

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ  
НАУКИ ИНСТИТУТ СИСТЕМНОГО ПРОГРАММИРОВАНИЯ ИМ. В. П.  
ИВАННИКОВА РАН

На правах рукописи

Аветисян Арам Арутюнович

**Методы и программные средства для выявления и анализа  
патологий по электрокардиограммам**

Специальность 2.3.5 —

«Математическое и программное обеспечение вычислительных систем,  
комплексов и компьютерных сетей»

Диссертация на соискание учёной степени кандидата технических наук

Научный руководитель:  
кандидат техн. наук  
Маркин Юрий Витальевич

Москва — 2026

## Оглавление

|                                                                                                                        | Стр.      |
|------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Введение . . . . .</b>                                                                                              | <b>5</b>  |
| <b>Глава 1. Обзор подходов к анализу электрокардиограмм . . . . .</b>                                                  | <b>12</b> |
| 1.1 Основные понятия и определения . . . . .                                                                           | 12        |
| 1.2 Анализ электрокардиограмм . . . . .                                                                                | 18        |
| 1.2.1 Классические методы обработки и анализа ЭКГ . . . . .                                                            | 19        |
| 1.2.2 Глубокое обучение для классификации патологий . . . . .                                                          | 25        |
| 1.2.3 Нейросетевые методы сегментации PQRST . . . . .                                                                  | 31        |
| 1.2.4 Самообучение для ЭКГ . . . . .                                                                                   | 34        |
| 1.3 Федеративное обучение . . . . .                                                                                    | 38        |
| 1.4 Выводы . . . . .                                                                                                   | 43        |
| <b>Глава 2. Методы и модели автоматизированного анализа</b><br><b>    электрокардиограмм в 12 отведениях . . . . .</b> | <b>46</b> |
| 2.1 Классификация электрокардиограмм в 12 отведениях по<br>патологическим состояниям . . . . .                         | 46        |
| 2.1.1 Анализируемые наборы данных для классификации . . . . .                                                          | 48        |
| 2.1.2 Сравнение моделей . . . . .                                                                                      | 50        |
| 2.1.3 Анализ моделей для классификации . . . . .                                                                       | 52        |
| 2.1.4 Дообучение моделей и оценка междатасетной устойчивости                                                           | 56        |
| 2.2 Сегментация комплексов PQRST . . . . .                                                                             | 61        |
| 2.2.1 Постановка задачи сегментации ЭКГ . . . . .                                                                      | 61        |
| 2.2.2 Используемые наборы данных . . . . .                                                                             | 62        |
| 2.2.3 Модель сегментации комплексов PQRST . . . . .                                                                    | 63        |
| 2.2.4 Схема обучения и метрики оценки качества . . . . .                                                               | 65        |
| 2.2.5 Постановка вычислительных экспериментов . . . . .                                                                | 66        |
| 2.2.6 Результаты экспериментов . . . . .                                                                               | 67        |
| 2.3 Выводы . . . . .                                                                                                   | 70        |
| <b>Глава 3. Самообучение и перенос представлений в задачах анализа</b><br><b>    электрокардиограмм . . . . .</b>      | <b>74</b> |

|                 |                                                                                                      |           |
|-----------------|------------------------------------------------------------------------------------------------------|-----------|
| 3.1             | Использование неразмеченных данных в задачах анализа электрокардиограмм . . . . .                    | 74        |
| 3.2             | Самообучение для 12-канальных электрокардиограмм . . . . .                                           | 75        |
| 3.2.1           | Подходы к самообучению моделей для ЭКГ . . . . .                                                     | 76        |
| 3.2.2           | Контрастное обучение для 12-канальных электрокардиограмм . . . . .                                   | 77        |
| 3.2.3           | Использование метаданных в контрастном обучении . . . . .                                            | 79        |
| 3.2.4           | Постановка экспериментов и результаты . . . . .                                                      | 81        |
| 3.3             | Перенос представлений с 12-канальных электрокардиограмм на ЭКГ с меньшим числом отведений . . . . .  | 84        |
| 3.3.1           | Особенности анализа 1-канальных ЭКГ . . . . .                                                        | 84        |
| 3.3.2           | Используемые наборы данных . . . . .                                                                 | 85        |
| 3.3.3           | Контрастное обучение для переноса знаний между конфигурациями с различным числом отведений . . . . . | 87        |
| 3.3.4           | Постановка экспериментов и результаты . . . . .                                                      | 91        |
| 3.4             | Выводы . . . . .                                                                                     | 94        |
| <b>Глава 4.</b> | <b>Федеративное обучение . . . . .</b>                                                               | <b>97</b> |
| 4.1             | Федеративное обучение в задачах анализа электрокардиограмм . . . . .                                 | 97        |
| 4.2             | Практическая система федеративного обучения для анализа ЭКГ . . . . .                                | 98        |
| 4.2.1           | Архитектура сервиса . . . . .                                                                        | 98        |
| 4.2.2           | Постановка эксперимента на данных ISP . . . . .                                                      | 100       |
| 4.2.3           | Разбиение данных между клиентами . . . . .                                                           | 102       |
| 4.2.4           | Результаты на различном числе клиентов . . . . .                                                     | 102       |
| 4.2.5           | Динамика достижения целевого качества . . . . .                                                      | 104       |
| 4.3             | Оценка клиентских обновлений при ошибочной разметке в федеративном обучении . . . . .                | 105       |
| 4.3.1           | Проблема ненадёжной разметки в медицинской федерации . . . . .                                       | 105       |
| 4.3.2           | Доверенная выборка для повышения устойчивости . . . . .                                              | 107       |
| 4.3.3           | Эксперименты . . . . .                                                                               | 110       |
| 4.3.4           | Эксперимент с распределённой ошибочной разметкой . . . . .                                           | 111       |
| 4.3.5           | Возможное использование доверительных оценок для контроля качества локальных данных . . . . .        | 114       |
| 4.4             | Выводы . . . . .                                                                                     | 115       |

|                          |     |
|--------------------------|-----|
| <b>Заключение</b>        | 117 |
| <b>Список литературы</b> | 122 |
| <b>Список рисунков</b>   | 132 |
| <b>Список таблиц</b>     | 133 |

## Введение

**Актуальность темы исследования.** Электрокардиография является одним из наиболее доступных, распространенных и клинически значимых методов функциональной диагностики сердечно-сосудистой системы. ЭКГ применяется при диагностике заболеваний, профилактических обследованиях, массовом скрининге и длительном мониторинговании состояния пациентов. Рост числа регистрируемых цифровых электрокардиограмм приводит к увеличению объема данных, требующих анализа специалистами, и повышает потребность в автоматизации их обработки и интерпретации. Автоматические методы анализа позволяют ускорить обработку больших объемов ЭКГ-исследований и могут использоваться для поддержки врача при выявлении электрокардиографических синдромов и патологических состояний.

Одной из основных проблем разработки методов автоматической интерпретации ЭКГ является отсутствие наборов данных, объем и разнообразие которых достаточны для обучения нейросетевых моделей и объективной оценки их качества на различных группах пациентов и диагностических состояниях. Для этого требуется большое число записей с согласованными экспертными заключениями, формирование которых требует участия квалифицированных специалистов и значительных временных затрат. В результате доступные наборы данных не всегда обеспечивают необходимое качество и разнообразие разметки.

При этом объём накопленных цифровых записей существенно превышает объем данных с качественными экспертными заключениями. Значительная часть таких записей не может непосредственно использоваться в традиционном обучении с учителем, несмотря на содержащуюся в них информацию о структуре и вариативности сигналов. Это делает актуальной задачу разработки методов, способных использовать неразмеченные данные на этапе предварительного обучения и тем самым снижать зависимость качества моделей от объема экспертной разметки.

Отдельной задачей автоматического анализа является определение временных границ основных морфологических элементов сигнала, необходимых для последующего вычисления интервалов и формирования дополнительных характеристик записи. Получение такой разметки является более трудоемкой задачей,

чем присвоение диагностической метки всей записи, поскольку требует анализа отдельных фрагментов сигнала. Существующие выборки с разметкой границ, как правило, имеют ограниченный объём, а сама разметка может различаться между наборами данных и экспертами. Это ограничивает возможности обучения и объективной оценки моделей сегментации и повышает значимость методов, позволяющих использовать дополнительные данные без ручной поэлементной разметки.

Дополнительное ограничение возникает при использовании сокращённых конфигураций регистрации. Полная многоканальная запись содержит больше информации о наблюдаемом процессе, однако в портативных и носимых устройствах часто доступно меньшее число каналов. Снижение размерности входного сигнала приводит к потере части информации и может ухудшать качество классификации. Поэтому актуальной является задача переноса знаний, полученных на более информативных многоканальных данных, на модели, работающие с уменьшенным числом доступных каналов.

Значительная часть клинических данных хранится распределённо в информационных системах отдельных организаций, а их передача и централизованное объединение могут быть ограничены требованиями к защите персональной информации и особенностями инфраструктуры. Вместе с тем совместное использование данных нескольких источников позволяет увеличить объём и разнообразие обучающей выборки. При распределённом обучении дополнительной проблемой становится неоднородность локальных данных и качества их разметки: систематические ошибки отдельных участников могут приводить к формированию некорректных клиентских обновлений и снижению качества общей модели.

Таким образом, актуальными проблемами являются ограниченность экспертно размеченных данных, большие трудозатраты детальной разметки, различия между источниками данных, снижение информативности при уменьшении числа доступных каналов и невозможность прямого объединения распределённых клинических выборок. Перспективными направлениями решения этих проблем являются использование неразмеченных данных при обучении моделей, перенос знаний между различными конфигурациями входного сигнала и федеративное обучение с оценкой качества клиентских обновлений. Разработка и исследование таких методов составляют актуальную задачу автоматического анализа медицинских временных рядов.

**Степень разработанности темы исследования.** В работах отечественных и зарубежных исследователей широко представлены методы автоматического анализа ЭКГ. Для выявления электрокардиографических синдромов активно применяются сверточные, рекуррентные и трансформерные архитектуры нейронных сетей, позволяющие извлекать необходимые для автоматического анализа признаки непосредственно из временных рядов ЭКГ. Для сегментации сигналов используются как классические алгоритмы обработки, так и современные нейросетевые подходы.

Значительное развитие получили методы предварительного обучения на данных без экспертных диагностических меток, в частности методы контрастного обучения, в которых представления формируются за счет использования внутренней структуры данных и различных преобразований исходных объектов. Такие методы позволяют использовать большие массивы неразмеченных записей и уменьшать зависимость последующего обучения от объема размеченной выборки. Одновременно развиваются нейросетевые методы сегментации, однако их эффективность по-прежнему существенно зависит от объема и качества доступной экспертной разметки.

Для работы с распределенными данными активно исследуются методы федеративного обучения, позволяющие формировать общую модель без передачи исходных записей между участниками. Наряду с базовыми алгоритмами усреднения разработаны методы устойчивой агрегации, предназначенные для снижения влияния аномальных и недостоверных клиентских обновлений. Часть методов использует геометрические свойства обновлений, другие — небольшую доверенную серверную выборку для оценки их полезности.

Несмотря на большое число существующих работ, сохраняются задачи эффективного использования неразмеченных данных, повышения качества сегментации при ограниченной экспертной разметке, переноса знаний при уменьшении числа доступных каналов и устойчивого обучения на распределенных данных с неоднородным качеством разметки.

**Объект исследования** — методы автоматического анализа медицинских временных рядов на примере цифровых электрокардиографических сигналов.

**Предмет исследования** — методы и алгоритмы обучения нейросетевых моделей классификации и сегментации медицинских временных рядов в усло-

виях ограниченного объема размеченных данных, изменения числа доступных каналов входного сигнала и распределенного хранения обучающих выборок.

**Целью** диссертационной работы является разработка и реализация методов и программных средств автоматического анализа медицинских временных рядов на примере анализа цифровых электрокардиограмм.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. Проанализировать актуальные задачи автоматического анализа цифровых электрокардиограмм, рассмотреть существующие методы их решения и выявить ограничения этих методов.
2. Провести сравнительное исследование нейросетевых моделей классификации патологических состояний по цифровым электрокардиограммам и выполнить обучение выбранной модели на клинических данных.
3. Разработать и реализовать метод сегментации цифровых ЭКГ с определением границ основных морфологических элементов в условиях ограниченного объема экспертно размеченных записей.
4. Разработать и реализовать метод контрастного предобучения нейросетевых моделей классификации, учитывающий метаданные пациентов при формировании признаков представлений.
5. Разработать и реализовать метод переноса знаний, полученных при предобучении на многоканальных данных, для повышения качества классификации при уменьшении числа доступных каналов входного сигнала.
6. Разработать и реализовать алгоритм федеративного обучения нейросетевых моделей классификации, основанный на оценивании полезности клиентских обновлений для идентификации аномальных обновлений и их учета при агрегации параметров модели.

**Основные положения, выносимые на защиту:**

1. Метод сегментации цифровых электрокардиографических сигналов с определением границ основных морфологических элементов, позволяющий повышать качество сегментации в условиях ограниченного объема экспертной разметки за счет использования автоматически сформированной псевдоразметки дополнительных записей.
2. Методы контрастного предобучения нейросетевых моделей, позволяющие повышать качество классификации за счет использования допол-

нительной информации о структуре данных: индивидуальных характеристик объектов при формировании признаков представлений и зависимостей между различными каналами одного сигнала при переносе знаний на задачу с уменьшенным числом доступных каналов.

3. Метод федеративного обучения нейросетевых моделей классификации, основанный на оценивании полезности клиентских обновлений, позволяющий выявлять и снижать влияние аномальных обновлений при агрегации параметров глобальной модели.

### **Научная новизна:**

1. Разработаны архитектура и метод сегментации цифровых ЭКГ, позволяющие использовать автоматически сформированную псевдоразметку дополнительных записей для повышения качества определения границ основных морфологических элементов при ограниченном объеме экспертно размеченных данных.
2. Разработан метод контрастного предобучения нейросетевых моделей классификации, основанный на использовании индивидуальных характеристик пациента при формировании признаков представлений.
3. Разработан метод переноса знаний, полученных при контрастном предобучении на многоканальных сигналах, на задачу классификации при уменьшенном числе доступных каналов.
4. Разработан алгоритм федеративного обучения нейросетевых моделей, позволяющий ограничивать влияние аномальных клиентских обновлений при формировании глобальной модели за счет их идентификации на основе оценки полезности.

**Теоретическая значимость** работы заключается в развитии методов анализа биомедицинских временных рядов в условиях ограниченного объема размеченных данных, изменения конфигурации входного сигнала и распределённого хранения обучающих выборок. В диссертации развиваются подходы к контрастному предобучению с использованием дополнительной информации об объектах, подходы с использованием автоматически сформированной псевдоразметки дополнительных записей и методы переноса знаний между моделями, работающими с различным числом каналов. Полученные результаты также развивают методы федеративного обучения за счет оценивания полезности клиентских об-

новлений для идентификации аномальных обновлений и использования полученных оценок при агрегации параметров глобальной модели.

**Практическая значимость** работы определяется возможностью применения разработанных методов и нейросетевых моделей в системах автоматического анализа цифровых электрокардиограмм и поддержки принятия врачебных решений. Разработанные подходы позволяют:

1. Выполнять автоматическую классификацию патологических состояний по многоканальным цифровым сигналам;
2. Автоматически определять границы основных морфологических элементов сигнала для последующего расчета интервалов и интерпретации записи;
3. Использовать наборы данных без диагностических заключений для предварительного обучения нейросетевых моделей и повышения качества классификации электрокардиографических синдромов;
4. Выполнять перенос знаний, полученных при обучении на многоканальных данных, для повышения качества классификации при уменьшенном числе доступных каналов;
5. Оценивать полезность клиентских обновлений для идентификации аномальных обновлений и учитывать полученные оценки при федеративном обучении нейросетевых моделей классификации электрокардиографических синдромов.

**Апробация работы.** Основные результаты диссертации были изложены на следующих конференциях:

1. Открытая конференция ИСП РАН, 2021;
2. Открытая конференция ИСП РАН, 2023;
3. XXXI Российский национальный конгресс «Человек и лекарство», 2024;
4. 2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC 2024);
5. Digital Innopolis Days x AI IN 2024;
6. 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2025);
7. AAAI Conference on Artificial Intelligence (AAAI 2026);

**Личный вклад.** Все представленные в рамках диссертационной работы результаты получены лично автором.

**Публикации.** По теме диссертации опубликовано 8 работ, в том числе 4 статьи – в рецензируемых журналах из перечня рекомендованных ВАК РФ. Также получено 2 свидетельства о государственной регистрации программы для ЭВМ. В публикациях [1; 3; 4; 7] автором сформулированы постановка задачи и методика её решения, разработан план проведения экспериментов, принято участие в реализации программного кода, выполнены экспериментальные исследования, анализ и интерпретация полученных результатов, а также подготовка текста публикаций. В публикациях [2; 5] автор совместно с соавторами участвовал в постановке задачи, проведении экспериментальных исследований, проверке и валидации результатов, а также в подготовке материалов к публикации. В публикациях [6; 8] совместно с соавторами сформулирована постановка задачи, разработан план экспериментов, выполнены проверка и валидация полученных результатов и подготовка текста публикации.

**Объем и структура работы.** Диссертация состоит из введения, четырёх глав и заключения. Полный объём диссертации составляет 134 страницы с 5 рисунками и 20 таблицами. Список литературы содержит 99 наименований.

## Глава 1. Обзор подходов к анализу электрокардиограмм

### 1.1 Основные понятия и определения

Сердечно-сосудистые заболевания занимают первое место среди причин смертности в мире и оказывают значительное влияние на систему здравоохранения и качество жизни пациентов. Раннее выявление нарушений работы сердца по электрокардиограммам (ЭКГ) является неотъемлемым этапом своевременной диагностики, выбора тактики лечения и профилактики неблагоприятных исходов. При этом большой объём регистрируемых ЭКГ-исследований создаёт потребность в надёжных автоматизированных системах анализа электрокардиограмм.

Анализ электрокардиограмм является одним из наиболее распространенных методов диагностики заболеваний сердца, поскольку позволяет быстро и неинвазивно оценивать электрическую активность сердца и выявлять клинически значимые изменения, включая нарушения ритма, проводимости и морфологии отдельных компонентов ЭКГ [11]. В клинической практике ЭКГ остаётся одним из наиболее доступных и информативных методов первичной оценки состояния сердца. Данный тип исследования стандартизован, является повторяемым во времени и применяется как в экстренных ситуациях, так и при плановой диагностике. Это делает задачу автоматизированного анализа ЭКГ актуальной в том числе в контексте масштабируемых сценариев скрининга и мониторинга.

ЭКГ-запись представляет собой последовательность значений, полученных путем разности электрических потенциалов, измеряемых на поверхности тела, которая комплексно отражает электрическую активность сердца в промежутке времени. Типичная форма ЭКГ-сигнала включает характерные компоненты — зубцы *P*, *Q*, *R*, *S*, *T*, а также интервалы и сегменты между ними. Морфология этих элементов и их временные соотношения используются в ходе клинической интерпретации (рис. 1.1).

**12-канальная ЭКГ и система отведений.** Стандартная электрокардиограмма в 12 отведениях представляет собой совокупность биоэлектрических сигналов, регистрируемых с различных точек поверхности тела, что обеспечивает про-

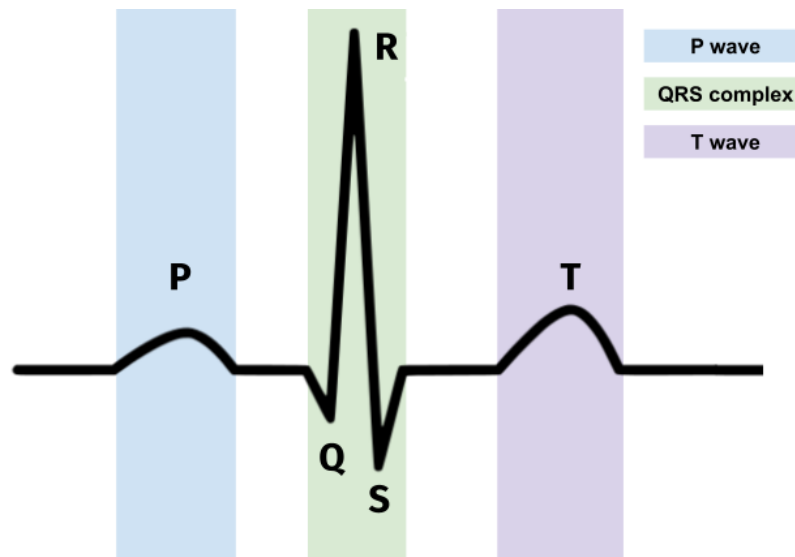


Рисунок 1.1 — Элементы кардиоцикла

странственную визуализацию электрической активности миокарда [12]. Стандартная ЭКГ-запись включает 3 биполярных отведения (I, II, III) с конечностей, 3 усиленных однополюсных отведения (aVR, aVL, aVF) с конечностей, а также 6 грудных отведений (V1–V6). Данный набор отведений стандартизован и является основой для современной клинической интерпретации ЭКГ [11].

Зависимость ЭКГ-сигнала от положения электродов наглядно иллюстрируется на примере отведений, регистрируемых на конечностях. В схеме Эйнтховена три электрода (правая рука, левая рука, левая нога) образуют треугольник, а стандартные отведения I, II и III соответствуют разностям потенциалов между его вершинами (рис. 1.2).

Далее формально будем считать, что в момент времени  $t$  на электродах с конечностей определены потенциалы  $\varphi_{RA}(t)$ ,  $\varphi_{LA}(t)$ ,  $\varphi_{LL}(t)$  (правая рука, левая рука, левая нога). Тогда три стандартных отведения определяются как разности потенциалов между соответствующими парами потенциалов:

$$I(t) = \varphi_{LA}(t) - \varphi_{RA}(t), \quad II(t) = \varphi_{LL}(t) - \varphi_{RA}(t), \quad III(t) = \varphi_{LL}(t) - \varphi_{LA}(t). \quad (1.1)$$

Поскольку упомянутые три отведения определяются с помощью трех потенциалов, между ними возникает простая алгебраическая связь. Из приведённых выше определений непосредственно следует закон Эйнтховена:

$$II(t) = I(t) + III(t), \quad (1.2)$$

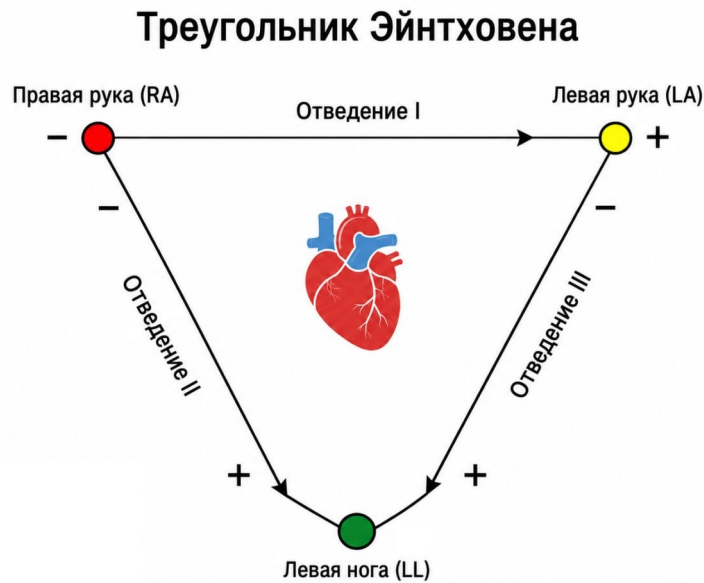


Рисунок 1.2 — Треугольник Эйнтховена и направления стандартных отведений на конечностях

который отражает согласованность перечисленных трех стандартных отведений.

