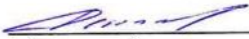


Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина»
Институт радиоэлектроники и информационных технологий – РТФ
Школа профессионального и академического образования

ДОПУСТИТЬ К ЗАЩИТЕ ПЕРЕД
ГЭК

Директор ШПиАО
 Д.В. Денисов
(подпись) (Ф.И.О.)
« 03 » июня 2024 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

АНАЛИЗ ГЕНЕРАТИВНЫХ МОДЕЛЕЙ ДЛЯ РАСПОЗНАВАНИЯ МОТОРИКИ
ПО СИГНАЛАМ ЭЛЕКТРОЭНЦЕФАЛОГРАФИИ

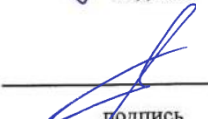
Научный руководитель: Ронкин Михаил Владимирович
к.т.н., доцент


подпись

Нормоконтролер: Огуренко Егор Владимирович


подпись

Студент группы: РИМ-220962 Толстых Андрей Андреевич


подпись

Екатеринбург
2024

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение высшего образования
**«Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина»**

Институт радиоэлектроники и информационных технологий – РТФ
Школа профессионального и академического образования
Направление подготовки 09.04.01 Информатика и вычислительная техника
Образовательная программа 09.04.01/33.03 Инженерия машинного обучения

ЗАДАНИЕ

на выполнение выпускной квалификационной работы

студента Толстых Андрея Андреевича группы РИМ-220962
(фамилия, имя, отчество)

1. Тема выпускной квалификационной работы Анализ генеративных моделей для распознавания моторики по сигналам электроэнцефалографии

Утверждена распоряжением по институту от «4» декабря 2023 г. № 33.02-05/298

2. Научный руководитель

Ронкин Михаил Владимирович, кандидат технических наук, доцент

(Ф.И.О., должность, ученая степень, ученое звание)

3. Исходные данные к работе

Техническое задание, научная статья

4. Перечень демонстрационных материалов

Презентация

5. Календарный план

№ п/п	Наименование этапов выполнения работы	Срок выполнения этапов работы	Отметка о выполнении
1.	<i>1 раздел (глава)</i>	до 23.03.2024 г.	✓
2.	<i>2 раздел (глава)</i>	до 29.04.2024 г.	✓
3.	<i>3 раздел (глава)</i>	до 20.05.2024 г.	✓
4.	<i>ВКР в целом</i>	до 24.05.2024 г.	✓

Научный руководитель Ронкин Михаил Владимирович

Ф.И.О.

(подпись)

Студент задание принял к исполнению 12.02.24

дата

(подпись)

6. Допустить Толстых Андрея Андреевича к защите выпускной квалификационной работы в экзаменационной комиссии

Директор ШПиАО

(подпись)

Д.В. Денисов

Ф.И.О.

РЕФЕРАТ

Объектом исследования являются сигналы электроэнцефалограммы, полученные неинвазивным методом. Предметом исследования – модели диффузионных нейронных сетей в задаче классификации моторики по сигналам электроэнцефалограммы. В качестве информационной базы исследования использованы: открытая база для построения интерфейсов «мозг-компьютер» openBCI и собственный набор данных.

Пояснительная записка к выпускной квалификационной работе состоит из 60 страниц, 25 рисунков, 57 источников.

Научная новизна работы заключается в анализе применимости моделей на основе диффузионных нейронных сетей для решения задачи распознавания моторики (реальной и воображаемой) по сигналам электроэнцефалограммы.

Целью работы является исследование эффективности моделей распознавания моторики (реальной и воображаемой) по сигналам электроэнцефалограммы на основе диффузионных нейронных сетей. Для достижения поставленной цели необходимо выполнить следующие задачи:

1. Литературный обзор методов распознавания моторики по сигналам электроэнцефалограммы.
2. Подготовка исследования, заключающаяся в обзоре математических моделей диффузионных нейронных сетей, первичном анализе данных и разработке плана постановки эксперимента.
3. Сравнительный анализ диффузионных нейросетевых архитектур для классификации моторики по сигналам ЭЭГ, представленных многопараметрическими временными рядами.

Ключевые слова: цифровая обработка сигналов, искусственные нейронные сети, обработка электроэнцефалограмм, интерфейс мозг-компьютер, диффузионные нейронные сети.

СОДЕРЖАНИЕ

Реферат	3
Определения, обозначения и сокращения	6
Введение.....	7
1. Обзор методов распознавания моторики по сигналам электроэнцефалограммы	11
1.1 Электроэнцефалография: общий обзор	11
1.2 Обзор методов классификации моторики по сигналам ЭЭГ	13
1.3 Диффузионные искусственные нейронные сети	16
1.3.1 Прямой диффузионный процесс	17
1.3.2 Обратный диффузионный процесс.....	19
1.3.3 Диффузионные модели условной генерации	23
1.3.4 Диффузионные модели повышения разрешения	25
1.3.5 Модели на основе оценки градиента вероятности целевого распределения.....	26
1.4 Диффузионные сети в классификации отрезков временных рядов	27
1.4.1 TimeGrad.....	28
1.4.2 ScoreGrad.....	28
1.4.3 D-Va	30
1.5 Выводы	30
2. Подготовка исследования	32
2.1 Первичный анализ наборов данных	32
2.1.1 Датасет EEG Motor Movement / Imagery.....	32
2.1.2 Собственный датасет	37
2.2 Постановка эксперимента	40

2.3 Выводы	44
3. Вычислительный эксперимент.....	45
3.1 Датасет EEG Motor Movement / Imagery	45
3.1.1 Дикторозависимые модели.....	45
3.1.2 Дикторозависимые модели.....	49
3.2 Собственный датасет	52
3.3 Выводы	54
Заключение.....	56
Список использованных источников.....	58

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

ИНС	Искусственная нейронная сеть.
ЭЭГ	Электроэнцефалография.
ДНС	Диффузионная нейронная сеть.
SOTA	State-of-the-Art, наиболее эффективный метод или модель.
MVR	Многопеременный временной ряд.
НКИ	Нейрокомпьютерный интерфейс, интерфейс «мозг-компьютер».
ПДП	Прямой диффузионный процесс.
ОДП	Обратный диффузионный процесс.
РНС	Реккурентная нейронная сеть.
ИВАЭ	Иерархический вариационный автоэнкодер
БПФ	Быстрое преобразование Фурье
ОПФ	Обратное преобразование Фурье

ВВЕДЕНИЕ

Успехи искусственных нейронных сетей (ИНС) в задачах распознавания образов повлияли на многие отрасли прикладной науки. Благодаря возможности аппроксимации нелинейных взаимосвязей на основе большого количества прецедентов, задачи, в настоящее время, формально не имеющие реализуемого алгоритмического решения перешли в область «стандартных задач» для нейросетей. Одним из перспективных направлений в области взаимодействия человека с техническими системами является разработка интерфейса «мозг-компьютер». Данное направление носит мультидисциплинарный характер ввиду того, что помимо распознавания сигналов от человека (решения задачи классификации), необходимо решить задачи снятия сигналов с мозга человека (задача конструирования) и первичной обработки этих сигналов (задача фильтрации). Практической реализации получения сигналов от мозга человека неинвазивными методами в области разработки интерфейсов «мозг-компьютер» является электроэнцефалография (ЭЭГ), так как необходимое оборудование может быть переносимым (в отличие от ближайшей альтернативы – томографии).

С другой стороны, развитие нейросетевых подходов привело к появлению диффузионных нейронных сетей (ДНС), которые наряду с сетями архитектуры «трансформер» являются наиболее эффективными (SOTA) в достаточно большом спектре задач. ДНС могут быть использованы для решения задач классификации сегментов многопараметрического временного ряда (МВР), которым является ЭЭГ.

Работа посвящена сравнительному анализу ДНС в задаче распознавания моторики по сигналам ЭЭГ. Под моторикой понимается три действия – сжатие в кулак правой руки; левой и расслабленное положение обеих рук. Действия могут быть как реальные, так и воображаемые (испытываемый только представляет необходимое действие, но не производит его физически).

Источником данных служит открытая база данных эксперимента EEG Motor Movement/Imagery [1] и собственная база данных¹.

Современные подходы к классификации моторики по сигналам ЭЭГ основываются на применении ИНС, преимущественно свёрточной архитектуры (свёрточные нейронные сети, СНС), входными данными для которых являются двумерные спектральные характеристики сигналов. Однако большинство работ по анализу сигналов ЭЭГ посвящено медицинской количественной оценке, которая необходима для диагностирования различных заболеваний, и в меньшей степени интерфейсам «мозг-компьютер». В связи с этим, применение передовых нейросетевых методов, таких как ДНС практически не исследовано, с точки зрения решения задач классификации. ДНС чаще используются для аугментации и генерации данных, что является перспективным направлением, однако его исследование выходит за рамки настоящей работы.

Целью работы является исследование эффективности моделей распознавания моторики (реальной и воображаемой) по сигналам электроэнцефалограммы на основе диффузионных нейронных сетей. Для достижения поставленной цели необходимо выполнить следующие задачи:

1. Литературный обзор методов распознавания моторики по сигналам электроэнцефалограммы.
2. Подготовка исследования, заключающаяся в обзоре математических моделей диффузионных нейронных сетей, первичном анализе данных и разработке плана постановки эксперимента.
3. Сравнительный анализ диффузионных нейросетевых архитектур для классификации моторики по сигналам ЭЭГ, представленных многопараметрическими временными рядами.

Объектом исследования являются сигналы электроэнцефалограммы, полученные неинвазивным методом.

¹ Данные получены до написания настоящей работы в соавторстве с Голубинским Андреем Николаевичем (ИППИ РАН им. А. А. Харкевича, Москва) и Журавлёвым Дмитрием Владимировичем (ВГТУ, Воронеж). Автор выражает благодарность за возможность использования полученных данных в работе.

Предметом исследования являются модели диффузионных нейронных сетей в задаче классификации моторики по сигналам электроэнцефалограммы.

Научная новизна исследования заключается в анализе применимости моделей на основе диффузионных нейронных сетей для решения задачи распознавания моторики (реальной и воображаемой) по сигналам электроэнцефалограммы.

Публикации. Результаты анализа современных подходов к построению и обучению диффузионных нейронных сетей отражены в материалах конференции «Радиолокация, навигация и связь» (2024).

Структура и содержание работы. Выпускная квалификационная работа состоит из введения, трех глав и списка литературы. Общий объем работы 60 страниц.

Во введении отражена актуальность рассматриваемых проблем, приводятся формулировки объекта, предмета, цели и задач исследования, описание структуры работы.

Первая глава посвящена теоретическому описанию данных электроэнцефалографии, анализу существующих подходов распознавания моторики на основе искусственных нейронных сетей, а также теоретическому описанию моделей диффузионных нейронных сетей.

Вторая глава содержит описание используемых датасетов, методику предобработки и разработку методики постановки вычислительного эксперимента.

Третья глава посвящена анализу результатов вычислительного эксперимента.

В заключении приведены основные результаты, выносимые на защиту выпускной квалификационной работы, проведен анализ и сравнение с другими методами и алгоритмами, также намечены направления для дальнейшего развития данного исследования.

Ограничения исследования. В рамках работы рассматривались три ДНС TimeGrad, ScoreGrad и DVa в качестве моделей выделения признаков

(*deep features*) временного ряда, задача классификации решалась с помощью полносвязных слоёв. Эксперименты проводились с несколькими наборами электродов (3 и 8 из 64), данное ограничение было обусловлено вычислительной сложностью обработки всех 64 каналов.

1. ОБЗОР МЕТОДОВ РАСПОЗНАВАНИЯ МОТОРИКИ ПО СИГНАЛАМ ЭЛЕКТРОЭНЦЕФАЛОГРАММЫ

Задача построения интерфейсов «мозг-компьютер» получила большой интерес в последнее время, ввиду широкого медийного освещения проекта neuralink [2], однако подобные проекты предполагают инвазивное взаимодействие (необходимо хирургическое вмешательство). Вместе с тем существует достаточно широкое поле применения неинвазивных методов построения интерфейсов «мозг-компьютер» (нейрокомпьютерный интерфейс, НКИ) с целью управления робототехническими или виртуальными системами. С точки зрения практической применимости, целесообразно использования неинвазивных ЭЭГ-датчиков, которые закрепляются на голове и снимают информацию через кожный покров. Использование других методик, например томографии, технологически нецелесообразно – стационарное оборудование налагает слишком строгие ограничения на практическое применение НКИ, построенных на его основе.

В главе рассмотрены вопросы снятия сигналов ЭЭГ, основные способы обработки, а также проведён анализ работ по распознаванию моторики на основе классификации сигналов ЭЭГ.

1.1 Электроэнцефалография: общий обзор

Электроэнцефалография – метод исследования головного мозга основанный на регистрации его электрических потенциалов [3]. Данный метод известен с 30х годов XX века и широко используется в различных разделах современных нейронаук, что позволяет выделить отдельные подвиды ЭЭГ: эпилептологическая; ЭЭГ критических состояний, физиологическая и др. [3]. В рамках настоящей работы рассматривается физиологическая ЭЭГ без семиотики – системы специальных терминов, описывающих,

классифицирующих и передающих информацию о паттернах суммарной биоэлектрической активности головного мозга [3].

Снятие и обработка сигналов ЭЭГ происходит с помощью электроэнцефалографа, который представляет собой многоканальное регистрирующее устройство, позволяющее регистрировать активность нескольких пар электродов. В настоящее время, практически все операции фильтрации и обработки сигналов выполняются в цифровом виде на компьютеризированных платформах.

Существует несколько схем расположения электродов на голове испытуемого: «10–20», «10–10» и др. В работе использовалась схема «10–10» (как в открытой базе OpenBCI [1] так и в собственной базе, детали реализации электроэнцефалографа приведены в [4]). На рисунке 1 приведена схема «10–10» для 64-канального электроэнцефалографа [1].

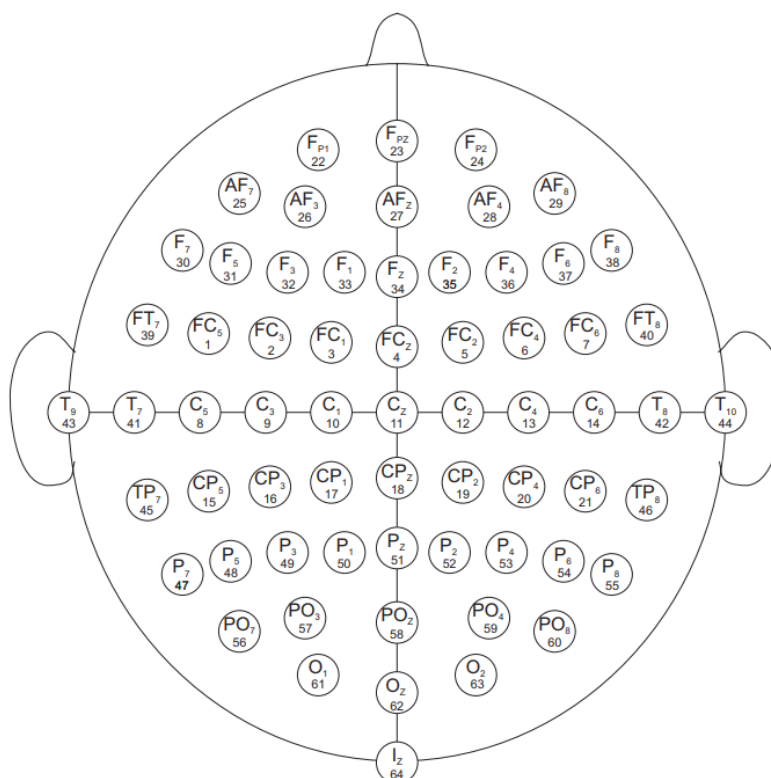


Рисунок 1. – Схема расположения электродов «10–10».

