
ВОССТАНОВЛЕНИЕ СНИМКОВ фМРТ ПО ПРОСМАТРИВАЕМОМУ ВИДЕОРЯДУ

Дорин Даниил
dorin.dd@phystech.edu

Киселев Никита
kiselev.ns@phystech.edu

Грабовой Андрей
grabovoy.av@phystech.edu

26 января 2024 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. Метод строится в предположении наличия постоянной задержки во времени между снимком и видеорядом. Каждый воксель снимка фМРТ аппроксимируется линейной моделью независимо. Кроме того, используется предположение марковости для последовательности снимков фМРТ. Для анализа предложенного метода проводится вычислительный эксперимент на выборке, полученной при томографическом обследовании большого числа испытуемых. На экспериментальных данных проводится анализ зависимости качества работы метода от задержки — времени гемодинамической ответной реакции зависимости уровня кислорода в крови. Проверяются гипотезы об инвариантности весов модели относительно человека и обобщающей способности построенного метода.

Ключевые слова: нейровизуализация · фМРТ · видеоряд · зависимость между данными · проверка гипотез · линейная модель · линейная регрессия

1 Введение

Совокупность методов, визуализирующих структуру и функции человеческого мозга, называется *нейровизуализацией*. Методы нейровизуализации [1], 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 [2]. Эти изменения происходят не моментально, а с некоторой задержкой, которая составляет 4–8 с [3]. Она возникает из-за того, что сосудистая система достаточно долго реагирует на потребность мозга в глюкозе [4, 5, 6].

При получении снимков фМРТ используются последовательности эхопланарных изображений (EPI) [7, 8, 9]. Обработка участков с изменяющейся интенсивностью сигнала в зависимости от способа активации, вида артефактов и длительности проводится с помощью специальных методов и программ [3, 10, 11]. Обработанные результаты оформляются в виде карт активации, которые совмещаются с локализацией анатомических образований головного мозга.

Метод фМРТ играет большую роль в нейровизуализации, однако имеет ряд важных ограничений. В работах [12, 13] 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 [14] предлагаются методы подавления вышеперечисленных шумов на основе графов и демонстрируется их эффективность в задаче выявления эпилепсии и депрессии.

При проведении фМРТ испытуемому дают различные тест-задания и применяют внешние раздражители, вызывающие активацию определенных локальных участков головного мозга, ответственных за выполнение соответствующих его функций. Применяются различные тест-задания: движения пальцами и конечностями [15, 16], поиск изображения и рассмотрение шахматной доски [17, 18], прослушивание неспецифичных шумов, единичных слов или связного текста [19, 20]. Причиной изменения активности человеческого мозга во время фМРТ-обследования также может служить просмотр видеоматериала [21], что является предметом исследования настоящей работы.

Наиболее известные методы обработки видео основаны на 3D свертках [22]. Отличие 3D от 2D свертки заключается в одновременной работе с пространственной и временной частью информации. Существенный недостаток данных методов — сильное увеличение числа параметров модели и большие вычислительные затраты. Одной из наиболее современных и улучшаемых архитектур нейронных сетей для обработки изображений является остаточная нейронная сеть ResNet [23]. Она позволяет обучать глубокие нейронные сети (до 152 слоев) с высокой точностью, преодолевая проблему затухания градиента, которая возникает при обучении глубоких сетей.

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 [6]. 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.

Согласно исследованию [24], при фМРТ-обследовании пациентов, просматривающих видеоряд, активируется определенная корковая сеть. Она включает в себя в том числе задние центральные и фронтальные области. Находится эта сеть преимущественно в правом полушарии. В настоящей работе рассматривается подход, использующий для анализа времени задержки именно эти части головного мозга.

Данные, на которых проводятся проверка гипотезы зависимости и демонстрация работы построенного метода, представлены в работе [25]. 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 [26] и [27].