Усиленные отведения, регистрируемые на конечностях, определяются относительно усреднённого потенциала двух других электродов с конечностей:

$$aVR(t) = \varphi_{RA}(t) - \frac{\varphi_{LA}(t) + \varphi_{LL}(t)}{2}, \quad (1.3)$$

$$aVL(t) = \varphi_{LA}(t) - \frac{\varphi_{RA}(t) + \varphi_{LL}(t)}{2}, \quad (1.4)$$

$$aVF(t) = \varphi_{LL}(t) - \frac{\varphi_{RA}(t) + \varphi_{LA}(t)}{2}. \quad (1.5)$$

Отведения V1–V6 регистрируются с помощью электродов, расположенных на груди. С их помощью обеспечивается оценка электрической активности миокарда в дополнительных плоскостях, не охватываемых отведениями от конечностей. 12-канальная ЭКГ даёт более полное представление о форме и распространении электрического сигнала, чем одно отдельное отведение [11]. Формальное обоснование расположения электродов и геометрии отведений подробно рассматривается в классических руководствах по электрокардиографии и биоэлектромагнетизму [12].

**ЭКГ как цифровой временной ряд.** В контексте анализа данных ЭКГ является дискретным временным сигналом. Стандартная 12-канальная ЭКГ представ-

ляется в виде матрицы

$$\mathbf{X} \in \mathbb{R}^{L \times N}, \quad \mathbf{X} = \begin{bmatrix} x_1[1] & \dots & x_1[N] & \vdots & \ddots & \vdots & x_L[1] & \dots & x_L[N] \end{bmatrix}, \quad (1.6)$$

где  $x_\ell[n]$  — дискретный сигнал  $\ell$ -го отведения,  $L$  — число отведений,  $N$  — число измерений, а дискретизация осуществляется с частотой  $f_s$ . В стандартной 12-канальной ЭКГ  $L = 12$ .

Результат ЭКГ-регистрации может быть представлен как на бумажном носителе, так и в цифровом виде. В настоящей работе рассматриваются цифровые ЭКГ, поскольку они непосредственно представлены временными рядами и позволяют применять методы цифровой предобработки и машинного обучения без промежуточной оцифровки изображения. Это важно, так как при работе с бумажными ЭКГ дополнительные искажения вносятся качеством печати, освещением, геометрическими деформациями и артефактами сканирования.

Важным параметром цифровой ЭКГ является частота дискретизации  $f_s$ . В реальных регистрирующих системах она принимает различные значения (как правило, от 250 до 1000 Гц и выше) в зависимости от режима регистрации и типа аппаратуры. Данный параметр определяет временное разрешение сигнала и, соответственно, точность оценки временных интервалов, длительностей зубцов и прочих морфологических характеристик; нормативные требования к диагностическому качеству регистрации ЭКГ закреплены в рекомендациях и стандартах [11]. В настоящей работе представление ЭКГ как временного ряда применяется как в задачах классификации патологий, так и в задачах детекции характерных зубцов, интервалов и сегментов.

**Сокращённые конфигурации отведений.** Наряду со стандартной 12-канальной регистрацией на практике применяются сокращённые конфигурации отведений. Так, 6-канальная ЭКГ, как правило, ограничивается отведениями от конечностей и не предполагает наложения грудных электродов. Подобные схемы отличаются простотой применения и могут оказаться предпочтительными в условиях, где определяющими являются скорость, доступность и воспроизводимость измерения. В то же время уменьшение числа отведений неизбежно сокращает объём пространственной информации о распространении электрической активности миокарда, вследствие чего сокращённые записи не могут рассматриваться как полная замена стандартной 12-канальной ЭКГ.

Особый интерес для скрининга представляет одноканальная ЭКГ. В простейшем случае её можно рассматривать как сигнал одного отведения, например отведения I:

$$\mathbf{x}^{(1)} = x_I[1:N] \in \mathbb{R}^N. \quad (1.7)$$

С практической точки зрения значимость одноканальной ЭКГ обусловлена прежде всего тем, что данный формат широко применяется в компактных и потребительских устройствах, включая смарт-часы и портативные регистраторы: большинство современных носимых кардиомониторов, встроенных в смарт-часы и портативные ЭКГ-регистраторы, физически регистрируют лишь одно отведение (как правило, аналог отведения I). Масштабные клинические исследования, охватывающие десятки и сотни тысяч пользователей, подтвердили практическую применимость такого формата для скрининга нарушений ритма [13]. Регистрация такой ЭКГ может выполняться оперативно, многократно и вне клинических условий, практически без участия медицинского персонала. Вследствие этого посредством одноканальных ЭКГ могут быть решены задачи длительного наблюдения, первичного скрининга, персонального мониторинга.

Вместе с тем одноканальной ЭКГ присущ ряд ограничений. Во-первых, в ней содержится меньший объём информации по сравнению с многоканальной записью, поскольку такая ЭКГ отражает электрическую активность миокарда лишь в одном направлении. Во-вторых, условия регистрации в потребительских устройствах существенно отличаются от клинических: применяются сухие электроды, контакт кожа-электрод характеризуется меньшей стабильностью, а двигательная активность пользователя во время измерения повышает вероятность возникновения артефактов. Вследствие этого подобные ЭКГ-записи зачастую оказываются более зашумлёнными и неоднородными по качеству.

Таким образом, одноканальная ЭКГ представляет собой более простую и доступную форму регистрации, однако уступает стандартной 12-канальной записи по объёму информации. В то же время одноканальный сигнал может рассматриваться как составная часть многоканальной ЭКГ, что позволяет исследовать взаимосвязь между отдельными отведениями и полной системой регистрации. Такой подход открывает возможность использовать взаимосвязи между отведениями для повышения качества автоматической интерпретации ЭКГ.

**Наборы данных и экспериментальная база.** В данной работе используются как открытые, так и закрытые клинические наборы данных.

Для задач классификации патологий используется открытый набор *PTB-XL* [14], включающий 21,799 клинических 12-канальных ЭКГ-записей с многоуровневой диагностической разметкой и сопутствующими метаданными. Разметка в *PTB-XL* основана на SCP-кодах: SCP представляет собой стандартизованную систему описания электрокардиографических заключений, позволяющую формализовать диагностические, ритмологические и морфологические утверждения в виде кодов. Благодаря открытому доступу, размеру и стандартизованной разметке *PTB-XL* широко используется как бенчмарк для задач автоматической классификации ЭКГ патологий.

Также для задач классификации патологий используется закрытый клинический архив *ISP*, содержащий более 1,5 миллиона ЭКГ-записей. Этот архив является более неоднородным по сравнению с *PTB-XL*: ЭКГ-записи получены из разных медицинских учреждений с применением различных регистрирующих устройств. Разметка данного архива также представлена в терминах SCP-кодов, что делает возможным сопоставление его диагностических классов с диагностическими классами в *PTB-XL*.

Для задач сегментации характерных элементов ЭКГ-сигнала используется открытый набор *LUDB* [15], содержащий 200 клинических 12-канальных записей с разметкой границ и пиков волн *P*, комплексов *QRS* и волн *T*. Наличие детальной покомпонентной разметки делает этот набор данных удобным для обучения и валидации алгоритмов делинеации ЭКГ.

Кроме того, в работе используется закрытый набор данных *ISPdeli*, состоящий из 475 ЭКГ-записей с разметкой границ *P*, *QRS* и *T*. Разметка этого набора выполнялась поэтапно с валидацией качества разметки независимым экспертом.

Характеристики используемых в работе наборов данных приведены в таблице 1.

Таблица 1 — Наборы данных, используемые в работе

| Набор   | Записей   | $f_s$ , Гц | Тип разметки           |
|---------|-----------|------------|------------------------|
| PTB-XL  | 21,799    | 500        | Диагнозы, SCP коды     |
| ISP     | 1,500,000 | 500/1000   | Диагнозы, SCP коды     |
| LUDB    | 200       | 500        | Границы и пики P/QRS/T |
| ISPdeli | 475       | 500/1000   | Границы P/QRS/T        |

Перечисленные наборы данных сопровождаются разметкой, позволяющей решать задачи классификации ЭКГ-патологий, а также задачи сегментации характерных элементов сигнала. Важно подчеркнуть, что медицинские ЭКГ-данные, как правило, относятся к категории конфиденциальной информации и не могут свободно передаваться между учреждениями либо аккумулироваться в едином централизованном хранилище. Данное обстоятельство препятствует прямому объединению данных из разных источников и определяет востребованность методов обучения, принимающих во внимание распределённый характер хранения медицинских данных.

## 1.2 Анализ электрокардиограмм

Автоматизированный анализ ЭКГ эволюционировал от детерминированных алгоритмов цифровой обработки сигналов к методам машинного обучения и современным нейросетевым моделям сквозного анализа. На ранних этапах преобладали алгоритмы, основанные на фильтрации, адаптивных порогах и явно заданных правилах выделения характерных элементов сигнала, — в частности, детекторы QRS-комплексов и методы определения границ зубцов на основе вейвлет-преобразований. Подобные подходы обладают высокой интерпретируемостью, вычислительной эффективностью и применимы даже при ограниченных объёмах данных. Однако они требуют ручной настройки параметров и, как правило, оказываются чувствительными к изменению условий регистрации, характеристик оборудования и особенностей целевой популяции.

Следующим этапом стало появление методов обучения с учителем с опорой на заранее сконструированные признаки: амплитудные, временные, морфологические и спектральные характеристики ЭКГ. В этом случае часть экспертных правил заменяется статистической моделью, обучаемой на размеченных данных. Такой подход позволяет учитывать более сложные закономерности, нежели простые пороговые правила, однако качество результата по-прежнему во многом определяется выбранным набором признаков и полнотой экспертной разметки.

Развитие глубокого обучения обеспечило переход к моделям, извлекающим информативные признаки непосредственно из исходного сигнала. Это существенно расширило возможности автоматической классификации патологий и выделения характерных элементов ЭКГ, особенно при наличии достаточно больших размеченных выборок. Вместе с тем такие модели сохраняют чувствительность к различиям между обучающей и тестовой выборками. Одна и та же архитектура нейронной сети, обученная на данных одного медицинского учреждения, может демонстрировать иные результаты на данных из другого источника. Эта проблема обычно описывается как сдвиг распределения и является одним из ключевых ограничений при практическом внедрении нейросетевых моделей анализа ЭКГ.

Самостоятельное направление образуют методы обучения, использующие внутреннюю структуру данных без необходимости их полной экспертной разметки. В задачах анализа ЭКГ перспективным видится контрастное обучение, при котором обучаемая нейросетевая модель формирует внутренние представления сигналов таким образом, чтобы в пространстве признаков семантически близкие ЭКГ-сигналы располагались на расстоянии, существенно меньшем по сравнению с ЭКГ-сигналами, семантика которых различна. В дополнение к ЭКГ-записи при обучении нейросетевой модели в данном случае могут использоваться структура её отведений, зависимости между отдельными отведениями, клинически значимые метаданные о пациенте, например, возраст. Такой подход позволяет задействовать большие массивы данных даже в условиях их ограниченной диагностической разметки.

Обозначенные подходы и основные задачи анализа ЭКГ подробнее рассматриваются в последующих подразделах: 1.2.1, 1.2.2, 1.2.3, 1.2.4.

### **1.2.1 Классические методы обработки и анализа ЭКГ**

До широкого распространения глубокого обучения автоматизированный анализ ЭКГ строился преимущественно как последовательность явно заданных этапов: предварительная обработка сигнала, детекция опорных точек, выделение характерных элементов ЭКГ, извлечение признаков и последующая класси-

фикация на основе правил или сравнительно простых статистических моделей. Такая схема отражала как клиническую логику интерпретации ЭКГ, так и ограничения вычислительных методов того времени: алгоритм должен был обладать интерпретируемостью, устойчивостью к базовым помехам и пригодностью для работы в условиях ограниченных вычислительных ресурсов.

Классические методы могут быть условно разделены на три группы. Первая группа связана с цифровой обработкой сигнала и включает фильтрацию, адаптивные пороги, дифференцирование, интегрирование по скользящему окну и иные преобразования, усиливающие отдельные элементы ЭКГ. Вторая группа основана на математических преобразованиях сигнала, таких как вейвлет-анализ и математическая морфология. Третья группа объединяет признаковые модели и ранние методы машинного обучения, в которых решение принимается не непосредственно по исходному сигналу, а по заранее сконструированному набору амплитудных, временных, морфологических и ритмологических признаков.

**Детекция QRS-комплексов и  $R$ -пиков.** Одной из центральных задач классического анализа ЭКГ является детекция QRS-комплексов и локализация  $R$ -пиков: от положения  $R$ -пиков зависят расчёт RR-интервалов, оценка частоты сердечных сокращений, анализ вариабельности ритма, разбиение сигнала на отдельные кардиоциклы и последующее выделение волн  $P$  и  $T$ .

Наиболее известным классическим алгоритмом детекции QRS-комплексов является метод Пана–Томпкинса [16]. В нём QRS-комплекс выделяется с помощью последовательности простых, но хорошо интерпретируемых операций: полосовой фильтрации, дифференцирования, нелинейного преобразования, интегрирования по скользящему окну и адаптивного порогового решения. Полосовая фильтрация снижает влияние дрейфа изолинии и высокочастотных помех, дифференцирование подчёркивает быстрые изменения сигнала, характерные для QRS-комплекса, а адаптивные пороги позволяют учитывать изменения амплитуды и формы сигнала во времени.

Дальнейшее развитие QRS-детекторов было связано с уточнением правил принятия решения и повышением устойчивости к вариативности ЭКГ. В работах Гамильтона и Томпкинса анализировались отдельные элементы правил детекции и способы их оптимизации на эталонных базах данных [17]. В более поздних алгоритмах применялись комбинированные критерии, учитывающие амплитуду и

скорость нарастания сигнала QRS-комплекса, локальную шумовую обстановку, а также регулярность RR-интервалов [18]. В обзорных работах по программной детекции QRS подчёркивается, что большинство классических детекторов может рассматриваться как различные варианты одной общей идеи: преобразовать сигнал таким образом, чтобы QRS-комплекс стал наиболее выраженным событием, после чего применить правило обнаружения с адаптивным порогом [19]. Впоследствии данная идея была систематизирована и апробирована при анализе различных баз ЭКГ, что, в свою очередь, позволило выработать оптимизированный набор эвристических правил детекции QRS-комплексов [20].

Наряду с пороговыми схемами применялись методы математической морфологии и вейвлет-преобразования. В рамках математической морфологии используются операции эрозии, дилатации, открытия и закрытия, позволяющие подавлять медленные изменения изолинии и выделять структуры заданной ширины [21]. Вейвлет-преобразование, в свою очередь, эффективно для анализа ЭКГ как нестационарного сигнала: оно позволяет рассматривать локальные события на различных масштабах и тем самым отделять QRS-комплекс от более медленных волн  $P$  и  $T$ , дрейфа изолинии и части артефактов [22; 23].

**Выделение границ волн и интервалов кардиоцикла.** В задаче определения границ волн и интервалов необходимо определить начало и конец волны  $P$ , QRS-комплекса и волны  $T$ , а также вычислить продолжительность клинически значимых интервалов и сегментов, включая PR, QRS и QT. Эта задача является более сложной, чем поиск  $R$ -пиков, поскольку зависят от формы конкретной ЭКГ-записи и даже могут неоднозначно определяться в ходе экспертной разметки.

Классические алгоритмы выделения границ волн ЭКГ-сигнала, как правило, опирались на уже найденный QRS-комплекс. На первом этапе определялось положение  $R$ -пика и ориентировочные границы QRS-комплекса, после чего в окрестностях кардиоцикла выполнялся поиск волн  $P$  и  $T$ . Такой подход позволял использовать физиологическую составляющую ЭКГ: волна  $P$  ожидается перед QRS-комплексом, а волна  $T$  — после него. Благодаря этому задача сводилась к локальному анализу заранее заданных временных окон.

В работах, основанных на применении вейвлет-преобразований, характерные точки ЭКГ связывались с максимумами модулей вейвлет-коэффициентов на

разных масштабах [22]. Позднее был предложен вейвлет-делинеатор, в котором QRS-комплексы, пики волн, начала и окончания волн определяются на основе локальных экстремумов вейвлет-коэффициентов и набора правил согласованности [24]. Такие методы естественным образом отражают локальную природу ЭКГ-событий и позволяют работать с различными временными масштабами: быстрым QRS-комплексом и более широкими волнами  $P$  и  $T$ .

Другой подход к выделению границ волн и интервалов основан на вероятностных моделях последовательностей. Например, скрытые марковские модели позволяют представить кардиоцикл как последовательность состояний, соответствующих изолинии, волне  $P$ , QRS-комплексу и волне  $T$ . Поэтому в данном случае задача выделения границ может быть сформулирована как поиск наиболее вероятной последовательности состояний по наблюдаемым признакам сигнала [25]. Преимущество такого подхода состоит в применении априорных знаний о порядке следования элементов ЭКГ и типичных длительностях отдельных фаз кардиоцикла.

**Предварительная обработка сигнала.** Причинами искажений ЭКГ-сигнала являются дрейф изолинии, сетевые наводки, мышечная активность и артефакты движения электродов. Для устранения таких искажений большинство классических алгоритмов анализа ЭКГ включают этап предварительной обработки сигнала: перед детекцией QRS-комплексов или извлечением признаков применяются полосовые, режекторные, медианные, адаптивные или морфологические фильтры.

В то же время удаление помех не должно приводить к искажению клинически значимых фрагментов сигнала. В частности, чрезмерное подавление низкочастотных компонент может изменить форму ST-сегмента или волны  $T$ , а агрессивное сглаживание может повлиять на амплитуду и длительность отдельных элементов ЭКГ. Поэтому в классических методах анализа ЭКГ предварительная обработка является компромиссом между повышением качества сигнала и сохранением его диагностически значимой морфологии [23; 26; 27].

**Признаковые модели и методы машинного обучения.** Следующим этапом развития автоматизированного анализа ЭКГ стали методы машинного обучения, основанные на заранее сконструированных признаках. В отличие от экспертных

правил, такие методы позволяли обучать классификатор по размеченным данным, однако сами признаки по-прежнему задавались вручную. Типичный конвейер алгоритма анализа ЭКГ включал детекцию *R*-пиков, разбиение ЭКГ-записи на кардиоциклы, нормализацию отдельных фрагментов сигнала, извлечение признаков и классификацию каждого сердечного сокращения или всей записи.

Значительную группу признаков составляли временные характеристики сигнала. К ним относятся RR-интервалы, отношения смежных RR-интервалов, локальная частота ритма, длительности PR, QRS и QT, а также относительное положение волн внутри кардиоцикла. Эти признаки особенно важны для анализа нарушений ритма, поскольку многие аритмии проявляются не только изменением формы отдельного комплекса, но и изменением последовательности сокращений во времени.

Другую существенную группу составляли морфологические признаки, характеризующие форму и амплитудно-временную структуру сигнала. К ней относятся амплитуды зубцов, длительность QRS-комплекса, площади под отдельными участками сигнала, характеристики наклона восходящих и нисходящих сегментов, отношения амплитуд отдельных элементов, а также показатели сходства исследуемого сокращения с эталонным или усреднённым шаблоном нормального кардиоцикла. Морфологические признаки применялись для дифференциации нормальных и эктопических сокращений, выявления нарушений внутрижелудочковой проводимости, включая блокады ножек пучка Гиса, идентификации желудочковых комплексов и других патологических состояний, сопровождающихся изменением формы ЭКГ-сигнала.

Отдельную категорию составляли признаки, формируемые на основе преобразований исходного сигнала. В частности, вейвлет-преобразование позволяло получать коэффициенты, характеризующие структуру кардиоцикла на различных временных масштабах, что обеспечивало возможность одновременного анализа локальных и глобальных особенностей сигнала. Спектральные признаки использовались для описания частотного состава сигнала, тогда как авторегрессионные характеристики позволяли учитывать статистические зависимости между последовательными отсчётами временного ряда. Для формирования компактного представления морфологии кардиоцикла также применялись методы снижения размерности, в частности метод главных компонент и метод независимых компонент. Использование указанных преобразований позволяло перейти

от исходного одномерного временного ряда к векторному представлению фиксированной размерности, пригодному для последующего применения стандартных алгоритмов машинного обучения.

В качестве классификаторов применялись линейные модели, метод  $k$  ближайших соседей, деревья решений, байесовские классификаторы, метод опорных векторов, ансамблевые методы, а также ранние нейросетевые и нечеткие модели. Например, рассматривалась пациент-адаптируемая классификация сердечных сокращений с использованием смеси экспертов [28]. Идея такой адаптации состояла в том, что универсальная модель может быть дополнена небольшим количеством индивидуальной информации о пациенте, что повышает качество распознавания при выраженной межпациентной вариативности ЭКГ. В других работах исследовались SVM-подходы к распознаванию сердечных сокращений по признаковому описанию [29].

Одной из наиболее показательных работ этого направления является исследование, где классификация сердечных сокращений строилась на сочетании морфологических признаков и признаков интервалов между сокращениями [30]. Такой подход хорошо иллюстрирует общую логику признаковых моделей: форма отдельного комплекса и ритмологический контекст рассматриваются совместно, поскольку оба источника информации важны для различения классов сокращений. Позднее похожая идея развивалась в работах, объединяющих морфологические и динамические признаки, в том числе признаки, полученные с помощью вейвлет-преобразования и анализа временной последовательности сокращений [31].

Преимущество признаковых моделей состоит в том, что их решения часто можно связать с клинически понятными характеристиками сигнала: шириной комплекса, изменением RR-интервала, амплитудой зубца или формой отдельного сокращения. Кроме того, такие методы требуют меньше данных, чем современные нейросетевые модели, и могут быть эффективны в узко заданных сценариях. Однако их качество существенно зависит от корректности предыдущих этапов обработки: ошибки детекции  $R$ -пиков, неточная сегментация или неудачно выбранный набор признаков напрямую ограничивают итоговую классификацию. Кроме того, признаки, подобранные под один набор данных, не всегда хорошо переносятся на другие записи, полученные с использованием другого оборудования, в другой популяции или при иных условиях регистрации.

**Ограничения классических подходов.** Классические методы сыграли ключевую роль в развитии автоматизированного анализа ЭКГ и до сих пор остаются важной частью многих практических систем. Их сильными сторонами являются интерпретируемость, вычислительная эффективность и возможность явно контролировать отдельные этапы обработки. Такие алгоритмы особенно удобны там, где требуется прозрачное объяснение решения или работа в условиях ограниченных вычислительных ресурсов.

В то же время классическая парадигма имеет принципиальные ограничения. Во-первых, качество системы зависит от ручной настройки параметров и набора эвристик. Во-вторых, ошибки ранних этапов, например детекции  $R$ -пиков или выделения границ волн, накапливаются и влияют на последующую классификацию. В-третьих, признаковые модели с трудом описывают сложные многоканальные зависимости 12-канальной ЭКГ без явного включения дополнительных экспертных правил. Эти ограничения стали одной из причин перехода к методам глубокого обучения, в которых признаки могут извлекаться непосредственно из исходного сигнала, а отдельные этапы анализа объединяются в единую обучаемую модель.

### 1.2.2 Глубокое обучение для классификации патологий

Развитие глубокого обучения существенно изменило постановку задачи автоматической интерпретации электрокардиограмм. В классических алгоритмах анализ ЭКГ строился как последовательность заранее заданных этапов: фильтрация сигнала, детектирование характерных точек, вычисление морфологических и интервальных признаков и последующее применение правил или классических классификаторов. Такой подход обладает высокой интерпретируемостью, однако его качество во многом определяется полнотой ручного описания сигнала и устойчивостью промежуточных алгоритмов. Нейросетевые модели, напротив, позволяют обучать представления непосредственно по исходным или минимально предобработанным временным рядам. В результате модель способна одновременно учитывать локальные морфологические признаки, такие как форма комплекса  $QRS$ , зубцов  $P$  и  $T$ , изменения сегмента  $ST$ , а также более

глобальные закономерности: ритмические паттерны, соотношения между отведениями и повторяемость патологических изменений во времени.

