

Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»
Физтех-школа прикладной математики и информатики
Кафедра интеллектуальных систем

Дорин Даниил Дмитриевич

**ПРОСТРАНСТВЕННО-ВРЕМЕННЫЕ
ХАРАКТЕРИСТИКИ В ЗАДАЧЕ ДЕКОДИРОВАНИЯ
ВРЕМЕННЫХ РЯДОВ**

03.03.01 — Прикладные математика и физика

Выпускная квалификационная работа бакалавра

Научный руководитель:
Грабовой Андрей Валериевич
канд. физ.-мат. наук

Москва — 2024

Аннотация

Исследуются пространственно-временные характеристики в задаче декодирования временных рядов с дискретным представлением времени. Проводится анализ временной зависимости между последовательностью снимков функциональной магнитно-резонансной томографии, далее фМРТ, и видеорядом, про-сматриваемым человеком. Предлагается метод аппроксимации показаний фМРТ по просматриваемому видеоряду. Метод строится в предположении наличия не изменяющегося времени гемодинамической ответной реакции зависимости уровня кислорода в крови. Для каждого voxеля снимка фМРТ независимо строится линейная модель. Используется предположение марковости последовательности снимков фМРТ. Для анализа предложенного метода проводится вычислительный эксперимент на выборке, полученной при томографическом обследовании большого числа испытуемых. Проводится анализ областей мозга, реагирующих на визуальные стимулы. Предлагается метод снижения пространственной размерности временных рядов фМРТ, основанный на взвешивании активных областей мозга. Построен метод классификации сегментов временных рядов фМРТ, основанный на снижении пространственной размерности и применении римановой геометрии. Проводится вычислительный эксперимент для проверки работоспособности модели. Анализируется влияние отдельных компонент метода на качество классификации.

Содержание

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. В рамках данной работы, основным объектом исследования являются временные ряды фМРТ.

Функциональная магнитно-резонансная томография или *фМРТ* является разновидностью магнитно-резонансной томографии и основана на изменениях в токе крови, вызванных нейронной активностью мозга [?]. Эти изменения происходят не моментально, а с некоторой задержкой, которая составляет 4–8 с [?]. Она возникает из-за того, что сосудистая система достаточно долго реагирует на потребность мозга в глюкозе [? ? ?].

При получении снимков фМРТ используются последовательности эхопланарных изображений (ЕПИ) [? ? ?]. Обработка участков с изменяющейся интенсивностью сигнала в зависимости от способа активации, вида артефактов и длительности проводится с помощью специальных методов и программ [? ? ?]. 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 [? ?] рассматривается временное и пространственное разрешение 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 флюктуациями самого прибора и т. д. В работе [?] 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 функций. Среди тест-заданий зачастую выделяют движения пальцами и конечностями [? ?], просмотр изображений различных категорий и рассмотрение шахматной доски [? ?], прослушивание неспецифичных шумов, единичных слов или связного текста [? ?]. Причиной изменения активности человеческого мозга

во время фМРТ-обследования также может служить просмотр видеоматериала [?], что является одним из объектов исследования настоящей работы.

Наиболее известные методы обработки видео основаны на 3D свертках [?]. Отличие 3D от 2D сверток 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 [?]. Она позволяет обучать глубокие нейронные сети (до 152 слоев) с высокой точностью, преодолевая проблему затухания градиента, которая возникает при обучении глубоких сетей.

В последнее время активно ведутся научные исследования, посвященные методам анализа активности головного мозга и декодирования информации [? ? ?]. Основу применения этих методов составляют технологии нейрокомпьютерных интерфейсов. Интерфейсы мозг-компьютер или ИМК являются относительно новой, но быстро развивающейся областью исследований методов нейровизуализации [?]. Технология ИМК открывает новый способ взаимодействия мозга с компьютером. Система собирает сигналы мозга, анализирует их и преобразует сигналы в команды, которые могут быть отправлены на устройство вывода для выполнения определенного действия. Это позволяет осуществлять прямую связь между мозгом и компьютером. Ключевой целью исследований ИМК является разработка нового канала связи, который позволяет людям с тяжелыми нервно-мышечными нарушениями напрямую передавать сообщения из своего мозга путем анализа умственной активности. Классификация моторных образов в задаче ИМК может быть использована, например, для управления протезами или другими устройствами с помощью мысленных команд [? ? ?]. Такой подход позволяет людям с ограниченными возможностями компенсировать потерю или отсутствие нормальной моторной функции.

Активные области мозга, ответственные за конкретную задачу, называются регионами интереса (англ. ROI — Region of Interest). Выделение регионов интереса является важным шагом в предварительной обработке данных ЭЭГ, фМРТ и других методов нейровизуализации, используемых в нейрокомпьютерных интерфейсах [?]. Данные фМРТ представляют собой объемные многомерные наборы данных. При этом значительная доля вокселей фМРТ соответствует фоновому

изображению, которое может вносить существенный шум при решении задачи декодирования сигнала. Кроме того, необходимо учитывать, что конкретные области мозга отвечают за выполнение определенных зрительных, когнитивных и моторных задач [?]. Таким образом, при анализе снимков фМРТ важным аспектом является корректное выделение регионов интереса, ответственных за выполнение конкретной задачи.

Согласно исследованию [?], 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. Также рассматривается задача восстановления 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 [?]. 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.

Кроме того, в настоящей работе в качестве задачи декодирования временных рядов рассматривается задача классификации. Внимание сосредоточено на анализе данных фМРТ для задачи классификации отрезков временного ряда, что может быть полезным для диагностики некоторых неврологических заболеваний [?]. Изучение активности мозга при обработке визуальной информации позволяет глубже понять механизмы работы мозга и процессы, происходящие в нем. Это, в свою очередь, способствует развитию нейронаук и когнитивной психологии. В работе предлагается метод взвешивания вокселей фМРТ, учитывающий пространственно-временные зависимости в данных. На основе предложенного метода строится модель классификации отрезков временного ряда фМРТ. Модель учитывается пространственные характеристики за счет трансформация пространства признаков в терминах римановой геометрии [? ?].

2. Постановка задачи

2.1. Восстановление снимков фМРТ по видеоряду

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

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

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

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

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

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

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

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

$$k_\ell = \frac{\ell \cdot \nu}{\mu}.$$

2.2. Классификация временных рядов фМРТ

2.2.1. Сегментация временного ряда фМРТ

Определение 1. *Фрагментом временного ряда $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_T]$ будем называть любую его подпоследовательность $\hat{\mathbf{X}}$:*

$$\hat{\mathbf{X}} = [\mathbf{x}_{(t_1)}, \dots, \mathbf{x}_{(t_k)}], \quad 1 \leq t_1 < \dots < t_k \leq T.$$

Определение 2. *Сегментом временного ряда $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_T]$ будем называть его непрерывный фрагмент $\hat{\mathbf{X}}$:*

$$\hat{\mathbf{X}} = [\mathbf{x}_{(t)}]_{t=t_0}^{t_1}, \quad 1 \leq t_0 < t_1 \leq T.$$

Тогда длиной сегмента называется число $\tau = t_1 - t_0 + 1$.

Определение 3. Под сегментацией временного ряда будем понимать отображение \mathcal{G} , сопоставляющее временному ряду \mathbf{X} множество его сегментов $\mathcal{G}(\mathbf{X}) \in 2^{\mathcal{S}(\mathbf{X})}$, где $\mathcal{S}(\mathbf{X})$ множество всех сегментов временного ряда \mathbf{X} .

При анализе медицинских данных важно учитывать уникальные особенности анатомии и реакции каждого человека на стимулы. Для более точного декодирования данных необходимо обучать модель индивидуально под конкретного индивида. В медицинских наборах данных обычно представлено только одно большое измерение для каждого человека. Для решения конкретных задач необходимо разбить или сегментировать временной ряд для дальнейшего анализа.

Определение 4. Под стимулами будем понимать зрительные, моторные или когнитивные задания различных категорий, которые выполняет человек в ходе процедуры фМРТ.

Формализуем задачу сегментации. Имеется одно непрерывное измерение фМРТ с дискретным представлением времени:

$$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_T], \quad \mathbf{x}_t \in \mathbb{R}^{X \times Y \times Z},$$

где X, Y и Z — размерности тензора снимка, T — длина временного ряда. Имеется дискретный временной ряд стимулов:

$$\mathbf{s} = [s_1, \dots, s_T], \quad s_t \in \{0, \dots, C\},$$

где $\{1, \dots, C\}$ — множество классов стимулов. Классом 0 обозначены моменты отдыха человека.

Необходимо построить сегментацию $\mathcal{G}(\mathbf{X}, \mathbf{s})$ временного ряда \mathbf{X} на сегменты фиксированной длины τ , чтобы в сегментах преобладала определенная категория.