При соблюдении стандартных условий регистрации (закрытые глаза, состояние спокойного бодрствования, эмоциональный комфорт, отсутствие внешних раздражителей) у большинства (85–90%) взрослых здоровых людей в норме на ЭЭГ регистрируется альфа-паттерн, представляющий собой организованный альфа-ритм [5]. Альфа-ритм или альфа-волны – нормальная ритмическая активность ЭЭГ, регистрирующаяся в области первичных или вторичных зон сенсорной, которая угнетается при активации сенсорных областей коры. У здорового человека альфа-ритм является доминирующим в области теменной, затылочной, а также сенсомоторной коры. Обычно, альфа-ритм имеет частоту от 8 до 13 Гц и амплитуду обычно около 30–50 мкВ, блокируется или ослабляется при повышении внимания (в особенности зрительного) или мыслительной активности [5]. При формировании базы OpenBCI [1] авторы утверждают, что у испытуемых перед сбором данных не было выявлено отклонений. При сборе собственной базы были привлечены здоровые люди (без значимых медицинских отклонений). Таким образом, в работе используются данные, полученные регистрацией сигналов ЭЭГ взрослых здоровых людей.

1.2 Обзор методов классификации моторики по сигналам ЭЭГ

В разделе проведён краткий анализ работ, посвящённых распознаванию моторики по сигналам ЭЭГ. В большинстве работ предлагается использование СНС, однако существуют работы, рассматривающие иные архитектуры.

В работе [6] используется СНС для классификации 6 действий, которые выполняли 13 человек. В результате авторы заявляют о том, что добились точности классификации 92,6%, что значительно превосходит классические методы классификации двумерных тензоров (авторы использовали спектральное представление в качестве входных данных). Однако, полученные результаты не оценивают обобщающую способность СНС – авторы проводили эксперименты на замкнутой выборке, без разделения на обучающую,

валидационную и тестовую. Результаты на замкнутой выборке подтверждают факт универсальности ИНС в задачах аппроксимации, однако, подобные результаты не несут информации о практической применимости СНС в задачах распознавания моторики по сигналам ЭЭГ. Развитием работы [6] стала работа [7], в которой авторы предложили использование ансамбля СНС, улучшив тем самым результат [6] на 2% – до точности классификации в 94,6%. Однако, так же, как и в предыдущей работе авторы не проанализировали обобщающую способность сети, исключая тем самым вопросы практического применения их решения. Завершающей работой в серии стала работа [8], в которой авторы протестировали рекуррентные нейронные сети, немного повысив точность на замкнутой выборке, до 94,8%. Основным недостатком данной серии работ является ошибка в постановке эксперимента, а именно игнорирование вопросов степени обобщения построенных классификаторов. Данное замечание учтено при постановке эксперимента в настоящей работе.

В работе [9] авторы используют комбинированный подход к решению задачи классификации с использованием плоской ИНС (многослойного персептрона [10]) и вариационного автоэнкодера. Датасет представляет собой сигналы ЭЭГ, для 3 и 5 электродов (см. рисунок 1, электроды $\{C_3, C_z, C_4\}$ и $\{C_3, C_1, C_z, C_2, C_4\}$) с частотой дискретизации 250 Гц. Для применения ИНС авторы преобразовали сигналы ЭЭГ в спектральное представление и сгруппировали. На первом этапе был обучен вариационный автокодировщик без использования трюка параметризации [11], для векторизации входных данных, а затем обучение многослойного персептрона с 5 скрытыми слоями для итоговой классификации. При обработке результатов авторы использовали каппа-статистику [12] для оценки точности классификации и другие статистические критерии для сравнения своих результатов с другими работами. Авторы продемонстрировали методику оценки значимости результатов классификации и получили важный результат: рассмотренные ими (в том числе, предложенный) классификаторы статистически не отличаются по

точности классификации. Таким образом, целесообразно включить данную методику в постановку эксперимента.

Исследования возможности применения иных архитектур ИНС² демонстрируют снижение точности классификации. Авторы работы [13] предлагают использование сетей радиальных базисных функций [14] и многослойные персептроны с одним скрытым слоем, а также метод опорных векторов [10]. Точность классификации перечисленных методов не превышает 70% на тестовых данных. Однако авторы показали, что если использовать предварительную фильтрацию, то точность возрастает до уровня 90%.

Вопросам предобработки также посвящена работа [15], в которой предлагается механизм предобработки на основе матрицы Грамма и Марковской матрицы перехода. Авторы показали, что использование матрицы Грамма (по 8 испытуемым) повышает точность работы СНС до 90%, а предобработка в виде Марковской матрицы перехода – до 96%. Авторы близкой по предмету работы [16] исследуют вопрос предобработки с помощью вейвлет анализа. Сущность предобработки заключается в построении усреднённого коэффициента на основе вейвлет-энергии сигналов от каждого электрода. Однако в работе не было показано влияние подобной предобработки на точность классификации и сравнения с другими видами предобработки.

Наиболее подробно процесс построения двумерного тензора из сигналов ЭЭГ описан в работе [17]. Например, при использовании в качестве входных параметров спектров сигналов ЭЭГ целесообразно использовать шаг дискретизации в 0,5 Гц. Синтезированная СНС обладает точностью классификации не менее 95% по каждому из 5 классов действий. Особенностью используемой СНС является наличие остаточных связей (архитектура ResNet [18]) и достаточно большое количество весов – порядка

² Авторы всех предыдущих работ использовали СНС в том или ином виде.

2×10^6 (у рассмотренных ранее СНС порядок количества весов находится в пределах $1 \times 10^4 - 2 \times 10^5$).

Работы, использующие современные архитектуры, такие как [19], в основном решают задачу аугментации данных для дальнейшего обучения классических СНС. В [19] авторы добились практически полного соответствия точности классификации реальных и сгенерированных данных, полученных с помощью ДНС. Однако, исследования прямого применения ДНС к задаче распознавания моторики по сигналам ЭЭГ не проводилось.

1.3 Диффузионные искусственные нейронные сети

ДНС представляют собой передовой тип ИНС, основанный на итеративном удалении шума – декомпозиции процесса генерации данных (*sampling*) используя множество относительно небольших шагов удаления шума (*denoising*) [20]. В настоящее время ДНС широко применяются в задачах генерации изображения [21–24] и звука [25,26], также ведутся исследования по использованию ДНС в анализе временных рядов [27–29,29–38].

Концептуально, работа сети описывается следующим образом: на обучающие данные итеративно накладывается гауссовский шум, гиперпараметром сети служит число шагов T . После этого сеть итеративно производит удаление шума на t ($t < T$) обучающих шагах, используя в качестве корректного ответа зашумлённые данные с $t + 1$. Однако, базовая идея ДНС имеет достаточно много вариантов реализации, вследствие чего уместно ввести следующую классификацию ДНС:

1. Классические ДНС на основе обратного диффузионного процесса.
2. ДНС с условной генерацией.
3. ДНС на основе повышения разрешения генерируемых данных
4. ДНС с оценкой градиента вероятности целевого распределения.

Приведённая выше классификация типов ДНС получена автором в ходе анализа литературы. В разделе рассматривается математический аппарат, лежащий в основе диффузионных процессов, а также кратко описаны отличительные черты каждого из типов ДНС.

Следует отметить, что область исследований, посвящённых разработке и модификации ДНС находится в процессе становления и приведённая классификация должна восприниматься как «моментальный срез», актуальный во время написания настоящей работы.

1.3.1 Прямой диффузионный процесс

ДНС могут быть рассмотрены как модели с латентным пространством (*feature space*). Под латентным пространством³ понимается многомерное пространство, в которые могут быть спроецированы входные данные (а также возможна проекция из латентного пространства в пространство входных данных). С теоретической точки зрения, ДНС представляет собой марковскую цепь для T шагов [20]. Формализация с помощью марковской цепи подразумевает, что шаг t зависит только от шага $t - 1$ [39].

Пусть дана точка входного N -мерного пространства, распределение точек описывается $x_0 \in \mathbb{R}^N \sim \mathcal{M}(N, \mu_R, \Sigma_R)^4$, которая получена из распределения $q(x)$, тогда прямой диффузионный процесс (ПДП) определяется добавлением шума к x_0 на каждом из T шагов.

На каждом шаге $t \in [1, T]$ к x_{t-1} добавляется гауссовский шум со стандартным отклонением β_t , порождая латентную⁵ переменную x_t , распределение которой является условным $q(x_t|x_{t-1})$. Формально ПДП описывается следующим образом [20]:

$$q(x_t|x_{t-1}) = \mathcal{N}(N, \mu_t = \sqrt{1 - \beta_t}x_{t-1}, \Sigma_t = \beta_t I), \quad (1)$$

³ Под пространством здесь и далее подразумевается $\mathbb{R}^N$, где N – размерность пространства

⁴ Индекс R отражает принадлежность параметров μ и Σ к исходному распределению данных ($\mathcal{M}(\cdot)$)

⁵ Под латентной переменной понимается результат отображения входных данных в латентное пространство

где $\mathcal{N}(\cdot)$ – многомерный гауссовский шум с размерностью N ; μ_t – среднее значение; Σ_t – автоковариационная матрица; I – единичная матрица. Использование единичной матрицы в вычислении автоковариационной матрицы позволяет гарантировать, что распределение будет иметь стандартное отклонение β_t по всем измерениям [40]. Таким образом, распределение на выходе ПДП примет вид:

$$q(x_{1:T}|x_0) = \prod_{t=1}^T q(x_t|x_{t-1}), \quad (2)$$

где индекс $1:T$ при x подразумевает повторение с шага 1 до шага T , в [20] авторы назвали данную последовательность индексов «траекторией» (*trajectory*).

Однако, в таком подходе для шага t необходимо рассчитать все предыдущие шаги, что достаточно затратно с точки зрения вычислений. В практике генеративных нейронных сетей (в частности, вариационных автоэнкодеров) широко применяется так называемый «приём репараметризации» (*reparameterization trick*) [11]. Его суть заключается во введении случайности в процесс генерации: в нейронную сеть вводится слой, который на основе оценки среднего значения и стандартного отклонения производит выборку (*sampling*) из многомерного нормального распределения, которая используется как входной тензор для последующих слоёв сети. Данный подход был использован в [20]. В контексте ПДП приём репараметризации определяется следующим образом:

$$\begin{aligned} a_t &= 1 - \beta_t; \\ \bar{a}_t &= \prod_{s=1}^t a_s; \\ q(x_t|x_0) &= \mathcal{N}(N, \mu_t = \sqrt{\bar{a}_t}x_0, \Sigma_t = (1 - \bar{a}_t)I). \end{aligned} \quad (3)$$

Подробное теоретическое обоснование введения (3) приведено в [20]. Описанная выше реализация ПДП позволяет заранее вычислить $\bar{a}_t$ для всех $t \in [1, T]$, что значительно сокращает вычислительные затраты. Также следует отметить, что ПДП имеет 2 гиперпараметра: $B = \{\beta_1, \dots, \beta_T\}$ и T , в

оригинальной работе [20] было использовано линейное возрастающее множество вида $B = \{\beta_1 = 10^{-4}, \dots, \beta_T = 0.02\}$ (*linear schedule*).

Ещё одно улучшение ПДП было предложено в [41] – вид множества B , значения множества изменялись в соответствии с функцией косинуса (*cosine schedule*).

Таким образом, ПДП представляет собой марковский процесс, моделируемый с помощью выражения (3), параметризованный числом шагов T и множеством значений стандартного отклонения $B = \{\beta_1, \dots, \beta_T\}$.

1.3.2 Обратный диффузионный процесс

При $T \rightarrow \infty$ латентная переменная x_T близка к изотропному⁶ гауссовскому распределению. Задача обратного диффузионного процесса (ОДП) состоит в восстановлении распределения $q(x_{t-1}|x_t)$, для этого можно выбрать (*sample*) x_T из распределения $\mathcal{N}(N, 0, I)$, затем провести вычисление ОДП для $q(x_0)$, тем самым, получив новые данные из домена x_n ⁷.

Фактически, распределение $q(x_{t-1}|x_t)$ неизвестно, до оценки исходного распределения $\mathcal{M}(N, \mu_R, \Sigma_R)$. Однако, возможно произвести аппроксимацию $q(x_{t-1}|x_t)$ с помощью параметрической модели P_θ . С точки зрения практики построения параметризованных моделей P_θ , наиболее эффективными являются ДНС [11,20,41]. Учитывая, что $q(x_{t-1}|x_t)$ является гауссовским распределением, для достаточно малых β_t достаточно оценить среднее значение и стандартное отклонение в качестве θ [20]:

$$P_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}, \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)). \quad (4)$$

Если применить (4) к траектории $(p_\theta(x_{T:0}))$, можно аппроксимировать новые данные из домена x_n [20]:

⁶ Ковариационная матрица изотропного распределения описывается как $\Sigma = \sigma^2 I$, где I – единичная матрица

⁷ Под новыми данными из домена x_n понимается точка из исходного распределения $x_n \in \mathbb{R}^N \sim \mathcal{M}(N, \mu_R, \Sigma_R)$

$$P_{\theta}(x_{T:0}) = P_{\theta}(x_T) \prod_{t=T}^0 P_{\theta}(x_{t-1}|x_t). \quad (4)$$

Таким образом, необходимо оценить параметры ДНС – произвести обучение в терминологии ИНС. Для обучения ДНС нужно оптимизировать отрицательный логарифм функции максимального правдоподобия (*negative log-likelihood*) для обучающих примеров данных $X \in \mathbb{R}^N \sim \mathcal{M}(N, \mu_R, \Sigma_R)$. Набор обучающих данных $x_0^k, k \in [0, K]$ представляет собой датасет⁸ из K примеров. В работе [42] получена обобщённая форма нижней доказательной оценки (ELBO), являющаяся функцией потерь [10]. В общем случае ELBO для ДНС определяется как:

$$\begin{aligned}
& \log p(x) \geq \quad \square \\
& \geq \mathbb{E}_q(x_1|x_0)[\log P_\theta(x_0|x_1)] - \quad (I) \\
& \quad - D_{KL}(q(x_T|x_0), P_\theta(x_T)) - \quad (II) \\
& - \sum_{t=2}^T \mathbb{E}_q(x_t|x_0)[D_{KL}(q(x_{t-1}|x_t, x_0), P_\theta(x_{t-1}|x_t))] = \quad (III) \quad (5) \\
& = L_0 - L_T - \sum_{t=2}^T L_{t-1}, \quad \square
\end{aligned}$$

где, $\mathbb{E}_{q(\cdot)}$ – усреднение по распределению; $D_{KL}(\cdot)$ – дивергенция Кульбака-Лейблера между истинным распределением и распределением, восстановленными параметризованной моделью P_θ .

Рассмотрим подробнее смысл каждого члена в (5) [20]:

- (I) определяет среднюю ошибку реконструкции x_0 по x_1 , являющийся ELBO для вариационного автоэнкодера;
- (II) является мерой расстояния между x_T и гауссовским распределением, данный член игнорируется в процессе обучения;
- (III) определяет среднее расстояние между оценкой распределения с помощью ДНС и аппроксимацией ПДП на шаге t .

⁸ Здесь и далее употребляется терминология, принятая в области ИНС, основные термины даны в [10]

В работе [43] было показано, что необходимы дополнительные ограничения на x_0 для корректного построения процесса обучения ИНС. Оно заключается в следующем: используя (3) можно показать, что:

$$\begin{aligned} q(x_{t-1}|x_t, x_0) &= \mathcal{N}(N, \tilde{\mu}_t(x_t, x_0), \tilde{\beta}_t I); \\ \tilde{\beta}_t &= \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} \beta_t; \\ \tilde{\mu}_t(x_t, x_0) &= \frac{\sqrt{\bar{\alpha}_{t-1}} \beta_t}{1 - \bar{\alpha}_t} x_0 + \frac{\sqrt{\bar{\alpha}_t} (1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} x_t \end{aligned} \quad (6)$$

Исходя из (6) для корректного процесса обучения необходимо произвести приём репараметризации, налагающий ограничение на вид x_0 :

$$x_0 = \frac{1}{\sqrt{\bar{\alpha}_t}} (x_t - \sqrt{1 - \bar{\alpha}_t} \epsilon), \quad \epsilon \sim \mathcal{N}(N, 0, I) \quad (7)$$

Таким образом, вместо параметризованной модели $P_\theta(x_{t-1}|x_t)$ можно перейти к использованию $E_\theta(x_t, t)$, которая аппроксимирует ϵ , используя данный переход, с учётом (6) и (7) оценка среднего значения будет иметь вид:

$$\tilde{\mu}_t(x_t, t) = \frac{1}{\sqrt{\bar{\alpha}_t}} \left(x_t - \frac{\beta_t}{1 - \bar{\alpha}_t} E_\theta(x_t, t) \right). \quad (8)$$

Используя результат (8) член суммы (III) уравнения (5) – L_t , измениться следующим образом:

$$\begin{aligned} L_t &= \mathbb{E}_{x_0, t, E} \left[\frac{1}{2 \|\Sigma_\theta(x, t)\|_2^2} \|\tilde{\mu}_t - \mu_\theta(x_t, t)\|_2^2 \right] = \\ &= \mathbb{E}_{x_0, t, E} \left[\frac{\beta_t^2}{w \alpha_t (1 - \bar{\alpha}_t) \|\Sigma_\theta(x, t)\|_2^2} \|\epsilon_t - E_\theta(\sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, t)\|_2^2 \right]. \end{aligned} \quad (9)$$

Выражение (9) позволяет перейти от прогнозирования среднего значения распределения к прогнозированию шума ϵ на каждом шаге $t \in T$. В работе [1] показан метод упрощения практически используемой ошибки, игнорирующий коэффициент нормировки (первый множитель в формуле (9)). Использование упрощённой функции потерь позволяет повысить эффективность обучения, итоговая форма функции потерь выглядит следующим образом:

$$L_t^s = \mathbb{E}_{x_0, t, E} \left[\|\epsilon_t - E_\theta(\sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, t)\|_2^2 \right]. \quad (10)$$

Формальный вывод формул (9) и (10), их теоретическое обоснование и эмпирическая эффективность (10) приведены в работе [44]. Базовый алгоритм обучения ДНС приведен на рисунке 1 [20].