2 Постановка задачи

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

$$\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_{\nu t}], \quad \mathbf{p}_\ell \in \mathbb{R}^{W \times H \times C}, \quad (1)$$

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.

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. С целью сокращения времени работы метода можно использовать сжатие снимков фМРТ путем уменьшения размерностей. Сжатие в 2 раза представляется в виде отображения

$$\chi : \mathbb{R}^{X \times Y \times Z} \rightarrow \mathbb{R}^{X/2 \times Y/2 \times Z/2}.$$

Сжатие в 2^k раз получается применением χ последовательно k раз. Далее для простоты сохраняем обозначения размерностей снимков $X \times Y \times Z$.

Предположим, что для последовательности снимков выполнено марковское свойство, т.е. каждый снимок зависит только от одного изображения и предыдущего снимка. Тогда соответствующее отображение можно записать в виде

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

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.

Отображение $\mathbf{g} : \mathbf{P} \rightarrow \mathbf{S}$ можно представить в виде композиции двух других:

$$\mathbf{g} = \varphi \circ \psi,$$

$\psi : \mathbf{P} \rightarrow \mathbb{R}^d$ — векторизация изображения,

$\varphi : \mathbb{R}^d \rightarrow \mathbf{S}$ — восстанавливаемое отображение.

Для каждого изображения из видеоряда имеем вектор признакового 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 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

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

$$f_{ijk}(\mathbf{x}, \mathbf{w}_{ijk}) = \langle \mathbf{x}, \mathbf{w}_{ijk} \rangle. \quad (7)$$

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

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

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

Требуется найти параметры, доставляющие минимум функционалу потерь $\mathcal{L}_{ijk}(\mathbf{w}_{ijk})$ при заданных гиперпараметрах Δt и α :

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

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

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

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

$$\Delta_{ijk} = [\delta_{ijk}^2, \dots, \delta_{ijk}^N]^\top \in \mathbb{R}^{N-1}. \quad (11)$$

Решение можно записать в виде

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

Получим формулу для восстановленных снимков фМРТ. Введем матрицу весов

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

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

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

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 и проверки гипотез проведен вычислительный эксперимент.

В качестве данных использовалась выборка, представленная в работе [25]. 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.

Характеристики выборки: продолжительность обследования, частота кадров видеоряда и снимков фМРТ, размерности изображений и снимков приведены в Таблице 1

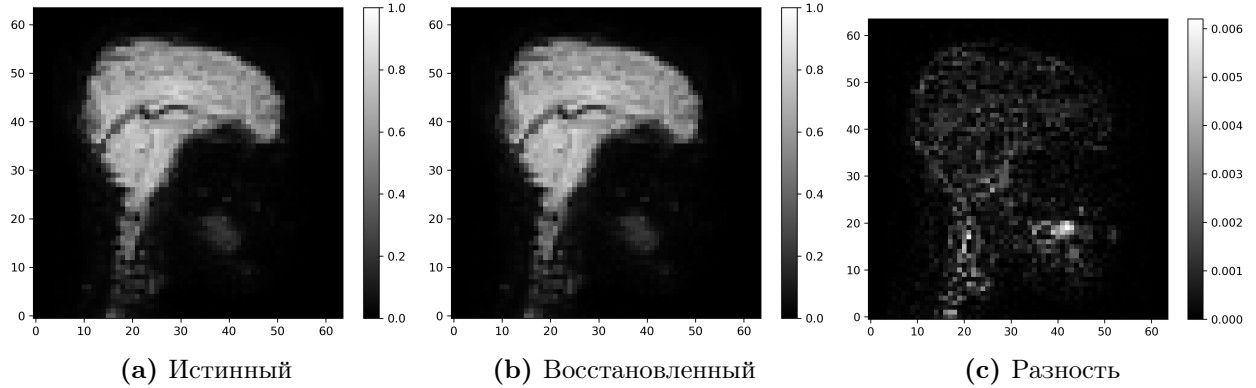
Произведено разделение выборки на тренировочную и тестовую в соотношении 70% и 30% соответственно. Критерием качества восстановления снимка фМРТ служит MSE — сумма квадратов отклонений между истинным и восстановленным снимками, усредненная по всем вокселям каждого снимка из тестовой выборки.