2.2.2. Классификация сегментов временного ряда фМРТ

Предположим, что нам дано N сегментов временного ряда фМРТ, каждый из которых имеет длину T . Обозначим эти сегменты как:

$$\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\},$$

где каждое наблюдение \mathbf{X}_i представляет собой последовательность воксельных изображений:

$$\mathbf{X}_i = [\mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_T^i],$$

где \mathbf{x}_t^i — это воксельное изображение в момент времени t для набора i .

Кроме того, каждому наблюдению соответствует метка класса y_i , которая может принимать значения от 1 до C , где C — общее число классов. Обозначаем вектор меток классов как:

$$\mathbf{y} = [y_1, y_2, \dots, y_N]^T.$$

Требуется построить модель классификации, которая учитывает пространственно-временные характеристики временных рядов фМРТ. Формально, имеется выборка

$$\mathcal{D} = \{(y_i, \mathbf{X}_i) \mid i = 1, \dots, N\}.$$

Требуется построить такое отображение g , что:

$$g : \mathbf{X} \rightarrow \{1, \dots, C\}.$$

3. Исследование и построение решения задачи

3.1. Восстановление снимков фМРТ по видеоряду

Схема предлагаемого метода восстановления снимков фМРТ приведена на Рис. ??.

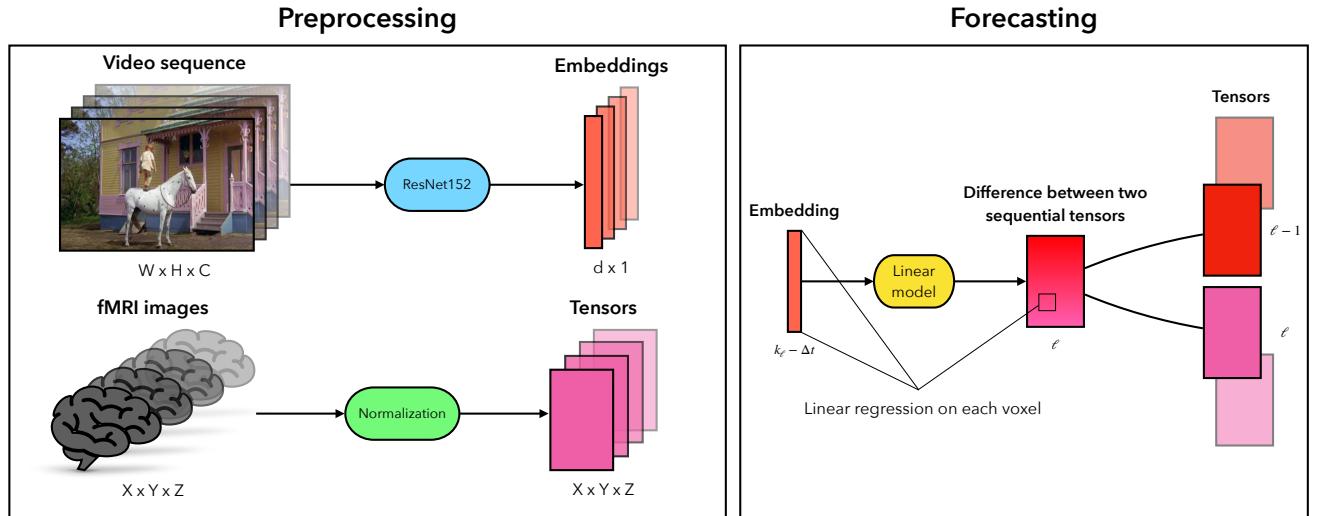


Рис. 1: Схема метода

Обозначим снимок фМРТ как $\mathbf{x}_\ell = [v_{ijk}^\ell] \in \mathbb{R}^{X \times Y \times Z}$, где $v_{ijk}^\ell \in \mathbb{R}_+$ — значение соответствующего voxеля. С целью сокращения времени работы метода предлагается 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 записывается в виде

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

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 $\mathbf{g} : \mathbf{P} \rightarrow \mathbf{X}$ представляется в виде композиции двух других:

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

$$\begin{aligned}\psi : \mathbf{P} \rightarrow \mathbb{R}^d &— векторизация изображения, \\ \varphi : \mathbb{R}^d \rightarrow \mathbf{X} &— восстанавливаемое отображение.\end{aligned}$$

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

$$\mathbf{z}_\ell = [z_1^\ell, \dots, z_d^\ell]^\top \in \mathbb{R}^d, \quad \ell = 1, \dots, \nu t.$$

Используется архитектура нейронной сети ResNet152 без последнего линейного слоя.

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

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

Поставлена задача восстановления регрессии

$$y_{ijk} : \mathbb{R}^d \rightarrow \mathbb{R}.$$

Используется линейная модель с вектором параметров

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

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

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

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

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 :

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

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

$$\mathbf{Z} = [\mathbf{z}_2, \dots, \mathbf{z}_N]^\top = [z_j^i] \in \mathbb{R}^{(N-1) \times d}$$

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

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

Решение записывается в виде

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

Получим формулу для восстановленных снимков фМРТ. Введем матрицу весов

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

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

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

Тогда вектор восстановленного снимка находится по формуле

$$\hat{\mathbf{x}}_\ell^R = \mathbf{x}_{\ell-1}^R + \hat{\boldsymbol{\delta}}_\ell^R = \mathbf{x}_{\ell-1}^R + \hat{\mathbf{W}} \mathbf{z}_\ell.$$

3.2. Классификация временных рядов фМРТ

3.2.1. Сегментация временного ряда фМРТ

Имеется временной ряд фМРТ \mathbf{X} и соответствующий дискретный временной ряд стимулов \mathbf{s} :

$$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_T], \quad \mathbf{x}_t \in \mathbb{R}^{X \times Y \times Z},$$

$$\mathbf{s} = [s_1, \dots, s_T], \quad s_t \in \{0, \dots, C\},$$

где X, Y и Z — размерности тензора снимка, T — длина временного ряда.

Временной ряд стимулов \mathbf{s} содержит информацию о том, какой стимул был представлен испытуемому в каждый момент времени t . Значения стимулов s_t кодируются числами от 0 до C . Кроме того, моменты отдыха испытуемого от стимула обозначены как $s_t = 0$.

Предположение 1. В работе предполагается, что испытуемый имел достаточное время для отдыха между демонстрацией стимулов разных категорий. На практике, данное предположение обычно выполняется, что позволяет избежать пересечения эффектов от различных стимулов.

Построим сегментацию $\mathcal{G}(\mathbf{X}, \mathbf{s})$ временного ряда \mathbf{X} . Для каждой категории $k \in \{1, \dots, C\}$ выделим из временного ряда \mathbf{X} сегменты фиксированной длины τ следующим образом:

1. Находим индексы начала и конца отрезков, соответствующих классу стимула k . Обозначим индексы в порядке возрастания $\{b_1, \dots, b_{n_k}\}$ и $\{e_1, \dots, e_{n_k}\}$ соответственно, где n_k — число отрезков, отвечающих категории k . Формально данные индексы определяются как:

$$\{b_1, \dots, b_{n_k}\} = \{t \mid s_t = k, s_{t-1} = 0\},$$

$$\{e_1, \dots, e_{n_k}\} = \{t \mid s_t = k, s_{t+1} = 0\},$$

2. Для каждого сегмента $[b_j, e_j] \subset [1, T]$ длины $\tau_j = e_j - b_j + 1$ добавим слева и справа сегменты, отвечающие моментам отдыха испытуемого, чтобы

получить отрезок длины τ :

$$\mathbf{X}_j = [\mathbf{x}_{b_j-\delta_1}, \dots, \mathbf{x}_{e_j+\delta_2}],$$

$$\mathbf{s}_j = [s_{b_j-\delta_1}, \dots, s_{e_j+\delta_2}],$$

где $j = 1, \dots, n_k$, $\delta_1 = \lfloor (\tau - \tau_j)/2 \rfloor$, $\delta_2 = \lceil (\tau - \tau_j)/2 \rceil$. Здесь $\lfloor \cdot \rfloor$ и $\lceil \cdot \rceil$ обозначают операции округления в меньшую и большую сторону соответственно.

Обратим внимание, что для каждой категории k предполагается $\tau_j < \tau$, $j = 1, \dots, n_k$.