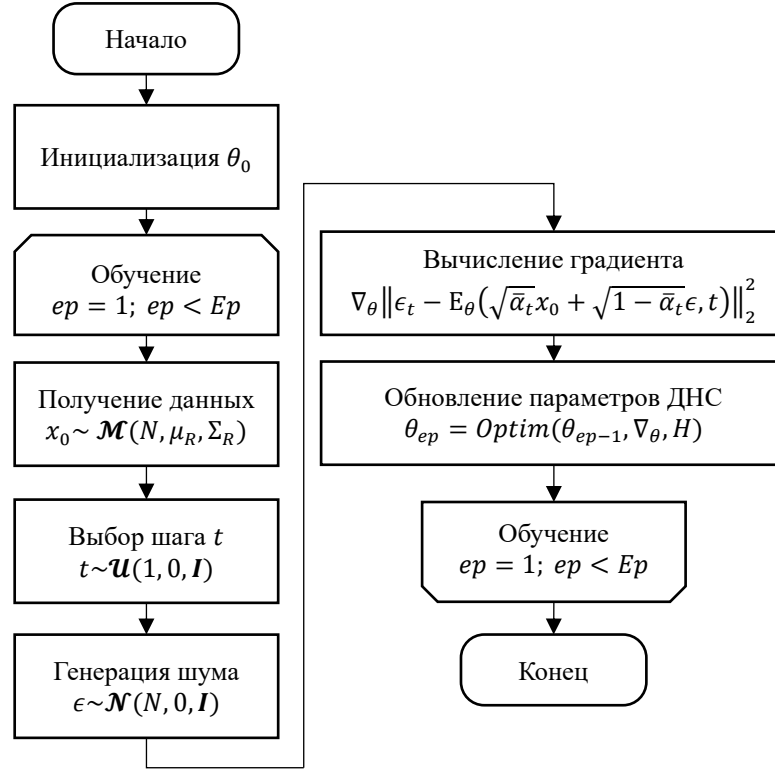


Рисунок 1 – Алгоритм обучения ДНС

ep – текущая эпоха обучения; Ep – количество эпох обучения;

$\mathcal{U}(\cdot)$ – равномерное распределение; $Optim$ – оптимизатор⁹;

H – параметры оптимизатора

На рисунке 2 приведён алгоритм генерации (*sampling*).

⁹ В настоящее время, де-факто, стандартом является подход Adam [45], подробнее о процессе оптимизации [10]

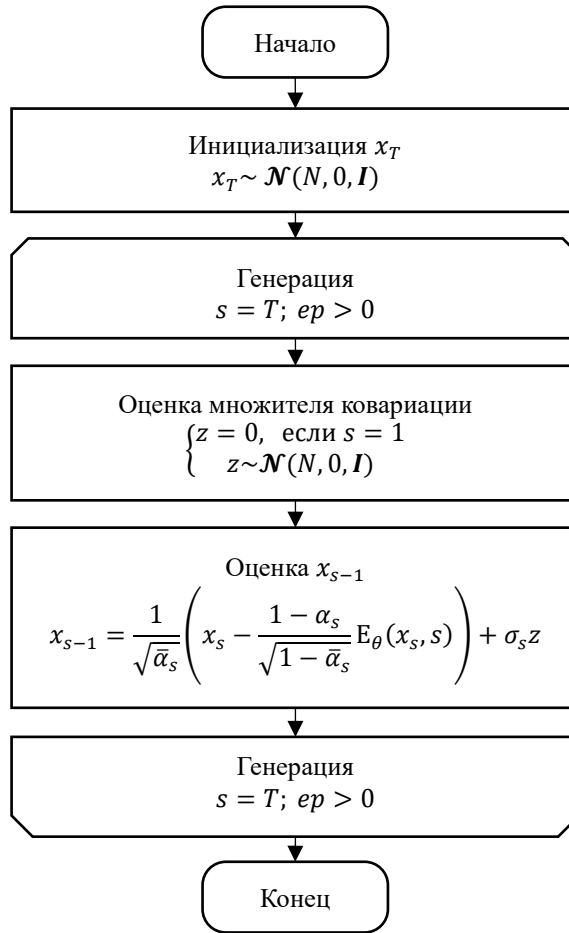


Рисунок 2 – Алгоритм генерации данных с помощью ДНС

s – текущий шаг генерации; z – множитель¹⁰ матрицы ковариации

1.3.3 Диффузионные модели условной генерации

Условная генерация необходима для внесения в модель дополнительных ограничений на генерацию: например, соответствие сгенерированных данных введённому текстовому описанию. Задача условной генерации формулируется следующим образом: на каждом шаге ПДП в модель добавляется условия генерации y^{11} , таким образом выражение (4) будет иметь вид:

$$P_\theta(x_{T:0}|y) = P_\theta(x_T) \prod_{t=T}^0 P_\theta(x_{t-1}|x_t, y). \quad (11)$$

¹⁰ Является матрицей размерностью $\mathbb{R}^N$

¹¹ Подразумевается векторизация (*encoding*) информации, например [46]

У данного типа ДНС функция потерь модифицируется с учётом информации о заданном условии y : $\log P_\theta(x_t|y)$ – информации о корректности генерации x_{t-1} и соответствии заданному условию y . Используя Байесовское правило [40] функцию потерь, можно представить как:

$$\log P_\theta(x_t|y) = \log P_\theta(x_t) + s \cdot \log P_\theta(y|x_t), \quad (12)$$

где s – скалярный множитель, подбираемый эмпирически, определяющий штраф за неточность соответствия заданным условиям y .

ДНС с условной генерацией можно разделить на два типа: с использованием дополнительной модели классификации (*classifier guidance diffusion*) и без её использования (*classifier-free guidance diffusion*) первые статьи по каждому из типов появились практически в одно время, и в дальнейшем развитие обоих типов ДНС с условной генерацией было тесно связано. Рассмотрим подробно оба типа ДНС с условной генерацией.

В [21] было предложен метод использования классификатора $f_\phi(y|x_t, t)$ для оценки класса y в ПДП и ОДП. Метод предполагает обучение классификатора $f_\phi(y|x_t, t)$ для предсказания класса y на зашумленных данных x_t , а затем, использование градиентов параметров модели $\nabla \log f_\phi(y|x_t, t)$ в ОДП. Используя обозначения из раздела 1.3.2 и выражение (12) оценка среднего значения ОДП будет иметь вид [21]:

$$\mu(x_t|y) = E_\theta(x_t|y) + s \cdot \Sigma_\theta(x_t|y) \nabla_{x_t} \log f_\phi(y|x_t, t). \quad (13)$$

Усовершенствованием данного подхода является использование CLIP-эмбеддингов для условной генерации [22]. Архитектура CLIP состоит из энкодера данных g и энкодера текста h [24]. Оценка среднего значения ОДП в подходе CLIP формулируется следующим образом:

$$\hat{\mu}(x_t|y) = E_\theta(x_t|y) + s \cdot \Sigma_\theta(x_t|y) \nabla_{x_t} g(x_t) \cdot h(c), \quad (14)$$

где c – текстовое описание (*prompt*), используемое в качестве условий генерации.

Тип условной генерации без использования дополнительного классификатора был предложен в [47]. Вместо обучения отдельной модели

классификации $f_\phi(y|x_t, t)$, предлагается обучать ДНС с условной генерацией $E_\theta(x_t|y)$ совместно с ДНС $E_\theta(x_t|0)$. Фактически, авторы в процессе обучения ДНС с условной генерацией случайно устанавливали метку класса $y = 0$, получая таким образом ДНС с условной генерацией и обычную ДНС:

$$\hat{\mu}(x_t|y) = s \cdot E_\theta(x_t|y) + (1 - s) \cdot E_\theta(x_t|0). \quad (15)$$

Данный подход имеет два преимущества: использование только одной модели для условной генерации и упрощение использования информации об условиях генерации¹².

1.3.4 Диффузионные модели повышения разрешения

Проблема использования описанных выше моделей заключается в высокой вычислительной сложности при большой размерности входных данных. Решение этой проблемы лежит в области увеличения разрешения исходной генерации $x_n \in \mathbb{R}^U \xrightarrow{\pi_\delta} x_n \in \mathbb{R}^N$, $U \ll N$, где U – размерность выхода ДНС; π_δ – параметрическая модель повышения разрешения данных. На данный момент существует два основных подхода: каскадная ДНС и латентная ДНС¹³. В [48] была предложена каскадная ДНС, для генерации изображений в высоком разрешении. Идея каскадной ДНС состоит в последовательном использовании множества ДНС, повышающих разрешение.

Работа [23] была посвящена латентным ДНС. Идея, лежащая в основе данного типа ДНС достаточно проста: перед применением ПДП к высокоразмерному входу $x_0 \in \mathbb{R}^N \sim \mathcal{M}(N, \mu_R, \Sigma_R)$, производится проецирование $x_0 \rightarrow \tilde{x}_0$ ($\tilde{x}_0 \in \mathbb{R}^M, M \ll N$). Процесс проецирования производится с помощью сети-энкодера¹⁴. ДНС применяется к данным в

¹² В контексте задачи условной генерации изображений по текстовому описанию достаточно сложно построить классификатор, предсказывающий текстовые эмбединги

¹³ Более известная как *Stable diffusion*

¹⁴ В общем случае сеть-энкодер преобразует входные данные в тензорное представление фиксированной размерности, концепция сетей-энкодеров пришла из VAE [11] (первая версия статьи – 2013 год)

низкой размерности $\widetilde{x}_0$, после чего применяется каскадная ДНС для повышения разрешения данных.

1.3.5 Модели на основе оценки градиента вероятности целевого распределения

Модели на основе оценки (*score-based models*) представляют собой семейство моделей использующие два блока: оценочный и блок динамического отжига Ланжевена.

Первый блок реализует функцию оценки Штена (*Stein score*) для функции распределения $q(x)$, определённую как $\nabla_x \log(q(x))$, которая имеет смысл векторного поля, указывающего в направлении, в котором логарифмическая вероятность целевого распределения растёт быстрее всего (градиент вероятности целевого распределения). Ввиду того, что истинное распределение вероятностей $q(x)$ неизвестно, данный блок аппроксимируется ИНС с параметрами γ (O_γ). В литературе встречаются два типа аппроксимации [49] и [50], отличающиеся нюансами реализации и обучения ИНС. Функция потерь ИНС для аппроксимации оценки Штена определяется как:

$$L_t^s = \mathbb{E}_{x_0, x_t, t} \left[\left\| O_\gamma(x_t, \sigma_t) - \nabla_{x_t} \log(q_t(x_t)) \right\|_2^2 \right]. \quad (16)$$

Используя предположение о нормальности распределений на каждом шаге траектории ПДП (16) преобразуется в:

$$L_t^s = \mathbb{E}_{x_0, x_t, t} \left[\pi_t \left\| O_\gamma(x_t, \sigma_t) + \frac{x_t - x_0}{\sigma_t^2} \right\|_2^2 \right], \quad (17)$$

где, π_t – константа, зависящая от стандартного отклонения σ_t ($\pi_t > 0$).

Второй блок представляет собой итеративный алгоритм выбора (*sampling*) образцов генерации с помощью метода Ланжувена-Монте-Карло [51] для каждого шага ОДП. Алгоритм выбора начинается с инициализации последовательности возрастающих σ_t и начинается в точке

$x_k \sim \mathcal{N}(x, 0, I)$, для каждого шага $t \in [T, \dots, 0]$ значение x_k преобразовывается F раз, согласно следующей формуле:

$$x_{t,f} = x_{t,f-1} + \frac{1}{2}\eta_t O_\gamma(x_{t,f-1}, \sigma_t) + \sqrt{\eta_t}\omega, \quad (18)$$

где, индекс $f \in [0, F]$ – показывает текущий шаг преобразования; η_t – константа скорости преобразования; $\omega \sim \mathcal{N}(1, 0, I)$. После проведения внутреннего цикла (по f) полученные данные $x_{t,f}$ будут являться исходными для следующего цикла $x_{t,f} = x_{t+1,0}$. Результатом генерации являются данные, полученные после T шагов описанного ОДП – $x_{0,f}$. Введение дополнительной случайности посредством ω необходимо для того, чтобы ДНС не генерировала образцы данных, в точности соответствующие образцам из обучающих данных.

Модификация ПДП в данном подходе состоит в том, что на каждом шаге ПДП (3) выполняется оценка Штена для $q_{\sigma_t}(x_t)$. После получения данной оценки производится выбор образцов генерации с помощью метода Ланжувена-Монте-Карло, постепенно смещая случайный гауссовский шум к областям с высокой плотностью целевого распределения, по направлениям, полученным на шаге оценки.

1.4 Диффузионные сети в классификации отрезков временных рядов

Область исследований ДНС в настоящее время бурно развивается, в связи с этим существует достаточно большое количество ДНС, применяемых для классификации отрезков временных рядов. В работе [52] приведен достаточно подробный обзор современных ДНС для решения задач анализа временных рядов. Настоящая работа сконцентрирована на исследовании конкретных решений: TimeGrad [27], ScoreGrad [28] и D-Va [53]. Выбор данных ДНС обусловлен их высокой точностью в задачах классификации и новизной. Подраздел посвящён краткому описанию особенностей архитектур выбранных ДНС.

1.4.1 TimeGrad

Модель TimeGrad является авторегрессивной моделью для вероятностного предсказания MBP [27]. По приведённой в разделе 1.3 классификации относится к латентным ДНС с условной генерацией. Основной особенностью TimeGrad является использование рекуррентной нейронной сети (РНС) в качестве параметрической модели P_θ , а скрытое состояние РНС h_t [10] является условием генерации: $P_\theta(x_{T:0}) = P_\theta(x_T) \prod_{t=T}^0 P_\theta(x_{t-1}|h_{t-1})$.

Рассмотрим подробнее особенности архитектуры P_θ . Она представляет собой 2 слоя LSTM [10] с размерностью скрытого состояния $h_t \in \mathbb{R}^{40}$, позиционные индексы кодируются с помощью трансформерной архитектуры [54]. Для выделения признаков из входного сигнала используется WaveNet [55] и DiffWave [56].

Авторы [27] рекомендуют использовать следующие параметры обучения:

- оптимизатор: Adam [45] (темп обучения¹⁵ 1×10^{-3});
- число шагов ПДП (T): 200;
- вид множества β (раздел 1.3.1): линейно возрастающее $\{10^{-4}, \dots, 0.1\}$;
- размер батча: 64;
- метод выбора батча: случайное окно с перекрытием.

Авторы [27] провели тестирование разработанной архитектуры на 6 наборах данных, в 5 из которых предложенная ДНС показала меньшую (статистически значимо) ошибку регрессии.

1.4.2 ScoreGrad

Модель ScoreGrad [28] относится к моделям на основе оценки градиента целевого распределения и использует в своей основе математический аппарат

¹⁵ Learning rate.

стохастических дифференциальных уравнений. Авторы разделяю архитектуру на два блока:

1. Блок выделения признаков из МВР, на основе РНС.
2. Блок условной оценки на основе стохастических дифференциальных уравнений, аппроксимированный с помощью ИНС.