Переход к глубокому обучению стал возможен прежде всего благодаря росту объёма доступных цифровых ЭКГ и увеличению вычислительных ресурсов. Ранние открытые базы, такие как MIT-BIH Arrhythmia Database, сыграли важную роль в развитии алгоритмов анализа ритма, однако по современным меркам они имели ограниченный объём и не покрывали всё разнообразие клинических сценариев [32; 33]. Поэтому первые убедительные результаты глубокого обучения в ЭКГ появились в работах, использовавших существенно более крупные частные выборки. Одним из наиболее показательных примеров стала работа, в которой сверточная сеть обучалась для классификации аритмий по одноканальным амбулаторным ЭКГ-записям и достигала качества, сопоставимого с кардиологами, на специально размеченной тестовой выборке [13]. Эта линия исследований показала, что end-to-end модели способны извлекать диагностически значимые признаки даже из одного отведения, что особенно важно для носимых устройств, длительного мониторинга и сценариев, где регистрация полной 12-канальной ЭКГ невозможна.

Следующий этап был связан с переносом подходов глубокого обучения на стандартную клиническую 12-канальную ЭКГ. В отличие от одноканальных записей, 12-канальная ЭКГ содержит пространственно распределённую информацию об электрической активности сердца: разные отведения отражают разные проекции одного и того же процесса. Поэтому часть патологий проявляется не только во временной форме отдельных комплексов, но и в согласованных изменениях между группами отведений. В работе [34] была обучена нейронная сеть для автоматической диагностики шести классов нарушений по коротким 12-отведённым ЭКГ, собранным в крупной телемедицинской сети; обучение проводилось на массиве более чем из двух миллионов записей. Эта работа важна не только высоким качеством, но и демонстрацией практической применимости нейросетевых моделей к стандартной клинической форме ЭКГ.

Параллельно с ростом объёма частных клинических данных стали появляться открытые наборы данных и воспроизводимые протоколы сравнения. Это было принципиально важно для академического развития области, поскольку качество моделей, обученных на закрытых данных, сложно независимо проверить и сопоставить. Соревнования PhysioNet/Computing in Cardiology, в частности

Challenge 2020 по классификации 12-отведённых ЭКГ, стимулировали развитие методов многометочной классификации и привели к появлению большого числа воспроизводимых решений [35]. Однако ещё более важным опорным ресурсом для последующих исследований стал РТВ-XL — крупный открытый набор 12-отведённых клинических ЭКГ с иерархической системой диагностических утверждений SCP, рекомендованными разбиениями и доступными метаданными пациентов [14].

РТВ-XL закрепил типовую постановку задачи для современных нейросетевых моделей, обучаемых с учителем. В нём каждая запись представляет собой 12-канальную ЭКГ, а разметка имеет многометочный характер: одной записи может соответствовать несколько диагностических утверждений. Это делает задачу ближе к реальной клинической практике, где у пациента могут одновременно присутствовать нарушения ритма, проводимости и морфологические изменения. Кроме того, РТВ-XL поддерживает несколько уровней агрегации меток: от детальных SCP-утверждений до диагностических супер-классов. Благодаря этому один и тот же набор может использоваться как для грубой классификации крупных групп патологий, так и для более детальной диагностики.

На РТВ-XL был выполнен систематический анализ моделей [36]. В этой работе сравнивались классические признаковые методы и различные архитектуры нейросетевых моделей, а также рассматривались несколько типов задач: диагностическая классификация, предсказание ритмовых и формальных утверждений, оценка качества сигнала, а также задачи, связанные с возрастом и полом пациента. Важный вывод этого исследования состоит в том, что одномерные сверточные сети, в частности ResNet [37] и Inception-подобные [38] архитектуры, являются сильными базовыми моделями для анализа ЭКГ и превосходят методы, основанные на ручных признаках. Этот результат важен для настоящей работы, поскольку обосновывает использование компактных одномерных сверточных архитектур как сильной и воспроизводимой отправной точки для дальнейших экспериментов.

Наиболее распространённый способ применения нейросетевых моделей к 12-канальной ЭКГ состоит в том, чтобы рассматривать отведения как входные каналы, а временную ось — как основное направление свёртки. В такой постановке одномерные CNN извлекают локальные паттерны в каждом отведении и постепенно агрегируют информацию на всё больших временных масштабах.

Остаточные связи, используемые в ResNet-подобных моделях, позволяют увеличивать глубину сети без деградации обучения, а многомасштабные блоки, характерные для Inception-подобных архитектур, позволяют одновременно анализировать короткие события, такие как комплекс *QRS*, и более протяжённые изменения, такие как сегмент *ST* и зубец *T* [36; 39]. Поэтому такие модели стали естественной базовой постановкой для задач классификации 12-канальной ЭКГ.

Наряду со сверточными моделями развивались гибридные архитектуры, объединяющие CNN с рекуррентными слоями или механизмами внимания. CNN-часть в таких моделях обычно отвечает за извлечение локальных признаков, тогда как LSTM/GRU или attention-модули используются для агрегации более длинных временных зависимостей. Одним из ранних и широко цитируемых примеров такого подхода является модель, объединяющая свёрточные слои с LSTM-сетью для классификации нарушений ритма по одноканальным записям соревнования PhysioNet/CinC [40]. Такие подходы особенно естественны для длительных записей, где диагноз определяется не только формой отдельного комплекса, но и динамикой ритма во времени. В то же время для стандартных 10-секундных 12-канальных записей сильные сверточные модели часто оказываются конкурентоспособными и более простыми в обучении, что делает их удобными для систематических сравнений и последующих модификаций [41; 42].

Отдельное направление исследований связано с тем, как именно использовать информацию из 12 отведений. Простейшая стратегия состоит в объединении всех отведений в общий многоканальный вход. Однако с клинической точки зрения отведения не являются независимыми: часть из них отражает близкие анатомические области или связанные проекции электрической активности сердца. Поэтому в литературе рассматриваются модели, которые анализируют отдельные отведения или группы отведений, а затем объединяют их признаки с помощью взвешивания, механизмов внимания или специализированных блоков агрегации. Например, механизм многоканального внимания (multi-lead attention) может автоматически оценивать вклад каждого из 12 отведений в итоговое решение при детекции инфаркта миокарда [43], а группировка отведений по анатомически близким областям в сочетании с многоветвевой сверточной сетью повышает качество классификации по многим классам [44]. Такой подход поз-

воляет приблизить архитектуру к клинической логике интерпретации ЭКГ, при которой врач сопоставляет изменения в нескольких согласованных отведениях.

Вопрос числа отведений имеет и практическое значение. Полная 12-канальная ЭКГ является стандартом клинической диагностики, но во многих прикладных сценариях доступно только одно или несколько отведений: носимые устройства, домашний мониторинг, длительные амбулаторные записи. Исследования одноканальных моделей показывают, что для ряда задач, прежде всего связанных с ритмом, одного отведения может быть достаточно для высокого качества классификации [13]. Однако для морфологических диагнозов и нарушений проводимости информация из нескольких отведений часто критична. В работе [45] было показано, что модели, обученные на отдельных отведениях, в среднем уступают моделям, использующим все 12 отведений одновременно, хотя вклад разных отведений может существенно различаться в зависимости от патологии. Это наблюдение важно для дальнейшей постановки задач: с одной стороны, 12-канальная ЭКГ обеспечивает наиболее полную диагностическую информацию; с другой стороны, перенос знаний с 12 отведений на меньшее число каналов является практически значимой задачей для мобильных и носимых сценариев.

Ещё одним направлением развития моделей, обучаемых с учителем, стало использование метаданных пациентов. В клинической практике интерпретация ЭКГ редко выполняется изолированно от контекста: возраст, пол, клинические показания и сопутствующие состояния влияют как на априорную вероятность диагноза, так и на допустимые границы нормы. В открытых наборах, таких как РТВ-XL, доступны базовые метаданные пациента, что позволяет исследовать не только классификацию патологий, но и связь между сигналом и демографическими характеристиками [14; 36]. На практике метаданные пациентов непосредственно включаются в архитектуру модели: например, при предсказании смертности по 12-канальной ЭКГ в течение года возраст и пол пациента конкатенируются с признаками, извлечёнными свёрточной сетью из сигнала, что повышает качество прогноза по сравнению с использованием одного лишь сигнала [46]. Более того, отдельные работы показывают, что нейросетевые модели способны предсказывать возраст и пол непосредственно по 12-канальной ЭКГ [47]. Это указывает на то, что демографическая информация частично закодирована в самом сигнале и может влиять на обучение диагностических моделей. По-

этому включение метаданных может быть полезным, но требует осторожности: модель может улучшать формальные метрики за счёт демографических корреляций, не всегда связанных с причинными электрофизиологическими признаками. Для клинического применения важно проверять, повышает ли использование метаданных устойчивость модели на внешних данных, а не только качество на фиксированном тестовом разбиении.

Несмотря на значительный прогресс, глубокое обучение с учителем для ЭКГ сохраняет ряд ограничений. Во-первых, качество моделей сильно зависит от объёма и разнообразия обучающей выборки: увеличение числа записей и числа источников данных часто оказывается не менее важным, чем усложнение архитектуры. Во-вторых, модели, обученные на одном наборе, могут терять качество при переносе на другой из-за различий в устройствах регистрации, частоте дискретизации, фильтрации, популяции пациентов и правилах разметки. Эта проблема междатасетного сдвига особенно существенна для медицинских данных и требует отдельной экспериментальной проверки. В-третьих, дисбаланс классов и шум в экспертных метках усложняют как обучение, так и интерпретацию метрик качества.

**Предобученные foundation-модели для ЭКГ.** Параллельно с развитием архитектур нейросетевых моделей, обучаемых с учителем на отдельных наборах данных, в последние годы активно развивается направление масштабного предобучения моделей для ЭКГ. Общая идея таких подходов состоит в том, чтобы предварительно обучить модель на большом массиве электрокардиограмм, а затем адаптировать полученное представление к различным downstream-задачам.

Одним из примеров предварительного обучения с учителем является ECGFounder [48]. Модель на основе архитектуры RegNet была предварительно обучена более чем на 10 миллионах ЭКГ с диагностическими метками, после чего оценивалась на нескольких downstream-задачах.

Другую группу составляют модели, использующие самообучение. ECG-JERA [49] следует архитектуре joint-embedding predictive architecture (JERA): модель учится предсказывать скрытые представления одних фрагментов записи по другим непосредственно в латентном пространстве, без явной реконструкции исходного сигнала. ST-MEM [50] применяет пространственно-временное маскированное моделирование, вынуждая модель учитывать как временную структуру

отдельных отведений, так и зависимости между ними. NuBERT-ECG [51] адаптирует подход NuBERT [52], первоначально предложенный для речевых сигналов: дискретные псевдометки, полученные посредством кластеризации, используются как цели контекстуального предсказания. Отдельное направление представляют мультимодальные методы: например, MERL [53] сопоставляет ЭКГ-сигналы с текстовыми медицинскими заключениями, что позволяет использовать семантическую информацию из клинических текстов и выполнять zero-shot классификацию.

Несмотря на концептуальные преимущества, эффективность foundation-моделей в конкретной экспериментальной постановке не гарантирована. Качество зависит от близости данных предобучения к целевому домену, выбранного способа адаптации, числа доступных размеченных примеров и вычислительной сложности модели.

Таким образом, развитие нейросетевых моделей для анализа ЭКГ можно представить как последовательное движение от ручных признаков к end-to-end обучению, от небольших открытых баз к крупным клиническим коллекциям, от одноканальных сценариев умных часов к полноценной 12-канальной диагностике, а затем к масштабному предобучению foundation-моделей. Этот обзор непосредственно подводит к экспериментам следующей главы: необходимо не только сравнить современные архитектуры нейросетевых моделей и foundation-модели, но и выбрать воспроизводимую базовую модель, проверить её устойчивость при переносе между наборами данных, оценить эффект увеличения обучающей выборки и дообучения, а также исследовать дополнительные задачи анализа ЭКГ, такие как сегментация характерных элементов сигнала.

### 1.2.3 Нейросетевые методы сегментации PQRST

Классические методы детекции QRS-комплексов и выделения границ волн рассмотрены в п. 1.2.1; настоящий подраздел посвящён современным нейросетевым подходам к сегментации PQRST. В отличие от детекции QRS, полная сегментация предполагает согласованное определение начала и конца волны *P*, комплекса *QRS* и волны *T* и является более чувствительной к шумам, ва-

риативности морфологии и неоднозначности границ (особенно для окончания  $T$ -волны).

Практическая значимость сегментации определяется её ролью промежуточного этапа: на её основе вычисляются интервалы PR, QRS и QT, выполняется нормализация по ударам, формируются морфологические признаки и входы для downstream-задач, например оценки риска удлинения QT.

**Наборы данных для сегментации: QTDB и LUDB.** Для разработки и сопоставимой валидации алгоритмов сегментации ключевую роль играют открытые базы с ручной разметкой.

*QT Database (QTDB).* QTDB содержит 105 пятнадцатиминутных двухканальных записей, для которых вручную выставлены onset/peak/end маркеры волн  $P$ , QRS,  $T$  (и  $U$ , где присутствует) для поднабора ударов (порядка 30–50 ударов на запись) [54]. Такая структура делает QTDB удобной для оценки алгоритмов измерения QT и других интервалов, но ограничивает объём «плотной» разметки (аннотируется не весь сигнал, а выбранные удары), что важно учитывать при обучении нейросетей.

*Lobachevsky University Database (LUDB).* LUDB включает 200 записей длительностью 10 секунд, зарегистрированных в 12 стандартных отведениях с частотой дискретизации 500 Гц. Для каждой записи размечены границы и пики  $P$ , QRS и  $T$ ; при этом разметка выполнена кардиологами отдельно для каждого из 12 отведений [15]. В отличие от QTDB, LUDB обеспечивает полную разметку по всем отведениям, но имеет иные ограничения: малую длительность записей и относительно небольшой размер набора данных.

Несмотря на то что оба набора являются открытыми и широко применяются в литературе по делинеации, в настоящей работе используется только LUDB. Основная причина связана с характером разметки QTDB: аннотация в этом наборе выполнена лишь для ограниченного подмножества ударов на запись (порядка 30-50 из значительно большего числа комплексов в 15-минутной записи), а не для сигнала целиком, и представлена только в двух отведениях.

Классические алгоритмы делинеации, включая вейвлет-методы [24] и модели на скрытых марковских цепях [25], рассмотрены в п. 1.2.1, далее основное внимание уделяется нейросетевым подходам.

**Нейросетевые подходы: постановка задачи семантической сегментации.** Современная линия работ рассматривает делинеацию как задачу семантической сегментации: каждому отсчёту  $x[n]$  ставится в соответствие класс (например,  $\{\text{background}, P, QRS, T\}$ ), а границы волн извлекаются постобработкой из полученной разметки. Эта формулировка естественно приводит к архитектурам типа энкодер-декодер, например, U-Net [55], адаптированным к одномерным сигналам.

Показано, что U-Net-подобные модели могут обучаться даже на сравнительно небольших наборах (QTDB), если должным образом организовать разбиение и использовать регуляризацию/аугментации; такие работы демонстрируют конкурентоспособность нейросетевого подхода относительно классических подходов без использования нейросетей при корректной постановке [56]. Для задач сегментации также предлагаются компактные полностью сверточные сети, заявляющие устойчивость к различным частотам дискретизации и типам мониторинга [57].

С ростом интереса к автоматическому анализу многоканальных ЭКГ появляются нейросетевые модели, использующие LUDB как основной тестовый стенд: например, постановки 12-канальной «семантической сегментации» по LUDB и варианты пост-обработки, где учитывается морфологическая согласованность внутри удара сердца [58]. Более новые подходы комбинируют сегментацию нейросетью и последующую «морфологическую» коррекцию границ, чтобы уменьшить влияние локальных ошибок классификации отсчётов [59].

**Ограничения и практические вызовы.** Сегментация осложняется тем, что (i) границы  $P$  и конца  $T$  неоднозначны, (ii) реальные шумы (движение, дрейф изолинии) приводят к смещениям границ, и (iii) объём «золотой» разметки ограничен (QTDB не размечает весь сигнал, LUDB относительно мал по числу записей). Это приводит к необходимости либо комбинировать нейросетевые и другие идеи (постобработка, ограничения согласованности), либо использовать дополнительные источники данных и слабые формы разметки, чтобы повысить устойчивость к доменным сдвигам и помехам.

### 1.2.4 Самообучение для ЭКГ

Одним из основных ограничений подходов обучения с учителем является зависимость от экспертной разметки. В клинических архивах число неразмеченных ЭКГ обычно значительно превосходит объём данных с качественными диагностическими метками и тем более с точной разметкой волн и сегментов (P, QRS, T). Поэтому самообучение стало одним из перспективных направлений для анализа ЭКГ. В рамках самообучения модель обучается на вспомогательных задачах, среди которых условно можно выделить восстановление скрытых фрагментов, предсказание будущих представлений для временных данных, сопоставление разных представлений одного объекта и различение положительных и отрицательных пар.

Среди различных направлений самообучения в дальнейшем изложении теоретической и экспериментальной частей диссертации основное внимание уделяется контрастному обучению. В такой постановке обучающий сигнал определяется способом формирования положительных и отрицательных пар, что позволяет непосредственно включать в него различные виды сходства между ЭКГ-записями. Положительные пары могут формироваться не только с помощью искусственных аугментаций одного сигнала, но и на основе структуры многоканальной ЭКГ, связей между различными конфигурациями отведений и метаданных пациента. Именно расширение способов формирования положительных пар рассматривается далее в третьей главе.

Контрастное обучение является одним из наиболее известных классов самообучения. Общая идея состоит в том, чтобы сближать представления положительных пар и разделять представления различных примеров. В компьютерном зрении этот подход получил широкое распространение в SimCLR [60]; методы без явных отрицательных пар, такие как BYOL, показали, что устойчивые представления также могут обучаться с использованием асимметричных online- и target-сетей [61]. Для временных рядов важным направлением является Contrastive Predictive Coding, в котором модель учится предсказывать будущие латентные представления на основе предшествующего контекста [62].

Контрастные методы существенно зависят от выбора аугментаций, поскольку используемые преобразования определяют, к каким изменениям входно-

го сигнала обучаемое представление должно быть инвариантно. В компьютерном зрении набор допустимых аугментаций достаточно хорошо изучен, однако для физиологических временных рядов выбор преобразований требует дополнительной осторожности: аугментация должна изменять сигнал, не разрушая его клинически значимые свойства. Для временных рядов в целом и для ЭКГ в частности применяются масштабирование амплитуды, временной сдвиг, добавление гауссова шума, временное маскирование фрагментов и перестановка временных сегментов [63; 64]. При этом такие преобразования, как инверсия сигнала или агрессивное временное сжатие и растяжение, способны изменять диагностически значимые характеристики ЭКГ.

Применительно к ЭКГ контрастная постановка позволяет использовать не только универсальные аугментации временных рядов, но и структуру самого сигнала. Характерным примером является CLOCS — *contrastive learning of cardiac signals across space, time and patients* [65]. В этом подходе положительные пары формируются с учётом пространственных и временных отношений между ЭКГ-сигналами, а также принадлежности записей одному пациенту. Другой пример — метод 3KG [66], в котором для 12-канальной ЭКГ используются физиологически мотивированные пространственные преобразования, основанные на представлении отведений как различных проекций электрической активности сердца. Эти преобразования применяются совместно с временными аугментациями при контрастном обучении. Таким образом, структура ЭКГ сама по себе может использоваться для определения различных представлений одной записи без привлечения диагностической разметки.

Параллельно развиваются контрастные подходы, специально адаптированные к временной структуре сигналов. Метод TS-TCC (*Time-Series Representation Learning via Temporal and Contextual Contrasting*) [67] для каждой исходной записи формирует два представления с помощью слабой и сильной аугментаций. Оба представления проходят через общий энкодер, после чего обучение включает две взаимодополняющие составляющие. *Temporal contrasting* использует кросс-предсказание между двумя представлениями: контекст, построенный по одному из них, используется для предсказания будущих латентных представлений другого, что стимулирует модель учитывать временные зависимости. *Contextual contrasting*, в свою очередь, обеспечивает согласованность контекстных представлений двух аугментированных версий одной записи относительно других

записей. Таким образом, TS-TCC расширяет стандартную контрастную постановку за счет явного моделирования временных зависимостей сигнала.

Метод TF-C (Time-Frequency Consistency) [68] использует представления сигнала одновременно во временной и частотной областях. В процессе предобучения обеспечивается согласованность между представлениями этих двух типов, что позволяет учитывать свойства временного ряда, проявляющиеся как непосредственно в форме сигнала, так и в его частотных характеристиках.

Другой важный класс методов самообучения составляют маскированное моделирование и реконструктивные подходы. Их идея состоит в том, чтобы скрыть часть входных данных и обучить модель восстанавливать недостающую информацию [69]. Для 12-канальной ЭКГ маскирование может выполняться как по временной оси, так и по отведениям, что позволяет модели изучать временные и межотведенческие зависимости. ST-MEM является примером такого пространственно-временного маскированного моделирования для ЭКГ [50]. В отличие от контрастных методов, здесь не требуется явно задавать положительные и отрицательные пары, а структура представления формируется через задачу восстановления скрытых частей сигнала.

Несмотря на развитие специализированных методов, выбор способа формирования положительных пар в контрастном обучении остаётся самостоятельной исследовательской задачей. В большинстве базовых контрастных методов положительные пары строятся из различных аугментаций одной и той же записи. Вместе с тем клинические ЭКГ-архивы содержат дополнительные структурированные метаданные, например возраст и пол пациента, которые также могут отражать физиологическое сходство между записями. Исследования показывают, что нейросетевые модели способны предсказывать возраст и пол непосредственно по ЭКГ [47], что указывает на наличие соответствующей информации в самом сигнале. Существующие методы уже используют отдельные виды специфичной для ЭКГ структуры [65; 66; 70], однако возможность формировать положительные пары из записей *разных* пациентов на основании их демографической близости, не используя диагностические метки, исследована значительно меньше.

Другой практически значимой задачей является перенос информации между различными конфигурациями отведений. Полная 12-канальная ЭКГ содержит больше пространственной информации, чем сокращённые конфигурации,

тогда как в носимых и мобильных устройствах часто доступно только одно или несколько отведений. Поэтому возникает задача использовать знания, доступные при анализе полной 12-канальной записи, для улучшения моделей, работающих с меньшим числом каналов.

В общем случае перенос знаний от более информативной модели к модели с ограниченным входом может выполняться с помощью дистилляции знаний, при которой модель-ученик обучается воспроизводить выходы или внутренние представления более сложной модели-учителя. Близкие по смыслу задачи возникают и при работе с ЭКГ: в литературе рассматриваются методы анализа записей с различным числом отведений, а также архитектурные решения, позволяющие учитывать структуру отдельных отведений и их групп [42; 44]. В таких подходах перенос информации обычно задаётся через архитектуру модели, адаптацию между доменами или непосредственное согласование представлений.

Контрастное обучение предоставляет альтернативный способ постановки этой задачи. Различные наборы отведений одной 12-канальной записи можно рассматривать как несколько связанных представлений одного и того же сердечного цикла и использовать их в качестве положительных пар. В этом случае модель обучается формировать представление, устойчивое к изменению доступного набора отведений. В отличие от стандартных аугментаций, положительная пара здесь формируется не искусственным преобразованием значений сигнала, а выбором реально существующих и физически связанных отведений одной ЭКГ. Такой подход позволяет использовать структуру полной многоканальной записи на этапе предобучения, а затем переносить полученное представление на задачи с сокращённым числом отведений. Эта постановка рассматривается далее в третьей главе для переноса между 12-канальными и 1-канальными ЭКГ.

