипотеза инвариантности весов модели относительно человека.

Список литературы

- [1] Julia Berezutskaya, Zachary V. Freudenburg, Luca Ambrogioni, Umut Güçlü, Marcel A. J. van Gerven, and Nick F. Ramsey. Cortical network responses map onto data-driven features that capture visual semantics of movie fragments. *Scientific Reports*, 10(1), July 2020.
- [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] Paul Boersma and David Weenink. Praat: doing phonetics by computer [computer program]. version 6.0. 37. Retrieved February, 3:2018, 2018.
- [4] Gary H. Glover. Overview of functional magnetic resonance imaging. *Neurosurgery Clinics of North America*, 22(2):133–139, April 2011.
- [5] Nikos K. Logothetis. The underpinnings of the BOLD functional magnetic resonance imaging signal. *The Journal of Neuroscience*, 23(10):3963–3971, May 2003.
- [6] Nikos K Logothetis. What we can do and what we cannot do with fmri. *Nature*, 453(7197):869–878, 2008.
- [7] Ravi S Menon and Seong-Gi Kim. Spatial and temporal limits in cognitive neuroimaging with fmri. *Trends in cognitive sciences*, 3(6):207–216, 1999.
- [8] Maxim Sharaev, Alexander Andreev, Alexey Artemov, Alexander Bernstein, Evgeny Burnaev, Ekaterina Kondratyeva, Svetlana Sushchinskaya, and Renat Akzhigitov. fmri: preprocessing, classification and pattern recognition, 2018.
- [9] Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, and Manohar Paluri. Learning spatiotemporal features with 3d convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015.
- [10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017.
- [11] Shen Yan, Xuehan Xiong, Anurag Arnab, Zhichao Lu, Mi Zhang, Chen Sun, and Cordelia Schmid. Multiview transformers for video recognition, 2022.