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.

Таблица 1: Описание выборки

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

Демонстрация работы метода. На Рис. 1 представлены срезы истинного и восстановленного снимков из тестовой выборки. На Рис. 1.(с) можно наблюдать разность между ними. Для демонстрации работы алгоритма был выбран 4-ый испытуемый, $\Delta t = 5$ с, коэффициент сжатия 1, коэффициент регуляризации $\alpha = 1000$. Рассмотрен 20-ый срез по первой координате 100-го снимка в последовательности. Так как значения вокселей нормированы на отрезок $[0; 1]$, то ошибка порядка $10^{-3} \div 10^{-2}$ свидетельствует о достаточно точном предсказании.


Рис. 1: Срез снимка фМРТ из тестовой выборки

Анализ времени задержки. Исследована зависимость качества восстановления от времени задержки. Для примера был выбран 47-ой испытуемый. На левом графике Рис. 3 представлена зависимость метрики MSE от времени задержки Δt . Исследование [24] подтверждает, что наиболее активная часть мозга при рассматриваемом обследовании — затылочная доля. Остальные части вносят шум в рассматриваемую зависимость. В настоящей работе проведена локализация вышеупомянутой области, что можно наблюдать на Рис. 2. Для локализации области отсекаются нижняя треть и правые две трети объемного томографического изображения. Красным цветом выделена та зона, в которую попадают 3% наиболее изменяющихся вокселей затылочной доли. Для этого все воксели локализованной области были упорядочены по убыванию суммарного абсолютного изменения значений. Далее были выбраны 3% вокселей с наибольшими изменениями. Произведен пересчет метрики MSE именно на этой части снимка. Соответствующий график приведен справа на Рис. 3.