Самообучение не является единственным способом использования больших массивов неразмеченных медицинских данных. Смежное направление образует обучение с частичным привлечением учителя, при котором предполагается наличие небольшой размеченной выборки и значительно большего объёма неразмеченных данных. Одним из базовых подходов здесь является псевдоразметка [71]: модель, обученная на доступных размеченных данных, используется для получения меток неразмеченных примеров, после чего наиболее надёжные из этих предсказаний включаются в дальнейшее обучение. Эта идея получила развитие в подходах типа «учитель-ученик». Например, Noisy Student [72]

использует модель-учителя для получения псевдометок и обучает модель-ученика с дополнительными аугментациями и шумом. Mean Teacher формирует более стабильную модель-учителя с помощью экспоненциального усреднения весов [73]. FixMatch сочетает псевдоразметку по слабо аугментированному примеру с требованием согласованности предсказания для его сильно аугментированной версии [74].

Для задач сегментации ЭКГ псевдоразметка представляет особый интерес из-за существенного различия между объёмами размеченных и неразмеченных данных. Открытые базы с экспертной разметкой границ волн сравнительно невелики: LUDB содержит 200 записей [15], а QTDB — 105 записей [54]. В то же время клинические архивы могут содержать на несколько порядков больше ЭКГ-данных без разметки для задачи сегментации. Обученная на экспертной разметке модель может использоваться для автоматического определения границ волн на таких записях, после чего полученные псевдометки могут быть использованы для расширения обучающей выборки. Основным ограничением этого подхода является возможность накопления систематических ошибок исходной модели, поэтому качество псевдоразметки и способ отбора автоматически размеченных примеров непосредственно влияют на результат последующего обучения. Использование такой схемы для расширения данных при сегментации ЭКГ рассматривается далее во второй главе.

### 1.3 Федеративное обучение

Цифровые медицинские данные, как правило, не сосредоточены в едином датацентре. ЭКГ-записи хранятся в информационных системах отдельных клиник, диагностических центров, региональных архивов и устройств мониторинга. Централизованное объединение таких данных часто ограничено законодательством, требованиями конфиденциальности, локальными регламентами и техническими различиями инфраструктуры. Федеративное обучение (federated learning, FL) предлагает альтернативную постановку централизованному обучению нейросетевых моделей: данные остаются у клиентов, а общий алгоритм обучает модель за счёт обмена параметрами или обновлениями модели [75—77].

В федеративном обучении принято различать две принципиально разные постановки: cross-device, где число клиентов может исчисляться миллионами мобильных устройств с нестабильным сетевым подключением, и cross-silo, где число клиентов сравнительно невелико (единицы–десятки), а каждый клиент представляет собой организацию с относительно надёжной инфраструктурой (например, отдельную клинику или больницу) и располагает большим набором локальных данных [77]. Медицинские приложения федеративного обучения относятся именно к cross-silo сценарию: число участвующих клиник ограничено, а требования к надёжности и защите канала связи выше, чем в массовых cross-device системах. В классической постановке задачи федеративного обучения нейросетевой модели имеется  $M$  клиентов, каждый из которых хранит локальный набор данных  $\mathcal{D}_m$ . Задача обучения состоит в минимизации функции

$$\min_{\theta} F(\theta) = \sum_{m=1}^M p_m F_m(\theta), \quad F_m(\theta) = \frac{1}{|\mathcal{D}_m|} \sum_{(x,y) \in \mathcal{D}_m} \ell(f_{\theta}(x), y), \quad (1.8)$$

где  $p_m$  обычно пропорционален размеру локального набора,  $f_{\theta}$  — модель, а  $\ell$  — функция потерь.

Практическая постановка федеративного обучения получила широкое распространение после работы FedAvg [76]. В работе сформулирована практичная схема решения задачи (1.8) без необходимости в централизованном хранении данных и остаётся отправной точкой практически для всех последующих методов, включая рассматриваемые далее в этой главе подходы к повышению устойчивости. Федеративное обучение нейросетевой модели в работе FedAvg выполняется в рамках клиент-серверной архитектуры: сервер рассылает текущие параметры (веса) нейросетевой модели клиентам, клиенты выполняют несколько локальных шагов оптимизации на своих данных и возвращают на сервер обновленные значения параметров. После этого сервер выполняет агрегацию полученных от клиентов обновлений, то есть объединяет их в новую глобальную модель. В FedAvg такая агрегация представляет собой взвешенное усреднение, где результирующие веса определяются размерами локальных выборок.

$$\theta^{t+1} = \sum_{m \in S_t} \frac{n_m}{\sum_{j \in S_t} n_j} \theta_m^{t+1}, \quad (1.9)$$

где  $S_t$  — подмножество клиентов, участвующих в раунде  $t$ , а  $\theta_m^{t+1}$  — параметры модели клиента  $m$  после выполнения им нескольких локальных шагов градиентного спуска, стартовавших с текущих глобальных параметров  $\theta^t$ . Эта схема проста, хорошо масштабируется и стала базовой точкой отсчёта для большого числа FL-исследований.

Для медицины федеративное обучение представляет непосредственную практическую ценность: оно позволяет использовать разнообразие данных отдельных клиентов без необходимости их передачи на агрегирующий сервер. Обзорные работы подчёркивают, что такая схема особенно важна для решения задач цифрового здравоохранения [78—81]. Для ЭКГ это означает возможность обучать модель на данных нескольких клиник, сохраняя локальный контроль над записями. Недавние исследования по анализу ЭКГ на данных нескольких больниц показывают, что федеративное обучение может достигать качества, сопоставимого с централизованным обучением, и приносить выгоду клиникам с меньшими локальными выборками [82; 83].

Однако федеративное обучение не сводится к техническому усреднению весов моделей. Важным аспектом федеративного обучения является статистическая гетерогенность данных клиентов, обозначаемая как non-IID (non-independent and identically distributed): данные клиентов, как правило, получены из различных распределений. В медицинской постановке это проявляется в различиях характеристик пациентов, распространённости патологий или качества регистрируемых сигналов. Вследствие этого локальные целевые функции клиентов различаются, а соответствующие им градиентные обновления могут иметь существенно расходящиеся направления. Это явление, известное как дрейф клиентов (client drift), снижает эффективность агрегации и замедляет сходимость глобальной модели [77; 84].

Качество обучения глобальной модели также зависит от того, можно ли доверять обновлениям параметров нейронной сети, получаемым от клиентов. В классической постановке федеративного обучения предполагается, что клиенты корректно выполняют локальное обучение и передают обновления. В то же время в литературе рассматриваются атаки на процесс федеративного обучения, при которых один или несколько клиентов намеренно изменяют локальные данные, метки или непосредственно передаваемые обновления с целью снижения качества глобальной модели. Более общая модель угроз описывает так называ-

емых «византийских клиентов» — участников, способных отправлять серверу произвольные обновления [85].

Некорректные обновления параметров обучаемой нейронной сети могут возникать и без злонамеренного поведения. В медицинской среде их причиной могут быть ошибки разметки данных, различия в локальной предобработке, неисправности оборудования, программные сбои. Такие клиенты также могут формировать обновления, существенно отличающиеся от обновлений остальных участников. Поэтому методы устойчивости федеративного обучения представляют интерес не только в вопросах защиты от целенаправленных атак, но и в качестве средства снижения влияния недостаточно качественных локальных данных.

Для снижения влияния византийских клиентов на результирующую нейросетевую модель предложено большое число методов устойчивой агрегации. В отличие от FedAvg, где все клиентские обновления участвуют во взвешенном усреднении, такие методы изменяют правило агрегации таким образом, чтобы отдельные аномальные или злонамеренные обновления не могли существенно повлиять на глобальную модель. Эти подходы можно условно разделить на методы, основанные на взаимном расположении клиентских обновлений, статистиках оценки устойчивости и дополнительной оценке надёжности каждого обновления.

Одним из первых методов является Krum [85]. Для каждого клиентского обновления он оценивает расстояния до обновлений остальных участников и выбирает то обновление, которое наиболее близко к группе других клиентов. Тем самым предполагается, что корректные клиенты формируют согласованную группу, тогда как злонамеренные обновления находятся дальше от неё. Другой класс подходов выполняет устойчивую агрегацию отдельно по координатам. Coordinate-wise median заменяет обычное среднее медианой значений каждой координаты, а trimmed mean предварительно исключает несколько наибольших и наименьших значений и усредняет оставшиеся [86]. Метод Bulyan [87] сочетает предварительный отбор клиентских обновлений с последующей устойчивой по координатам агрегацией. Такие методы позволяют ограничить влияние части византийских клиентов без использования дополнительного доверенного набора данных, однако их эффективность зависит от предположений о числе атакующих клиентов и характере распределения корректных обновлений. При стати-

стической неоднородности данных честные клиентские обновления сами могут существенно различаться, что усложняет отделение их от аномальных.

Другую группу составляют методы, которые дополнительно оценивают полезность или надёжность каждого клиентского обновления и используют эту оценку при агрегации. В Zeno [88] сервер оценивает клиентские обновления по тому, насколько применение соответствующего шага уменьшает функцию потерь на небольшой серверной выборке; обновления с низкими оценками исключаются из агрегации. Такой подход позволяет не полагаться исключительно на геометрическое сходство клиентских обновлений.

В FLTrust [89] небольшой доверенный набор данных хранится непосредственно на сервере. На каждом раунде сервер самостоятельно вычисляет эталонное обновление, после чего сравнивает с ним направления клиентских обновлений. Неотрицательная косинусная близость используется как вес при агрегации: чем сильнее направление клиентского обновления согласуется с серверным, тем больше его вклад в глобальное обновление. Дополнительно нормы клиентских обновлений приводятся к норме серверного обновления, что ограничивает возможность атакующего увеличить своё влияние простым масштабированием вектора.

Более поздние методы стремятся противодействовать адаптивным атакам отравления модели (model poisoning), при которых атакующий специально формирует передаваемое обновление так, чтобы изменить поведение глобальной модели. Например, RECESS [90] не ограничивается анализом уже присланных обновлений: сервер формирует специальные проверочные обновления и анализирует реакцию клиентов на них, после чего накопленная оценка надёжности клиента используется при последующей агрегации. Другой класс подходов, включая centered clipping [91], ограничивает влияние обновлений, существенно отклоняющихся от текущего центра агрегации.

Таким образом, существующие методы устойчивого федеративного обучения в основном рассматривают ситуацию, в которой часть клиентов формирует аномальные обновления либо вследствие целенаправленной атаки, либо вследствие нарушения предположений стандартной процедуры локального обучения. Для медицинского федеративного обучения интерес представляет случай, когда клиент в целом корректно выполняет локальное обучение, однако в разметке его локальных данных содержатся систематические ошибки.

## 1.4 Выводы

В первой главе рассмотрено развитие подходов к автоматизированному анализу электрокардиограмм — от классических алгоритмов цифровой обработки сигналов и методов, основанных на ручном выделении признаков, до современных нейросетевых моделей и методов самообучения и федеративного обучения. Классические методы, основанные на предварительно сконструированных признаках, обладают преимуществами в виде интерпретируемости и сравнительно низкой вычислительной сложности, однако существенно зависят от ручной настройки и качества отдельных этапов обработки сигнала. Современные модели глубокого обучения, напротив, позволяют извлекать признаки непосредственно из исходного сигнала и достигать высокого качества в задачах классификации патологий и сегментации ЭКГ, но предъявляют более высокие требования к объёму и качеству обучающих данных и могут быть чувствительны к различиям между источниками данных.

Для задачи классификации патологий по ЭКГ-сигналу рассмотрены как сверточные архитектуры нейронных сетей, так и современные предобученные модели. Предобучение позволяет задействовать большие объёмы медицинских данных и формировать универсальные представления сигнала, однако само по себе не гарантирует достижения достаточного уровня качества при решении конкретной целевой задачи. Качество результирующей нейросетевой модели определяется в том числе близостью распределений данных предобучения и целевой выборки и объёмом доступных размеченных данных. Вследствие этого непосредственное сопоставление результатов, получаемых с помощью предобученных нейросетевых моделей и моделей, обученных на целевом клиническом наборе данных, составляет одну из задач настоящей работы.

Существенным ограничением при обучении нейросетевых моделей может стать объем обучающей выборки, в частности, при решении задачи сегментации ЭКГ. Открытые наборы данных LUDB и QTDB содержат лишь несколько сотен размеченных записей, тогда как объём доступных клинических архивов с частичной разметкой может быть существенно больше. Анализ и разработка подхода для работы с ограниченным набором размеченных данных для задачи сегментации ЭКГ рассматриваются во второй главе.

Отдельное направление связано с использованием неразмеченных данных в рамках самообучения нейросетевых моделей. Расширение обучающей выборки может осуществляться не только с использованием универсальных методов аугментации временных рядов, применимых к сигналам различной природы, но и с применением специализированных преобразований, учитывающих структурные и физиологические особенности ЭКГ. Вместе с тем сохраняются возможности для дальнейшего расширения способов формирования положительных пар. В частности, демографические метаданные о пациентах могут применяться в качестве дополнительного критерия физиологического сходства между ЭКГ-записями пациентов. Исследованию данного подхода к формированию положительных пар 12-канальных ЭКГ в рамках самообучения посвящена первая часть третьей главы.

Практически значимой задачей является перенос информации между различными конфигурациями отведений. 12-канальная ЭКГ содержит более полную пространственную информацию об электрической активности сердца, тогда как в мобильных и носимых устройствах часто доступно только одно отведение. Существующие методы автоматического анализа ЭКГ используют различные способы работы с малым числом отведений, включая архитектурные решения, адаптацию представлений и дистилляцию знаний. В настоящей работе рассматривается другой подход к использованию знания о структуре многоканальной ЭКГ: различные наборы отведений одной записи формируют связанные представления одного и того же состояния сердца и могут использоваться как положительные пары при контрастном обучении.

Отдельный круг задач возникает при обучении нейросетевых моделей в случае распределённого хранения обучающих медицинских данных. Федеративное обучение позволяет обучать общую модель без непосредственной передачи ЭКГ между медицинскими организациями, однако результат такого обучения зависит, в том числе, от статистической неоднородности локальных выборок и невозможности абсолютного доверия к клиентским обновлениям. Существующие методы устойчивого федеративного обучения используют различные механизмы снижения влияния клиентов с «аномальным» поведением. В рамках медицинского домена «аномалия» на клиенте может возникать не вследствие его злонамеренного поведения, а из-за ошибок в разметке его локального набора данных.

Проведённый обзор определяет основные направления последующего исследования. Во второй главе выполняется обоснованный выбор и экспериментальная оценка нейросетевых моделей классификации и сегментации ЭКГ, их переносимость между наборами данных, а также подходы к обучению в условиях сравнительно небольшого объема размеченных данных. В третьей главе исследуются способы применения сведений о пациенте и многоканальной структуры ЭКГ при контрастном обучении. В четвертой главе рассматривается федеративное обучение в условиях невозможности абсолютного доверия к обновлениям, поступающим от клиентов.

## **Глава 2. Методы и модели автоматизированного анализа электрокардиограмм в 12 отведениях**

В первой части второй главы выполняется сравнительный анализ нейросетевых моделей классификации, осуществляется выбор базовой модели и исследуется её устойчивость при варьировании объёма и источника обучающих данных. Данная часть формирует экспериментальную основу для последующего исследования. Во второй части главы рассматривается задача сегментации, излагаются первое положение, выносимое на защиту, и первый пункт научной новизны. Предлагается метод определения границ основных морфологических элементов ЭКГ, основанный на модели ECG-CODE и self-training подходе с автоматически формируемой псевдоразметкой дополнительных записей, что позволяет повышать качество сегментации при ограниченном объёме экспертной разметки.

### **2.1 Классификация электрокардиограмм в 12 отведениях по патологическим состояниям**

В настоящем разделе рассматривается задача автоматической классификации электрокардиограмм в 12 отведениях по патологическим состояниям на основе нейросетевых методов. Цель раздела — выбор базовой архитектуры модели для последующих вычислительных экспериментов, а также анализ устойчивости выбранной модели на наборах данных различного происхождения.

В последние годы в области анализа ЭКГ активно развиваются foundation-модели, предварительно обученные на больших массивах сигналов и ориентированные на последующую адаптацию к различным диагностическим задачам. К таким моделям относятся, в частности, ECGFounder, ECG-JERA, ST-MEM и HuBERT-ECG [48—51]. Предполагается, что масштабное предварительное обучение позволяет формировать более универсальные представления ЭКГ-сигналов и за счёт этого повышать качество решения прикладных задач после дообучения на целевых данных.

Однако наличие предварительного обучения на большом наборе данных само по себе не гарантирует превосходства модели в конкретной экспериментальной постановке. Итоговое качество классификации может зависеть от близости данных предварительного обучения к целевому набору, выбранного способа дообучения, числа доступных размеченных примеров, архитектурных ограничений и вычислительной сложности модели. Поэтому в данной работе foundation-модели рассматриваются не как безусловно лучший вариант, а как отдельная группа подходов, требующая прямого экспериментального сравнения с нейросетевыми моделями, обучаемыми с учителем.

В качестве альтернативных моделей сравнения используются две нейросетевые архитектуры меньшего размера: `resnet1d50` и `densenet1d121`. Обе модели относятся к классу одномерных сверточных сетей и обучаются непосредственно на целевых данных. Их включение в сравнение обусловлено не только меньшим числом параметров по сравнению с foundation-моделями, но и тем, что сверточные архитектуры на сегодняшний день остаются одним из наиболее проверенных и эффективных решений для анализа ЭКГ-сигналов [34; 36]. Сверточные фильтры естественным образом настраиваются на выделение локальных морфологических закономерностей и делают это устойчиво независимо от точного положения паттерна во времени, что хорошо согласуется со структурой сердечного цикла. На архитектурах этого класса, таких как одномерные варианты ResNet и DenseNet [37; 92], а также их модификациях, в литературе неоднократно получены результаты, сопоставимые по качеству распознавания болезней сердца с уровнем эксперта-кардиолога [13]. Использование указанных архитектур позволяет оценить, насколько специализированные модели, обучаемые с нуля на конкретной задаче, уступают или, напротив, превосходят более крупные предварительно обученные модели при классификации 12-канальных ЭКГ. Для одномерных ЭКГ-сигналов такие свойства особенно важны, поскольку модель должна устойчиво извлекать локальные морфологические паттерны зубцов, интервалов и комплексов.

Таким образом, в разделе последовательно решаются три задачи. Во-первых, выполняется сравнение foundation-моделей и компактных сверточных архитектур в единой экспериментальной постановке и выбирается базовая архитектура. Во-вторых, для выбранной базовой архитектуры анализируется зависимость качества классификации от размера обучающей выборки. Наконец, исследуется

устойчивость выбранной модели при переносе между наборами данных различного происхождения и возможность её дообучения для повышения качества классификации.

### **2.1.1 Анализируемые наборы данных для классификации**

Для решения задачи классификации электрокардиограмм в 12 отведениях в работе используются два набора данных: открытый набор РТВ-XL [14] и закрытый клинический набор ISP. Выбор указанных наборов обусловлен необходимостью исследовать как поведение моделей на широко используемом публичном датасете, так и их применимость к более крупным и гетерогенным клиническим данным. РТВ-XL представляет собой один из наиболее распространённых открытых наборов данных для анализа 12-канальных ЭКГ, тогда как ISP отражает более сложные условия клинической практики, где записи отличаются по устройствам регистрации, параметрам сигнала и условиям получения данных.

Набор РТВ-XL включает 21 799 записей ЭКГ длительностью 10 секунд, зарегистрированных с частотой дискретизации 500 Гц. Записи этого набора предварительно обработаны полосовым фильтром для уменьшения уровня шума и размечены не более чем двумя кардиологами. Благодаря открытости, стандартизованному формату и опубликованной схеме разбиения на обучающую, валидационную и тестовую части РТВ-XL широко используется как один из основных бенчмарков для сравнения алгоритмов классификации 12-канальных ЭКГ [36]. Набор данных имеет сравнительно небольшой объём, фиксированные параметры регистрации и относительно однородные условия сбора данных. Эти особенности необходимо учитывать при анализе обобщающей способности моделей.

Набор данных ISP, напротив, представляет собой значительно более крупную клиническую коллекцию 12-канальных ЭКГ, собранных в нескольких медицинских учреждениях в течение более чем двух лет. Полный массив содержит более 1,5 млн записей. В отличие от РТВ-XL, данные ISP изначально не подвергались унифицированной предварительной обработке и характеризуются выраженной неоднородностью по типам регистрирующих устройств, частоте дискретизации, длительности записей и клиническим условиям получения сигнала.

Существенным преимуществом набора ISP является способ аннотирования: заболевания указываются врачами непосредственно в процессе описания ЭКГ из стандартизованного списка, что позволяет избежать дополнительного этапа извлечения меток из текстовых заключений и связанных с этим ошибок.

Для обеспечения корректного сравнения моделей форматы записей двух наборов были приведены к единому виду. Из набора ISP были отобраны записи длительностью от 9 до 10 секунд, зарегистрированные с частотой 500 или 1000 Гц. После этого все сигналы были передискретизированы к частоте 500 Гц и дополнены нулями до длины 10 секунд. В тестовую часть ISP включались только 10-секундные записи с частотой 500 Гц. В анализ также включались только пациенты старше 18 лет. Записи от одного и того же пациента попадали только в одну из выборок: обучающую, валидационную или тестовую. Для исключения технически некорректных записей на этапе предобработки удалялись ЭКГ, в которых хотя бы в одном отведении наблюдалось постоянное значение сигнала. Наконец, для устранения различий в амплитудном масштабе между наборами данных к обоим наборам применялась  $z$ -нормализация. В результате в наборе данных ISP были отобраны 900 000 записей для обучения и 100 000 записей для теста.

Сравнение используемых наборов данных показывает, что PTB-XL и ISP различаются по нескольким принципиальным характеристикам: объёму выборки, степени гетерогенности, числу врачей, участвовавших в аннотировании, а также по условиям регистрации и предварительной обработке сигналов. Эти различия делают выбранную пару наборов данных удобной для исследования междатасетного сдвига. Основные характеристики наборов данных приведены в таблице 2.

Таблица 2 — Сравнение наборов данных PTB-XL и ISP, используемых в задаче классификации ЭКГ

| Характеристика                      | PTB-XL               | ISP                                      |
|-------------------------------------|----------------------|------------------------------------------|
| Тип набора данных                   | открытый             | закрытый клинический                     |
| Рассматриваемый объём данных        | 21 799 записей       | более 1,5 млн записей                    |
| Использовано в обучении (train+val) | 19 601               | 900 000                                  |
| Использовано в тестировании (test)  | 2 198                | 100 000                                  |
| Предварительная обработка           | полосовая фильтрация | исходно отсутствует единая предобработка |
| Число врачей-аннотаторов            | до 2                 | более 130                                |

### 2.1.2 Сравнение моделей

В данном подразделе выполняется сравнительный анализ нейросетевых архитектур для бинарной классификации патологических состояний по электрокардиограммам в 12 отведениях. Сравнение необходимо не только для определения модели с наилучшим качеством, но и для выбора базовой архитектуры, которая далее используется в экспериментах по анализу устойчивости к междатасетному сдвигу и дообучению на целевых данных в последующих разделах и главах.

В литературе для автоматической классификации электрокардиограмм в 12 отведениях используются различные классы нейросетевых архитектур: рекуррентные сети (LSTM, GRU) [93; 94], ориентированные на моделирование временной динамики сигнала; сверточные сети, включая остаточные (ResNet) и сети с плотными связями (DenseNet), эффективно выделяющие локальные морфоло-

гические признаки ЭКГ; трансформерные архитектуры, применяющие механизм self-attention для анализа длительных зависимостей в последовательности [95]; а в последние годы – foundation-модели, предварительно обученные на больших массивах ЭКГ и адаптируемые под конкретную задачу дообучением (см. также обзор в главе 1). Поскольку в рамках данной работы решается прикладная задача выбора базовой архитектуры нейросетевой модели для дальнейшей модификации и развития алгоритма, сравнение сфокусировано на двух группах подходов: компактных сверточных моделях, обучаемых непосредственно на целевых данных, и крупных предварительно обученных foundation-моделях.