Авторы вдохновлялись WaveNet [55] и DiffWave [56], их архитектура выделения признаков фактически повторяет аналогичный блок TimeGrad с незначительными изменениями.

Второй блок аппроксимируется с помощью условной ДНС. Фактически рассмотрение стохастических дифференциальных уравнений сводится к использованию условной ДНС, основная формула, описывающая стохастические дифференциальные уравнения для решения задачи регрессии МВР из работы [28] – (16)¹⁶ совпадает (по существу) с формулой (5) из раздела 1.3.2 настоящей работы.

Авторы [28] использовали тот же набор датасетов что и авторы TimeGrad [28], показав меньшую ошибку регрессии МВР по сравнению с TimeGrad, однако прирост метрик не превышает $7 \pm 1\%$.

Таким образом, основное отличие TimeGrad и ScoreGrad в методике обучения ДНС. Авторы [28] рекомендуют использовать следующие параметры обучения:

- горизонт предсказания: 10, 20, 40 и 60 отсчётов (обучение и инференс);
- оптимизатор: Adam [45] (темп обучения¹⁷ 5×10^{-4});
- размер батча: 16
- длительность обучения: 20 эпох на инструмент (один из 110).

¹⁶ Нумерация по оригинальной статье <https://arxiv.org/pdf/2106.10121.pdf>

¹⁷ Learning rate.

1.4.3 D-Va

Модель D-Va [53] представляет собой гибридную модель, использующую ДНС и иерархический вариационный автоэнкодер (ИВАЭ) [57]. Основная идея, стоящая за архитектурой D-Va следующая: на каждом шаге ПДП полученные элементы x_t формируются с помощью ИВАЭ с фиксированными весами (веса ИВАЭ не изменяются для каждого $t \in [1, T]$). ДНС, лежащая в основе D-Va, относится к ДНС с условной генерацией.

Целевым набором данных для оценки эффективности D-Va выступали данные о 110 финансовых инструментах в различные периоды времени (2014–2016; 2017–2019; 2020–2022). Сравнение производилось с 5 моделями предсказаний поведения рынка (среди которых не было ДНС). Авторы [53] показали, что D-Va (статистически значимо) более качественно, с точки зрения метрик, решает задачу регрессии по данным инструментам по всем наборам данных.

1.5 Выводы

В разделе проведён анализ современного состояния систем распознавания моторики по сигналам ЭЭГ с помощью аппарата ИНС, а также детально разобран аппарат ДНС. На основе проведённого анализа можно сделать следующие выводы:

1. Исследователи в основном используют классические нейросетевые архитектуры, в связи с чем актуально исследование эффективности перспективных архитектур ИНС.

2. Практически все исследования строятся вокруг спектрального представления данных, их фильтрации и предобработки, что гипотетически может быть заменено сплошной (*end-to-end*) архитектурой.

3. Применяемые сети имеют относительно небольшую внутреннюю размерность, что позволяет сделать предположение о их применимости только в дикторозависимых¹⁸ задачах.

4. Современные архитектуры, такие как ДНС используются только для генерации или аугментации данных, но не для непосредственно классификации.

Исходя из вышеизложенных пунктов, целесообразно исследовать применение ДНС в задаче классификации моторики по сигналам ЭЭГ. Вместе с тем, целесообразно учесть при разработке плана постановки эксперимента следующие моменты:

1. Провести первичный анализ открытого и собственного датасета, сформировать единые критерии: набор действий и параметры сигналов (частота дискретизации, фильтрация и т. д.).

2. Использовать подвыборки электродов, определить оптимальное количество электродов для каждого из датасетов.

3. Обрабатывать результаты в соответствии с методикой из [9], заключающейся в децимации и фильтрации сигналов с датчиков. В ходе вычислительного эксперимента целесообразно проанализировать влияние фильтрации на качество классификации.

4. Сравнить результаты для дикторозависимых и дикторонезависимых наборов данных.

¹⁸ Используются термины из области обработки речевых сигналов. Под дикторозависимостью автор понимает возможность корректной работы системы классификации только в рамках данных, полученных от одного человека. Дикторонезависимость подразумевает корректность работы системы на всех данных.

2. ПОДГОТОВКА ИССЛЕДОВАНИЯ

Для построения классификаторов моторной активности по сигналам ЭЭГ необходимо провести первичный анализ данных в датасетах, определить подгруппы используемых электродов и наборов действий. После проведения первичного анализа данных необходимо сформировать обучающие, валидационные и тестовые выборки, определить метрики унифицировать методику оценки результатов.

2.1 Первичный анализ наборов данных

Первичный анализ данных состоит в описании датасетов, выборе метода предобработки и выделения моментов времени, в которые начинаются интересующие события (сжатие правой или левой руки, воображение сжатия правой или левой руки), а также длительность событий.

2.1.1 Датасет EEG Motor Movement / Imagery

Датасет [1] представляет собой записи сигналов ЭЭГ для 109 испытуемых с помощью 64 электродов. Данные представлены в виде edf-файлов, содержащих отсчёты сигнала с частотой дискретизации 160 Гц. На рисунке 3 приведены отсчёты для датчика Cz (легенда расположения электродов – рисунок 1, раздел 1).

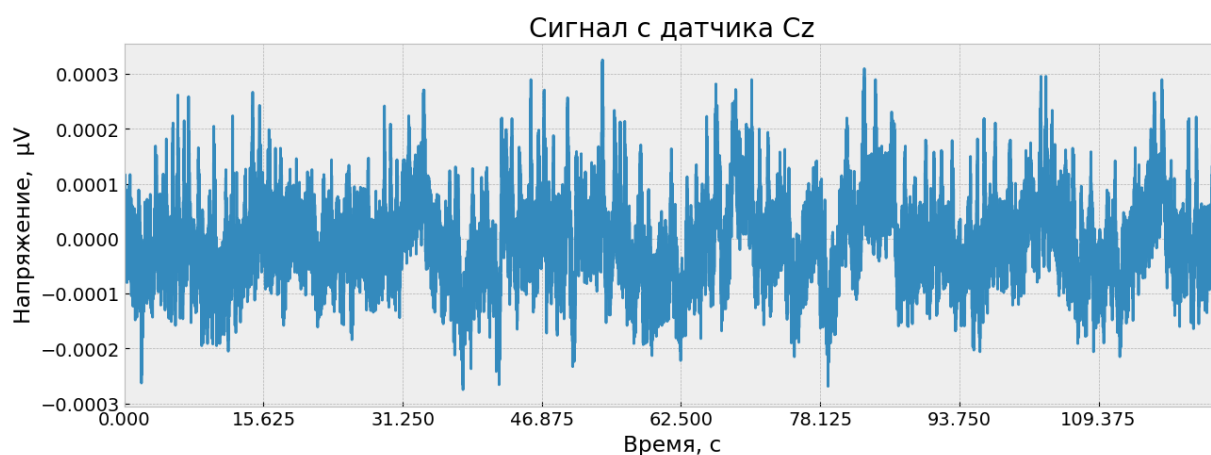


Рисунок 3 – Отсчёты записанного сигнала с датчика ЭЭГ

Для пердобработки предлагается использовать следующую последовательность преобразований:

1. Быстрое преобразование Фурье (БПФ);
2. Фильтр Баттерворта с параметрами $f_H = 5$ Гц, $f_B = 40$ Гц, порядок 5;
3. Децимация (порядок 2).

Подобные преобразования позволяют снизить влияние шумов датчиков. На рисунке 4 приведён предобработанный и исходный сигналы для датчика Cz.



Рисунок 4 – Соотнесение исходного и фильтрованного сигнала датчика Cz

Расположение и значение сигнала (усреднённое и нормированное) для датчиков приведено на рисунке 5.

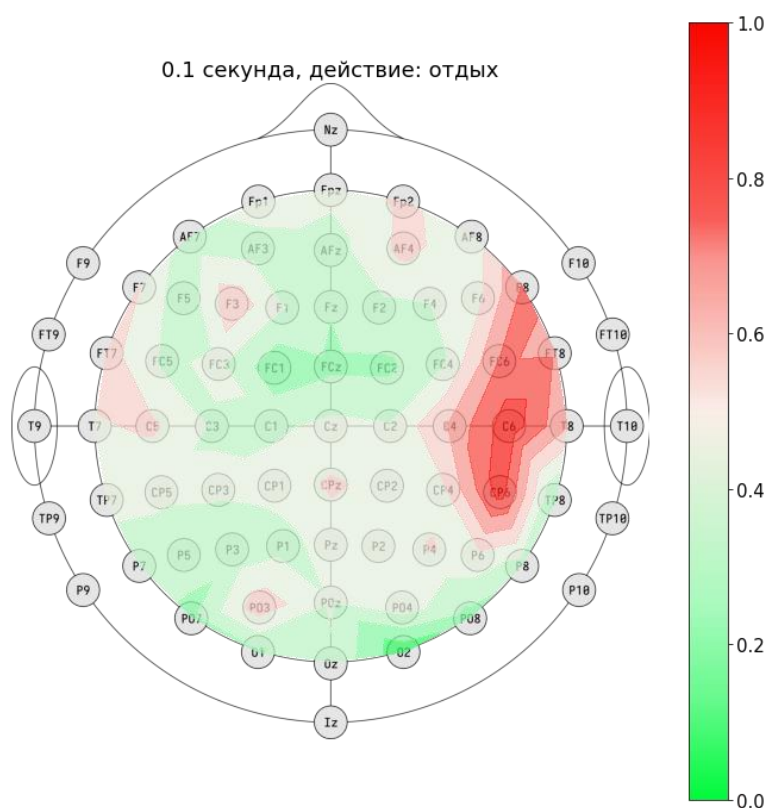


Рисунок 5 – Расположение и значение сигнала для всех датчиков

Использование информации со всех датчиков для обучения ДНС – вычислительно сложная задача. Ввиду этого, была выбрана подгруппа электродов $\{C3, C4, Cz\}$, которая наиболее часто используется в публикациях по распознаванию моторики [6,7,9]. Использование ограниченной подгруппы позволит произвести обучение ДНС на пользовательском GPU. Помимо этого, данные из собственного датасета получены с помощью 9 датчиков, что позволяет использовать одну и ту же подгруппу в экспериментах. Второй подгруппой выбран набор $\{FP1, FP2, C3, C4, P7, P8, O1, O2\}$ электродов, включающий все датчики из собственного датасета. Таким образом произведён выбор 2 подгрупп для сравнения эффективности ДНС в зависимости от количества используемых электродов. Учитывая, что собственный датасет не содержит опорного электрода Cz (вместо него

используется два опорных электрода на ушах), подгруппа $\{C3, C4, Cz\}$ при обучении на собственном датасета использована не будет.

Визуализация паттернов для различных действий (для одной записи) приведена на рисунке 6.

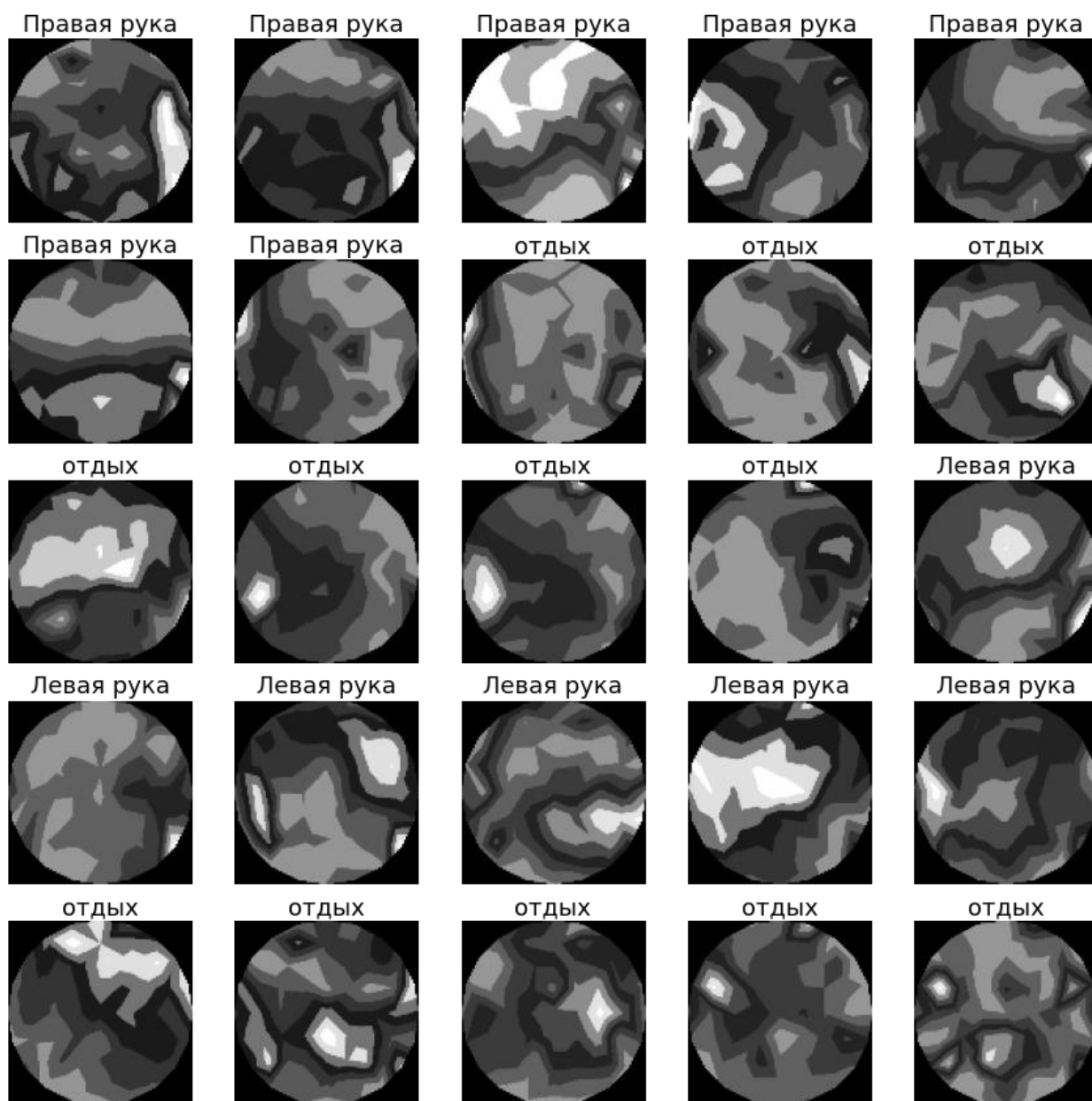


Рисунок 6 – Визуализация паттернов для различных действий для одной записи

Датасет имеет следующую нотацию: S<число> – идентификатор испытуемого, R<число> – идентификатор эксперимента, всего существует 14 типов экспериментов, в рамках работы интересуют эксперименты {3,4,7,8,11,12}, в которых происходит запись для сжатия и разжимания правого и левого кулака (реальное и воображаемое). Например, запись S001R03 соответствует эксперименту с реальным сжатием кулаков для испытуемого №1. Реальная моторика регистрировалась в экспериментах {3,7,11}, воображаемая – {4,8,12}. На рисунке 7 приведена визуализация периодов активности для испытуемого S001 (электрод Cz, эксперимент R03).