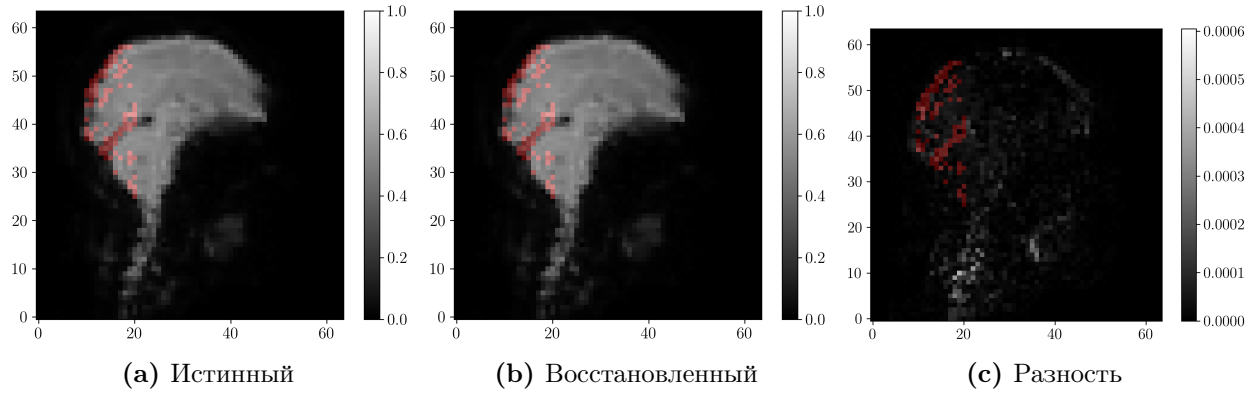


Рис. 2: Локализация наиболее активной зоны

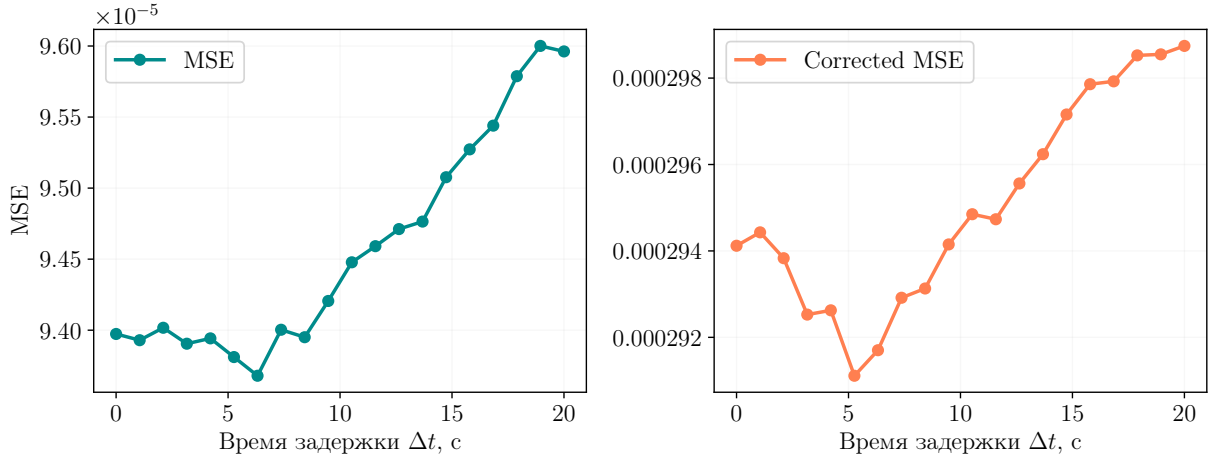


Рис. 3: Зависимость метрики MSE от времени задержки

Подбор оптимального коэффициента регуляризации. Проведен анализ зависимости MSE от коэффициента регуляризации α . Рассматривались коэффициенты сжатия 1, 2, 4 и 8. Соответствующие графики приведены на Рис. 4. Для построения графика производилось усреднение по испытуемым. Обозначены границы среднеквадратичного отклонения. Из графиков видно, что оптимальное значение коэффициента $\alpha \approx 1000$. Вид кривой сохраняется независимо от коэффициента сжатия снимков фМРТ.

Влияние коэффициента сжатия снимков на время работы метода.

Анализ распределения весов модели. Построен график распределения значений компонент вектора весов модели. Для построения производилось усреднение по всем вокселям для 4-го испытуемого. Результат представлен на Рис. 5. Веса модели не лежат в окрестности какого-то конкретного значения, то есть их распределение не вырождено. Этот результат вполне согласуется с реальностью, поскольку человек во время просмотра обращает внимание на определенные части кадра, например, персонажей или другие детали.

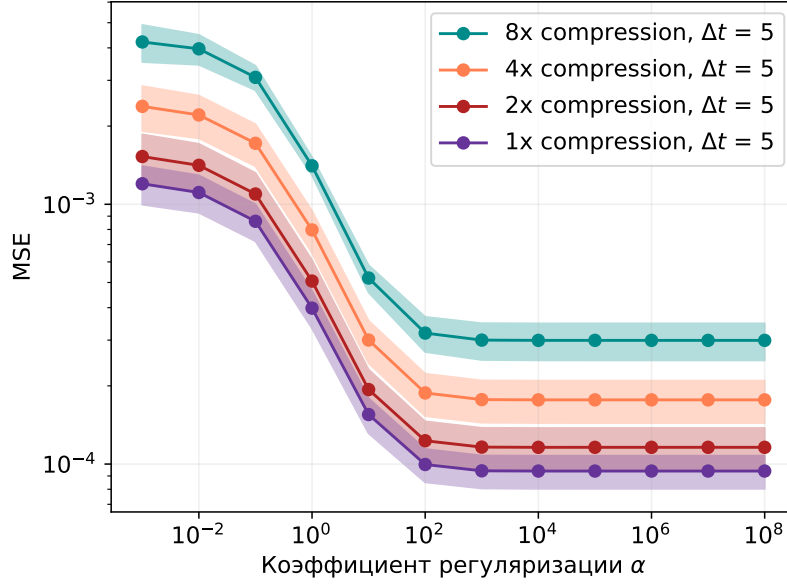


Рис. 4: Зависимость метрики MSE от коэффициента регуляризации α на снимках из тестовой выборки