В сравнительный анализ включены две группы архитектур нейросетевых моделей, различающиеся способом получения признакового представления ЭКГ-сигнала. Первая группа — foundation-модели: ECGFounder, ECG-JERA, ST-MEM и HuBERT-ECG. Эти модели используют различные стратегии предварительного обучения, включая обучение представлений на больших массивах ЭКГ, маскирование частей входного сигнала и подходы на основе совместного пространства представлений. Общей особенностью данных подходов является стремление получить универсальное представление ЭКГ-сигнала, пригодное для последующей адаптации к различным диагностическим задачам. Вторая группа — одномерные сверточные модели resnet1d50 и densenet1d121. В отличие от foundation-моделей, эти архитектуры не опираются на внешнее предварительное обучение и обучаются непосредственно в рамках рассматриваемой задачи классификации. При этом они имеют меньшее число параметров и более простую схему применения, что делает их удобными кандидатами на роль базовой архитектуры нейросетевой модели для дальнейших исследований.

Сравнение проводится в постановке бинарной классификации отдельных патологических состояний. Такой выбор обусловлен двумя причинами. Во-первых, бинарная постановка позволяет модели сосредоточиться на распознавании конкретной патологии без дополнительной сложности, связанной с многоклассовой или мультиметочной постановкой. Во-вторых, это обеспечивает более интерпретируемое и наглядное сравнение архитектур нейросетевых моделей по отдельным клинически значимым состояниям. Поскольку спектр сердечных патологий широк, для эксперимента были выбраны четыре патологии, отражающие разные типы нарушений сердечной деятельности и разный характер диагностических признаков: фибрилляция предсердий (AFIB) — нарушение рит-

ма, распознаваемое по нерегулярности сердечных сокращений; полная блокада левой ножки пучка Гиса (CLBBB) — нарушение проводимости с выраженной и устойчивой деформацией морфологии комплекса  $QRS$ ; атриовентрикулярная блокада первой степени (1AVB) — нарушение проводимости, диагностируемое по удлинению интервала  $PR$  и потому требующее точной оценки временных, а не только амплитудных характеристик сигнала; и желудочковая экстрасистолия (PVC) — эктопический комплекс, отличающийся по форме и происхождению от синусовых сокращений. Такой набор позволяет проверить, насколько устойчиво работает каждая архитектура как на выраженных амплитудно-морфологических, так и на временных диагностических признаках.

Все модели обучаются на одних и тех же обучающих данных и оцениваются на одной и той же тестовой выборке в едином протоколе. Данные разбиваются на обучающую, валидационную и тестовую части без пересечения по пациентам, что исключает утечку информации между подвыборками. Используются единая процедура предобработки сигналов, бинарная кросс-энтропия в качестве функции потерь с учётом дисбаланса классов за счёт взвешивания положительного класса, а также единые метрики качества. Такой протокол позволяет минимизировать влияние различий в данных и сопоставить выбранные модели в единой постановке.

В предварительном сравнительном эксперименте все модели обучались на одной и той же случайной подвыборке, составляющей около 10% обучающей части набора ISP. Такой объём был выбран как компромисс между вычислительной стоимостью сравнения и возможностью получить устойчивые оценки качества. Отдельного пояснения требует режим адаптации foundation-моделей. Для них предобученный энкодер фиксировался, а обучение на целевой задаче проводилось только для классификационной головы. Компактные модели `resnet1d50` и `densenet1d121`, напротив, обучались только на целевых данных.

### 2.1.3 Анализ моделей для классификации

Ниже приведены результаты сравнения двух компактных сверточных моделей (`resnet1d50`, `densenet1d121`) и четырёх foundation-моделей

(ECGFounder, ECG-JEPA, ST-MEM, HuBERT-ECG), описанных в предыдущем подразделе. Поскольку в последующих разделах работы предполагается модификация и развитие базового алгоритма, требуется выбрать архитектуру нейросетевой модели, которая, с одной стороны, обеспечивает наилучшее качество классификации в рассматриваемой постановке, а с другой — сохраняет достаточную универсальность и удобство для дальнейших экспериментальных исследований.

В качестве основной метрики качества используется геометрическое среднее чувствительности и специфичности:

$$G-mean = \sqrt{Se \cdot Sp},$$

где  $Se$  — чувствительность, а  $Sp$  — специфичность. Выбор данной метрики обусловлен особенностями медицинской диагностики. В задачах автоматического анализа ЭКГ критически важно одновременно обеспечивать выявление пациентов с патологией и не допускать избыточного числа ложноположительных срабатываний. В отличие от общей точности классификации, которая может быть завышена при выраженном дисбалансе классов, метрика  $G-mean$  более корректно отражает баланс между чувствительностью и специфичностью. Именно поэтому при выборе базовой архитектуры нейросетевой модели основное внимание уделяется указанной метрике.

Тестирование всех рассматриваемых моделей проводилось на тестовой части ISP. Результаты сравнения по каждой из выбранных патологий представлены в таблице 3.

Таблица 3 — Сравнение архитектур нейросетевых моделей на тестовой части набора данных ISP

| Модель        | AFIB  |       |              | CLBBB |       |              | 1AVB  |       |              | PVC   |       |              |
|---------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|
|               | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       |
| resnet1d50    | 0.875 | 0.988 | 0.930        | 0.908 | 0.994 | 0.950        | 0.821 | 0.920 | 0.869        | 0.822 | 0.981 | 0.898        |
| densenet1d121 | 0.981 | 0.987 | <b>0.984</b> | 0.972 | 0.968 | 0.970        | 0.958 | 0.899 | <b>0.928</b> | 0.944 | 0.974 | <b>0.959</b> |
| ECGFounder    | 0.958 | 0.990 | 0.974        | 0.893 | 0.995 | 0.943        | 0.836 | 0.942 | 0.887        | 0.844 | 0.817 | 0.830        |
| ECG-JEPA      | 0.975 | 0.945 | 0.960        | 0.956 | 0.987 | <b>0.971</b> | 0.888 | 0.839 | 0.863        | 0.766 | 0.975 | 0.864        |
| ST-MEM        | 0.961 | 0.991 | 0.976        | 0.937 | 0.992 | 0.964        | 0.898 | 0.920 | 0.909        | 0.869 | 0.977 | 0.921        |
| HuBERT-ECG    | 0.883 | 0.951 | 0.917        | 0.945 | 0.991 | 0.967        | 0.596 | 0.774 | 0.679        | 0.853 | 0.914 | 0.883        |

**Выводы по сравнению архитектур нейросетевых моделей.** По результатам экспериментов (табл. 3) наилучшее качество по метрике G-mean на большинстве рассматриваемых патологий продемонстрировала модель `densenet1d121`: она лидирует по AFIB (0.984), 1AVB (0.928) и PVC (0.959) и лишь незначительно уступает лучшему результату на CLBBB, где сопоставимый показатель (0.970) даёт ECG-JERA (0.971). При этом вторая компактная модель, `resnet1d50`, показывает более низкие результаты на всех четырёх патологиях. В рамках сравниваемых компактных сверточных архитектур это свидетельствует о существенности выбора базовой архитектуры нейросетевой модели для конкретной задачи. Более высокое качество `densenet1d121` может быть связано, в частности, с плотными межслойными связями и переиспользованием признаков на разных уровнях глубины, что оказывается полезным при выделении локальных морфологических паттернов ЭКГ.

Разброс качества между моделями существенно различается и по патологиям. Для AFIB, ключевым диагностическим признаком которой является нерегулярность сердечного ритма, большинство моделей демонстрирует высокое качество. Для CLBBB аналогично высокие значения связаны с хорошо выраженной и устойчивой деформацией комплекса *QRS*. Напротив, для 1AVB, диагностика которой опирается на точную оценку длительности интервала *PR*, разброс качества максимален: от 0.679 у HuBERT-ECG до 0.928 у `densenet1d121`. В использованном режиме адаптации foundation-модели не обеспечили для данной патологии того же качества, что и `densenet1d121`, обученная непосредственно на целевой задаче. С практической точки зрения это означает, что при выборе архитектуры для конкретной группы патологий недостаточно ориентироваться только на усреднённые по всем классам метрики: качество может существенно различаться в зависимости от диагностических особенностей рассматриваемого состояния.

В связи с этим именно `densenet1d121` была выбрана в качестве базовой архитектуры для дальнейших исследований: она демонстрирует не только наивысшее качество на большинстве рассматриваемых патологий, но и наиболее стабильное качество по всем четырём типам нарушений.

**Зависимость качества от размера обучающей выборки.** Одним из факторов, существенно влияющих на качество классификации, является объём доступных

размеченных данных. Для анализа данной зависимости проводится отдельный эксперимент, в котором базовая модель `densenet1d121` обучается на различных долях обучающей выборки. Подобная постановка позволяет оценить, насколько быстро модель достигает приемлемого качества и сохраняется ли потенциал дальнейшего роста метрик при увеличении числа обучающих примеров.

В данном эксперименте фиксируется одно разбиение данных на обучающую, валидационную и тестовую выборки. После этого из обучающей части последовательно отбираются подмножества различного размера, например 10%, 25%, 50% и 100% от полного обучающего набора. Для обеспечения сопоставимости результатов доли формируются с сохранением соотношения классов. На каждом подмножестве модель обучается независимо, а затем оценивается на одном и том же тестовом наборе.

Модель обучается на подвыборках набора ISP, а оценивается как на тестовой части ISP, так и на внешнем наборе PTB-XL, чтобы дополнительно проверить, одинаково ли влияет объём обучающих данных на качество внутри домена обучения и на устойчивость к междатасетному сдвигу. Результаты представлены в таблицах 4 и 5.

Таблица 4 — Зависимость качества классификации от размера обучающей выборки для модели `densenet1d121` (Тест на ISP)

| Доля train | AFIB  |       |              | CLBBB |       |              | 1AVB  |       |              | PVC   |       |              |
|------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|
|            | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       |
| 10%        | 0.981 | 0.987 | 0.984        | 0.972 | 0.968 | 0.970        | 0.958 | 0.899 | 0.928        | 0.944 | 0.974 | 0.959        |
| 25%        | 0.987 | 0.982 | 0.984        | 0.969 | 0.979 | 0.973        | 0.955 | 0.906 | 0.930        | 0.981 | 0.975 | 0.978        |
| 50%        | 0.982 | 0.990 | 0.986        | 0.967 | 0.982 | <b>0.975</b> | 0.946 | 0.927 | 0.937        | 0.981 | 0.984 | <b>0.983</b> |
| 100%       | 0.981 | 0.995 | <b>0.988</b> | 0.955 | 0.995 | <b>0.975</b> | 0.937 | 0.939 | <b>0.938</b> | 0.988 | 0.976 | 0.982        |

Таблица 5 — Зависимость качества классификации от размера обучающей выборки для модели `densenet1d121` (Тест на PTB-XL)

| Доля train | AFIB  |       |              | CLBBB |       |              | 1AVB  |       |              | PVC   |       |              |
|------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|
|            | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       |
| 10%        | 0.930 | 0.948 | 0.939        | 1.000 | 0.880 | 0.938        | 0.921 | 0.840 | 0.880        | 0.954 | 0.941 | 0.948        |
| 25%        | 0.958 | 0.927 | 0.943        | 1.000 | 0.902 | 0.949        | 0.908 | 0.849 | 0.878        | 0.991 | 0.929 | 0.959        |
| 50%        | 0.944 | 0.959 | 0.952        | 1.000 | 0.928 | 0.963        | 0.921 | 0.867 | 0.894        | 0.991 | 0.967 | 0.979        |
| 100%       | 0.923 | 0.983 | <b>0.953</b> | 1.000 | 0.969 | <b>0.984</b> | 0.934 | 0.884 | <b>0.909</b> | 0.982 | 0.980 | <b>0.981</b> |

Результаты показывают, что зависимость качества классификации от объёма обучающей выборки носит нелинейный характер и различается в зависимо-

сти от того, тестируется ли модель на данных того же источника, что и обучение, или на внешнем наборе. При тестировании на ISP (табл. 4) качество в основном возрастает по мере увеличения обучающей выборки, при этом основной прирост достигается уже на сравнительно небольшой доле выборки: для AFIB G-mean увеличивается всего на 0.4 п.п. в диапазоне от 10% до 100% ( $0.984 \rightarrow 0.988$ ), для CLBBB и PVC максимум достигается уже при 50% (0.975 и 0.983 соответственно) и далее не растёт. Наиболее выраженная, хотя и умеренная, положительная динамика наблюдается для 1AVB ( $0.928 \rightarrow 0.938$ ). При этом уже при 10% обучающей выборки модель демонстрирует на ISP высокое качество классификации.

Иная закономерность наблюдается при тестировании той же серии моделей на внешнем наборе PTB-XL (табл. 5). Здесь увеличение объёма обучающей выборки сопровождается заметным улучшением качества вплоть до 100%: эффект наиболее выражен для CLBBB (G-mean увеличивается с 0.938 при 10% до 0.984 при полном обучающем наборе, то есть более чем на 4.5 п.п.) и для PVC (с 0.948 до 0.981). Для одной и той же модели увеличение объёма обучающих данных оказывает более заметное влияние на качество при переносе на внешний источник, чем на качество внутри домена обучения. Более крупная выборка может способствовать формированию признаковых представлений, менее привязанных к особенностям конкретного источника данных и лучше переносимых на данные с иным распределением.

Полученный результат показывает, что расширение обучающей выборки может быть более эффективной стратегией именно для повышения надёжности модели в клинически значимом сценарии – при переносе на данные, отличающиеся от обучающих, – чем для простого улучшения качества на “своих” данных, прирост на которых для большинства патологий уже исчерпан на сравнительно небольшой доле выборки.

#### **2.1.4 Дообучение моделей и оценка междатасетной устойчивости**

Одной из ключевых проблем при применении нейросетевых моделей к классификации электрокардиограмм является ограниченная междатасетная переносимость. Модель, обученная на одном наборе данных, может демонстриро-

вать высокое качество на собственном тестовом разбиении, но заметно ухудшать результаты при применении к данным иного происхождения. Причиной этого являются различия в параметрах регистрации ЭКГ, типах оборудования, методах предварительной обработки, характеристиках пациентских выборок и клиническом контексте сбора записей.

В настоящей работе исследование междатасетной устойчивости проводится на базе архитектуры `densenet1d121`, выбранной на предыдущем этапе. В качестве источников данных используются два набора: PTB-XL и ISP. В рассматриваемой постановке сравниваются три базовых сценария обучения. В первом сценарии модель обучается только на PTB-XL и затем тестируется как на PTB-XL, так и на закрытом наборе ISP. Во втором сценарии модель обучается только на ISP и без дополнительной адаптации тестируется на обоих наборах; этот сценарий характеризует междатасетную переносимость модели. В третьем сценарии реализуется подход трансферного обучения: модель сначала обучается на большом и гетерогенном ISP, после чего дообучается на целевом наборе PTB-XL.

Для корректного сравнения все модели обучаются в таком же формате, как и в предыдущих экспериментах. В анализ включаются только пациенты старше 18 лет. Используются 10-секундные записи ЭКГ с частотой 500 Гц; записи иной длины и частоты приводятся к этому формату путём фильтрации, передискретизации и дополнения нулями. После преобразования к обоим наборам применяется  $z$ -нормализация. Разбиение на обучающую, валидационную и тестовую части выполняется по пациентам, чтобы исключить попадание записей одного и того же пациента в разные части выборки.

При дообучении архитектура базовой модели `densenet1d121` сохраняется без изменений. В использованной конфигурации замораживались первые семь блоков сети, обученных на ISP, а незамороженные верхние слои переинициализировались и обучались на целевом наборе PTB-XL. В качестве рабочей гипотезы такой схемы использовалось предположение, что ранние блоки сохраняют переносимые низко- и среднеуровневые морфологические признаки, тогда как верхняя часть сети адаптируется к особенностям целевого домена.

Для оценки качества используются чувствительность, специфичность и метрика G-mean. Основное внимание уделяется именно G-mean, как и в предыдущих экспериментах.

Первый эксперимент в рамках данного подраздела направлен на прямое сравнение междатасетной устойчивости моделей, обученных либо на PTB-XL, либо на ISP. Для этого каждая из двух моделей тестируется как на “своём” тестовом наборе, так и на альтернативном наборе данных. Такая постановка позволяет количественно оценить снижение качества при переносе между датасетами. Результаты эксперимента представлены в таблице 6.

Второй эксперимент посвящён оценке эффективности дообучения (далее – эксперимент по эффективности дообучения). Здесь сравниваются три варианта модели на целевом наборе PTB-XL: модель, обученная только на PTB-XL; модель, обученная только на ISP; а также модель, предварительно обученная на ISP и затем дообученная на PTB-XL. Основным интерес представляет не только вопрос, позволяет ли fine-tuning объединить преимущества крупного и разнообразного исходного набора данных с особенностями целевого домена, но и то, насколько важна именно специфика целевых данных: если дообучение всего на сравнительно небольшом наборе PTB-XL даёт прирост качества по сравнению с моделью, обученной на существенно большем по объёму наборе ISP, но без адаптации к целевому домену, это будет означать, что доменная специфичность дообучающих данных играет не менее важную роль, чем их общий объём.

Третий эксперимент направлен на проверку того, в какой степени эффект fine-tuning определяется количеством данных. Для этого дополнительно рассматривается сценарий, в котором предварительное обучение проводится не на полном ISP, а на его уменьшенной вдвое части с сохранением долей классов (модель `ISP_half_tuned_dn`), после чего модель также дообучается на PTB-XL. Результаты эксперимента приведены в общей таблице 7 вместе с результатами эксперимента по эффективности дообучения.

В совокупности такие эксперименты позволяют ответить на три вопроса: насколько устойчива модель к смене источника данных, как увеличение объёма исходной выборки влияет на последующий перенос и способен ли fine-tuning повысить качество на целевом наборе данных по сравнению с обучением «с нуля».

Таблица 6 — Оценка междатасетной устойчивости модели densenet1d121

| Обучение | Тест   | AFIB  |       |              | CLBBB |       |              | 1AVB  |       |              | PVC   |       |              |
|----------|--------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|
|          |        | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       |
| PTB-XL   | PTB-XL | 0.937 | 0.955 | 0.946        | 1.000 | 0.947 | 0.973        | 0.908 | 0.871 | 0.889        | 0.936 | 0.963 | 0.949        |
| ISP      | PTB-XL | 0.923 | 0.983 | <b>0.953</b> | 1.000 | 0.969 | <b>0.984</b> | 0.934 | 0.884 | <b>0.909</b> | 0.982 | 0.980 | <b>0.981</b> |
| PTB-XL   | ISP    | 0.975 | 0.966 | 0.970        | 0.955 | 0.920 | 0.937        | 0.765 | 0.921 | 0.839        | 0.887 | 0.880 | 0.883        |
| ISP      | ISP    | 0.981 | 0.995 | <b>0.988</b> | 0.955 | 0.995 | <b>0.975</b> | 0.937 | 0.939 | <b>0.938</b> | 0.988 | 0.976 | <b>0.982</b> |

Таблица 7 — Влияние дообучения на качество классификации на целевом наборе PTB-XL

| Модель            | AFIB  |       |              | CLBBB |       |              | 1AVB  |       |              | PVC   |       |              |
|-------------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|
|                   | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       |
| PTB-XL            | 0.937 | 0.955 | 0.946        | 1.000 | 0.947 | 0.973        | 0.908 | 0.871 | 0.889        | 0.936 | 0.963 | 0.949        |
| ISP_full          | 0.923 | 0.983 | 0.953        | 1.000 | 0.969 | <b>0.984</b> | 0.934 | 0.884 | 0.909        | 0.982 | 0.980 | 0.981        |
| ISP_half_tuned_dn | 0.951 | 0.967 | 0.959        | 1.000 | 0.964 | 0.982        | 0.961 | 0.882 | 0.920        | 1.000 | 0.977 | <b>0.989</b> |
| ISP_full_tuned_dn | 0.940 | 0.981 | <b>0.960</b> | 1.000 | 0.968 | <b>0.984</b> | 0.974 | 0.919 | <b>0.946</b> | 1.000 | 0.978 | <b>0.989</b> |

**Выводы по дообучению и междатасетной устойчивости.** Результаты первого эксперимента (табл. 6) демонстрируют выраженную асимметрию в переносимости моделей между наборами данных. Модель, обученная на ISP, не только показывает наилучшее качество на собственном тестовом наборе (G-mean от 0.938 до 0.988 в зависимости от патологии), но и при переносе на PTB-XL превосходит модель, обученную непосредственно на PTB-XL и протестированную на нём же, причём сразу по всем четырём патологиям (например, для CLBBB – 0.984 против 0.973, для PVC – 0.981 против 0.949). Иными словами, модель, ни разу не видевшая обучающих примеров PTB-XL, оказывается точнее модели, специально под него обученной.

В обратном направлении перенос выражен значительно слабее, и это особенно заметно при сопоставлении модели, обученной на PTB-XL и тестируемой на ISP, с моделью, обученной непосредственно на ISP. При таком сравнении на общей тестовой выборке ISP разница оказывается существенной по всем без исключения патологиям: AFIB – 0.970 против 0.988 (разница 1.8 п.п.), CLBBB – 0.937 против 0.975 (3.8 п. п.), 1AVB – 0.839 против 0.938 (9.9 п. п.) и PVC – 0.883 против 0.982 (9.9 п. п.). Таким образом, модель, обученная только на PTB-XL, теряет от 2 до почти 10 п.п. G-mean при переносе на ISP по сравнению с моделью, для которой ISP является исходным доменом, – то есть перенос в этом направлении выражен даже сильнее, чем можно было бы заключить только по

её собственной деградации относительно результата на PTB-XL. Такая асимметрия может быть связана с различиями между наборами данных: ISP включает записи разных устройств, учреждений и клинических сценариев, тогда как PTB-XL является более однородным по условиям сбора. Обучение на ISP тем самым предоставляет модели более широкий набор вариантов входных данных, что может способствовать её устойчивости при переносе на PTB-XL.

Результаты экспериментов по эффективности дообучения и влиянию объёма исходной выборки (табл. 7) показывают, что дообучение на целевом наборе PTB-XL улучшает качество по сравнению как с обучением исключительно на PTB-XL, так и с применением модели, обученной на ISP, без адаптации. Наибольший прирост наблюдается для 1AVB: G-mean возрастает с 0.889 для модели, обученной только на PTB-XL, и 0.909 для модели ISP без адаптации до 0.946 после предварительного обучения на полном ISP и последующего fine-tuning на PTB-XL. Можно отметить, что такой результат достигается именно для патологии, которая на всех предыдущих этапах главы (см. п. 2.1.3) оказывалась наиболее чувствительной к выбору архитектуры нейросетевой модели и объёму обучающих данных.

Сравнение двух дообученных вариантов, предварительно обученных соответственно на 50% и 100% ISP, показывает, что удвоение объёма исходного набора при неизменной процедуре адаптации даёт ограниченный дополнительный эффект: различие G-mean не превышает 0.3 п.п. для AFIB, CLBBB и PVC и составляет около 2.6 п.п. для 1AVB. Таким образом, в рассмотренной постановке существенный вклад в качество на целевом наборе вносит сама адаптация к PTB-XL, тогда как увеличение объёма предварительного обучения с половины до полного ISP даёт меньший дополнительный выигрыш для большинства исследованных патологий.