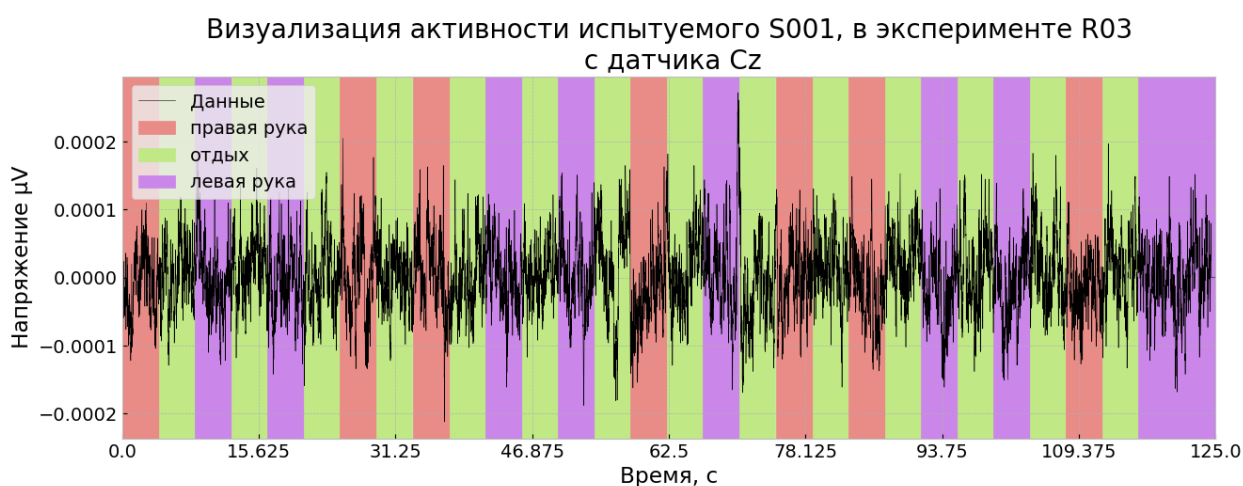


Рисунок 7 – Визуализация активности испытуемого S001, в эксперименте R03
с датчика Cz

Таким образом, из датасета EEG Motor Movement / Imagery сформировано 8 наборов данных:

1. Воображаемые движения, без фильтрации, 3 датчика.
2. Воображаемые движения, без фильтрации, 8 датчиков.
3. Воображаемые движения, с фильтрацией, 3 датчика.
4. Воображаемые движения, с фильтрацией, 8 датчиков.
5. Реальные движения, без фильтрации, 3 датчика.
6. Реальные движения, без фильтрации, 8 датчиков.
7. Реальные движения, с фильтрацией, 3 датчика.

8. Реальные движения, с фильтрацией, 8 датчиков.

Для удобства использования датасет был сохранен в единую иерархичную структуру данных, описываемая следующей схемой: «тип движения» – «тип фильтрации» – «номер испытуемого» – «номер испытания» – данные (все 64 канала).

Рассмотрим распределения длительностей активности для различных испытуемых и различных экспериментов. На рисунке 7 приведены сводные статистики по всему набору данных.

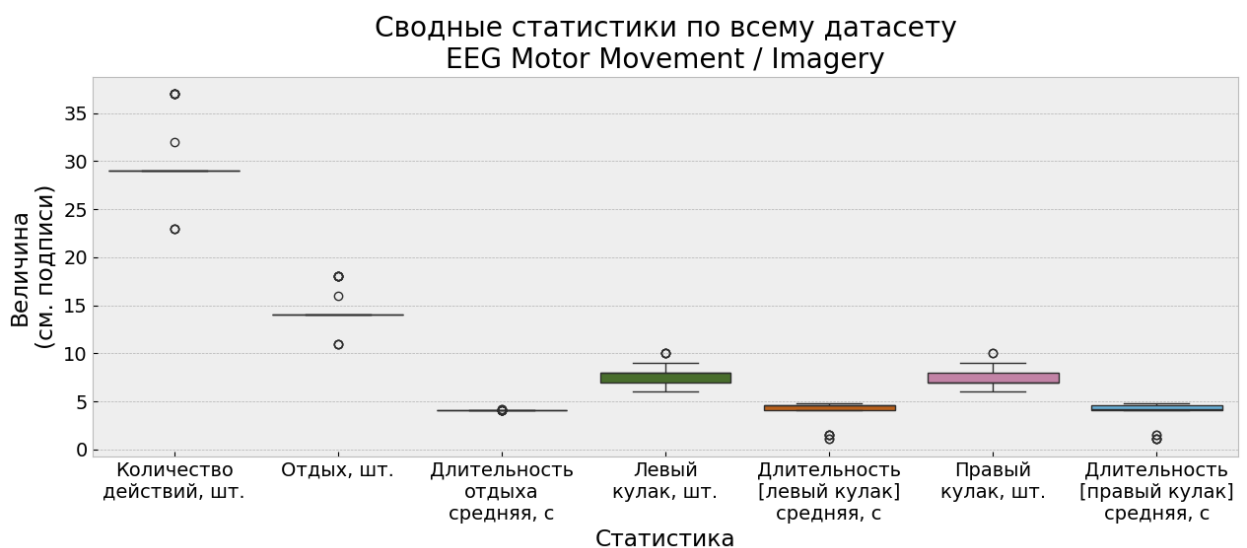


Рисунок 8 – Сводные статистики по всему датасету EEG Motor Movement / Imagery

Из рисунка 8 видно, что сбор данных происходил достаточно аккуратно с точки зрения контроля времени записи и количества периодов активности. Существуют незначительный разброс средней длительности действий, однако он достаточно незначительный.

2.1.2 Собственный датасет

Датасет представляет собой записи сигналов ЭЭГ одного испытуемого с помощью 8 электродов и содержит 10 сессий по 3–4 минуты воображаемой

активности. Частота дискретизации составляет 250 Гц, в конструкции шлема используются активные датчики¹⁹. Шлем содержит электроды {FP1, FP2, C3, C4, P7, P8, O1, O2} (легенда расположения электродов – рисунок 1, раздел 1), вместо опорного электрода Cz используется два опорных электрода на ушах, данные с которых используются для предобработки сигнала и не используются при анализе. Конструкция шлема подробно описана в [4].

Для получения записей использовалось собственное ПО. Каждой сессии предшествовала настройка и калибровка, а также 3 минутная тренировка испытуемого.

На рисунке 9 приведены отсчёты для датчика C3.

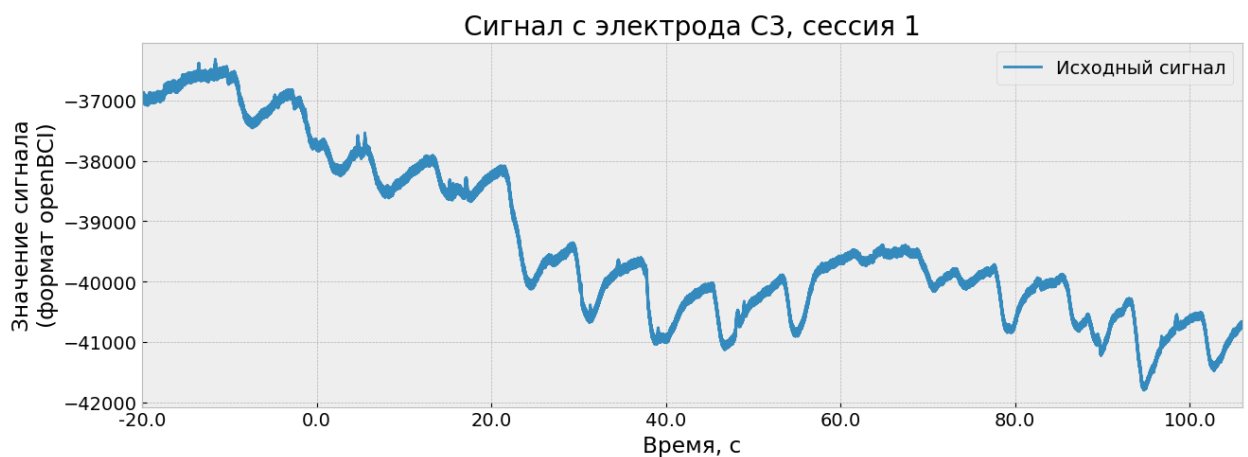


Рисунок 9 – Отсчёты записанного сигнала с датчика ЭЭГ

Для предобработки предлагается использовать следующую последовательность преобразований:

1. Быстрое преобразование Фурье (БПФ);
2. Фильтр Баттерворта с параметрами $f_n = 5$ Гц, $f_v = 40$ Гц, порядок 5;
3. Децимация (порядок 2).

¹⁹ Информации о типе датчиков в EEG Motor Movement / Imagery в открытом доступе нет.

Подобные преобразования позволяют снизить влияние шумов датчиков. На рисунке 10 приведён предобработанный и исходный сигналы для датчика С3.

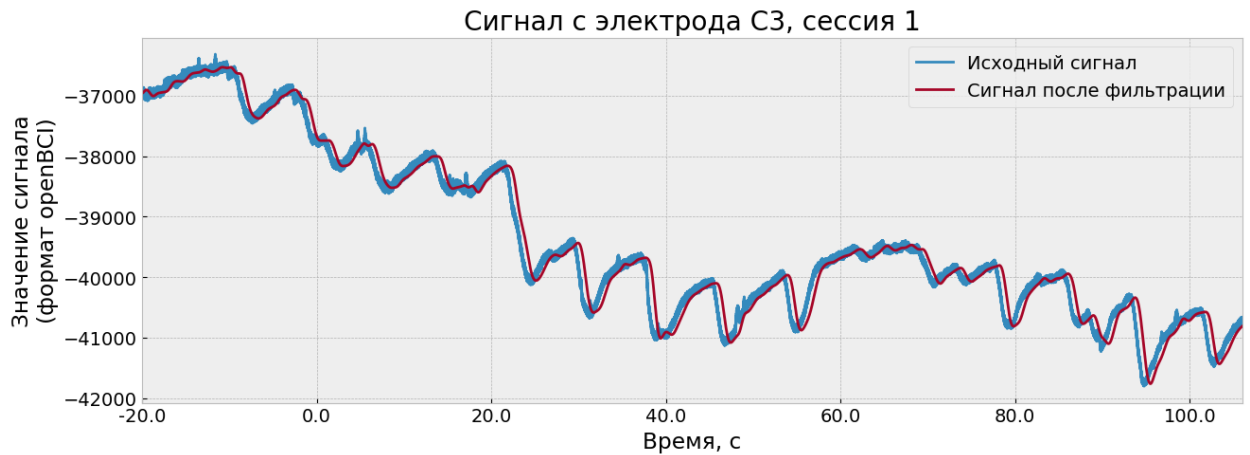


Рисунок 10 – Соотнесение исходного и фильтрованного сигнала датчика С3

На рисунке 11 приведена визуализация активностей для датчика С3, сессия 1.

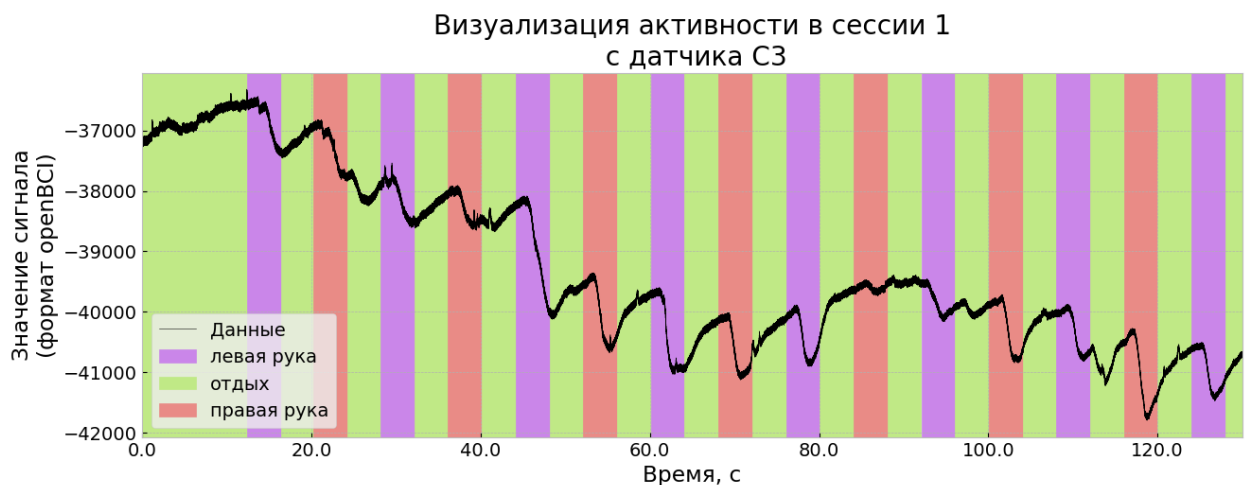


Рисунок 11 – Визуализация активности в сессии 1 с датчика С3

Следует отметить, что использование активных электродов позволяет лучше выделять активность и отдых, на рисунке 11 визуально достаточно

чётко отделяются периоды отдыха и мысленного сжатия правого или левого кулака.

Таким образом из данного датасета сформирован только один набор данных: дикторозависимый, для электродов $\{FP1, FP2, C3, C4, P7, P8, O1, O2\}$. Набор данных был сохранен в той же структуре что и данные из EEG Motor Movement / Imagery (см. раздел 2.1.1.).

Рассмотрим распределения длительностей активности различных сессий. На рисунке 12 приведены сводные статистики по всему набору данных.

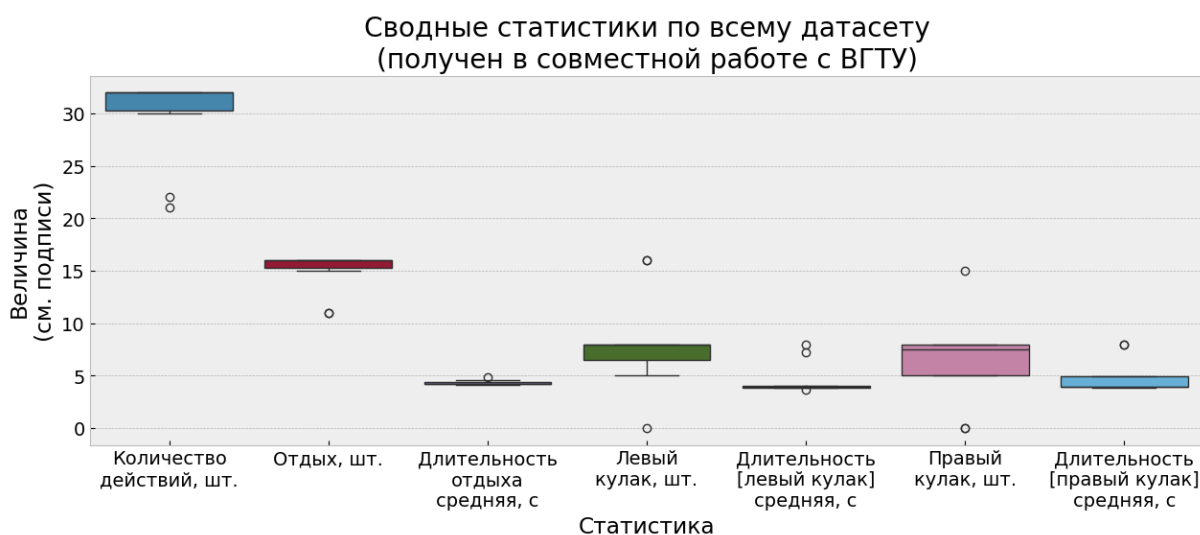


Рисунок 12 – Сводные статистики по всему собственному датасету

Из рисунка 12 видно, что существуют незначительный разброс средней длительности действий, однако он достаточно незначительный.

2.2 Постановка эксперимента

Рассмотренные в разделе 1.4 ДНС являются сетями для прогнозирования значений МВР. Однако в своей архитектуре все они содержат базовую часть, состоящую из ДНС, которая выделяет признаки (*deep features*) из МВР и прогностическую часть, которая преобразует признаки в прогноз.

В TimeGrad и ScoreGrad прогностическая часть представляет собой рекуррентную нейронную сеть с ячейками типа LSTM [10]. Для решения задачи классификации участка МВР целесообразно использовать скрытое состояние (*hidden state*) LSTM в качестве входных данных для полносвязной нейронной сети, которая будет выступать в качестве классификатора.

В D-Va признаки формируются непосредственно в ПДП, что позволяет использовать предобученную диффузионную часть D-Va в качестве генератора признаков МВР для последующей классификации с помощью полносвязной ИНС.

Общий конвейер (*pipeline*) построения классификаторов на основе ДНС состоит из следующих шагов:

1. Обучение ДНС на задачу прогнозирования (в соответствии с рекомендациями в работах [27,28,53]).
2. Использование предобученных ДНС для выделения признаков МВР и обучение полносвязной ИНС для классификации.

Целесообразно провести серии экспериментов для дикторозависимого датасета (EEG Motor Movement / Imagery) с целью оценки дисперсии оценки.

Для сравнения качества классификаторов на основе моделей TimeGrad, ScoreGrad и D-Va предлагается использовать классические метрики классификации, такие как «точность» (*accuracy*), ROC AUC с оценкой дисперсии метрики, где это возможно. Для оценки результатов в качестве реперных показателей используются показатели классификатора на основе свёрточной ИНС построенной в соответствии с [7].