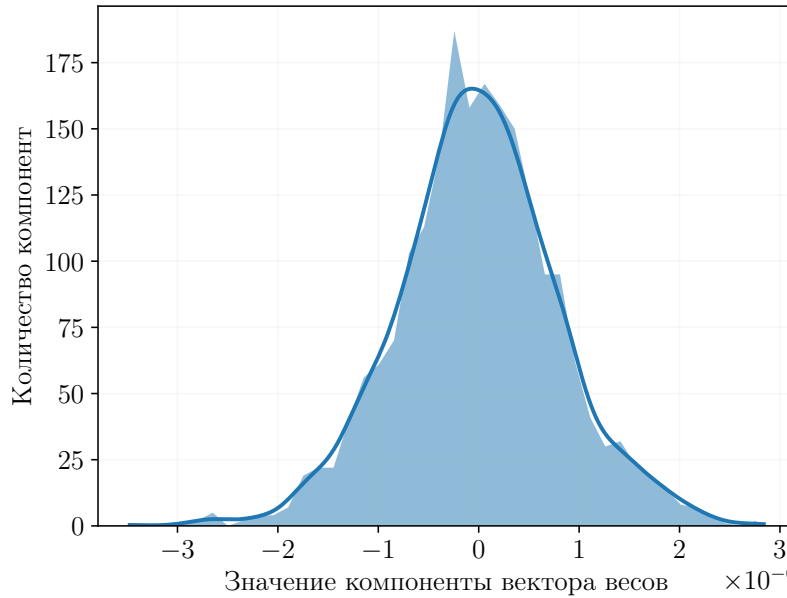


Рис. 5: Распределение значений компонент вектора весов

Гипотеза инвариантности весов модели относительно человека. Проведена проверка гипотезы инвариантности весов модели относительно человека: можно ли проводить 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.

Таблица 2: Проверка гипотезы инвариантности весов модели относительно человека

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

Аналогичный эксперимент проведен для каждой пары испытуемых. Полученные результаты можно наблюдать на Рис. 6, который был получен следующим образом. Рассматривается некоторый испытуемый (соответствует строке матрицы), для него вычисляется MSE — «истинный». Далее рассматривается другой испытуемый (соответствует столбцу матрицы), берется его матрица весов, и с помощью нее делается предсказание для первого испытуемого, затем вычисляется MSE — «подмешанный». Разность между полученными MSE в процентах от «истинного» заносится в матрицу. Положительное значение означает, что «подмешанный» MSE больше, чем «истинный». Отрицательное — что «подмешанный» меньше. Идеальная модель должна приводить только к положительным значениям отклонений, однако, как видно из Рис. 6, в матрице есть и отрицательные значения. Тем не менее, они достаточно малы, а именно соответствуют отклонениям порядка 1%. Это можно объяснить тем, что модель достаточно простая и подвержена переобучению. Однако это не мешает сделать вывод о том, что данные не противоречат гипотезе об инвариантности весов модели относительно человека.