3. В результате объединения по всем категориям получим $N = \sum_{k=1}^C n_k$ сегментов временного ряда длины τ . Формально получаем:

$$\mathbf{X} = \{\mathbf{X}_1, \dots, \mathbf{X}_N\},$$

$$\mathbf{X}_j = [\mathbf{x}_1^j, \dots, \mathbf{x}_\tau^j], \quad \mathbf{x}_t^j \in \mathbb{R}^{X \times Y \times Z},$$

$$\mathbf{s}_j = [s_1^j, \dots, s_\tau^j], \quad s_t \in \{0, 1\},$$

$$\mathbf{y} = [y_1, \dots, y_N], \quad y_j \in \{1, \dots, C\},$$

где y_j — метка класса временного ряда \mathbf{X}_j , \mathbf{s}_j — временной ряд стимулов, где s_t^j принимает ненулевое значение в моменты стимулирования категории y_j .

3.2.2. Взвешивание вокселей фМРТ снимков

Предлагается метод *Cross-Correlation Weighting* взвешивания активных областей мозга в задаче декодирования временных рядов фМРТ. Результатом работы алгоритма является бинарная маска $\mathcal{M} \in \mathbb{R}^{X \times Y \times Z}$, в которой ненулевым значениям соответствуют области человеческого мозга, реагирующие на стимул.

Имеется временной ряд фМРТ \mathbf{X} и соответствующий дискретный временной ряд стимулов \mathbf{s} :

$$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_\tau], \quad \mathbf{x}_t \in \mathbb{R}^{X \times Y \times Z},$$

$$\mathbf{s} = [s_1, \dots, s_\tau], \quad s_t \in \{0, 1\},$$

где X, Y и Z — размерности тензора снимка, T — длина временного ряда. Временной ряд стимулов \mathbf{s} содержит информацию о моментах стимулирования испытуемого. Моменты отдыха от стимула обозначены как $s_t = 0$, в то время как моменты стимулирования — $s_t = 1$. Обратим внимание, что предполагается однотипное стимулирование в ходе испытания.

Предположение 2. *Наличие неизменяющегося времени гемодинамической ответной реакции зависимости уровня кислорода в крови.*

Предположение 3. *Если область мозга реагирует на стимул, то значение кросс-корреляционной функции для voxелей из этой области ожидается большим.*

Предложенный метод состоит из следующих этапов:

1. На вход подается временной ряд фМРТ \mathbf{X} с частотой снимков μ и временной ряд стимулов \mathbf{s} .
2. 3D Average Pooling: ко всем изображениям фМРТ в \mathbf{X} применяется свертка $\text{AvgPool3d}(\cdot, k_s)$ с ядром k_s :

$$\begin{aligned}\mathbf{X}' &= [\mathbf{x}'_1, \dots, \mathbf{x}'_T], \\ \mathbf{x}'_t &= \text{AvgPool3d}(\mathbf{x}_t, k_s) \in \mathbb{R}^{X/k_s \times Y/k_s \times Z/k_s},\end{aligned}$$

где \mathbf{X}' — это сжатые данные фМРТ.

3. Z-нормализация временных рядов: временные ряды каждого voxеля в сжатых данных фМРТ \mathbf{X}' нормализуются путем вычитания их выборочного среднего и деления на несмещенную оценку стандартного отклонения. Аналогично нормализуется временной ряд \mathbf{s} . Обозначим нормализованные данные $\hat{\mathbf{X}'}$ и $\hat{\mathbf{s}}$ соответственно.
4. Вычисление кросс-корреляции: введем обозначение $v_t^{i,j,k}$ — значение voxеля снимка фМРТ $\hat{\mathbf{x}}'_t$ в позиции (i, j, k) . Тогда для каждой тройки индексов (i, j, k) определим функцию кросс-корреляции между временным рядом voxеля $\mathbf{v}^{i,j,k}$ и временным рядом стимулов $\hat{\mathbf{s}}$ следующим образом:

$$c_{i,j,k}(p) = (\hat{\mathbf{s}} * \mathbf{v}^{i,j,k})(p) = \frac{1}{\tau - 1} \sum_{t=1}^{\tau-p} \hat{s}_t \cdot v_t^{i,j,k}, \quad p = 0, \dots, \tau - 1, \quad \text{где}$$

$$\mathbf{v}^{i,j,k} = [v_1^{i,j,k}, \dots, v_\tau^{i,j,k}],$$

$$\hat{\mathbf{s}} = [\hat{s}_1, \dots, \hat{s}_\tau].$$

5. Воспользуемся предположениями. Формально есть задержка Δt , которая является параметром метода. Вычислим дискретный момент времени соответствующий задержке:

$$p_{\Delta t} = \lfloor \mu \Delta t \rfloor.$$

Находим значение кросс-корреляционной функции $c_{i,j,k}(p_{\Delta t})$ для каждой тройки индексов (i, j, k) . Рассмотрим нулевой тензор $\mathcal{M}_c \in \mathbb{R}^{X/k_s \times Y/k_s \times Z/k_s}$.

Теперь присваиваем единицы для значений тензора в позициях (i, j, k) , соответствующих h наибольшим значениям $c_{i,j,k}(p_{\Delta t})$. В итоге получаем маску, однако сжатую.

6. В завершение процедуры выполняем операцию обратной свертки, также известную как *upsample*, которая позволяет получить тензор \mathcal{M} с разрешением $X \times Y \times Z$.

В этой операции назначаем каждому элементу тензора \mathcal{M}_c с разрешением $X/k_s \times Y/k_s \times Z/k_s$ маленький тензор размером $k_c \times k_c \times k_c$. Значения этого маленького тензора совпадают со значениями элемента в \mathcal{M}_c .

Метод имеет три настраиваемых параметра: размер ядра k_c , время гемодинамической ответной реакции Δt и число активных областей h .

Статистический анализ локализованных областей мозга. Для проверки статистической значимости областей, полученных с помощью метода *Cross-Correlation Weighting*, предлагается проверить значимость корреляции между временным рядом стимула и временным рядом каждого voxelя фМРТ. Воспользуемся выборочным коэффициентом корреляции Кенделла.

Определение 5. Выборочный коэффициент корреляции Кендалла:

$$\hat{\tau}_{X_1 X_2} = 1 - \frac{4}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=1}^n [[X_{1i} < X_{1j}] \neq [X_{2i} < X_{2j}]] = \frac{C - D}{C + D},$$

где X_1, X_2 — выборки объема n , C — число согласованных пар, D — число несогласованных пар.

Рассмотрим временной ряд \mathbf{v} каждого voxеля фМРТ и временной ряд стимула \mathbf{s} , тогда применим следующий критерий:

выборки:	$\mathbf{v} = [v_1, \dots, v_\tau]$
	$\mathbf{s} = [s_1, \dots, s_\tau]$
	выборки связанные
нулевая гипотеза:	$H_0: \tau_{\mathbf{v}\mathbf{s}} = 0$
альтернатива:	$H_1: \tau_{\mathbf{v}\mathbf{s}} > 0$
статистика:	$\hat{\tau}_{\mathbf{v}\mathbf{s}}$
нулевое распределение:	табличное

При справедливости H_0 :

$$\mathbb{E}\hat{\tau}_{\mathbf{v}\mathbf{s}} = 0, \quad \mathbb{D}\hat{\tau}_{\mathbf{v}\mathbf{s}} = \frac{2(2\tau + 5)}{9\tau(\tau - 1)}.$$

Ставится задача множественного тестирования гипотез, так как критерий применяется для каждого voxеля фМРТ. Делаем поправку Холма на множественное тестирование гипотез. Последним шагом отбираем позиции voxелей, где гипотеза с учетом поправки отвергается. Корреляция полученных областей мозга со стимулом статистически значима.

3.2.3. Классификация сегментов временного ряда фМРТ

Схема предлагаемого метода классификация сегментов временного ряда фМРТ приведена на Рис. ??.

Пусть имеется N сегментов временного ряда фМРТ длины τ . Кроме того, имеются временные ряды стимулов и метки классов, соответствующие сегментам. Формально:

$$\mathbf{X} = \{\mathbf{X}_1, \dots, \mathbf{X}_N\},$$

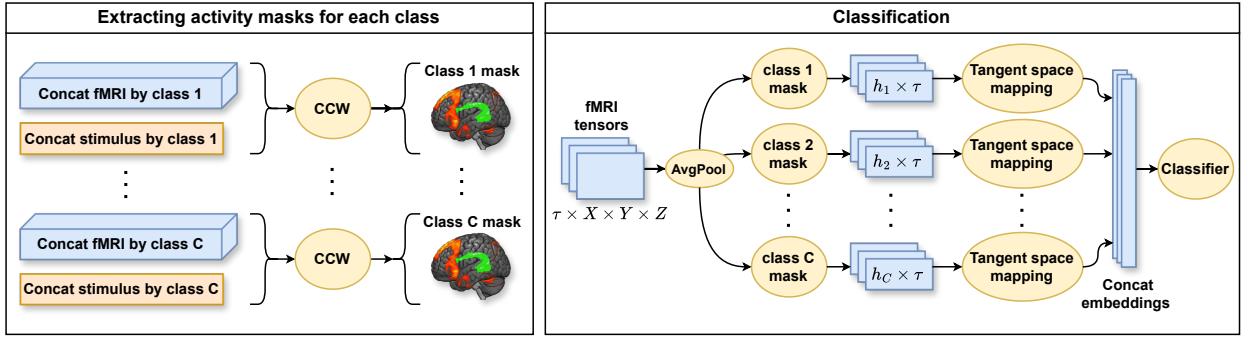


Рис. 2: Схема метода

$$\mathbf{X}_j = [\mathbf{x}_1^j, \dots, \mathbf{x}_\tau^j], \quad \mathbf{x}_t^j \in \mathbb{R}^{X \times Y \times Z},$$

$$\mathbf{s}_j = [s_1^j, \dots, s_\tau^j], \quad s_t \in \{0, 1\},$$

$$\mathbf{y} = [y_1, \dots, y_N], \quad y_j \in \{1, \dots, C\},$$

где y_j — метка класса временного ряда \mathbf{X}_j , \mathbf{s}_j — временной ряд стимулов, где s_t^j принимает ненулевое значение в моменты стимулирования категории y_j .