Была разработана единая архитектура классификатора, состоящая из 2 скрытых слоёв: 2048 нейронов на первом скрытом слое, 512 нейронов на втором скрытом слое. Функции активации в скрытых слоях – сигмоидальные [10], функция активации выходного слоя – softmax [10]. Таким образом, классификатор реализован в виде полносвязной ИНС (многослойный персептрон) [14]. Указанная архитектура выбрана исходя из серии экспериментов. Критерием выбора являлась возможность классификатора показывать 100% «точность» на замкнутой выборке (без валидации).

Размерность вектора признаков (после генерации признаков с помощью ДНС) – 1024. Данный размер признаков обусловлен особенностями реализации, TimeGrad и ScoreGrad при большем количестве признаков численно неустойчивые²⁰, по этой же причине ограничено количество ПДП 50 шагами. Для корректного сравнения архитектур указанные гиперпараметры были использованы для всех исследуемых ДНС.

Таким образом, для проведения вычислительного эксперимента необходимо получить следующие результаты (метрики качества):

1. Для дикторонезависимого датасета EEG Motor Movement / Imagery, без фильтрации, для 3 электродов.
2. Для дикторонезависимого датасета EEG Motor Movement / Imagery, без фильтрации, для 8 электродов.
3. Для дикторонезависимого датасета EEG Motor Movement / Imagery, с фильтрацией, для 3 электродов.
4. Для дикторонезависимого датасета EEG Motor Movement / Imagery, с фильтрацией, для 8 электродов.
5. Для дикторозависимого датасета EEG Motor Movement / Imagery, без фильтрации, для 3 электродов (статистика по 10 дикторам).
6. Для дикторозависимого датасета EEG Motor Movement / Imagery, без фильтрации, для 8 электродов (статистика по 10 дикторам).
7. Для дикторозависимого датасета EEG Motor Movement / Imagery, с фильтрацией, для 3 электродов (статистика по 10 дикторам).
8. Для дикторозависимого датасета EEG Motor Movement / Imagery, с фильтрацией, для 8 электродов (статистика по 10 дикторам).
9. Для дикторозависимого собственного датасета, без фильтрации для 8 электродов (статистика по 2 подвыборкам).
10. Для дикторозависимого собственного датасета, с фильтрацией для 8 электродов (статистика по 2 подвыборкам).

²⁰ На практике численная неустойчивость проявляется в появлении тезоров со значением NaN (либо деление на ноль, либо переполнение типа данных float32) после первой итерации

Для дикторозависимых датасетов в качестве валидационной выборки используются отдельные сессии (EEG Motor Movement / Imagery – R07 для реальных движений, R08 – для воображаемых; для собственного – 4 сессии используется для обучения, 2 для валидации, 2 для тестирования). Для дикторонезависимого датасета для обучения используется 70% записей, валидация – 15%, тестирование – 15%. На рисунке 13 приведена схема получения датасетов для обучения, валидации и тестирования.



Рисунок 13 – Схема разбиения датасетов для проведения вычислительного эксперимента

Следует отметить, что многие исследователи [6,7,13] ставят перед собой задачу бинарного распознавания: классификатор применяется только в момент активности. Данная задача более простая относительно поставленной в настоящей работе, так как выделение периода активности из фона вносит достаточно значительные коррективы: обычно активность наступает внезапно,

фон мало отличается от периодов активности. Более того, дополнительный класс в задаче классификации снижает численное значение для случайного классификатора – для бинарной классификации «точность» случайного классификатора составляет 50%, для классификации на 3 класса – 33%²¹.

Также при анализе результатов обучения и тестирования целесообразно учитывать следующие значения: точность случайной классификации – 0.33 (теоретическое математическое ожидание для точности при случайном присвоении меток класса в сбалансированном датасете) и метрики СНС из работы [9], выбранная в качестве baseline.

2.3 Выводы

В разделе проведён первичный анализ датасетов, а также разработан постановка эксперимента. На основе проведённого анализа датасетов можно сделать вывод о репрезентативности датасетов, следовательно на их основе возможно построить модель классификации активности испытуемого по сигналам ЭЭГ.

Помимо этого, в разделе предложена архитектура, позволяющая использовать рассмотренные ранее ДНС (TimeGrad, ScoreGrad и D-Va) в качестве модели параметризации отрезков МВР для последующей классификации. Подобная двусоставная архитектура (сеть для выделения признаков и классификатор) имеет несколько преимуществ: во-первых, обучение ДНС на более сложную задачу не требует ручной разметки (задачи прогнозирования возможно решать на основе исторических данных МВР без разметки); во-вторых, обученную ДНС возможно использовать в качестве основы для различных задач классификации, переобучая только полносвязную сеть классификатора.

²¹ Модель случайного классификатора предполагает случайный выбор из равномерного дискретного распределения количество значений, которого соответствует количеству классов. Таким образом, вероятность верной классификации $1/n$, где n – количество классов.

3. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Вычислительный эксперимент проводился на основе реализаций ДНС:

- TimeGrad – [zalandoresearch/pytorch-ts](https://github.com/zalandoresearch/pytorch-ts)²²;
- ScoreGrad – [yantijin/ScoreGradPred](https://github.com/yantijin/ScoreGradPred)²³;
- D-Va – [koa-fin/dva](https://github.com/koa-fin/dva)²⁴.

В качестве вычислительного устройства использовалась пользовательская GPU Nvidia RTX 3060 (12 Gb VRAM)

3.1 Датасет EEG Motor Movement / Imagery

Обучение TimeGrad и ScoreGrad не зависят от длинны входного отрезка, т. к. в них используются рекуррентные сети (реализация LSTM), поэтому в экспериментах зафиксирована длинна отрезка для всех сетей 640 отсчётов (4 с). С одной стороны, подобная длинна гарантирует, что не будут перекрыты отдых и активность (активность всегда не менее 4 с, отдых всегда более 5 с), с другой стороны, позволяет использовать всю информацию из периода активности.

Каждая модель обучалась (без учёта классификатора) по 20 минут – данное время выбрано по серии экспериментов на частях датасета, обучение дольше указанного времени приводило к параличу сети²⁵.

3.1.1 Дикторозависимые модели

Результаты обучения и валидации для дикторозависимых моделей обобщались на основе 10 случайных (фиксированных для всех сетей) испытуемых из датасета. На рисунке 14 приведены результаты обучения и валидации TimeGrad, ScoreGrad и D-Va для 3 и 8 электродов соответственно, данные без фильтрации, реальные движения.

²² <https://github.com/zalandoresearch/pytorch-ts/tree/version-0.7.0>

²³ <https://github.com/yantijin/ScoreGradPred>

²⁴ <https://github.com/koa-fin/dva>

²⁵ Эффект, при котором ошибка обучения выходит на плато – не изменяется в процессе дальнейшего обучения

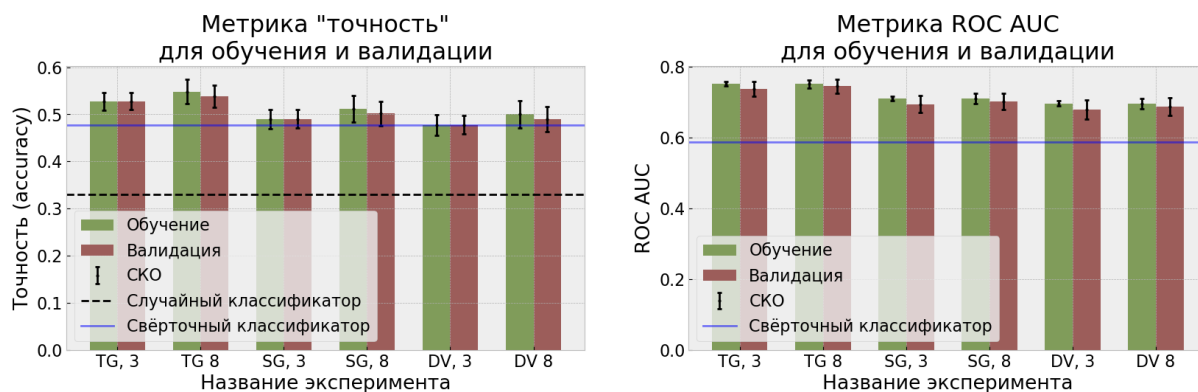


Рисунок 14 – Визуализация метрик обучения и валидации для моделей, построенных на данных реальных движений, без фильтрации

На рисунке 14 используются следующие подписи: TG – TimeGrad; SG – ScoreGrad; DV – D-Va; числа после запятой – число электродов. Из рисунка 13 видно, что все классификаторы на основе ДНС статистически превосходят случайный классификатор, помимо этого, незначительное отличие средних (столбцы) и среднеквадратических отклонений (черные вертикальные отрезки) между обучением и валидацией свидетельствует об отсутствии эффекта переобучения модели. Помимо этого, изучаемые модели немного превосходят классификатор на основе свёрточной ИНС.

На рисунке 15 приведены результаты обучения и валидации для реальных движений, при предобработке сигналов с датчиков использована фильтрация.

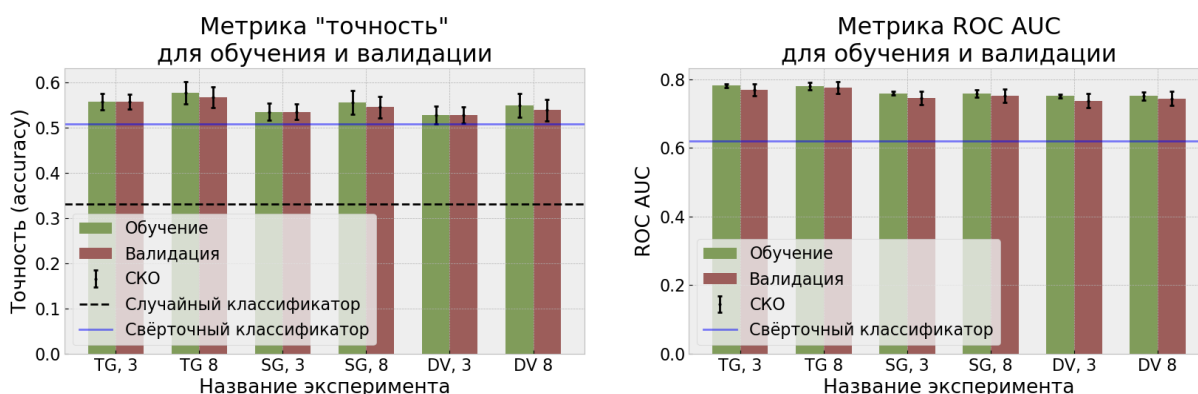


Рисунок 15 – Визуализация метрик обучения и валидации для моделей, построенных на данных реальных движений, с фильтрации

Из рисунка 15 видно, что в среднем, использование фильтрации в качестве предобработки улучшает метрик классификации. В среднем по моделям, прирост метрики «точность» на обучении и валидации составляет $8 \pm 2\%$ ²⁶, а ROC AUC вырастает на $6 \pm 1\%$. Таким образом, при предобработке данных целесообразно использовать фильтрацию, описанную в п. 2.1.1.

На рисунке 16 приведены результаты обучения и валидации для воображаемых движений, без использования фильтрации при предобработке.

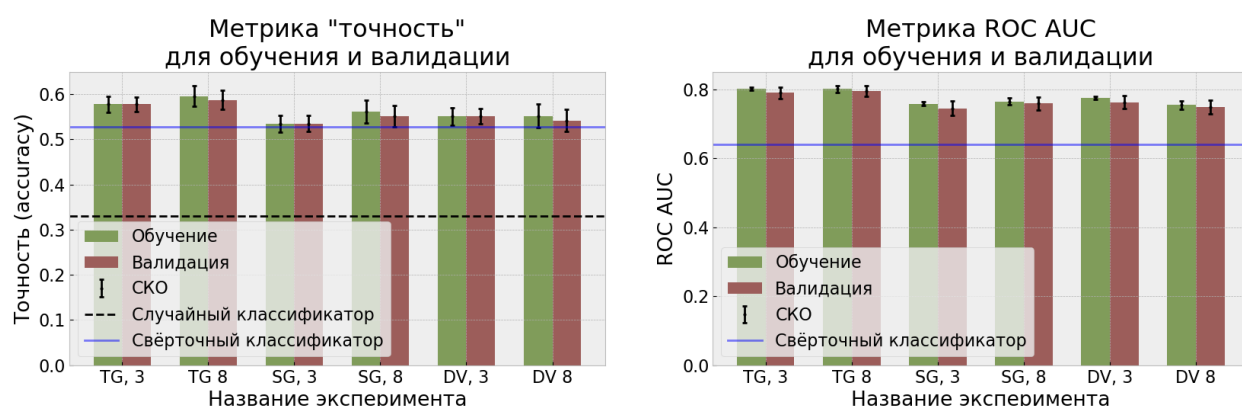


Рисунок 16 – Визуализация метрик обучения и валидации для моделей, построенных на данных воображаемых движений, без фильтрации

Сравнивая рисунок 16 и рисунок 14 можно сделать вывод о более точном распознавании воображаемых движений, нежели реальных. В численном выражении прирост по метрике «точность» в среднем составляет $3 \pm 1\%$, по метрике ROC AUC – $2 \pm 0,5\%$ ²⁷.

В качестве гипотезы объяснения эффекта прироста метрик при воображаемых движениях относительно реальных может служить различие в когнитивных нагрузках при выполнении реальных моторных и воображаемых задач. Возможно, сжатие кулака – полуавтоматический процесс, не требующий

²⁶ Здесь и далее расчёт процентов производится относительно меньшей метрики, а не в абсолютной шкале

²⁷ Подобный эффект характерен не только для классификаторов на основе ДНС, но также для любых других классификаторов. Он обсуждён с соавторами по работе над датасетом, однако опубликованных работ или обоснованных объяснений пока не найдено.

большого контроля со стороны мозга, а визуализация этого сжатия более сложная задача.

На рисунке 17 приведены результаты обучения и валидации для воображаемых движений, с использованием фильтрации при предобработке.

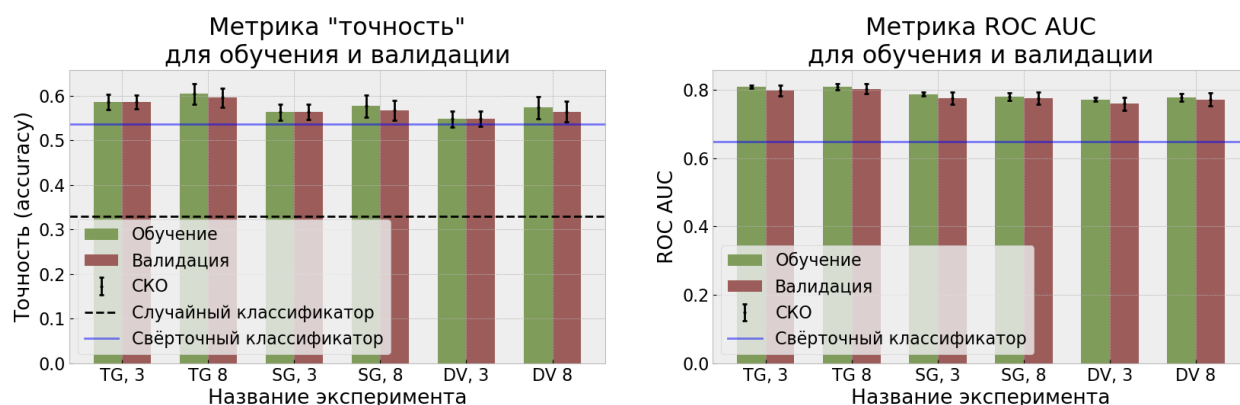


Рисунок 16 – Визуализация метрик обучения и валидации для моделей, построенных на данных воображаемых движений, с фильтрацией

Из рисунка 17 видно, что в среднем, использование фильтрации в качестве предобработки улучшает метрик классификации. В среднем по моделям, прирост метрики «точность» на обучении и валидации составляет $2 \pm 0,5\%$, однако метрика ROC AUC практически не изменяется.

Рассмотрим обобщённые показатели моделей для всех классов задач. На рисунке 18 приведены распределения валидационных метрик для моделей в различных задачах.