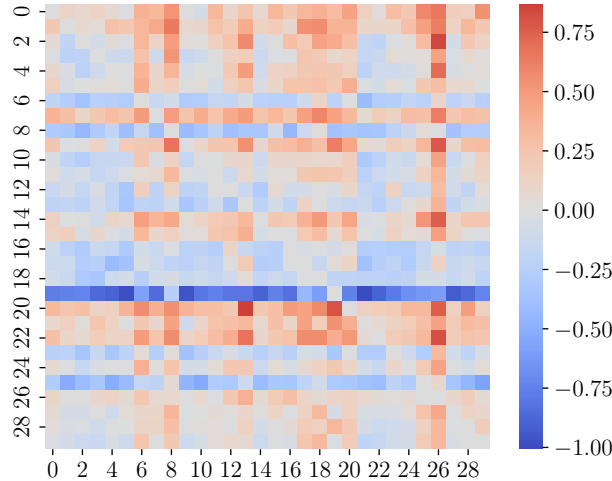


Рис. 6: Изменение MSE при предсказании на подмешанной матрице весов (в процентах)

Корректность метода. Рассмотрено качество работы метода на случайно сгенерированных данных. В качестве матрицы объекты-признаки \mathbf{X} взята матрица случайных чисел из равномерного распределения на $[0; 1]$. Произведено сравнение с результатами на настоящей матрице признаковового описания. К первому снимку 35-го испытуемого последовательно прибавляются все восстановленные изменения значений вокселей. В результате имеем последний снимок последовательности. На Рис. 7 представлены срезы последнего истинного и восстановленного снимков из тестовой выборки. На Рис. 7.(с) можно наблюдать разность между ними. Результаты на случайных данных продемон-

стрированы на Рис. 8. Можно видеть, что разность между истинным и восстановленным снимками при работе со случайными данными значительно выше, что подтверждает наличие зависимости между показаниями датчиков и изображениями из видеоряда. Численные результаты приведены в Таблице 3.