Извлечение масок активности мозга для каждой категории стимула.
Задана обучающая выборка:

$$\mathfrak{D} = \{(\mathbf{X}_j, \mathbf{s}_j, y_j) \mid j = 1, \dots, N\}.$$

Выделение маски активности мозга для каждой категории стимула позволяет выделить специфичные для каждого класса области мозга, которые активируются в ответ на соответствующий стимул.

Без ограничения общности рассмотрим произвольный класс $k \in \{1, \dots, C\}$. Выделим подвыборку \mathfrak{D}_k , соответствующую классу k , и перенумеруем для удобства объекты подвыборки:

$$\mathfrak{D}_k = \{(\mathbf{X}_j, \mathbf{s}_j, y_j) \in \mathfrak{D} \mid y_j = k\} = \{(\mathbf{X}_j, \mathbf{s}_j) \mid j = 1, \dots, N_k\}.$$

При использовании метода *Cross-Correlation Weighting* на отдельном объекте подвыборки нет гарантии, что полученная маска соответствует нужной области мозга, отвечающей стимулированию класса k . Это связано в первую очередь с ограничениями и недостатками процедуры фМРТ, о чем подробно написано в разделе ?? Поэтому для повышения статистической значимости анализа пред-

лагается брать усредненную активность по объектам подвыборки. Для этого рассмотрим конкатенацию временных рядов фМРТ и стимула вдоль временной координаты. Формально для подвыборки \mathfrak{D}_k получаем:

$$\mathbf{X}^k = \mathbf{X}_1 \oplus \mathbf{X}_2 \oplus \dots \oplus \mathbf{X}_{N_k},$$

$$\mathbf{s}^k = \mathbf{s}_1 \oplus \mathbf{s}_2 \oplus \dots \oplus \mathbf{s}_{N_k},$$

$$\mathbf{X}^k \in \mathbb{R}^{\tau N_k \times X \times Y \times Z}, \quad \mathbf{s}^k \in \{0, 1\}^{\tau N_k},$$

где операция $(\cdot \oplus \cdot)$ — конкатенация по первой координате. К полученным рядам \mathbf{X}^k и \mathbf{s}^k применяем метод *Cross-Correlation Weighting* с гиперпараметрами h_k, k_s , считаем, что задержка учтена $\Delta t = 0$. Решено не делать шаг с Upsample, так как применение Average Pooling к тензорам фМРТ позволяет уменьшить влияние аномальных или шумовых вокселей на анализ активности головного мозга. Получаем для категории стимула k сжатую бинарную маску активности $\mathcal{M}^k \in \{0, 1\}^{X/k_s \times Y/k_s \times Z/k_s}$.

Определение 6. *Формально определим операцию применения бинарной маски активности к временному ряду фМРТ. Пусть задан временной ряд фМРТ:*

$$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_\tau], \quad \mathbf{x}_t \in \mathbb{R}^{X \times Y \times Z}.$$

Задана бинарная маска активности головного мозга:

$$\mathcal{M} \in \{0, 1\}^{X \times Y \times Z}.$$

Если в вокселе (i, j, k) присутствует активность, то $\mathcal{M}(i, j, k) = 1$, иначе $\mathcal{M}(i, j, k) = 0$. Тогда операция применения маски \mathcal{M} кциальному ряду \mathbf{X} определяется как:

$$\mathbf{Z} = \mathcal{M} \bullet \mathbf{X} = [\mathbf{z}_1, \dots, \mathbf{z}_\tau],$$

где $\mathbf{z}_t \in \mathbb{R}^h$, h — число ненулевых элементов в маске \mathcal{M} .

$$\mathbf{z}_t(\ell) = \mathbf{x}_t(i, j, k), \quad \text{где } (i, j, k) = \text{ind}_\ell(\mathcal{M}).$$

$\text{ind}_\ell(\mathcal{M})$ — функция, возвращающая индексы ℓ -го ненулевого вокселя в маске \mathcal{M} . В результате операции применения маски получаем матрицу $\mathbf{Z} \in \mathbb{R}^{h \times \tau}$.

Классификация сегментов временного ряда. Задана обучающая выборка:

$$\mathfrak{D} = \{(\mathbf{X}_j, y_j), \mathbf{X}_j \in \mathbb{R}^{\tau \times X \times Y \times Z} \mid j = 1, \dots, N\}.$$

Кроме того, для каждой категории стимулов задана сжатая с ядром k_s бинарная маска активности:

$$\mathbf{M} = \{\mathcal{M}^k \in \{0, 1\}^{X/k_s \times Y/k_s \times Z/k_s} \mid k = 1, \dots, C\}.$$

Искомое отображение $g : \mathbf{X} \rightarrow \{1, \dots, C\}$ представляется в виде композиции трех других:

$$g := \varphi \circ \psi \circ \mathcal{A},$$

$$\mathcal{A} : \mathbf{X} \rightarrow \mathbb{R}^{\tau \times X/k_s \times Y/k_s \times Z/k_s} — \text{Average Pooling},$$

$$\psi : \mathbb{R}^{\tau \times X/k_s \times Y/k_s \times Z/k_s} \rightarrow \mathbb{R}^d — \text{векторизатор},$$

$$\varphi : \mathbb{R}^d \rightarrow \{1, \dots, C\} — \text{классификатор}.$$

Отображение \mathcal{A} — применение 3d Average Pooling с ядром k_s к каждому тензору фМРТ во временном ряду.

Построим отображение ψ и определим размерность эмбеддингов d . Обозначим число ненулевых элементов в \mathcal{M}^k как h_k , $k = 1, \dots, C$. Отображение ψ является конкатенацией отображений ψ_k , $k = 1, \dots, C$. Запишем формально:

$$\psi := \psi_1 \oplus \dots \oplus \psi_C,$$

$$\psi_k : \mathbf{X} \rightarrow \mathbb{R}^{d_k}, \quad k = 1, \dots, C,$$

$$d = \sum_{k=1}^C d_k.$$

Для каждой категории $k \in \{1, \dots, C\}$ отображение ψ_k представимо в виде композиции двух отображений:

$$\psi_k := \pi_k \circ \mathbf{f}_k.$$

Здесь $\mathbf{f}_k : \mathbb{R}^{\tau \times X/k_s \times Y/k_s \times Z/k_s} \rightarrow \mathbb{R}^{h_k \times \tau}$ — применение маски активности \mathcal{M}^k , см. Определение 6. Отображение $\pi_k : \mathbb{R}^{h_k \times \tau} \rightarrow \mathbb{R}^{d_k}$ — проекция на риманово

касательное пространство с последующей векторизацией.