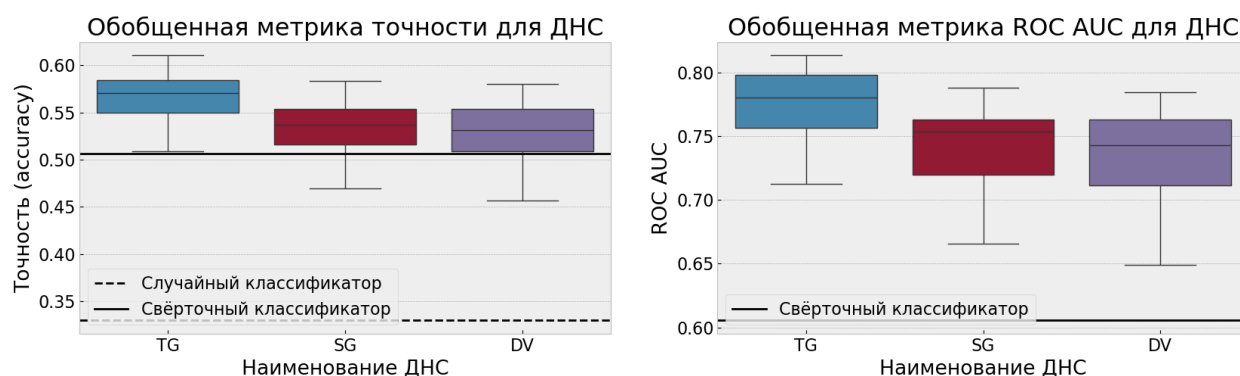


Рисунок 18 – Распределение метрик валидации для различных задач

Рисунок 18 демонстрирует тот факт, что в общем случае, не имеет значение какой именно ДНС из 3 рассмотренных использовать в качестве сети для выделения параметров МВР в задачах распознавания моторики. С другой стороны, независимо от задачи классификатор на основе ДНС демонстрирует статистически значимое отличие от случайного классификатора. Однако, модели на основе ДНС не значительно отличаются от свёрточного классификатора.

3.1.2 Дикторозависимые модели

Дикторонезависимые модели являются попыткой построения единой модели классификации для всех испытуемых. В ходе экспериментов валидационная выборка содержала данные о испытуемых, которые не входили в обучающую выборку. На рисунке 19 приведены метрики для дикторонезависимых моделей классификации реальной моторики, без использования фильтрации при предобработке.

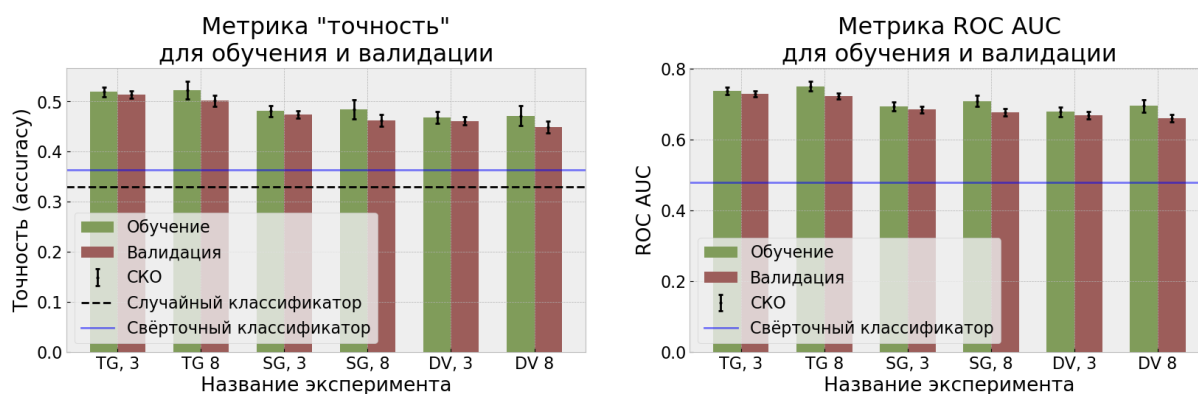


Рисунок 19 – Визуализация метрик обучения и валидации для моделей, построенных на данных реальных движений, без фильтрации

Рисунок 19 демонстрирует преимущества ДНС – их применение позволяет решать дикторонезависимую задачу классификации реальной моторики, в отличие от СНС, эффективность классификации которых практически не отличается от случайного классификатора. Распределение

метрик внутри группы исследуемых ДНС такое же, как и в дикторозависимых экспериментах, лучшей моделью является TimeGrad.

На рисунке 20 приведены результаты обучения и валидации для реальных движений, при предобработке сигналов с датчиков использована фильтрация.

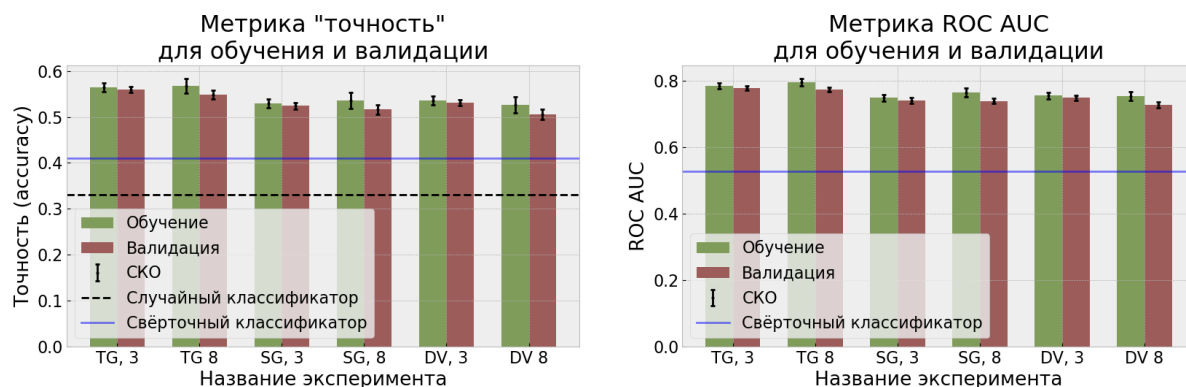


Рисунок 20 – Визуализация метрик обучения и валидации для моделей, построенных на данных реальных движений, с фильтрации

Как и в случае дикторозависимых моделей (п. 3.1.1, рисунок 15) использование фильтрации при предобработке даёт значительный прирост качества классификации: метрика «точность» вырастает на $10 \pm 2\%$, метрика ROC AUC – на $8 \pm 2\%$.

На рисунке 21 приведены результаты обучения и валидации для воображаемых движений, без использования фильтрации при предобработке.

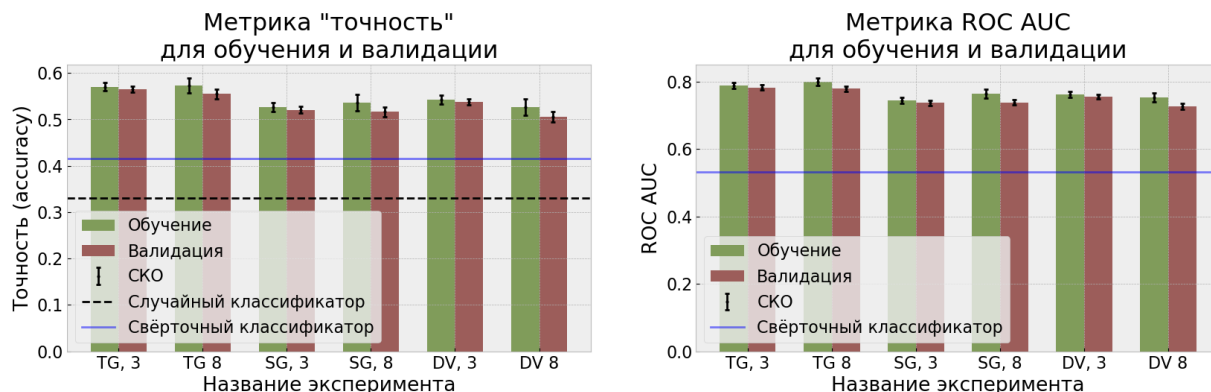


Рисунок 21 – Визуализация метрик обучения и валидации для моделей, построенных на данных воображаемых движений, без фильтрации

В отличие от дикторонезависимых моделей (п. 3.1.1, рисунок 16), значимого прироста метрик при замене реальных движений воображаемыми, не происходит. Гипотезой объяснения данного эффекта может являться «размытие» индивидуальных особенностей испытуемого в общем датасете. Однако этот вопрос требует более подробных исследований.

На рисунке 22 приведены результаты обучения и валидации для воображаемых движений, с использованием фильтрации при предобработке.

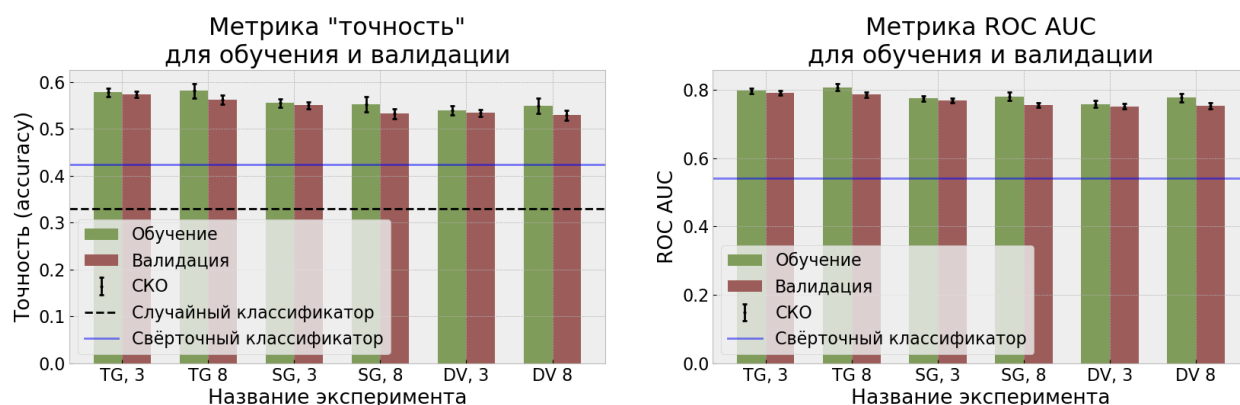


Рисунок 22 – Визуализация метрик обучения и валидации для моделей, построенных на данных воображаемых движений, с фильтрацией

В случае воображаемой моторики использования фильтрации при предобработке позволило повысить метрики «точность» – на $3 \pm 1\%$, ROC AUC – на $2 \pm 1\%$. Таким образом, проведённый дикторонезависимый эксперимент позволяет сделать следующие выводы: во-первых, использование ДНС в качестве модели выделения признаков МВР позволяет более качественно решать задачи дикторонезависимой классификации моторики (в отличие от использования свёрточного классификатора); во-вторых, использование фильтрации является эффективным способом повышения качества классификации (как дикторозависимой, так и дикторонезависимой).

Рассмотрим обобщённые показатели моделей для всех классов задач. На рисунке 23 приведены распределения валидационных метрик для моделей в различных задачах.

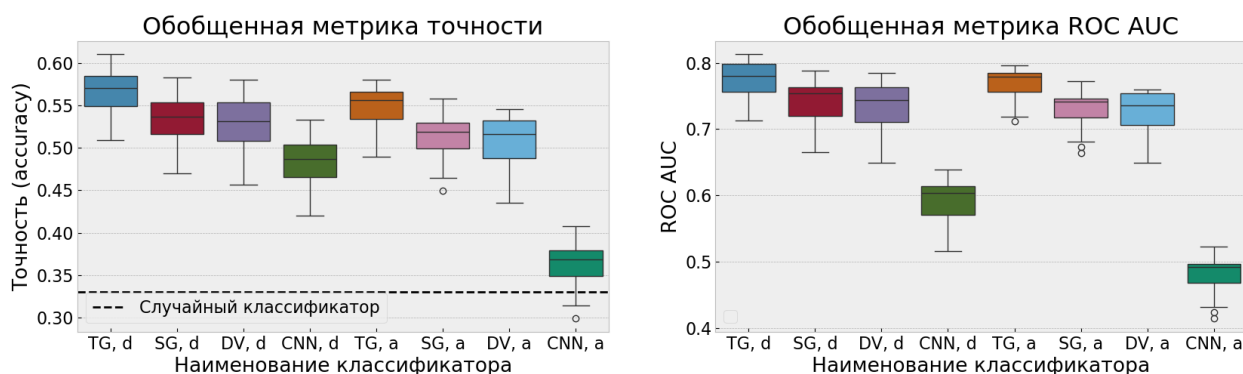


Рисунок 23 – Распределение метрик валидации для различных задач

На рисунке 23 используются следующие подписи: TG – TimeGrad; SG – ScoreGrad; DV – D-Va; CNN – свёрточный классификатор [9], выступающий в качестве реперной точки для сравнения результатов; d –дикторозависимый датасет; а – дикторонезависимый датасет. Рисунок 23 является обобщённой иллюстрацией качества классификации моторики (реальной и воображаемой) на датасете EEG Motor Movement / Imagery, из анализа рисунка можно заключить, что, во-первых исследуемые ДНС (TimeGrad, ScoreGrad и D-Va) не сильно отличаются при использовании в качестве моделей выделения признаков МВР; во-вторых, использование ДНС значительно повышает качество классификации при дикторонезависимом подходе (относительно свёрточных классификаторов).

3.2 Собственный датасет

На собственном датасете возможно воспроизвести только эксперименты по воображаемым движениям для дикторозависимого случая (п. 2.1.2).

В экспериментах зафиксирована длина отрезка для всех сетей 482 отсчёта (4 с, с учётом децимации в процессе предобработки, см. п. 2.1.2). Как и в случае с датасетом EEG Motor Movement / Imagery, каждая модель обучалась (без учёта классификатора) по 20 минут – данное время выбрано по серии экспериментов на частях датасета, обучение дольше указанного времени приводило к параличу сети.

На рисунке 24 приведены результаты обучения и валидации TimeGrad, ScoreGrad и D-Va для 8 электродов соответственно, данные без фильтрации, воображаемые движения.

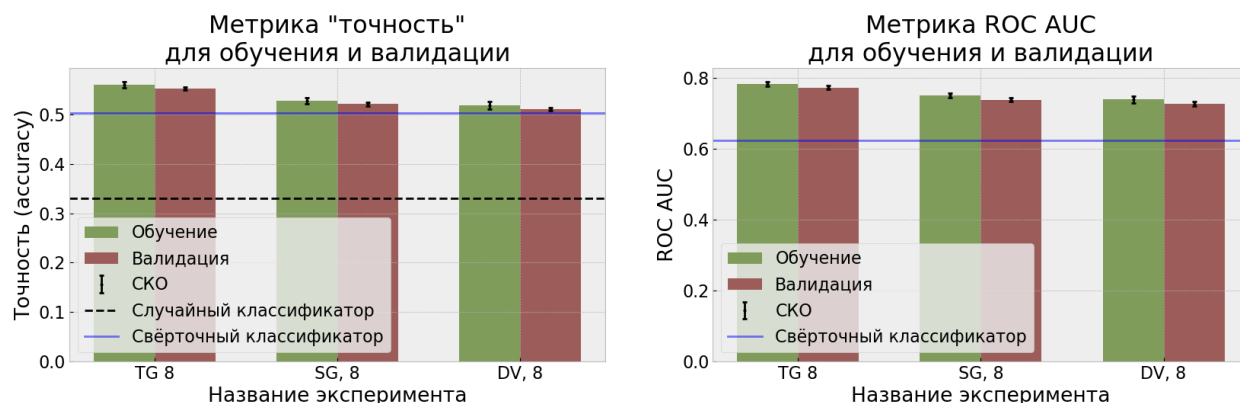


Рисунок 24 – Визуализация метрик обучения и валидации для моделей, построенных на данных воображаемых движений, без фильтрации

Из рисунка 24 видно, что все классификаторы на основе ДНС статистически превосходят случайный классификатор, помимо этого, незначительное отличие средних и среднеквадратических отклонений между обучением и валидацией свидетельствует об отсутствии эффекта переобучения модели. По сравнению с результатами для EEG Motor Movement / Imagery (п. 3.1.1, рисунок 15) оценки метрик показывают меньшую дисперсию. С другой стороны, средние значения метрик для двух различных датасетов достаточно близки.

На рисунке 24 приведены результаты обучения и валидации для воображаемых движений, при предобработке сигналов с датчиков использована фильтрация.