В совокупности эти результаты подтверждают, что модель `densenet1d121`, обученная на большом и гетерогенном клиническом наборе данных, обладает более высокой междатасетной устойчивостью, а дообучение такой модели на целевом наборе данных обеспечивает дополнительный прирост качества, особенно выраженный для патологий, чувствительных к точной временной структуре сигнала. Это подтверждает практическую целесообразность transfer learning в задаче анализа 12-канальных ЭКГ: вместо обучения отдельной модели “с нуля” для каждого нового медицинского учреждения целесообразно

переиспользовать представления, сформированные на большом разнородном наборе, адаптируя лишь верхние слои сети под особенности целевых данных.

## 2.2 Сегментация комплексов PQRS $T$

Задача сегментации электрокардиограммы состоит в автоматическом выделении ключевых волн и границ сердечного цикла, в частности начала и окончания зубца  $P$ , комплекса  $QRS$  и зубца  $T$ . Решение данной задачи представляет самостоятельный интерес, поскольку точная локализация характерных точек ЭКГ позволяет перейти от общей классификации записи к более детальному анализу морфологии сигнала и длительностей интервалов. Кроме того, результаты сегментации могут быть использованы для повышения интерпретируемости алгоритмов анализа ЭКГ.

В данном разделе рассматривается новый подход к сегментации комплексов  $PQRS\mathit{T}$  на 12-канальных электрокардиограммах, основанный на предлагаемой в настоящей работе нейросетевой модели ECG-CODE. Дополнительно в рамках предложенного метода исследуется возможность повышения качества сегментации за счёт self-training подхода, основанного на псевдоразметке большого массива неаннотированных данных. Такой подход особенно актуален для задачи делинеации ЭКГ, поскольку открытые наборы данных с ручной разметкой характерных точек остаются относительно небольшими, что ограничивает обобщающую способность моделей.

### 2.2.1 Постановка задачи сегментации ЭКГ

Пусть задана электрокардиографическая запись в 12 отведениях

$$X = \{x_1(t), x_2(t), \dots, x_{12}(t)\},$$

где  $x_i(t)$  — сигнал  $i$ -го отведения. Требуется для каждого сердечного цикла определить положения шести характерных точек:

$$P_{\text{on}}, P_{\text{off}}, QRS_{\text{on}}, QRS_{\text{off}}, T_{\text{on}}, T_{\text{off}}.$$

Таким образом, задача формулируется как задача временной сегментации, в которой необходимо выделить интервалы, соответствующие зубцу  $P$ , комплексу  $QRS$  и зубцу  $T$ , а также определить их границы.

### 2.2.2 Используемые наборы данных

Для обучения и оценки модели сегментации использовались два набора данных с ручной разметкой характерных точек: открытый набор LUDB и сформированный в рамках настоящей работы набор ISPdeli. Дополнительно для self-training использовался открытый набор PTB-XL, не содержащий разметки границ волн, но включающий большой объём 12-канальных ЭКГ.

Набор ISPdeli включает 475 размеченных 10-секундных 12-канальных записей ЭКГ. Разметка выполнена двумя кардиологами. Данный набор содержит существенно больше вручную размеченных характерных точек, чем открытые наборы данных, обычно используемые для задачи делинеации ЭКГ. Набор LUDB содержит 200 записей и также включает разметку волн  $P$ ,  $QRS$  и  $T$ . Набор QTDB в исследование не включался, поскольку содержит двухканальные записи, тогда как в работе рассматриваются именно 12-канальные ЭКГ.

Для этапа псевдоразметки использовался набор PTB-XL. Сводные характеристики наборов данных с ручной разметкой приведены в таблице 8.

Таблица 8 — Характеристики наборов данных с ручной разметкой, использованных для сегментации

| Набор данных | Число записей | Зубцы P | Комплексы QRS | Зубцы T |
|--------------|---------------|---------|---------------|---------|
| ISPdeli      | 475           | 5105    | 5948          | 5653    |
| LUDB         | 200           | 1393    | 1803          | 1621    |

### 2.2.3 Модель сегментации комплексов PQRST

В качестве архитектуры для сегментации предлагается новая сверточная нейросетевая модель ECG-CODE (ECG COmplex DElineation). На вход модели подаются все 12 отведений ЭКГ, которые предварительно преобразуются в mel-спектрограммы. Спектральное представление позволяет одновременно описывать локальное изменение частотного состава сигнала во времени и использовать двумерные сверточные блоки для совместного анализа временной и частотной структуры. Одинаковые параметры преобразования применяются ко всем наборам данных и на всех этапах эксперимента.

Внутри модели применяются блоки, основанные на нейросетевой модели MobileNet [96], обеспечивающие последовательное сжатие и расширение пространства признаков. Общая схема архитектуры представлена на рисунке 2.1.

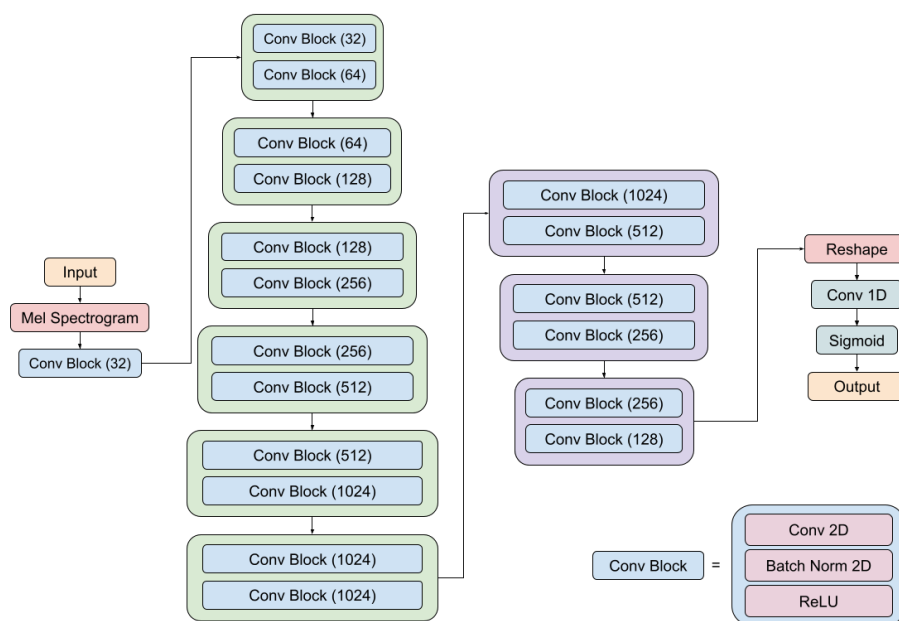


Рисунок 2.1 — Архитектура модели ECG-CODE

Выход модели формируется по множеству неперекрывающихся временных интервалов. Для каждого такого интервала модель предсказывает:

1. наличие или отсутствие объекта заданного типа;
2. координаты начала и конца сегмента, если соответствующий объект присутствует.

Рассматриваются три типа объектов: волна  $P$ , комплекс  $QRS$  и волна  $T$ . Таким образом, задача для каждого интервала сводится к определению наличия соответствующей волны и, при её наличии, к регрессии её граничных точек.

Функция потерь модели состоит из двух компонентов: потери для предсказания наличия объекта (*confidence loss*,  $CL$ ) и потери для предсказания его границ (*start-end loss*,  $SEL$ ):

$$CL = (P_C - T_C)^2, \quad (2.1)$$

$$SS = (P_S - T_S)^2 + (P_E - T_E)^2, \quad (2.2)$$

$$SEL = \begin{cases} 0, & \text{если } SS < \textit{accepted\_error}, \\ SS \cdot T_C, & \text{если } SS \geq \textit{accepted\_error}. \end{cases} \quad (2.3)$$

Здесь  $T_C$  и  $P_C$  обозначают истинную и предсказанную уверенность в наличии объекта,  $T_S$  и  $P_S$  — истинную и предсказанную относительные координаты начала сегмента внутри рассматриваемого интервала,  $T_E$  и  $P_E$  — соответствующие относительные координаты конца сегмента. В работе значение параметра *accepted\_error* было установлено равным 0.15. Порог вводит область допустимой погрешности локализации: если суммарная квадратичная ошибка границ меньше заданного значения, дополнительный регрессионный штраф не начисляется, что предотвращает избыточную оптимизацию уже достаточно близких предсказаний. Итоговая функция потерь вычислялась как сумма компонент  $CL$  и  $SEL$ .

После получения предсказаний выполнялась постобработка. На первом этапе объединялись соседние сегменты одного типа, если расстояние между ними было меньше заранее заданного порога. На втором этапе удалялись слишком короткие сегменты.

Для использования большого объёма неразмеченных данных был предложен self-training подход. Сначала базовая модель ECG-CODE, обученная на размеченном наборе, использовалась для псевдоразметки записей РТВ-XL. Для каждой записи сохранялись предсказанные классы, их координаты и значения *confidence score*. Затем записи ранжировались по специальной мере качества псевдоразметки:

$$\textit{delineation score} = |0.5 - \textit{predicted confidence}|. \quad (2.4)$$

Чем дальше значение *confidence* находится от 0.5, тем больше *delineation score* и тем более уверенным считается предсказание. Такая симметричная мера позволяет одинаково учитывать уверенные предсказания присутствия объекта (*confidence*, близкий к 1) и его отсутствия (*confidence*, близкий к 0). Для каждой записи оценки агрегировались по предсказанным интервалам отдельно для зубца *P*, комплекса *QRS* и зубца *T*; при ранжировании учитывались среднее значение и разброс *confidence* для соответствующего типа объекта. Отбор выполнялся отдельно для каждого из трёх типов объектов, после чего выбранные псевдометки использовались для формирования обучающего набора. На сформированном псевдоразмеченном наборе новая модель обучалась заново, а затем дообучалась на вручную размеченном наборе данных.

#### 2.2.4 Схема обучения и метрики оценки качества

На первом этапе модель ECG-CODE обучалась на обучающей части набора ISPdeli. Далее эта модель использовалась для self-training и дообучения с привлечением набора PTB-XL. Для каждого типа объекта (*P*, *QRS* и *T*) отдельно отбирались верхние 50% записей по соответствующей оценке уверенности псевдоразметки.

В ходе обучения применялась аугментация данных, включавшая последовательное применение полосового фильтра и режекторного фильтра; вероятность применения каждого преобразования составляла 0.5. Все записи были приведены к частоте дискретизации 1000 Гц и нормализованы с использованием *z*-нормализации.

На этапе постобработки максимальное расстояние для объединения сегментов было установлено равным 300 отсчётам, а минимальная допустимая длина сегмента — 50 отсчётам.

Качество сегментации оценивалось по шести характерным точкам:

$$P_{\text{on}}, P_{\text{off}}, QRS_{\text{on}}, QRS_{\text{off}}, T_{\text{on}}, T_{\text{off}}.$$

Предсказанная точка считалась истинно положительной, если она попадала в окно допуска (*window tolerance*) относительно соответствующей истинной

точки и принадлежала тому же классу. Предсказания без соответствующих истинных точек внутри окна допуска учитывались как ложноположительные, а отсутствующие предсказания для истинных точек — как ложноотрицательные.

В качестве основных метрик качества использовались:

- чувствительность (*Sensitivity*, *Se*);
- положительная прогностическая ценность (*Positive Predictive Value*, *PPV*);
- $F_1$ -мера;
- средняя ошибка локализации и её стандартное отклонение, выраженные в миллисекундах.

### 2.2.5 Постановка вычислительных экспериментов

Для проверки предложенного подхода были проведены два основных эксперимента.

В первом эксперименте сравнивалась устойчивость модели ECG-CODE при обучении на разных наборах данных. Для этого обучались две версии модели: одна на LUDB, другая на ISPdeli. Обе модели тестировались как на LUDB, так и на ISPdeli. Для обеспечения корректности сравнения использовались одинаковые параметры оценки, а ширина окна допуска составляла 150 мс — это значение широко используется в литературе по делинеации ЭКГ и, в частности, применяется в работах [57; 59; 97], с которыми предложенный подход сравнивается далее (табл. 10), что обеспечивает сопоставимость результатов. При тестировании на LUDB из рассмотрения исключались первая и последняя секунды записи, поскольку для этих фрагментов отсутствовала разметка.

Во втором эксперименте оценивалось влияние self-training на качество сегментации. Для этого сравнивались:

1. модель ECG-CODE, обученная только на ISPdeli;
2. модель Pseudolabeled ECG-CODE, предварительно обученная с использованием псевдоразмеченного PTB-XL и затем дообученная на ISPdeli;
3. три современных метода сегментации ЭКГ на основе нейросетевых моделей типа «энкодер-декодер»: компактная полностью сверточная сеть,

заявляющая устойчивость к разным частотам дискретизации и типам мониторинга [57]; ECG\_SegNet – энкодер-декодер с модулем много-масштабного извлечения признаков и BiLSTM-компонентой [97]; и метод, комбинирующий сегментацию сетью типа 1D-UNet с последующей морфологической коррекцией границ [59].

Для сопоставления на ISPdeli сравниваемые нейросетевые методы были воспроизведены и оценены на одной и той же тестовой части ISPdeli с единым окном допуска 150 мс. Во всех экспериментах оценка проводилась на исходных тестовых данных без дополнительной обработки. В качестве ключевой метрики для сравнения методов использовалась  $F_1$ -мера.

### 2.2.6 Результаты экспериментов

Результаты первого эксперимента представлены в таблице 9. Модель, обученная на LUDB, показывает высокие результаты на собственном тестовом наборе ( $F_1$  для  $P$ -волны – 0.961, для  $T$ -волны – 0.999/0.987), однако при переносе на ISPdeli ее качество заметно снижается:  $F_1$  для  $P_{on}$  падает до 0.932, а для  $T_{off}$  – до 0.919, то есть примерно на 3 п.п. для начала зубца  $P$  и на 7 п.п. для окончания зубца  $T$ . Комплекс  $QRS$  при этом переносится существенно лучше –  $F_1$  снижается не более чем на 0.4 п.п., что ожидаемо, поскольку  $QRS$ -комплекс обладает наибольшей амплитудой и в меньшей степени зависит от условий регистрации, тогда как зубцы  $P$  и  $T$  гораздо более чувствительны к уровню шума, дрейфу изолинии и различиям в разметке между аннотаторами двух наборов.

В то же время модель, обученная на ISPdeli, демонстрирует более устойчивые результаты как на собственном тестовом наборе, так и на LUDB: при переносе на LUDB её метрики по  $P$ - и  $T$ -волнам практически не уступают модели, обученной непосредственно на LUDB, а по отдельным точкам ( $P_{on}$ ,  $P_{off}$ ) даже превосходят её. Полученные результаты показывают, что обучение на более чем втрое превосходящем LUDB по числу размеченных объектов наборе ISPdeli (см. табл. 8) сопровождается более устойчивым переносом между рассматриваемыми источниками данных.

Таблица 9 — Результаты сегментации модели ECG-CODE при обучении на LUDB и ISPdeli

| Тестирование на LUDB    |                       |                  |                  |                 |                  |                 |                 |
|-------------------------|-----------------------|------------------|------------------|-----------------|------------------|-----------------|-----------------|
| Обучение                | Метрика               | P on             | P off            | QRS on          | QRS off          | T on            | T off           |
| LUDB                    | Se                    | 0.943            | 0.943            | 1.000           | 1.000            | 0.997           | 0.986           |
|                         | PPV                   | 0.979            | 0.979            | 1.000           | 1.000            | 1.000           | 0.988           |
|                         | F1-score              | 0.961            | 0.961            | <b>1.000</b>    | <b>1.000</b>     | <b>0.999</b>    | <b>0.987</b>    |
|                         | $\mu \pm \sigma$ , мс | $-7.3 \pm 17.4$  | $-0.3 \pm 14.1$  | $0.6 \pm 9.2$   | $0.6 \pm 15.3$   | $6.6 \pm 30.7$  | $-0.1 \pm 25.4$ |
| ISPdeli                 | Se                    | 0.970            | 0.977            | 1.000           | 1.000            | 0.988           | 0.968           |
|                         | PPV                   | 0.980            | 0.987            | 1.000           | 1.000            | 0.991           | 0.971           |
|                         | F1-score              | <b>0.975</b>     | <b>0.982</b>     | <b>1.000</b>    | <b>1.000</b>     | 0.990           | 0.970           |
|                         | $\mu \pm \sigma$ , мс | $-5.4 \pm 20.8$  | $-12.8 \pm 13.9$ | $8.1 \pm 10.5$  | $-10.8 \pm 16.8$ | $7.0 \pm 30.9$  | $-7.9 \pm 29.7$ |
| Тестирование на ISPdeli |                       |                  |                  |                 |                  |                 |                 |
| Обучение                | Метрика               | P on             | P off            | QRS on          | QRS off          | T on            | T off           |
| LUDB                    | Se                    | 0.907            | 0.910            | 1.000           | 0.999            | 0.972           | 0.927           |
|                         | PPV                   | 0.959            | 0.963            | 0.994           | 0.993            | 0.955           | 0.910           |
|                         | F1-score              | 0.932            | 0.936            | <b>0.997</b>    | 0.996            | 0.963           | 0.919           |
|                         | $\mu \pm \sigma$ , мс | $-10.7 \pm 21.8$ | $2.0 \pm 17.6$   | $-8.5 \pm 10.1$ | $12.5 \pm 12.2$  | $1.1 \pm 28.9$  | $3.6 \pm 32.3$  |
| ISPdeli                 | Se                    | 0.960            | 0.962            | 0.996           | 0.996            | 0.986           | 0.983           |
|                         | PPV                   | 0.978            | 0.981            | 0.998           | 0.998            | 0.980           | 0.977           |
|                         | F1-score              | <b>0.969</b>     | <b>0.971</b>     | <b>0.997</b>    | <b>0.997</b>     | <b>0.983</b>    | <b>0.980</b>    |
|                         | $\mu \pm \sigma$ , мс | $-3.1 \pm 16.4$  | $-5.1 \pm 14.2$  | $0.2 \pm 9.7$   | $0.4 \pm 11.3$   | $-3.0 \pm 27.3$ | $-0.4 \pm 26.4$ |

Результаты второго эксперимента приведены в таблице 10. Модель Pseudolabeled ECG-CODE, обученная с использованием self-training, улучшает  $F_1$ -меру по сравнению с базовой моделью ECG-CODE по всем шести характеристическим точкам, хотя и с разной величиной эффекта: наиболее заметно улучшение для зубцов  $P$  и  $T$  ( $P_{\text{on}}$ :  $0.969 \rightarrow 0.975$ ;  $T_{\text{off}}$ :  $0.980 \rightarrow 0.986$ ), тогда как сегментация комплекса  $QRS$  улучшается лишь на 0.2 п.п., поскольку уже для базовой модели она практически идеальна ( $F_1 = 0.997$ ). Это согласуется с ожиданием, что self-training приносит наибольшую пользу именно для более сложных и переменных объектов, для которых у исходной модели сохраняется больший запас для улучшения.

Сравнение с рассмотренными методами [57; 59; 97] показывает, что уже базовая модель ECG-CODE, обученная без self-training, превосходит все сравниваемые методы по  $F_1$ -мере для комплекса  $QRS$  (0.997 против 0.980–0.996) и зубца  $T$  (0.983/0.980 против 0.972–0.981), уступая лишь по зубцу  $P$ , где лучший результат показывает Chen et al. (0.975 против 0.969/0.971 у ECG-CODE). После добавления self-training модель Pseudolabeled ECG-CODE демонстрирует лучшие значения  $F_1$ -меры среди сравниваемых методов сразу по всем шести ха-

рактерным точкам одновременно (0.975–0.999). Таким образом, self-training не компенсирует ограничения базовой архитектуры, а усиливает изначально конкурентоспособную модель, нивелируя её единственное относительное отставание – по зубцу *P*.

Таблица 10 — Сравнение качества сегментации на ISPdeli с современными методами

| Метод                  | Метрика               | P on            | P off           | QRS on          | QRS off         | T on            | T off           |
|------------------------|-----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Moskalenko et al.      | Se                    | 0.962           | 0.962           | 0.999           | 0.999           | 0.972           | 0.972           |
|                        | PPV                   | 0.981           | 0.981           | 0.992           | 0.992           | 0.984           | 0.984           |
|                        | F1-score              | 0.971           | 0.971           | 0.996           | 0.996           | 0.978           | 0.978           |
|                        | $\mu \pm \sigma$ , мс | $-0.0 \pm 7.9$  | $0.7 \pm 9.0$   | $-0.3 \pm 6.6$  | $-0.5 \pm 8.6$  | $-3.2 \pm 20.7$ | $0.4 \pm 15.0$  |
| Liang et al.           | Se                    | 0.952           | 0.953           | 0.997           | 0.996           | 0.994           | 0.990           |
|                        | PPV                   | 0.952           | 0.953           | 0.988           | 0.987           | 0.967           | 0.963           |
|                        | F1-score              | 0.952           | 0.953           | 0.992           | 0.991           | 0.981           | 0.976           |
|                        | $\mu \pm \sigma$ , мс | $-2.0 \pm 16.5$ | $0.5 \pm 16.1$  | $0.7 \pm 7.4$   | $0.0 \pm 9.2$   | $-3.1 \pm 23.1$ | $0.0 \pm 18.1$  |
| Chen et al.            | Se                    | 0.964           | 0.964           | 0.977           | 0.977           | 0.968           | 0.968           |
|                        | PPV                   | 0.986           | 0.986           | 0.983           | 0.983           | 0.976           | 0.976           |
|                        | F1-score              | <b>0.975</b>    | 0.975           | 0.980           | 0.980           | 0.972           | 0.972           |
|                        | $\mu \pm \sigma$ , мс | $-0.2 \pm 5.4$  | $0.0 \pm 2.7$   | $0.2 \pm 9.6$   | $-0.2 \pm 10.2$ | $-0.6 \pm 4.8$  | $0.1 \pm 4.6$   |
| ECG-CODE               | Se                    | 0.960           | 0.962           | 0.996           | 0.996           | 0.986           | 0.983           |
|                        | PPV                   | 0.978           | 0.981           | 0.998           | 0.998           | 0.980           | 0.977           |
|                        | F1-score              | 0.969           | 0.971           | 0.997           | 0.997           | 0.983           | 0.980           |
|                        | $\mu \pm \sigma$ , мс | $-3.1 \pm 16.4$ | $-5.1 \pm 14.2$ | $0.2 \pm 9.7$   | $0.4 \pm 11.3$  | $-3.0 \pm 27.3$ | $-0.4 \pm 26.4$ |
| Pseudolabeled ECG-CODE | Se                    | 0.966           | 0.967           | 1.000           | 1.000           | 0.992           | 0.991           |
|                        | PPV                   | 0.985           | 0.986           | 0.999           | 0.999           | 0.983           | 0.982           |
|                        | F1-score              | <b>0.975</b>    | <b>0.976</b>    | <b>0.999</b>    | <b>0.999</b>    | <b>0.987</b>    | <b>0.986</b>    |
|                        | $\mu \pm \sigma$ , мс | $2.0 \pm 15.6$  | $-3.5 \pm 13.3$ | $-0.1 \pm 10.1$ | $-0.1 \pm 11.4$ | $2.4 \pm 27.4$  | $-0.8 \pm 24.6$ |

Пример сегментации, выполненной моделью Pseudolabeled ECG-CODE, приведен на рисунке 2.2.