Подробнее остановимся на методе π_k трансформации признакового пространства в терминах римановой геометрии [? ?]. В работе [?] метод называется *Tangent Space Mapping*. Первым шагом данного алгоритма является формирование пространства центрированных признаков. Введем следующие обозначения:

$$\mathbf{Y}_j^k = \mathbf{f}_k(\mathring{\mathcal{A}}(\mathbf{X}_j)), \quad j = 1, \dots, N,$$

$$\mathbf{Y}_j^k = \begin{bmatrix} y_{11}^{kj} & y_{12}^{kj} & \dots & y_{1\tau}^{kj} \\ y_{21}^{kj} & y_{22}^{kj} & \dots & y_{2\tau}^{kj} \\ \vdots & \vdots & \ddots & \vdots \\ y_{h_k 1}^{kj} & y_{h_k 2}^{kj} & \dots & y_{h_k \tau}^{kj} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1^{kj} \\ \mathbf{y}_2^{kj} \\ \vdots \\ \mathbf{y}_{h_k}^{kj} \end{bmatrix}, \quad j = 1, \dots, N,$$

где \mathbf{y}_i^{kj} , $i = 1, \dots, h_k$ — центрированные временные ряды. Тогда ковариационная матрица для \mathbf{Y}_j^k — многомерного временного ряда:

$$\mathbf{R}_j^k = \frac{1}{\tau - 1} \mathbf{Y}_j^k \mathbf{Y}_j^{k\top}, \quad \mathbf{R}_j^k \in \mathbb{R}^{h_k \times h_k}, \quad j = 1, \dots, N.$$

Известно, что пространство, состоящее из матриц ковариации, представляет собой Риманово многообразие [?]. В каждой точке риманова многообразия имеется касательная плоскость с определенным на ней скалярным произведением. Общая касательная плоскость для проекции матриц ковариации в выборке формируется в точке среднего геометрического по римановой метрике известных ковариационных матриц. Среднее геометрическое симметричных положительно определенных матриц [?] имеет вид:

$$\mathbf{R}_k = \mathfrak{G}(\mathbf{R}_1^k, \dots, \mathbf{R}_N^k) = \arg \min_{\mathbf{R}} \sum_{j=1}^N \delta_R^2(\mathbf{R}, \mathbf{R}_j^k),$$

где риманова метрика или геодезическое расстояние определяется следующим образом:

$$\delta_R(\mathbf{R}_1, \mathbf{R}_2) = \|\log(\mathbf{R}_1^{-1} \mathbf{R}_2)\|_F = \sqrt{\sum_{i=1}^{3N} \log^2 \lambda_i}.$$

Здесь λ_i — собственные значения матрицы $\mathbf{R}_1^{-1} \mathbf{R}_2$. В работе [?] получено, что для

каждой ковариационной матрицы \mathbf{R}_j^k существует проекция $\boldsymbol{\pi}_j^k$ на касательное пространство. Кроме того, определено отображение:

$$\text{Exp}_R(\boldsymbol{\pi}_j^k) = \mathbf{R}_j^k = \mathbf{R}_k^{\frac{1}{2}} \exp \left(\mathbf{R}_k^{-\frac{1}{2}} \boldsymbol{\pi}_j^k \mathbf{R}_k^{-\frac{1}{2}} \right) \mathbf{R}_k^{\frac{1}{2}}$$

$$\text{Log}_R(\mathbf{R}_j^k) = \boldsymbol{\pi}_j^k = \mathbf{R}_k^{\frac{1}{2}} \log \left(\mathbf{R}_k^{-\frac{1}{2}} \mathbf{R}_j^k \mathbf{R}_k^{-\frac{1}{2}} \right) \mathbf{R}_k^{\frac{1}{2}}$$

После проектирования каждой ковариационной матрицы \mathbf{R}_j^k с помощью $\text{Log}_R(\cdot)$, получаем матрицы $\boldsymbol{\pi}_j^k \in \mathbb{R}^{h_k \times h_k}$. Последним шагом метода является векторизация. Процесс векторизации каждой матрицы $\boldsymbol{\pi}_j^k$ в пространство с евклидовой метрикой — последовательная запись элементов верхнетреугольной матрицы от $\boldsymbol{\pi}_j^k$, где диагональные элементы имеют коэффициент 1, а недиагональные — коэффициент $\sqrt{2}$. Получаем эмбеддинги:

$$\hat{\boldsymbol{\pi}}_j^k = \left[\boldsymbol{\pi}_j^k(1, 1) \ \sqrt{2}\boldsymbol{\pi}_j^k(1, 2) \ \dots \ \sqrt{2}\boldsymbol{\pi}_j^k(1, h_k) \ \boldsymbol{\pi}_j^k(2, 2) \ \dots \ \boldsymbol{\pi}_j^k(h_k, h_k) \right]^T,$$

$$\hat{\boldsymbol{\pi}}_j^k \in \mathbb{R}^{d_k}, \quad d_k = \frac{h_k(h_k + 1)}{2}.$$

Важно подчеркнуть, что при таком подходе извлекаются пространственные характеристики, которые демонстрируют степень соответствия объекта маске активности конкретного класса.

В качестве классификатора φ рассматриваются традиционные варианты, такие как логистическая регрессия и перцептрон с двумя скрытыми слоями.

4. Вычислительный эксперимент

Для анализа работоспособности предложенных методов и проверки гипотез проведен вычислительный эксперимент.

4.1. Восстановление снимков фМРТ по видеоряду

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

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

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 voxelам каждого снимка из тестовой выборки.

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

В Таблице ?? приведены технические характеристики компьютера, на котором производился вычислительный эксперимент.

Демонстрация работы метода. На Рис. ?? представлены срезы истинного и восстановленного снимков из тестовой выборки. На Рис. ??.??.?? изображена

Таблица 2: Технические характеристики компьютера

Элемент	Описание
Процессор	Intel Core i7-7700 3.6 GHz
Оперативная память	16 GB 2400 MHz
Видеокарта	NVIDIA GeForce GTX 1060 3 GB
Жесткий диск	M.2 SSD
Операционная система	Windows 10

разность между ними. Для демонстрации работы алгоритма был выбран 7-ой испытуемый, $\Delta t = 5$ с, коэффициент сжатия 1, коэффициент регуляризации $\alpha = 1000$. Рассмотрен 20-ый срез по первой координате 37-го снимка в последовательности. Так как значения вокселей нормированы на отрезок $[0; 1]$, то ошибка порядка 10^{-3} свидетельствует о достаточно точном предсказании.

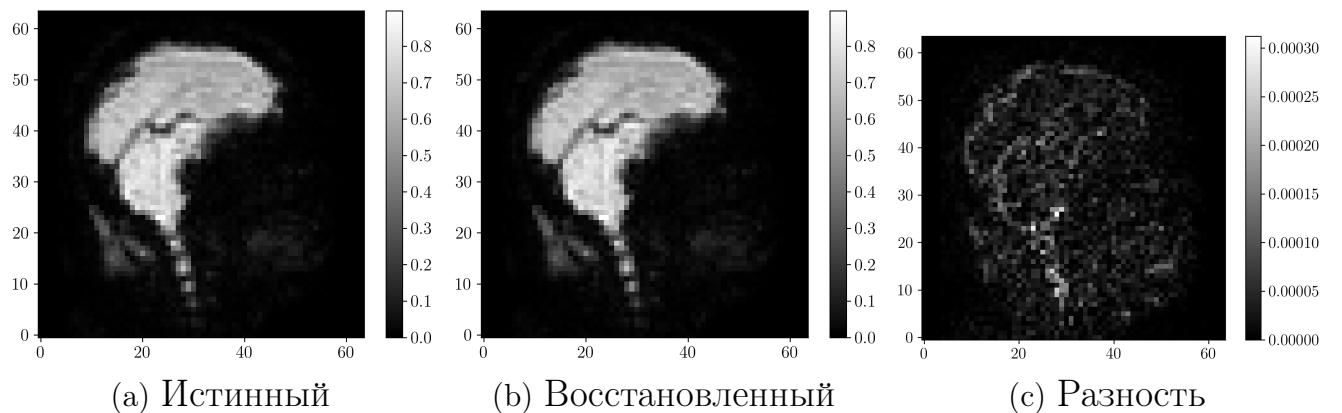


Рис. 3: Срез снимка фМРТ из тестовой выборки

Анализ времени задержки. Исследована зависимость качества восстановления от времени задержки. Для примера был выбран 47-ой испытуемый. На левом графике Рис. ?? представлена зависимость метрики MSE от времени задержки Δt . Исследование [?] 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 видно на Рис. ???. 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 абсолют-

ного 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. ??.