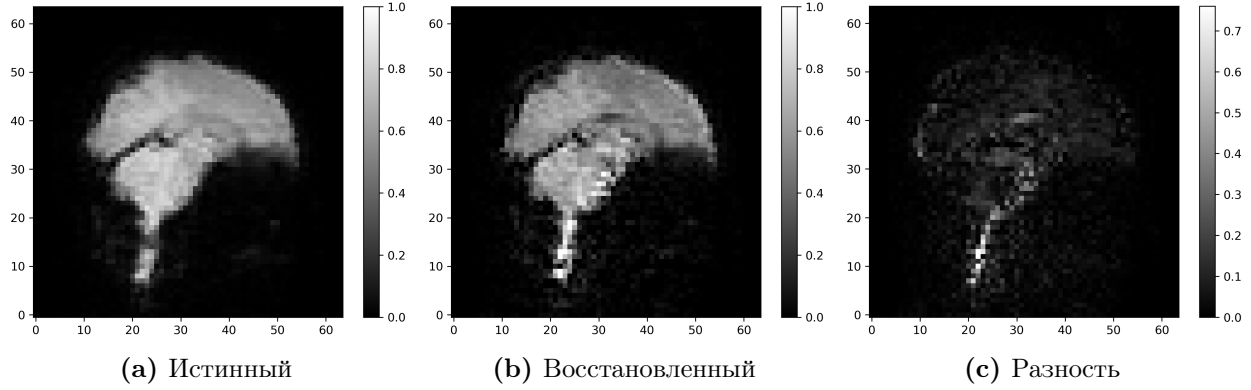


Рис. 7: Срез снимка фМРТ из тестовой выборки

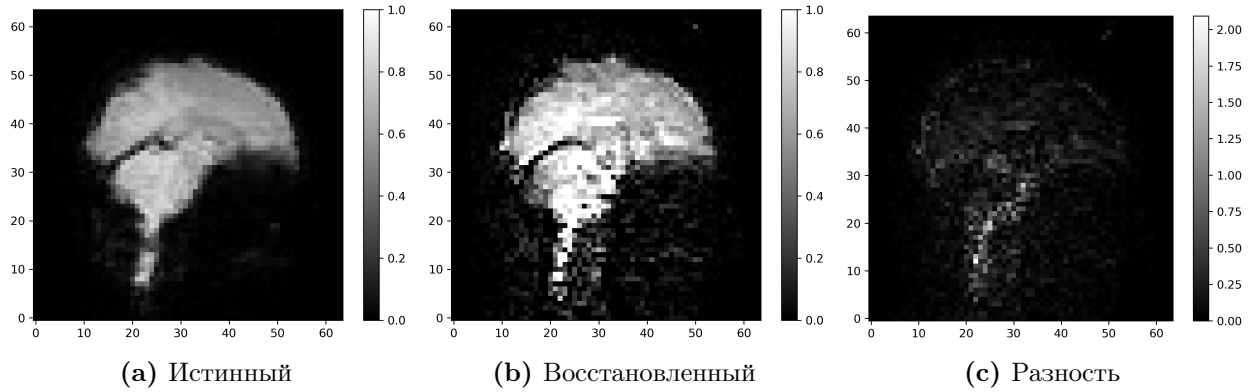


Рис. 8: Срез снимка фМРТ из тестовой выборки (на случайных данных)

Таблица 3: Качество работы метода на случайных данных

Выборка	Истинная	Случайные данные
MSE	$2 \cdot 10^{-3}$	10^{-1}

5 Заключение

В работе рассматривается задача восстановления зависимости между показаниями датчиков фМРТ и восприятием внешнего мира человеком. предлагается метод аппроксимации последовательности снимков фМРТ по видеоряду. Метод учитывает время гемодинамической ответной реакции — время задержки между изображениями из видеоряда и снимками фМРТ. Каждый воксель снимка фМРТ аппроксимируется независимо от других линейной моделью. Каждая линейная модель строится в предположении марковости последовательности снимков фМРТ. В ходе экспериментов для каждого испытуемого подбирается оптимальное значение времени задержки. Оптимальное значение находится из анализа графика зависимости MSE от времени задержки. Предполагается, что за информацию со зрительных органов отвечает затылочная доля мозга. Производится поправка MSE на основе локализации задней области мозга и выбора наиболее изменяющихся вокселей. При таком построении график имеет характерный минимум, отвечающий оптимальному значению времени задержки. Полученное значение времени задержки согласуется с нейробиологическими данными. Экспериментальные значения MSE малы, что говорит о наличии корреляции между данными. Учитывается изменение изображений в видеоряде, так как распределение весов модели не вырождено. Проверена гипотеза инвариантности весов модели относительно человека. Корректность метода подтверждается экспериментами с шумовыми данными.

Список литературы

- [1] Ju V Puras and EV Grigorieva. The neurovisualization methods in diagnostics of head injury. part 1. computer tomography and magnetic resonance imaging. *Russian journal of neurosurgery*, (2):7–16, 2014.
- [2] Gary H. Glover. Overview of functional magnetic resonance imaging. *Neurosurgery Clinics of North America*, 22(2):133–139, April 2011.
- [3] Peter A. Bandettini, Eric C. Wong, R. Scott Hinks, Ronald S. Tikofsky, and James S. Hyde. Time course epi of human brain function during task activation. *Magnetic Resonance in Medicine*, 25(2):390–397, June 1992.
- [4] S Ogawa, T M Lee, A R Kay, and D W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24):9868–9872, December 1990.
- [5] Denis Le Bihan and Avi Karni. Applications of magnetic resonance imaging to the study of human brain function. *Current Opinion in Neurobiology*, 5(2):231–237, 1995.
- [6] Nikos K. Logothetis. The underpinnings of the BOLD functional magnetic resonance imaging signal. *The Journal of Neuroscience*, 23(10):3963–3971, May 2003.
- [7] A Connelly, G D Jackson, R S Frackowiak, J W Belliveau, F Vargha-Khadem, and D G Gadian. Functional mapping of activated human primary cortex with a clinical mr imaging system. *Radiology*, 188(1):125–130, July 1993.
- [8] K K Kwong, J W Belliveau, D A Chesler, I E Goldberg, R M Weisskoff, B P Poncelet, D N Kennedy, B E Hoppel, M S Cohen, and R Turner. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences*, 89(12):5675–5679, June 1992.
- [9] S Ogawa, D W Tank, R Menon, J M Ellermann, S G Kim, H Merkle, and K Ugurbil. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, 89(13):5951–5955, July 1992.
- [10] Klaus Baudendistel, Lothar R. Schad, Michael Friedlinger, Frederik Wenz, Johannes Schröder, and Walter J. Lorenz. Postprocessing of functional mri data of motor cortex stimulation measured with a standard 1.5 t imager. *Magnetic Resonance Imaging*, 13(5):701–707, 1995.
- [11] Robert W. Cox. Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29(3):162–173, 1996.
- [12] 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.
- [13] Nikos K Logothetis. What we can do and what we cannot do with fmri. *Nature*, 453(7197):869–878, 2008.
- [14] Maxim Sharaev, Alexander Andreev, Alexey Artemov, Alexander Bernstein, Evgeny Burnaev, Ekaterina Kondratyeva, Svetlana Sushchinskaya, and Renat Akzhigitov. fmri: preprocessing, classification and pattern recognition, 2018.