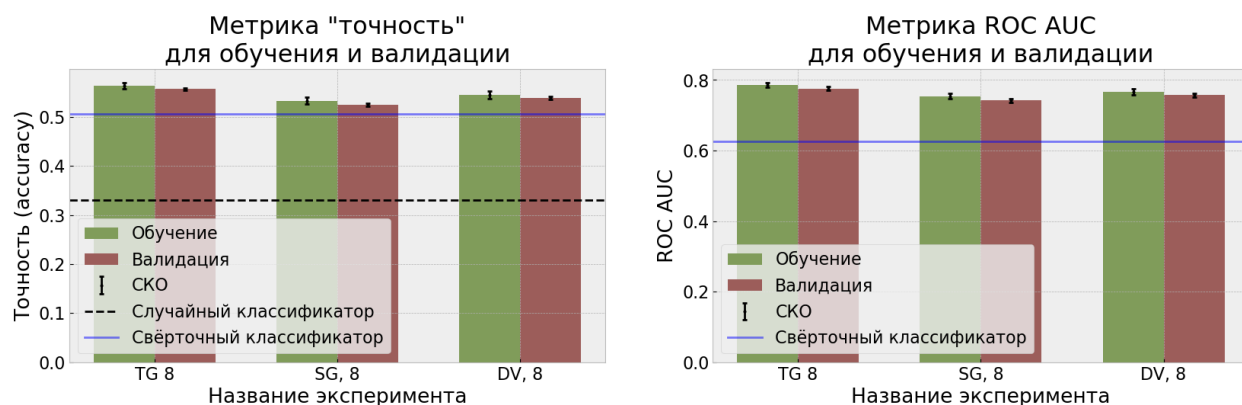


Рисунок 25 – Визуализация метрик обучения и валидации для моделей, построенных на данных воображаемых движений, с фильтрации

Как и в случае с EEG Motor Movement / Imagery (п. 3.1.1, рисунок 17) использование фильтрации при предобработке датасета позволяет поднять метрику «точность» (в среднем, по моделям) на $7 \pm 3\%$; ROC AUC – на $6 \pm 2\%$. Полученные результаты согласуются с результатами из раздела 3.1, и ещё раз подтверждают эффективность использования фильтрации на этапе предобработки данных при построении классификаторов на основе ДНС.

Таким образом, использование дополнительного датасета подтвердило ранее подтверждает выводы (раздел 2.1) и исключает предвзятость исследования – датасеты получены на различных программно-аппаратных платформах, в разное время и на различных испытуемых, однако, метрики качества эффективности классификаторов на основе ДНС значительно не отличаются.

3.3 Выводы

В разделе проведён анализ результатов вычислительного эксперимента по оценке качества классификации моторики (воображаемой и реальной) по сигналам ЭЭГ с помощью классификатор, использующих для выделения признаков МВР трех разных ДНС.

На основе проведенного вычислительного эксперимента можно сделать следующие выводы:

1. Использование фильтрации (фильтр Баттерворта с параметрами $f_n = 5$ Гц, $f_v = 40$ Гц, порядок 5) при всех вариантах эксперимента (дикторозависимый / дикторонезависимый, реальный / воображаемый) повышает метрики классификации. Метрика «точность» для дикторозависимого случая, реальные движения повышается на $8 \pm 2\%$; для дикторозависимого случая, воображаемые движения – на $2 \pm 0,5$; для дикторонезависимого случая, реальные движения – на $10 \pm 2\%$; для дикторонезависимого случая, воображаемые движения – на $3 \pm 1\%$. Эффект повышения точности при использовании фильтрации при предобработке подтверждён на двух различных датасетах.

2. Использование ДНС позволяет более качественно (с точки зрения метрик) решать задачу дикторонезависимой классификации движений, по сравнению с использованием СНС в качестве классификатора. В дикторозависимых задачах ДНС показывают сопоставимое СНС качество классификации.

3. Исследованные модели ДНС – TimeGrad, ScoreGrad и D-Va между собой отличаются незначительно, при их использовании в качестве моделей выделения признаков МВР в задачах классификации моторики по сигналам ЭЭГ. Таким образом, выбор конкретной модели должен производиться исходя из её технической реализации (потребление памяти, скорость обучения и т. д.).

4. Значительным недостатком ДНС является скорость их работы. В среднем, исследованные модели работают в 100 медленнее чем СНС. Использование подобных моделей в режиме «онлайн» невозможно, однако ДНС могут использоваться как модели-учителя для процессов дистилляции.

ЗАКЛЮЧЕНИЕ

В ходе анализа литературы было установлено, что классификация моторики по сигналам ЭЭГ является сложной прикладной задачей, единого устоявшегося решения которой на сегодняшний день не существует. С другой стороны, математический аппарат ДНС является перспективным направлением развития искусственных нейронных сетей и может быть применен к задаче классификации сегментов МВР. Исходя из результатов сравнительного анализа в [52] были отобраны три архитектуры ДНС: TimeGrad, ScoreGrad и DVa для анализа применимости ДНС в задачах распознавания моторики по сигналам ЭЭГ.

Затем был проведен анализ датасетов EEG Motor Movement / Imagery и собственного (полученного в рамках другой работы). Датасеты содержали записи сигналов с ЭЭГ-шлемов, различной частоты дискретизации и для различных испытуемых. Анализ датасетов показал, что данные собраны корректно – отсутствуют выбросы по длительности участков активности, разметка строго соответствует заявленной в эксперименте методике [1].

В ходе подготовки исследования был разработан план вычислительного эксперимента в соответствии с которым исходные датасеты были разбиты на 10 наборов данных, внутри которых происходило разделение на обучающую, валидационную и тестовую выборки. Помимо этого, для дикторозависимых датасетов модели обучались на 10 испытуемых для формирования статистического описания метрик, для дикторонезависимых датасетов статистики метрик формировались на основе переразбиения на обучающую, валидационную и тестовую выборки

Затем был проведён вычислительный эксперимент, в ходе которого получены следующие уникальные результаты:

1. Выявлена статистическая неразличимость результатов TimeGrad, ScoreGrad и DVa в задачах классификации моторики по сигналам ЭЭГ.

2. Получено эмпирическое подтверждение эффективности использования ДНС в дикторонезависимых задачах, по сравнению с использованием СНС в качестве классификатора. В дикторазависимых задачах ДНС показывают сопоставимое СНС качество классификации.

Таким образом, в ходе работы были выполнены все поставленные частные задачи и достигнута цель работы

Полученные результаты могут быть применены на практике для построения более качественных классификаторов моторики по сигналам ЭЭГ. Помимо этого, подтверждение положительного влияния фильтрации на точность классификации позволяет дать рекомендации по подготовке данных для решения задачи классификации моторики по сигналам ЭЭГ.

С другой стороны, в ходе работы возникли вопросы, позволяющие обозначить направление дальнейших исследований: а именно, обоснование эффекта повышения точности классификации при воображаемых движениях (относительно реальных); снижение эффекта от применения фильтрации при переходе от дикторозависимых моделей к дикторонезависимым. Более того, ДНС достаточно медленные и не могут обеспечить работу в режиме «онлайн», в связи с чем целесообразно провести дальнейшие исследования по дистилляции ДНС в СНС.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Schalk G. et al. BCI2000: a general-purpose brain-computer interface (BCI) system // IEEE Trans Biomed Eng. 2004. Vol. 51, № 6. P. 1034–1043.
2. Studio P. Neuralink [Electronic resource] // Neuralink. URL: <https://neuralink.com/?202308049001> (accessed: 24.02.2024).
3. Гнездицкий В.В. Обратная задача ЭЭГ и клиническая электроэнцефалография. Москва: МЕДпресс-информ, 2004. 624 p.
4. Журавлёв Д.В. Аппаратура для электроэнцефалографических исследований: монография. Воронеж: ФГБОУ ВО «Воронежский государственный технический университет»: Издательство ВГТУ, 2021. 259 p.
5. Александров М.В. et al. Электроэнцефалография: руководство. 3rd-е изд., перераб. и доп. ed. Санкт-Петербург: СпецЛит, 2020. 224 p.
6. Попов Е.Ю., Фоменков С.А. Детектирование событий движения руки в сигналах ЭЭГ головного мозга // Известия Волгоградского государственного технического университета. 2015. № 4 (32) P. 131–140.
7. Попов Е.Ю., Фоменков С.А. Детектирование событий движения руки в сигналах ЭЭГ головного мозга с помощью сверточных нейронных сетей // Известия Волгоградского государственного технического университета. 2016. № 1 (33) P. 62–67.
8. Popov E., Fomenkov S. Classification of hand motions in EEG signals using recurrent neural networks // 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM). 2016. P. 1–4.
9. Dai M. et al. EEG Classification of Motor Imagery Using a Novel Deep Learning Framework // Sensors. MDPI AG, 2019. Vol. 19, № 3. P. 551.
10. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение. Москва: «ДМК Пресс», 2017. 652 p.
11. Kingma D.P., Welling M. Auto-Encoding Variational Bayes: arXiv:1312.6114. arXiv, 2022. P. 1–12

12. Almutairi H., Hassan G.M., Datta A. Classification of sleep stages from EEG, EOG and EMG signals by SSNet: arXiv:2307.05373. arXiv, 2023. P. 1–20
13. Куркин С.А., Пицик Е.Н., Храмов А.Е. Использование искусственных нейронных сетей для классификации электрической активности головного мозга в процессе воображения движений у нетренированных испытуемых // Информационно-управляющие системы. 2020. № 6. P. 77–84.
14. Хайкин С. Нейронные сети. Полный курс. 2nd ed. Москва: Издательский дом “Вильямс,” 2006. 1104 p.
15. Брагин А.Д., Спицын В.Е. Распознавание моторных образов на электроэнцефалограммах с применением свёрточных нейронных сетей // Computer Optics. 2020. Vol. 44, № 3. P. 482–487.
16. Руннова А.Е. et al. Метод вейвлет-анализа паттернов двигательной активности на экспериментальных данных многоканальной электроэнцефалографии человека для управления внешними устройствами // Информационно-управляющие системы. 2019. № 1. P. 106–115.
17. Miao M. et al. Spatial-Frequency Feature Learning and Classification of Motor Imagery EEG Based on Deep Convolution Neural Network // Computational and Mathematical Methods in Medicine. Hindawi, 2020. Vol. 2020. P. 1–13.
18. He K. et al. Deep Residual Learning for Image Recognition // CoRR. 2015. Vol. abs/1512.03385. P. 1–12
19. Tosato G., Dalbagno C.M., Fumagalli F. EEG Synthetic Data Generation Using Probabilistic Diffusion Models: arXiv:2303.06068. arXiv, 2023. P. 1–16.
20. Ho J., Jain A., Abbeel P. Denoising Diffusion Probabilistic Models: arXiv:2006.11239. arXiv, 2020. P. 1–25.
21. Dhariwal P., Nichol A. Diffusion Models Beat GANs on Image Synthesis: arXiv:2105.05233. arXiv, 2021. P. 1–44.

22. Nichol A. et al. GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models: arXiv:2112.10741. arXiv, 2022. P. 1–20.
23. Rombach R. et al. High-Resolution Image Synthesis with Latent Diffusion Models: arXiv:2112.10752. arXiv, 2022. P. 1–45.
24. Saharia C. et al. Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding: arXiv:2205.11487. arXiv, 2022. P. 1–46.
25. Schneider F. ArchiSound: Audio Generation with Diffusion: arXiv:2301.13267. arXiv, 2023. P. 1–48.
26. Liu H. et al. AudioLDM: Text-to-Audio Generation with Latent Diffusion Models: arXiv:2301.12503. arXiv, 2023. P. 1–25.
27. Rasul K. et al. Autoregressive Denoising Diffusion Models for Multivariate Probabilistic Time Series Forecasting: arXiv:2101.12072. arXiv, 2021. P. 1–11.
28. Yan T. et al. ScoreGrad: Multivariate Probabilistic Time Series Forecasting with Continuous Energy-based Generative Models: arXiv:2106.10121. arXiv, 2021. P. 1–12.
29. Li Y. et al. Generative Time Series Forecasting with Diffusion, Denoise, and Disentanglement: arXiv:2301.03028. arXiv, 2023. P. 1–29.
30. Biloš M. et al. Modeling Temporal Data as Continuous Functions with Stochastic Process Diffusion: arXiv:2211.02590. arXiv, 2023. P. 1–19.
31. Shen L., Kwok J. Non-autoregressive Conditional Diffusion Models for Time Series Prediction: arXiv:2306.05043. arXiv, 2023. P. 1–14.
32. Wen H. et al. DiffSTG: Probabilistic Spatio-Temporal Graph Forecasting with Denoising Diffusion Models: arXiv:2301.13629. arXiv, 2023. P. 1–14.
33. Li R. et al. Graph Convolution Recurrent Denoising Diffusion Model for Multivariate Probabilistic Temporal Forecasting // ADMA. 2023. Vol. 14176. P. 661–676

34. Tashiro Y. et al. CSDI: Conditional Score-based Diffusion Models for Probabilistic Time Series Imputation: arXiv:2107.03502. arXiv, 2021. P. 1–30.
35. Alcaraz J.M.L., Strodthoff N. Diffusion-based Time Series Imputation and Forecasting with Structured State Space Models: arXiv:2208.09399. arXiv, 2023. P. 1–36.
36. Liu M. et al. PriSTI: A Conditional Diffusion Framework for Spatiotemporal Imputation: arXiv:2302.09746. arXiv, 2023. P. 1–13.
37. Lim H. et al. Regular Time-series Generation using SGM: arXiv:2301.08518. arXiv, 2023. P. 1–12.
38. Coletta A. et al. On the Constrained Time-Series Generation Problem: arXiv:2307.01717. arXiv, 2023. P. 1–39.
39. Волков И.К., Зуев С.М., Цветкова Г.М. Случайные процессы: Учеб. для вузов. МГТУ им. Баумана, 1999. 448 p.
40. Печинкин А.В., Тескин О.И., Цветкова Г.М. Теория вероятностей: Учеб. для вузов. 3rd-е изд., испр. ed. МГТУ им. Баумана, 2004. 456 p.
41. Nichol A., Dhariwal P. Improved Denoising Diffusion Probabilistic Models [Electronic resource] // arXiv.org. 2021. URL: <https://arxiv.org/abs/2102.09672v1> (accessed: 10.01.2024).
42. Sutter T.M., Daunhawer I., Vogt J.E. Generalized Multimodal ELBO: arXiv:2105.02470. arXiv, 2021. P. 1–23.
43. Sohl-Dickstein J. et al. Deep Unsupervised Learning using Nonequilibrium Thermodynamics: arXiv:1503.03585. arXiv, 2015. P. 1–18.
44. Luo C. Understanding Diffusion Models: A Unified Perspective: arXiv:2208.11970. arXiv, 2022. P. 1–23.
45. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization. // CoRR. 2014. Vol. abs/1412.6980. P. 1–15.
46. Radford A. et al. Learning Transferable Visual Models From Natural Language Supervision: arXiv:2103.00020. arXiv, 2021. P. 1–48.
47. Ho J., Salimans T. Classifier-Free Diffusion Guidance. arXiv: 2207.12598. arXiv, 2021. P. 1–14.

48. Ho J. et al. Cascaded Diffusion Models for High Fidelity Image Generation: arXiv:2106.15282. arXiv, 2021. P. 1–33.
49. Vincent P. A Connection Between Score Matching and Denoising Autoencoders // Neural Computation. 2011. Vol. 23, № 7. P. 1661–1674.
50. Song Y. et al. Sliced Score Matching: A Scalable Approach to Density and Score Estimation: arXiv:1905.07088. arXiv, 2019. P. 1–31.
51. Liu S. Langevin Quasi-Monte Carlo: arXiv:2309.12664. arXiv, 2023. P. 1–19.
52. Lin L. et al. Diffusion models for time-series applications: a survey: arXiv: 2305.00624. arXiv, 2023. P. 1–25.
53. Koa K.J.L. et al. Diffusion Variational Autoencoder for Tackling Stochasticity in Multi-Step Regression Stock Price Prediction // Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. 2023. P. 1087–1096.
54. Vaswani A. et al. Attention is All you Need // Advances in Neural Information Processing Systems. Curran Associates, Inc., 2017. Vol. 30. P. 1–11.
55. Oord A. van den et al. WaveNet: A Generative Model for Raw Audio: arXiv:1609.03499. arXiv, 2016. P. 1–15.
56. Kong Z. et al. DiffWave: A Versatile Diffusion Model for Audio Synthesis: arXiv:1609.03499. arXiv, 2020. P. 1–17.
57. Vahdat A., Kautz J. NVAE: A Deep Hierarchical Variational Autoencoder // Advances in Neural Information Processing Systems. Curran Associates, Inc., 2020. Vol. 33. P. 19667–19679.