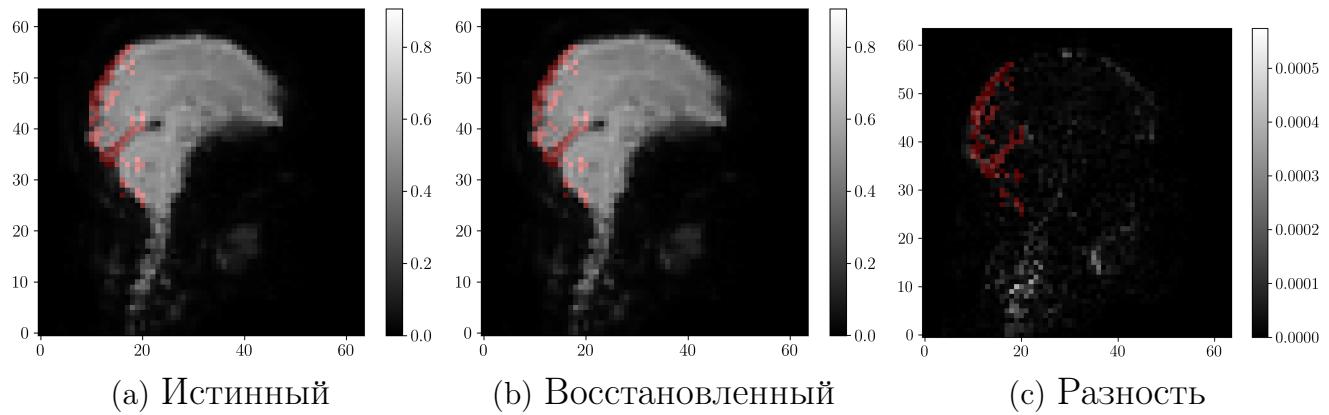


Рис. 4: Локализация наиболее активной зоны

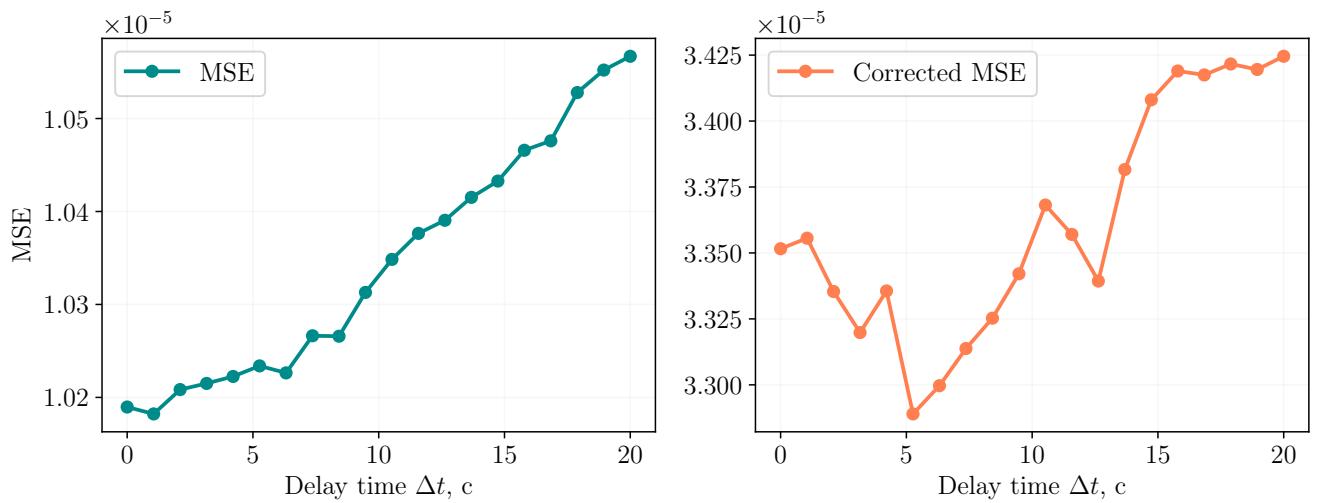


Рис. 5: Зависимость метрики MSE от времени задержки

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. ???. 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.

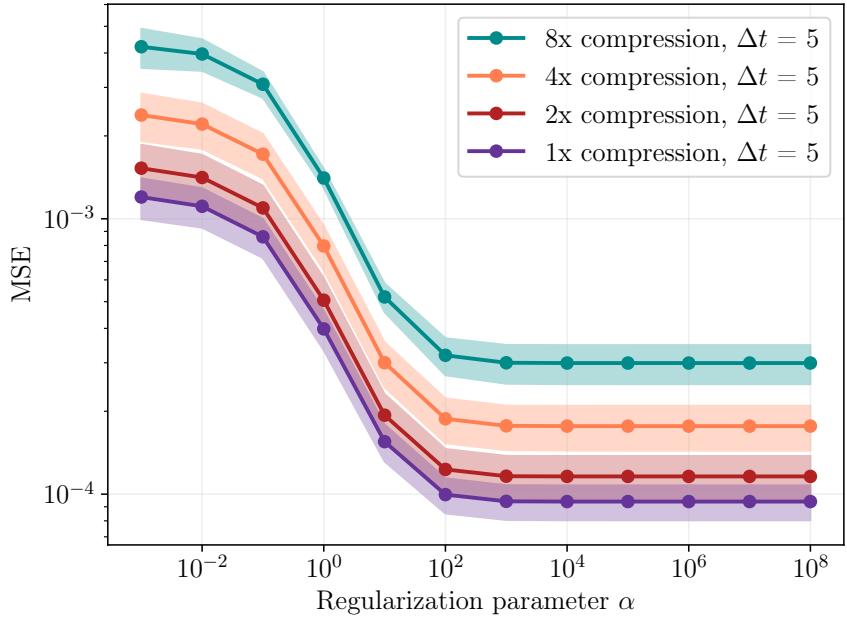


Рис. 6: Зависимость метрики MSE от коэффициента регуляризации α на снимках из тестовой выборки

Влияние коэффициента сжатия снимков на время работы метода. Производится сравнение времени обучения модели при использовании различных коэффициентов сжатия снимков фМРТ. Рассматриваются коэффициенты 1, 2, 4 и 8. Для каждого значения коэффициента сжатия подсчитывается среднее по всем испытуемым значение времени обучения модели. Вычисляется среднеквадратичное отклонение. Результаты эксперимента приведены в Таблице ???. Время работы метода существенно уменьшается при использовании предварительного сжатия снимков фМРТ. Эксперимент с подбором оптимального коэффициента регуляризации подтверждает то, что сжатие снимков не изменяет вид зависимостей.

Таблица 3: Зависимость времени обучения модели от коэффициента сжатия

Коэффициент сжатия	Среднее время, с	Среднеквадратичное отклонение, с
1	36.3	6.1
2	6.7	0.5
4	1.6	0.1
8	1.4	0.3

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. ??.

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.

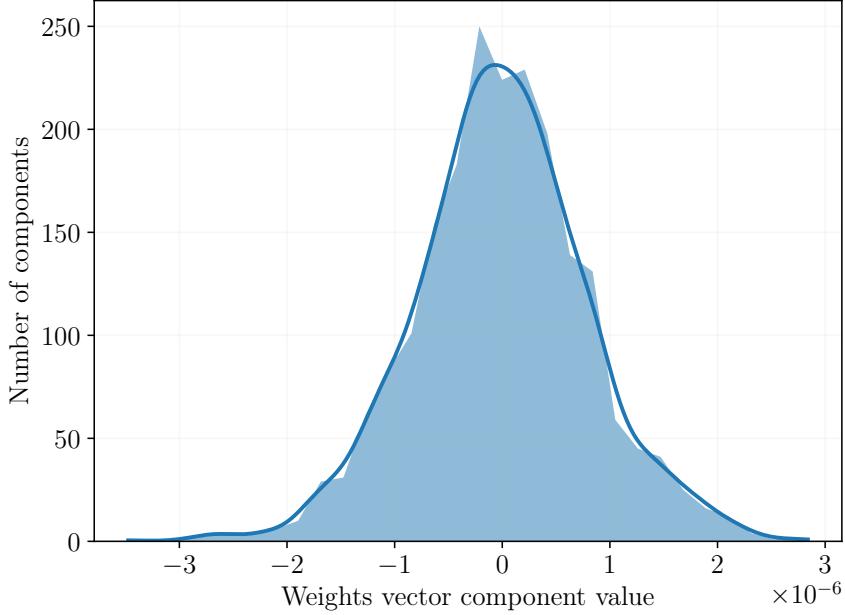


Рис. 7: Распределение значений компонент вектора весов

Гипотеза инвариантности весов модели относительно человека. Проведена проверка гипотезы инвариантности весов модели относительно человека: использование матрицы весов одного испытуемого для восстановления снимков фМРТ другого. Использовалась метрика MSE на тестовой выборке. Результаты представлены в Таблице ???. Рассмотрены 4-ый и 7-ый испытуемые. Матрица весов 4-го использовалась для восстановления снимков 7-го. Значения MSE практически совпадают.

Таблица 4: Проверка гипотезы инвариантности весов модели относительно человека

Матрица весов	Истинная	Подмешанная	Разность
MSE	$6.3912 \cdot 10^{-5}$	$6.3911 \cdot 10^{-5}$	$1.37 \cdot 10^{-9}$

Аналогичный эксперимент проведен для каждой пары испытуемых. Полученные результаты представлены на Рис. ???, 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. ??, в матрице есть и отрицательные значения. Тем не менее, они достаточно малы, а именно соответствуют отклонениям порядка 1%. Это объясняется тем, что модель достаточно простая, а потому обладает высокой обобщающей способностью. Однако это не мешает сделать вывод о том, что данные не противоречат гипотезе об инвариантности весов модели относительно человека.