- [15] F.E. Roux, J.P. Ranjeva, K. Boulanouar, C. Manelfe, J. Sabatier, M. Tremoulet, and I. Berry. Motor Functional MRI for Presurgical Evaluation of Cerebral Tumors. *Stereotactic and Functional Neurosurgery*, 68(1-4):106–111, 07 1998.
- [16] K. Papke, T. Hellmann, B. Renger, C. Morgenroth, S. Knecht, G. Schuierer, and P. Reimer. Clinical applications of functional mri at 1.0 t: motor and language studies in healthy subjects and patients. *European Radiology*, 9(2):211–220, 1999.
- [17] Stephen A. Engel, David E. Rumelhart, Brian A. Wandell, Adrian T. Lee, Gary H. Glover, Eduardo-Jose Chichilnisky, and Michael N. Shadlen. fmri of human visual cortex. *Nature*, 369(6481):525–525, 1994.
- [18] Walter Schneider, B. J. Casey, and Douglas Noll. Functional mri mapping of stimulus rate effects across visual processing stages. *Human Brain Mapping*, 1(2):117–133, January 1994.
- [19] J. R. Binder, S. M. Rao, T. A. Hammeke, F. Z. Yetkin, A. Jesmanowicz, P. A. Bandettini, E. C. Wong, L. D. Estkowski, M. D. Goldstein, V. M. Haughton, and J. S. Hyde. Functional magnetic resonance imaging of human auditory cortex. *Annals of Neurology*, 35(6):662–672, June 1994.
- [20] S. Dymarkowski, S. Sunaert, S. Van Oostende, P. Van Hecke, G. Wilms, P. Demaerel, B. Nuttin, C. Plets, and G. Marchal. Functional mri of the brain: localisation of eloquent cortex in focal brain lesion therapy. *European Radiology*, 8(9):1573–1580, 1998.
- [21] Jean Decety, Julie Grezes, Nicolas Costes, Daniela Perani, Marc Jeannerod, Emmanuel Procyk, Franco Grassi, and Ferruccio Fazio. Brain activity during observation of actions. influence of action content and subject’s strategy. *Brain: a journal of neurology*, 120(10):1763–1777, 1997.
- [22] 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*, pages 4489–4497, 2015.
- [23] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.
- [24] Daniel Anderson, Katherine Fite, Nicole Petrovich, and Joy Hirsch. Cortical activation while watching video montage: An fmri study. *Media Psychology - MEDIA PSYCHOL*, 8:7–24, 02 2006.
- [25] 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.
- [26] Paul Boersma and David Weenink. Praat: doing phonetics by computer [computer program]. version 6.0. 37. *Retrieved February*, 3:2018, 2018.
- [27] 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.