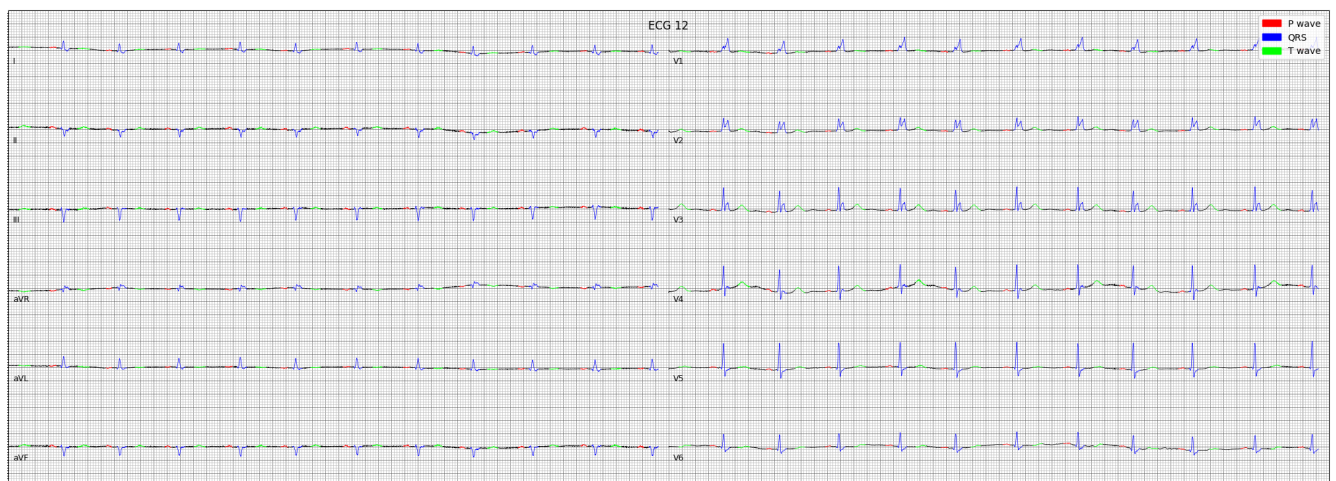


Рисунок 2.2 — Примеры сегментации ЭКГ моделью Pseudolabeled ECG-CODE

Таким образом, результаты проведённых экспериментов показывают, что качество сегментации существенно зависит как от объёма и состава размеченной обучающей выборки, так и от возможности привлекать дополнительные записи без экспертной разметки границ. Использование self-training позволяет повысить полноту обнаружения характерных точек, прежде всего для зубцов  $P$  и  $T$ , тогда как для комплекса  $QRS$ , для которого базовая модель уже обеспечивает высокие значения метрик, дополнительный эффект меньше. Полученный результат согласуется с общей логикой экспериментов по классификации: расширение доступных обучающих данных и адаптация к целевому источнику могут давать дополнительный прирост качества без изменения базовой архитектуры нейросетевой модели. Вместе с тем отмеченный компромисс между полнотой обнаружения точек и точностью их временной локализации указывает на возможное направление дальнейшего развития метода, например за счёт дополнительной регрессионной калибровки границ уже найденных сегментов. Это делает предложенный подход перспективным не только для задачи делинеации ЭКГ, но и для последующего использования результатов сегментации для повышения интерпретируемости алгоритмов анализа электрокардиограмм.

## 2.3 Выводы

В данной главе исследованы две задачи автоматического анализа 12-канальных электрокардиограмм: классификация патологических состояний и сегментация комплексов PQRST. Для задачи классификации выполнены выбор базовой архитектуры нейросетевой модели, анализ влияния объёма обучающих данных, оценка междатасетной устойчивости и исследование дообучения на целевом источнике данных. Для задачи сегментации разработана архитектура ECG-CODE и исследована схема self-training с использованием псевдоразметки дополнительных записей, позволяющая повысить качество выделения характерных элементов сигнала и обеспечить устойчивость моделей к вариативности входных данных.

На первом этапе был проведён сравнительный анализ современных архитектур для классификации патологических состояний по ЭКГ. В сравнение были

включены как предварительно обученные foundation-модели, так и нейросетевые модели, обучаемые непосредственно на целевых данных. Такой эксперимент был необходим, поскольку в современных работах предварительное обучение на больших массивах данных часто рассматривается как универсальный способ повышения качества, однако практическая эффективность таких моделей должна проверяться в конкретной клинической постановке и на конкретном наборе патологий.

Полученные результаты показали, что foundation-модели и наличие предварительного обучения на больших наборах данных не гарантируют наилучшего качества распознавания патологических состояний. На большом закрытом клиническом наборе данных компактная сверточная архитектура `densenet1d121` показала качество выше или сопоставимое с качеством рассмотренных предварительно обученных моделей для выбранного набора синдромов. Это означает, что при решении задачи классификации ЭКГ важны не только масштаб модели и объём данных предварительного обучения, но и соответствие архитектуры целевой задаче, характер обучающей выборки, процедура адаптации и особенности клинических меток.

На основании проведённого сравнения в качестве базовой модели для дальнейших экспериментов была выбрана архитектура `densenet1d121`. Данный выбор обусловлен высоким качеством классификации на большинстве рассматриваемых патологий и удобством последующей модификации архитектуры в дальнейших экспериментах. Кроме того, сверточная природа модели хорошо согласуется со структурой ЭКГ-сигнала, в котором диагностически значимые признаки часто имеют локальный морфологический характер и проявляются в форме зубцов, комплексов, интервалов и их временных соотношений. Поэтому дальнейшая работа в рамках главы была сосредоточена именно на этой архитектуре.

На следующем этапе был проведён эксперимент по анализу зависимости качества классификации от объёма обучающей выборки. Было показано, что зависимость качества от объёма обучающей выборки носит нелинейный характер и различается в зависимости от условий тестирования: при оценке на данных того же источника, что и обучение, качество по большинству патологий выходит на плато уже при сравнительно небольшой доле выборки, тогда как при оценке на внешнем наборе данных рост объёма обучающих данных продолжает при-

носить заметный прирост качества вплоть до использования полной выборки. Это означает, что расширение обучающих данных особенно важно не столько для улучшения качества на данных из той же выборки, сколько для повышения устойчивости модели к переносу на данные из другого источника, что напрямую связывает этот результат со следующим этапом главы.

Отдельное внимание в главе было уделено междатасетной устойчивости моделей. Эксперименты показали, что модель, обученная на одном источнике данных, может демонстрировать снижение качества при применении к другому набору ЭКГ, что связано с различиями в оборудовании, частоте дискретизации, протоколах регистрации, характеристиках пациентских выборок, предварительной обработке и структуре клинических меток. При этом асимметрия переноса оказалась выраженной: модель, обученная на большом и разнородном клиническом наборе, не только сохраняла высокое качество при переносе на открытый набор, но и превосходила на нём модель, обученную непосредственно на этом наборе, тогда как перенос в обратном направлении сопровождался заметной потерей качества. Дополнительный эксперимент показал, что при последующем дообучении на целевом наборе увеличение объёма исходных данных ISP в два раза обеспечивает лишь ограниченный дополнительный прирост качества. Таким образом, адаптация модели к данным целевого домена посредством дообучения является значимым этапом переноса алгоритма на новый источник данных. При этом для достижения высокого качества в рассмотренной постановке не требуется обязательного использования полного объёма исходного набора ISP.

В завершающей части главы была рассмотрена задача сегментации ЭКГ и предложена нейросетевая модель ECG-CODE, предназначенная для выделения границ зубцов и комплексов PQRST на 12-канальных записях. Для повышения качества сегментации был предложен метод self-training, основанный на псевдоразметке дополнительных записей, не имеющих экспертной разметки границ волн. Такой подход позволяет использовать дополнительные данные без необходимости полной ручной разметки, что особенно важно для задачи делинеации ЭКГ, где получение экспертных аннотаций является трудоёмким процессом. Эксперименты показали, что использование псевдоразмеченных данных повышает полноту обнаружения характерных точек и позволяет получить результаты по  $F_1$ -мере, превосходящие или сопоставимые с результатами лучших из сравниваемых существующих методов сразу по всем шести характерным точкам.

Тем самым первое положение, выносимое на защиту, реализуется в виде метода сегментации цифровых ЭКГ при ограниченном объёме экспертной разметки. Сначала предложенная нейросетевая модель ECG-CODE обучается на записях с экспертно заданными границами зубца  $P$ , комплекса  $QRS$  и зубца  $T$  и для каждого временного интервала оценивает наличие соответствующего объекта и координаты его начала и конца. Полученные предсказания проходят постобработку с объединением близких сегментов и удалением слишком коротких фрагментов.

Для расширения обучающей выборки обученная модель применяется к дополнительным записям, не имеющим экспертной разметки границ. Псевдометки ранжируются по уверенности отдельно для  $P$ ,  $QRS$  и  $T$ , после чего для каждого типа объекта отбираются записи с наиболее уверенной псевдоразметкой. На сформированном псевдоразмеченном наборе обучается новая модель, которая затем дообучается на экспертно размеченных данных.

Полученные результаты формируют основу для дальнейшего развития методов классификации и сегментации электрокардиограмм.

### **Глава 3. Самообучение и перенос представлений в задачах анализа электрокардиограмм**

Третья глава посвящена задачам предобучения моделей классификации и раскрывает второй и третий пункты научной новизны. Кроме того, в ней излагается второе положение, выносимое на защиту. Первый предлагаемый метод, рассматриваемый в первой части главы, использует метаданные пациентов при формировании признаков представлений. Вторым методом, LC-CL (Lead Combinations with Contrastive Learning), опирается на зависимости между различными отведениями одной ЭКГ для переноса знаний на модель с уменьшенным числом доступных каналов.

#### **3.1 Использование неразмеченных данных в задачах анализа электрокардиограмм**

Одной из ключевых проблем при разработке алгоритмов автоматического анализа электрокардиограмм является ограниченная доступность качественно размеченных медицинских данных. Несмотря на широкое распространение цифровых ЭКГ-систем, формирование крупных экспертно аннотированных выборок требует значительных временных затрат и участия специалистов высокой квалификации. Вследствие этого даже при наличии большого числа зарегистрированных сигналов объём данных, пригодных для полноценного обучения моделей с учителем, часто оказывается недостаточным.

Особую значимость данная проблема приобретает в задачах анализа ЭКГ, где наблюдаются выраженные различия между наборами данных по условиям регистрации, характеристикам устройств и структуре популяции пациентов. Модели, обученные исключительно с учителем, нередко демонстрируют снижение качества при переносе на записи иного происхождения, что ограничивает их применимость в реальной клинической практике.

Одним из перспективных направлений преодоления указанной проблемы является использование неразмеченных наборов данных. В отличие от ручной разметки, сбор больших массивов необработанных электрокардиографических

записей требует существенно меньших затрат. Это обуславливает возможность применения самообучения, в рамках которого модель обучается извлекать информативные представления из данных без явного использования экспертных меток. В дальнейшем такие представления могут быть адаптированы к конкретной диагностической задаче на сравнительно небольшом размеченном наборе данных.

Для временных рядов и ЭКГ-сигналов методы предобучения моделей представляются особенно востребованными, поскольку позволяют задействовать скрытую структуру сигнала, взаимосвязи между каналами, временные зависимости и доступные метаданные пациента. Среди существующих подходов особое место занимают методы контрастного обучения, продемонстрировавшие в последние годы высокую эффективность как в задачах анализа медицинских изображений, так и в задачах обработки биомедицинских сигналов.

В данной главе рассматриваются методы предобучения моделей на неразмеченных данных для анализа ЭКГ, а также исследуется возможность переноса представлений, полученных на 12-канальных электрокардиограммах, на записи с меньшим числом отведений. Основное внимание сосредоточено на контрастном обучении как одном из наиболее распространённых классов методов для работы с неразмеченными ЭКГ-данными.

### **3.2 Самообучение для 12-канальных электрокардиограмм**

В данном разделе рассматриваются подходы к предобучению моделей для анализа 12-канальных электрокардиограмм. Цель состоит в исследовании методов, позволяющих использовать большие массивы неразмеченных ЭКГ для построения более информативных представлений сигнала, которые затем могут быть применены в конкретных задачах классификации сердечных патологий.

Сначала приводится обзор основных направлений самообучения для временных рядов. Далее в разделе предлагается и исследуется возможность включения метаданных пациента в процесс самообучения. Эффективность предложенного подхода исследуется экспериментально в задаче классификации 12-канальных ЭКГ посредством сравнения базовых контрастных методов SimCLR и TS-

ТСС с их group-based модификациями, использующими метаданные пациента при формировании дополнительных положительных пар.

### 3.2.1 Подходы к самообучению моделей для ЭКГ

Самообучение включает несколько групп методов, различающихся по постановке вспомогательной задачи, используемой на этапе предобучения. Поскольку отдельные подходы могут сочетать признаки нескольких постановок, такое разделение носит условный характер. Для анализа электрокардиограмм можно выделить по меньшей мере три крупных направления.

К первому направлению относятся предиктивные методы, в которых модель обучается восстанавливать или предсказывать части сигнала на основе наблюдаемого контекста. Применительно к ЭКГ это может быть предсказание будущих фрагментов временного ряда, отдельных комплексов, пропущенных участков сигнала или представлений соседних временных окон. Такие методы ориентированы на изучение временной структуры сигнала и зависимостей между его сегментами [49].

Ко второму направлению относятся reconstruction-based методы, основанные на восстановлении сигнала из частично искажённой или зашумлённой версии. К этой группе относятся классические автоэнкодеры, вариационные автоэнкодеры, шумоустойчивые (denoising) автоэнкодеры, а также более современные masked-modeling подходы, в которых модель обучается восстанавливать замаскированные фрагменты сигнала по оставшемуся контексту [50]. Для ЭКГ такие методы позволяют модели изучать общую морфологию сигнала, устойчивые признаки волн и комплексов, а также особенности формы сигнала в разных отведениях. Ограничением подобных подходов является то, что качество восстановления сигнала не всегда напрямую коррелирует с диагностической полезностью формируемых представлений: модель может акцентировать внимание на точном воспроизведении шума и малозначимых деталей формы сигнала.

К третьему направлению относятся методы контрастного обучения и их многочисленные адаптации к временным рядам и биомедицинским сигналам. В таких подходах модель обучается формировать представления, в которых поло-

жительные пары объектов располагаются близко друг к другу, а отрицательные — далеко. Положительные пары в основном образуются из двух аугментированных версий одного и того же сигнала. Контрастное обучение особенно хорошо подходит для ЭКГ, поскольку позволяет использовать как временную структуру сигнала, так и многоканальную природу электрокардиограммы, не требуя при этом явного восстановления самого сигнала как промежуточной цели обучения [65; 67].

Несмотря на то что все три направления применимы к анализу ЭКГ и современные predictive- и reconstruction-based методы также демонстрируют высокое качество [49; 50], в настоящей работе основное внимание уделяется контрастному обучению. Эффективность контрастного предобучения для медицинских временных рядов и ЭКГ подтверждена в ряде современных исследований [65; 67]. При этом определяющим для настоящей работы является не предположение об универсальном превосходстве контрастных методов, а возможность непосредственно управлять способом формирования положительных пар. Такая постановка позволяет естественным образом включать в процесс предобучения дополнительную информацию, не связанную напрямую с формой сигнала, — в частности, различные комбинации отведений одной записи и метаданные пациента.

### **3.2.2 Контрастное обучение для 12-канальных электрокардиограмм**

В настоящей работе контрастное обучение используется в качестве базовой постановки самообучения. Для каждой исходной ЭКГ формируется связанное с ней представление — например, аугментированная версия того же сигнала, — после чего энкодер обучается сближать представления положительных пар и разделять представления различных записей.

В задачах анализа ЭКГ применяются различные подходы контрастного обучения. В экспериментальной части в качестве базовых методов рассматриваются SimCLR и TS-TCC. Первый представляет классическую постановку контрастного обучения на основе аугментированных представлений одного объекта, тогда как второй специально ориентирован на временные ряды и дополнитель-

но учитывает их временную структуру. Предлагаемый далее подход не является самостоятельной заменой указанных методов, а представляет собой расширение контрастной постановки за счёт дополнительного способа формирования положительных пар.

Для дальнейшего описания предлагаемых модификаций введём базовую контрастную функцию потерь, относительно которой будут формулироваться дополнительные компоненты.

Пусть  $x_i$  обозначает исходный объект,  $\tilde{x}_i$  — его аугментированную версию, а  $z_i$  и  $\tilde{z}_i$  — соответствующие представления, формируемые энкодером. Для положительной пары  $z_i$  и  $\tilde{z}_i$  определим функцию потерь

$$\ell(z_i, \tilde{z}_i) = -\log \frac{\exp(\text{sim}(z_i, \tilde{z}_i)/\tau)}{\sum_{k=1}^{2N} \mathbb{I}[k \neq i] \exp(\text{sim}(z_i, z_k)/\tau)}, \quad (3.1)$$

где  $\text{sim}(\cdot, \cdot)$  — мера сходства, в данной работе косинусное сходство, а  $\tau$  — температурный параметр. Аналогично определяется  $\ell(\tilde{z}_i, z_i)$ , где в качестве опорного выступает аугментированное представление  $\tilde{z}_i$ .

Тогда стандартная симметричная contrastive-постановка на основе NT-Xent может быть записана в виде

$$L_{aug} = \frac{1}{2N} \sum_{i=1}^N (\ell(z_i, \tilde{z}_i) + \ell(\tilde{z}_i, z_i)) \quad (3.2)$$

где каждое из  $2N$  представлений поочерёдно выступает в качестве опорного, а соответствующее ему представление того же исходного объекта образует положительную пару.

В SimCLR контрастное обучение выполняется непосредственно между представлениями исходного объекта и его аугментированной версии с использованием NT-Xent. В TS-TCC контрастное обучение организовано с учётом временной структуры сигнала и включает временные и контекстные компоненты, формирующие базовую функцию потерь метода.

### 3.2.3 Использование метаданных в контрастном обучении

Одной из особенностей медицинских данных является наличие дополнительной информации, не относящейся непосредственно к самому сигналу, но способной улучшить качество извлекаемых представлений. К такой информации относятся возраст пациента, идентификатор пациента, пол, клинические категории и другие признаки, доступные в наборе данных. В отличие от разметки патологий, требующей квалифицированного врача-эксперта и, как правило, просмотра всей записи, метаданные такого рода либо уже присутствуют в системе учёта пациентов, либо фиксируются автоматически в процессе регистрации ЭКГ. Это делает метаданные особенно удобным источником дополнительного сигнала именно для этапа предобучения, где предполагается использование больших объёмов данных, разметка для которых чаще всего отсутствует. В контексте контрастного обучения эти метаданные могут использоваться не только на этапе конечной классификации, но и непосредственно в процессе предобучения.

В контрастном обучении метаданные позволяют расширить понятие положительной пары. Помимо связи между исходным объектом и его аугментированной версией можно рассматривать как положительные пары различные записи с близкими значениями доступных характеристик пациента. Такой подход особенно актуален для медицинских задач, в которых семантическая близость объектов часто частично определяется именно клиническим контекстом.

Для числовых метаданных принадлежность к положительной паре определяется относительно текущего исходного объекта  $x_i$ . Для признака  $P_j$  введём множество допустимых положительных примеров

$$\mathcal{P}_j(i) = \{k \neq i : |P_j(x_i) - P_j(x_k)| \leq \text{diff}_j\}. \quad (3.3)$$

Таким образом, для числовых признаков речь идёт о формировании множества положительных примеров относительно каждого исходного объекта. В частности, если возраст трёх пациентов составляет 40, 42 и 43 года при  $\text{diff} = 2$ , то для пациента возрастом 40 лет запись пациента возрастом 42 года является допустимым положительным примером, а запись пациента возрастом 43 года — нет.

В экспериментах главы в качестве признака  $P_j$  используется возраст пациента. Выбор данного признака обусловлен его доступностью для значительной части записей клинического набора ISP, что позволяет использовать его при предобучении без привлечения диагностической разметки. Значение порога  $diff_j$  было выбрано по результатам предварительных экспериментов: наилучшее качество последующей классификации достигалось при  $diff_j = 2$  года, поэтому данное значение использовалось во всех дальнейших экспериментах. Для каждого исходного объекта  $x_i$  из множества  $\mathcal{P}_j(i)$  выбирается объект  $x_i^{P_j}$ , образующий с ним дополнительную положительную пару.

В предлагаемом подходе стандартная аугментационная контрастная компонента  $L_{aug}$  дополняется group-based составляющей

$$L_{group_j} = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(\text{sim}(z_i, z_i^{P_j})/\tau)}{\sum_{k=1}^{2N} \mathbb{I}[k \neq i] \exp(\text{sim}(z_i, z_k)/\tau)}. \quad (3.4)$$

Важно, что  $L_{aug}$  и  $L_{group_j}$  представляют собой отдельные контрастные слагаемые. Для каждого из них рассматриваются собственные  $2N$  представлений: в  $L_{aug}$  сопоставляются  $N$  исходных записей и  $N$  их аугментированных версий, а в  $L_{group_j}$  —  $N$  исходных записей и  $N$  выбранных для них положительных примеров по признаку  $P_j$ . Поэтому нормирующая сумма каждой NT-Xent компоненты содержит  $2N$ , а не  $3N$  элементов.

Итоговая функция потерь представляет собой взвешенную сумму аугментационной и group-based компонент:

$$L = \alpha L_{aug} + (1 - \alpha) \sum_{j=1}^M L_{group_j}, \quad (3.5)$$

где  $\alpha$  определяет относительный вклад стандартной и group-based компонент. В экспериментах настоящей главы используется один признак группировки ( $M = 1$ , возраст пациента), а коэффициент  $\alpha$  фиксируется на протяжении обучения и выбирается таким образом, чтобы обеспечить сопоставимый вклад двух компонент функции потерь. В данной работе коэффициент  $\alpha$  был выбран равным  $1/2$ .

Предложенная постановка требует специальной организации батчей. Функция потерь (3.4) максимизирует сходство представлений внутри объявлен-

ных положительных пар и минимизирует сходство для остальных пар объектов в батче, трактуемых как отрицательные примеры. Поэтому случайное присутствие в батче дополнительной записи, удовлетворяющей тому же критерию группировки, привело бы к появлению ложноотрицательной пары.

Батч формируется непосредственно из кортежей  $(x_i, \tilde{x}_i, x_i^{P_1}, \dots, x_i^{P_M})$ . Сначала выбирается исходная запись  $x_i$ , затем для каждого признака  $P_j$  из множества допустимых кандидатов  $\mathcal{P}_j(i)$  выбирается один объект  $x_i^{P_j}$  и вместе с аугментированной версией  $\tilde{x}_i$  резервируется как часть того же кортежа. При последующем заполнении батча запись, удовлетворяющая критерию группировки относительно уже выбранного исходного объекта, не добавляется в качестве независимого элемента, если она не является назначенным для него положительным примером. Таким образом, для каждого исходного объекта в соответствующей group-based компоненте присутствует один явно заданный положительный пример, а остальные записи этой же группы не используются как отрицательные.

Условие  $k \neq i$  означает, что положительная пара состоит из двух различных записей, однако не требует принадлежности этих записей разным пациентам. В экспериментах главы группировка выполняется только по возрасту, а дополнительное ограничение по идентификатору пациента не вводится. Поэтому в отдельных случаях две различные записи одного пациента также могут удовлетворять критерию положительной пары.

### 3.2.4 Постановка экспериментов и результаты

Для оценки эффективности контрастного предобучения на 12-канальных ЭКГ проведена серия экспериментов, направленных на сравнение различных вариантов обучения и оценку влияния предобучения на качество классификации ЭКГ.

В связи с этим представленные далее эксперименты сосредоточены на сравнении методов контрастного обучения между собой: базовых подходов SimCLR и TS-TCC и их group-based модификаций с группировкой по возрасту. В обоих случаях group-based составляющая добавляется к исходной функции

потерь базового метода на уровне общей функции потерь, не изменяя внутреннюю организацию соответствующего подхода контрастного обучения.

В качестве энкодера для всех сравниваемых методов используется архитектура `densenet1d121` – та же базовая модель, что была выбрана по результатам сравнения архитектур нейросетевых моделей в главе 2. Такой выбор позволяет напрямую сопоставлять результаты, полученные после контрастного обучения, с результатами обучения той же модели с учителем и с результатами дообучения на данных другого домена, рассмотренными в главе 2, поскольку во всех случаях используется одна и та же базовая сеть.