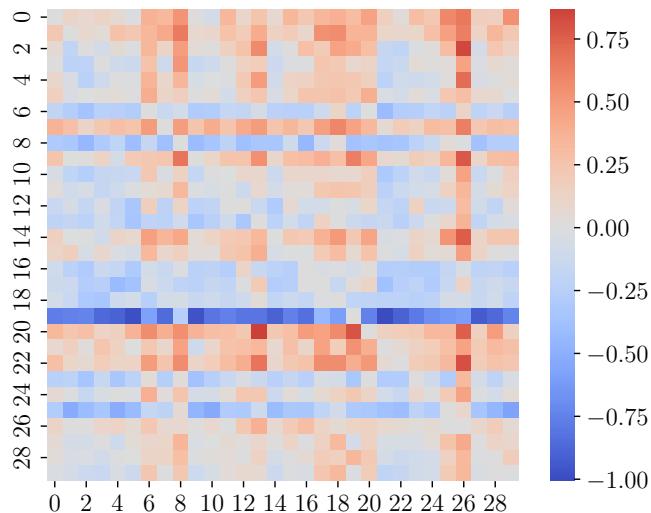


Рис. 8: МАРЕ изменения MSE при предсказании по подмешанной матрице весов (в процентах)

Корректность метода. Рассмотрено качество работы метода на неинформативных данных. В качестве матрицы объекты-признаки \mathbf{X} взята матрица, целиком состоящая из единиц. Произведено сравнение с результатами на настоящей матрице признакового описания. К первому снимку 35-го испытуемого последовательно прибавляются все восстановленные изменения значений вокселей. В результате имеем последний снимок последовательности. На Рис. ?? представлены срезы последнего истинного и восстановленного снимков из тестовой выборки. На Рис. ???.?? видна разность между ними. Результаты на неинформативных продемонстрированы на Рис. ?. Разность между истинным

и восстановленным снимками при работе с неинформативными данными значительно выше, что подтверждает наличие зависимости между показаниями датчиков и изображениями из видеоряда. Численные результаты приведены в Таблице ??.

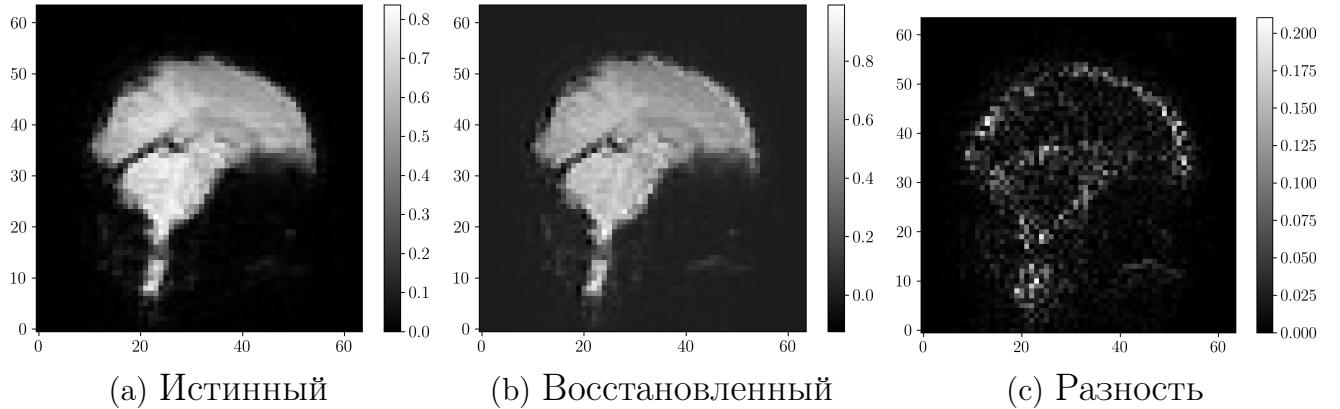


Рис. 9: Срез снимка фМРТ из тестовой выборки

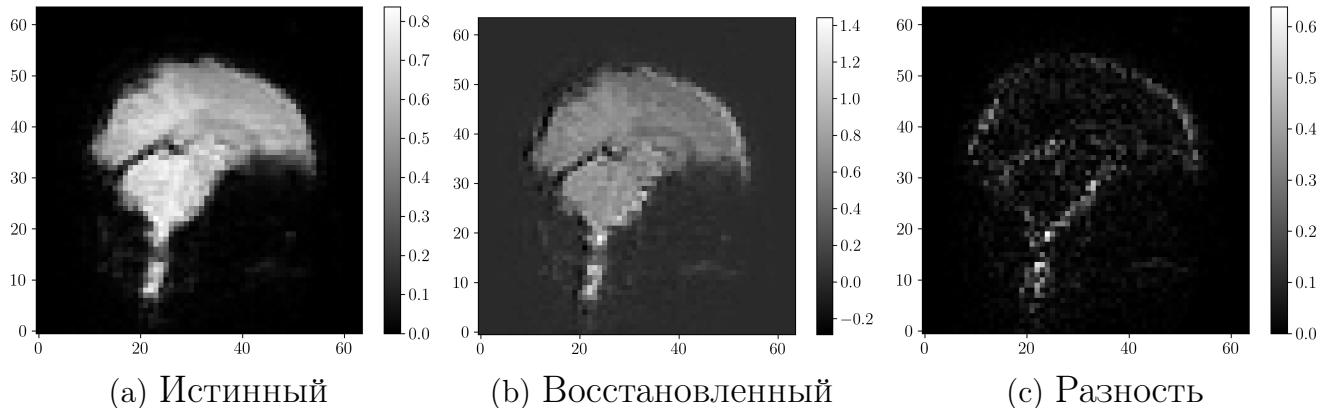


Рис. 10: Срез снимка фМРТ из тестовой выборки (на неинформативных данных)

Таблица 5: Качество работы метода на неинформативных данных

Выборка	Истинная	Неинформативные данные	Разность
MSE	$3.67 \cdot 10^{-4}$	$1.39 \cdot 10^{-3}$	$1.02 \cdot 10^{-3}$

4.2. Взвешивание voxелей фМРТ снимков

Для проведения вычислительного эксперимента используются данные фМРТ, собранные при изучении представлений лиц и объектов вентральной височной коре. Набор данных впервые использовался в исследовании [?] и

впоследствии стал известен как датасет Haxby. Данные широко используются в качестве сравнительной базы при изучении механизмов обработки визуальной информации в мозге.

Набор данных содержит показания фМРТ 6 испытуемых. Каждый испытуемый прошел 12 сеансов, в каждом из которых пассивно рассматривал черно-белые изображения восьми категорий, сгруппированные в блоки по 24 секунды, разделенные периодами отдыха. Каждое изображение демонстрировалось в течение 500 мс с последующим интервалом между стимулами 1500 мс. Данные фМРТ головного мозга были записаны с частотой 2.5 с^{-1} . Характеристики выборки приведены в Таблице ??.

Таблица 6: Описание выборки

Название	Обозначение	Значение
Количество снимков в сеансе	τ	121
Частота снимков фМРТ	μ	2.5 с^{-1}
Размерности снимка	X, Y, Z	40, 64, 64

Кроме того, набор данных содержит временные ряды стимулов и маски активных областей мозга, полученные нейробиологами. Эти области были определены для каждого участника эксперимента индивидуально. Маски активности были созданы путем объединения этих областей для всех участников эксперимента.

Демонстрация работы метода. Для демонстрации работы алгоритма был выбран 1-ый испытуемый, размер ядра $k_s = 4$, $h = 10$. Рассмотрен 28-ой срез по первой координате. Временной ряд стимула приведен в бинарный формат, где моменты наблюдения стимула отмечены единицей, а моменты отдыха — нулем. На Рис. ?? представлены срезы снимка из выборки с соответствующими масками активности. На Рис. ??..?? цветом отмечены области мозга, полученные методом *Cross-Correlation Weighting*. На Рис. ??..?? цветом показаны области мозга, полученные при статистическом анализе. Наконец, на Рис. ??..?? цветом представлена разметка нейробиологов.

На основе ?? экспериментально проверено, что корреляция областей мозга, полученных предложенным методом, со стимулом является статистически значимой. Кроме того, полученные области близки к разметке нейробиологов.

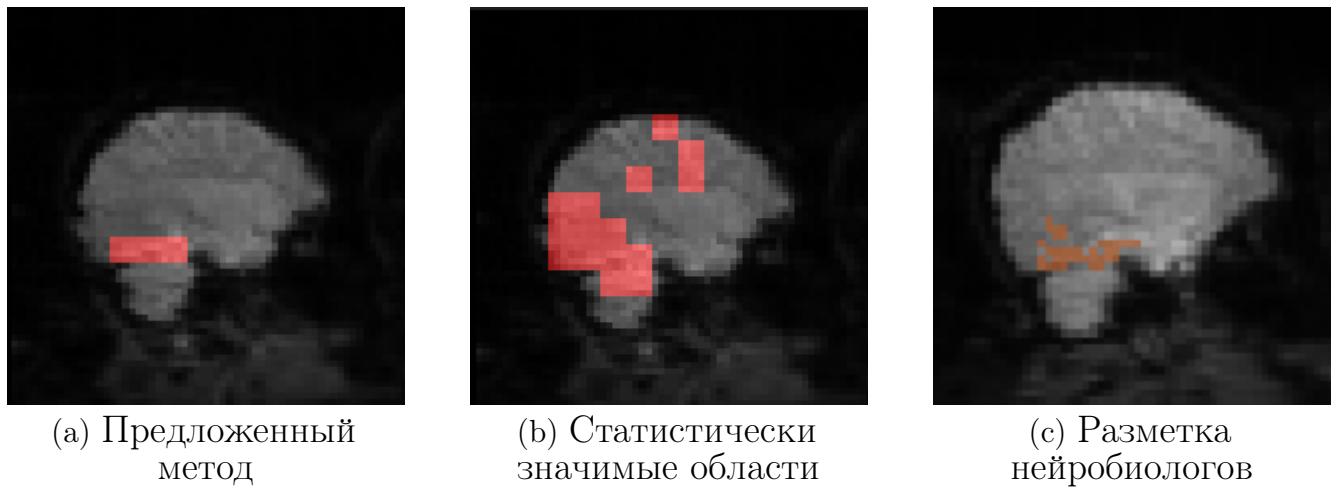


Рис. 11: Взвешенные области мозга на срезе фМРТ снимка

Корректность метода. Рассмотрено качество работы метода на неинформативных данных. Используется временной ряд фМРТ 1-го испытуемого и случайно сгенерированный бинарный ряд стимула. На Рис. ?? представлены срезы снимка фМРТ, цветом отмечены области мозга, предсказанные методом.

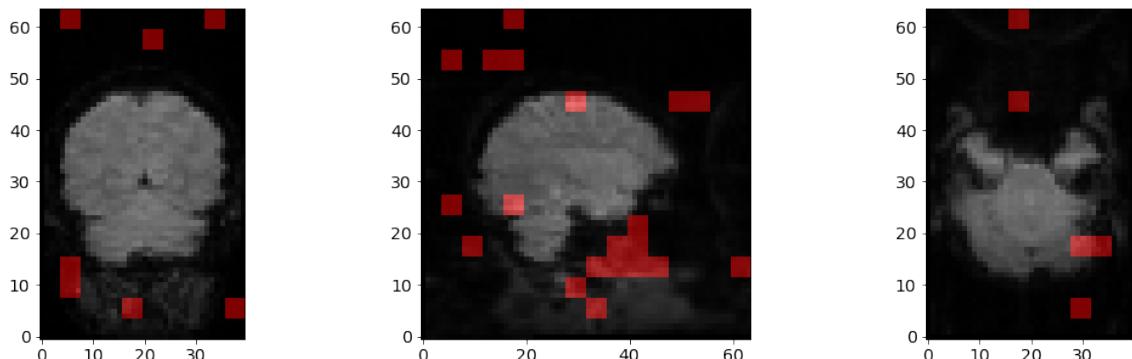


Рис. 12: Взвешенные области мозга, полученные по неинформативным данным

Метод предсказывает случайные области, не имеющие отношения к активности головного мозга. На основе ?? проверены гипотезы на значимость корреляции между выделенными областями и сгенерированным рядом стимула. Результаты статистического анализа не выявили значимости для какой-либо из полученных областей, что свидетельствует о наличии корреляции между активностью предсказанных областей мозга и истинным стимулированием.

4.3. Классификация сегментов временного ряда фМРТ

В рамках данного исследования проводится вычислительный эксперимент с использованием данных фМРТ Нахбы. Рассматриваются данные фМРТ первых трех испытуемых. Для получения сегментов ряда фМРТ был применен алгоритм сегментации, описанный в разделе ??

В Таблице ?? приведены параметры, рассчитанные для каждой из трех выборок, полученных в результате сегментации.

Таблица 7: Описание выборок, полученных в результате сегментации

Название	Обозначение	Значение
Число объектов	N	96
Длина каждого сегмента	τ	19
Число классов стимула	C	8
Число объектов на каждый класс	N_k	12

Произведено разделение каждой выборки на тренировочную и тестовую в соотношении 80% к 20% соответственно. Для оценки качества классификации сегментов фМРТ использовались метрики Accuracy, Macro-averaged F1 score и Micro-averaged F1 score.

В ходе экспериментов фиксировались следующие параметры предложенного метода: число выделяемых областей для каждого класса $h_k = 10$, $k = 1, \dots, C$, размер ядра $k_s = 4$. Использовалась логистическая регрессия со стратегией «один против всех» и единичным коэффициентом $l2$ -регуляризации. Кроме того, рассмотрен перцептрон, содержащий два скрытых слоя с размером, равным 100, и сигмоидной функцией активации.

В таблице ?? представлены усредненные значения метрик, полученные с помощью предложенного метода для двух испытуемых. Недостаточный объем

Таблица 8: Среднее арифметическое метрик по испытуемым

Классификатор	Accuracy	Macro F1 score	Micro F1 score
Логистическая регрессия	0.600	0.558	0.600
Перцептрон	0.700	0.636	0.636

выборки и большое число классов являются основными причинами переобучения и низких значений метрик. Тем не менее, у метода есть потенциал, на отдельных классах качество неплохое.

Для сравнения с предложенным методом классификации используются две упрощенные модели:

1. В первой модели отсутствует векторизация с помощью *Tangent Space Mapping*. После применения масок каждого класса временные ряды выпрямляются в вектор, конкатенируются и подаются на классификатор. Рассмотрены такие классификаторы, как логистическая регрессия и перцептрон.
2. Во второй модели отсутствует шаг с подсчетом масок активности для каждого класса стимула. Мaska подсчитывается аналогичным образом, но уже всреднем по всем временным рядам без учета категорий стимула. Кроме того, после применения маски используется векторизация с помощью *Tangent Space Mapping*. Рассмотрены аналогичные первой модели классификаторы.

В Таблице ?? представлены результаты работы трех методов на данных второго испытуемого. В качестве классификатора рассмотрена логистическая регрессия. После проведения анализа влияния различных компонентов модели на ее

Таблица 9: Влияние отдельных компонент метода на качество классификации

Метод	Accuracy	Macro F1 score	Micro F1 score
Предложенный	0.650	0.598	0.650
Без <i>Tangent Space Mapping</i>	0.150	0.121	0.150
Без масок активности	0.400	0.376	0.400

эффективность, было выявлено, что наибольшее снижение метрик происходит при исключении проекции на риманово касательное пространство. Кроме того, качество классификации также ухудшается, если не извлекаются маски активности для каждого класса стимула, а используется общая усредненная маска. Данные результаты свидетельствуют о том, что пространственно-временные характеристики, такие как проекция на риманово касательное пространство и маски активности, играют важную роль в задаче декодирования нейронной активности.

5. Заключение

В работе исследуются пространственно-временные характеристики в задаче декодирования временных рядов. Предлагается метод аппроксимации последовательности снимков фМРТ по видеоряду. Метод учитывает время гемодинамической ответной реакции — время задержки между изображениями из видеоряда и снимками фМРТ. Для каждого voxelя снимка фМРТ независимо строится линейная модель. 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. Подбирается коэффициент регуляризации. Исследуется влияние коэффициента сжатия снимков фМРТ на время обучения модели. Предполагается, что за информацию со зрительных органов отвечает затылочная доля мозга. Производится поправка MSE на основе локализации этой области и выбора наиболее изменяющихся voxelей. При таком построении график имеет характерный минимум, отвечающий оптимальному значению времени задержки. Полученное значение времени задержки согласуется с нейробиологическими сведениями. 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 со случайными данными.

Кроме того, предложен метод взвешивания активных областей мозга в задаче декодирования временных рядов фМРТ. Демонстрируется качество работы метода на реальных данных фМРТ. Полученные методом области близки к разметке нейробиологов. Проверяется статистическая значимость взвешенных voxelей. Корректность метода подтверждается экспериментом со случаемным рядом стимула. В работе предложен метод классификации сегментов временного ряда фМРТ. Данный метод учитывает пространственно-временные характеристики благодаря применению масок активности головного мозга и извлечению признаков с помощью метода *Tangent Space Mapping*, основанного на римановой геометрии. При проведении экспериментов было выявлено, что исключение отдельных компонент метода приводит к значительному снижению качества классификации. Это свидетельствует о том, что пространственно-временные

характеристики, извлекаемые данным методом, являются важными для достижения высокой точности классификации.