Предобучение всех моделей проводится на одном и том же наборе 12-канальных ЭКГ ISP – том же клиническом наборе, что и в главе 2, но без использования содержащихся в нём меток патологий: для базового контрастного обучения используются только сами сигналы и, для `group-based` вариантов, метаданные возраста пациента, тогда как разметка диагнозов на этом этапе не задействуется. Для SimCLR и дополнительной `group-based` компоненты используется NT-Xent loss с температурным параметром  $\tau = 0.2$ . В TS-TCC сохраняется исходная временная и контекстная постановка контрастного обучения, к которой добавляется `group-based` слагаемое. Для входных ЭКГ используется та же  $z$ -нормализация, что и в постановке классификации главы 2; поэтому параметры аугментаций задаются в нормализованном масштабе сигнала. В предварительных экспериментах рассматривались изменение амплитуды, добавление шума и временные сдвиги. По итоговому качеству последующей классификации для дальнейших экспериментов была выбрана аугментация типа `jitter`: к нормализованному сигналу добавляется шум, распределённый по нормальному закону с нулевым средним и стандартным отклонением 0.8 в единицах нормализованного сигнала.

Последующее дообучение выполняется на открытом наборе PTB-XL в постановке бинарной классификации отдельных патологий. Процедура дообучения повторяет схему, использованную для `densenet1d121` в главе 2 (п. 2.1.4): первые блоки предобученной сети замораживаются, а последние слои переинициализируются и дообучаются на целевой патологии.

В качестве основной задачи на дообучение в данном разделе рассматривается классификация сердечных патологий по 12-канальной ЭКГ. Из четырёх патологий, использованных в главе 2, для данного сравнения выбраны AFIB,

1AVB и PVC; CLBVB не включена в сравнение, поскольку уже базовая модель, обученная на размеченных данных, показывает по этой патологии высокое качество (см. табл. 3 главы 2, где G-mean по CLBVB для всех сравниваемых архитектур нейросетевых моделей превышал 0.94), вследствие чего диапазон возможного дополнительного улучшения в данной постановке существенно меньше. Для сравнения моделей используется метрика G-mean, а также чувствительность и специфичность как её составляющие. Результаты сравнения представлены в таблице 11.

Таблица 11 — Результаты классификации 12-канальных ЭКГ с использованием различных методов contrastive learning

| Метод               | AFIB  |       |              | 1AVB  |       |              | PVC   |       |              |
|---------------------|-------|-------|--------------|-------|-------|--------------|-------|-------|--------------|
|                     | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       | Se    | Sp    | G-mean       |
| Обучение с учителем | 0.937 | 0.955 | 0.946        | 0.908 | 0.871 | 0.889        | 0.936 | 0.963 | 0.949        |
| SimCLR              | 0.923 | 0.955 | 0.939        | 0.882 | 0.902 | 0.892        | 0.945 | 0.971 | 0.958        |
| SimCLR-GB (Age)     | 0.916 | 0.967 | 0.941        | 0.921 | 0.897 | <b>0.909</b> | 0.954 | 0.977 | <b>0.966</b> |
| TS-TCC              | 0.916 | 0.975 | 0.945        | 0.884 | 0.911 | 0.897        | 0.963 | 0.978 | 0.971        |
| TS-TCC-GB (Age)     | 0.923 | 0.979 | <b>0.951</b> | 0.895 | 0.928 | <b>0.911</b> | 0.972 | 0.982 | <b>0.977</b> |

Результаты различаются по патологиям. Для AFIB значения G-mean всех рассмотренных методов лежат в узком диапазоне 0.939–0.951 при значении 0.946 для модели, обученной с учителем. Добавление возрастной group-based компоненты повышает значение G-mean относительно соответствующего базового метода с 0.939 до 0.941 для SimCLR и с 0.945 до 0.951 для TS-TCC, при этом все значения остаются близкими к результатам модели, обученной с учителем.

Для 1AVB и PVC эффект group-based компоненты выражен сильнее. При парном сравнении SimCLR и SimCLR-GB G-mean увеличивается с 0.892 до 0.909 для 1AVB и с 0.958 до 0.966 для PVC. Для TS-TCC добавление той же компоненты повышает G-mean с 0.897 до 0.911 и с 0.971 до 0.977 соответственно. Таким образом, во всех шести парных сравнениях значение G-mean у group-based варианта выше. Для AFIB разница составляет 0.2 и 0.6 п.п., для 1AVB — 1.7 и 1.4 п.п., а для PVC — 0.8 и 0.6 п.п. для SimCLR и TS-TCC соответственно.

По результатам данного раздела можно заключить, что включение возраста пациента в качестве дополнительного критерия формирования положительных пар может повышать качество представлений, получаемых при контрастном преобучении.

### **3.3 Перенос представлений с 12-канальных электрокардиограмм на ЭКГ с меньшим числом отведений**

В данном разделе исследуется возможность переноса представлений, сформированных в ходе предобучения на 12-канальных электрокардиограммах, на ЭКГ с меньшим числом отведений. Такой перенос имеет практическую ценность, поскольку в реальных условиях всё более широкое распространение получают устройства, регистрирующие не полный набор из 12 отведений, а только одно или несколько отведений, в частности носимые устройства и упрощённые мобильные ЭКГ-системы.

По сравнению с 12-канальной ЭКГ одноканальные записи содержат меньше диагностической информации. Помимо этого, для них, как правило, доступен существенно меньший объём размеченных данных. Это обусловлено не только меньшей распространённостью открытых одноканальных наборов данных, но и тем, что стандартная клиническая регистрация ЭКГ обычно выполняется в 12 отведениях: клинические диагностические критерии исторически формировались прежде всего для 12-канальной записи, тогда как разметка одноканальных сигналов не относится к привычной рутинной практике кардиологов и требует дополнительных усилий по адаптации диагностических критериев к существенно менее информативному сигналу. В силу этого подходы самообучения приобретают особую значимость для задач анализа одноканальных ЭКГ. В данном разделе сначала рассматривается классификация одноканальной ЭКГ, после чего исследуется, в какой мере контрастное предобучение на 12-канальных данных позволяет повысить качество моделей классификации одноканальных ЭКГ.

#### **3.3.1 Особенности анализа 1-канальных ЭКГ**

Хотя стандартом клинической электрокардиографии остаётся 12-канальная запись, во многих прикладных сценариях используются ЭКГ с меньшим числом каналов. 1-канальные ЭКГ широко применяются в носимых устройствах, включая умные часы и портативные мониторы. Такие устройства обеспечивают про-

стую и быструю регистрацию сердечного сигнала, что делает их удобными для длительного мониторинга состояния пациента. Вместе с тем переход к одному отведению сопровождается заметной потерей диагностической информации.

Для задач анализа 1-канальных ЭКГ целесообразно выделить две основные проблемы. Во-первых, уменьшение числа каналов снижает количество информации о пространственной структуре электрической активности сердца. Во-вторых, для таких данных обычно имеется меньше размеченных наборов, чем для традиционной 12-канальной ЭКГ. Вследствие этого модели, обучаемые только на узком целевом наборе данных, могут демонстрировать ограниченное качество и слабую устойчивость.

В рамках данной работы показано, что даже в стандартной постановке обучения на размеченных данных классификация 1-канальных ЭКГ остаётся возможной и позволяет достигать приемлемого качества для ряда патологий. Однако дальнейшее использование контрастного обучения на 12-канальных ЭКГ-записях даёт возможность заметно улучшить результаты за счёт переноса знаний о структуре многоканальной ЭКГ в задачи с уменьшенным числом отведений. Предложенный принцип переноса не привязан жёстко к единственному отведению: если научиться устойчиво адаптировать представления, сформированные на 12-канальных записях, к 1-канальной постановке, аналогичная схема может быть применена и к промежуточным по информативности случаям – например, к записям с шестью отведениями, – расширяя область применимости подхода за пределы рассматриваемого в настоящей работе сценария анализа 1-канальных ЭКГ.

### **3.3.2 Используемые наборы данных**

Для исследования переноса представлений с 12-канальных ЭКГ на 1-канальные используются как многоканальные, так и 1-канальные варианты ЭКГ, различающиеся способом получения. В качестве источника данных для предобучения применяется набор 12-канальных электрокардиограмм ISP.

Для дообучения под конкретную задачу и оценки качества классификации применяются записи с одним отведением, полученные из двух принципиально

разных источников. Первый источник – редуцированные 1-канальные записи, полученные выделением первого отведения из стандартных 12-канальных ЭКГ открытого набора РТВ-XL: сигнал в них совпадает с одним из каналов исходной 12-канальной записи, зарегистрированной тем же прибором и в тех же условиях.

Второй источник – независимый тестовый набор WatchECG, содержащий 1000 1-канальных записей ЭКГ, зарегистрированных с помощью умных часов. Метки патологий для каждой одноканальной ЭКГ выставлялись врачом по данным анамнеза и дополнительному просмотру записи на предмет наличия заболевания. Каждая запись соответствует уникальному пациенту, поэтому набор включает данные 1000 различных пациентов. WatchECG используется исключительно для итогового тестирования: записи этого набора не участвуют ни в предобучении, ни в последующем дообучении моделей или подборе их гиперпараметров. Пациенты WatchECG не пересекаются с пациентами наборов ISP и РТВ-XL, используемых в экспериментах.

Основные характеристики используемых наборов данных приведены в таблице 12.

Таблица 12 — Наборы данных, используемые для переноса представлений с 12-канальных ЭКГ на 1-канальные

| Набор данных           | Число отведений | Этап использования    | Описание                                                                                                         |
|------------------------|-----------------|-----------------------|------------------------------------------------------------------------------------------------------------------|
| ISP                    | 12              | предобучение          | набор 12-канальных ЭКГ для предобучения (без разметки патологий)                                                 |
| PTB-XL (I отведение)   | 1               | дообучение, тест      | редуцированная 1-канальная запись – первое отведение, выделенное из 12-канальных записей открытого набора PTB-XL |
| WatchECG (I отведение) | 1               | дополнительная оценка | 1-канальные ЭКГ, зарегистрированные непосредственно с умных часов, а не выделенные из 12-канальных записей       |

### 3.3.3 Контрастное обучение для переноса знаний между конфигурациями с различным числом отведений

Ключевая идея переноса знаний с 12-канальных ЭКГ на сигналы с меньшим числом отведений состоит в том, что различные отведения одной и той же записи представляют собой взаимосвязанные представления одной и той же электрической активности сердца. Следовательно, многоканальная ЭКГ может

рассматриваться как описание одного объекта разными способами, что делает её естественной основой для контрастного обучения.

В отличие от подходов, рассмотренных в п. 3.2.2–3.2.3, где дополнительный положительный сигнал строится либо на аугментациях одного и того же сигнала, либо на близости метаданных различных записей, в настоящей работе предлагается постановка контрастного обучения, использующая в качестве источника дополнительной информации многоканальную структуру записи. Различные комбинации отведений одной и той же ЭКГ рассматриваются как согласованные представления одного объекта, к которым применяется принцип дополнительного контрастного сближения, аналогичный использованному для метаданных в п. 3.2.3, но уже внутри одной записи. Одна комбинация отведений, совпадающая с целевым сигналом для последующего дообучения, рассматривается как основная, а остальные выбранные комбинации — как вспомогательные. За счёт этого основной энкодер обучается выделять признаки, согласованные с представлениями, получаемыми из дополнительных отведений той же ЭКГ вспомогательными энкодерами, и тем самым использует многоканальную информацию ещё на этапе предобучения. Такой способ построения положительных пар между разными «проекциями» одного и того же физического процесса составляет основную идею предлагаемого метода LC-CL (*Lead Combinations with Contrastive Learning*).

Формально, пусть  $e \in \mathbb{R}^{12 \times L}$  — 12-канальная ЭКГ длины  $L$ , а  $k_1 = \{1\}, k_2, \dots, k_n$  — различные комбинации отведений, разбивающие подмножество отведений из 12 каналов (например,  $k_1$  может соответствовать первому отведению, а  $k_2$  — объединению всех оставшихся одиннадцати, или же отведения могут быть разбиты на несколько непересекающихся вспомогательных комбинаций). Основной энкодер получает на вход целевую для последующего дообучения комбинацию отведений  $e^{(k_1)}$  (в экспериментах настоящей главы — первое отведение), тогда как вспомогательные энкодеры  $E_{h_2}, \dots, E_{h_n}$  работают с остальными комбинациями отведений  $e^{(k_2)}, \dots, e^{(k_n)}$  независимо друг от друга. Для каждой комбинации, помимо исходного сигнала, строится и его аугментированная версия, так что каждый энкодер обрабатывает как исходные, так и аугментированные представления своей комбинации отведений.

В рамках данного подхода итоговая функция потерь складывается из трёх компонент, отражающих три типа согласованности представлений.

Основная компонента контрастного обучения задаётся как

$$L_{main} = L_{NTXent}(h_l^{(k_1)}, h_m^{(k_1)}), \quad (3.6)$$

где положительной парой являются исходный и аугментированный варианты сигнала первого отведения; эта компонента отвечает за качество представлений собственно целевого сигнала 1-канальной ЭКГ и по форме совпадает со стандартной аугментационной компонентой контрастного обучения  $L_{aug}$  (п. 3.2.2).

Для вспомогательных энкодеров аналогичным образом вводится helper loss:

$$L_{helper}^i = L_{NTXent}(h_l^{(k_i)}, h_m^{(k_i)}), \quad i = 2, \dots, n, \quad (3.7)$$

которая обеспечивает согласованность между исходными и аугментированными представлениями соответствующих вспомогательных комбинаций отведений и тем самым позволяет каждому вспомогательному энкодеру самостоятельно формировать содержательное представление «своих» отведений.

Наконец, ключевая для переноса знаний между отведениями компонента – entity loss, согласующая представления основного энкодера с представлениями каждого из вспомогательных:

$$L_{entity}^j = L_{NTXent}(h_l^{(k_1)}, h_m^{(k_j)}), \quad j = 2, \dots, n. \quad (3.8)$$

Именно эта компонента непосредственно реализует перенос знаний: она заставляет представление, построенное основным энкодером только по первому отведению, быть согласованным с представлениями, построенными вспомогательными энкодерами по другим отведениям той же записи, – то есть связывает 1-канальное представление с информацией, извлекаемой из выбранных дополнительных отведений той же записи.

Итоговая функция потерь имеет вид взвешенной суммы всех трёх компонент:

$$L = \lambda_1 L_{main} + \sum_{i=2}^n \lambda_i L_{helper}^i + \sum_{j=2}^n \eta_j L_{entity}^j, \quad (3.9)$$

где коэффициенты  $\lambda_i$  и  $\eta_j$  определяют относительный вклад каждой компоненты в общую функцию потерь и позволяют настраивать, в какой мере обучение приоритизирует качество представлений отдельных комбинаций отведений

$(L_{main}, L_{helper}^i)$  по сравнению с согласованностью между основным и вспомогательными энкодерами ( $L_{entity}^j$ ). В данной работе для всех компонент функции потерь используются одинаковые весовые коэффициенты  $\lambda_i = \eta_j = 0.33$ .

Схема предобучения LC-CL представлена на рис. 3.1. Основной энкодер  $E_m$  получает первое отведение, тогда как один или несколько вспомогательных энкодеров  $E_{h_i}$  обрабатывают выбранные комбинации остальных отведений. Для каждой ветви формируются исходное и аугментированное представления;  $L_{main}$  и  $L_{helper}$  обеспечивают контрастное согласование внутри соответствующих ветвей, а  $L_{entity}$  связывает представление первого отведения с представлениями дополнительных отведений той же 12-канальной записи.

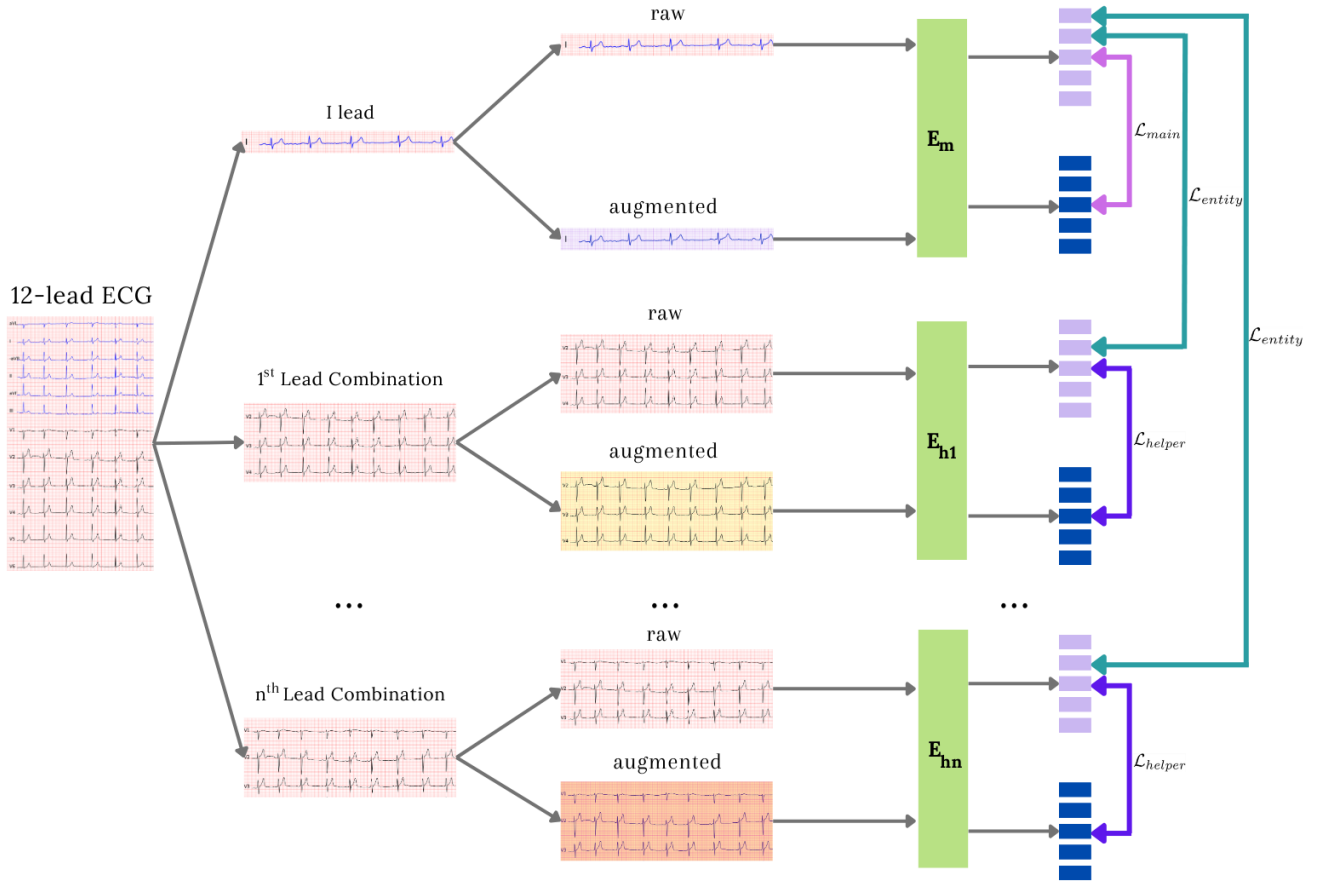


Рисунок 3.1 — Схема LC-CL предобучения для переноса представлений с 12-канальной ЭКГ на модель, использующую одно отведение.

В работе рассматриваются две конфигурации LC-CL, различающиеся распределением дополнительных отведений между вспомогательными энкодерами. В обеих конфигурациях основной энкодер получает первое отведение ЭКГ, представление которого после предобучения используется в целевой 1-канальной задаче.

В конфигурации 1vs7 используется один вспомогательный энкодер, на вход которого совместно подаются отведение II и шесть грудных отведений  $V_1-V_6$ . В конфигурации 1vs1vs6 используются два вспомогательных энкодера. Первый из них получает только отведение II, тогда как второй получает группу из шести грудных отведений  $V_1-V_6$ . После завершения предобучения вспомогательные энкодеры не используются: для последующего дообучения и классификации сохраняется только основной энкодер, принимающий на вход первое отведение.

### 3.3.4 Постановка экспериментов и результаты

Для оценки предлагаемого подхода проведена серия экспериментов по классификации 1-канальной ЭКГ. На первом этапе строится базовая классификация 1-канальных ЭКГ-записей с использованием размеченных данных, аналогично постановке предыдущей главы. Это позволяет показать, что классификация при уменьшении числа каналов до одного остаётся возможной, хотя и сопровождается заметным снижением качества по сравнению с 12-канальной постановкой (см. п. 3.3.1).

На втором этапе исследуется эффект контрастного предобучения на 12-канальных ЭКГ с последующим дообучением на записях с меньшим числом отведений. В рамках оценки сравниваются следующие варианты обучения:

- Supervised - обучение модели с учителем непосредственно на первом отведении, выделенном из размеченных 12-канальных ЭКГ, без предобучения;
- SimCLR - предварительное контрастное обучение на ISP с последующим дообучением основного энкодера на 1-канальной целевой задаче;
- TS-TCC - аналогичная схема с использованием TS-TCC в качестве метода контрастного предобучения;
- LC-CL - предлагаемый метод предобучения, в котором представление первого отведения дополнительно согласуется с представлениями других отведений той же 12-канальной ЭКГ.

Для LC-CL рассматриваются две конфигурации организации вспомогательных энкодеров:

- 1vs7 — основной энкодер получает первое отведение, а единственный вспомогательный энкодер — совместно отведение II и шесть грудных отведений  $V_1-V_6$ ;
- 1vs1vs6 — первое отведение также используется основным энкодером, отведение II и группа грудных отведений  $V_1-V_6$  обрабатываются двумя отдельными вспомогательными энкодерами.

В качестве целевых патологий в данном разделе выбраны AFIB, CLBBB, 1AVB и PVC — те же патологии, что и в главе 2. Такой набор клинически значимых состояний также различается по сложности детектирования на 1-канальных ЭКГ. Для оценки качества используются чувствительность, специфичность и G-mean.

Результаты сравнения методов для анализа 1-канальных ЭКГ на PTB-XL представлены в таблице 13, а на реальном наборе 1-канальных ЭКГ WatchECG — в таблице 14.

Таблица 13 — Результаты моделей классификации на первом отведении PTB-XL

| Модель                    | AFIB  |       |        | CLBBB |       |        | 1AVB  |       |        | PVC   |       |        |
|---------------------------|-------|-------|--------|-------|-------|--------|-------|-------|--------|-------|-------|--------|
|                           | Se    | Sp    | G-mean | Se    | Sp    | G-mean | Se    | Sp    | G-mean | Se    | Sp    | G-mean |
| Supervised (12 отведений) | 0.937 | 0.955 | 0.946  | 1.000 | 0.947 | 0.973  | 0.908 | 0.871 | 0.889  | 0.936 | 0.963 | 0.949  |
| Supervised (1 отведение)  | 0.902 | 0.897 | 0.900  | 0.939 | 0.959 | 0.949  | 0.921 | 0.772 | 0.843  | 0.844 | 0.905 | 0.874  |
| SimCLR                    | 0.937 | 0.936 | 0.937  | 1.000 | 0.936 | 0.967  | 0.908 | 0.837 | 0.872  | 0.936 | 0.938 | 0.937  |
| TS-TCC                    | 0.951 | 0.961 | 0.956  | 0.980 | 0.951 | 0.965  | 0.882 | 0.878 | 0.880  | 0.972 | 0.967 | 0.970  |
| LC-CL (1vs7)              | 0.909 | 0.962 | 0.935  | 0.980 | 0.978 | 0.979  | 0.842 | 0.882 | 0.862  | 0.936 | 0.959 | 0.947  |
| LC-CL (1vs1vs6)           | 0.930 | 0.949 | 0.940  | 1.000 | 0.952 | 0.976  | 0.934 | 0.848 | 0.890  | 0.945 | 0.945 | 0.945  |

Таблица 14 — Результаты моделей классификации 1-канальных ЭКГ на данных WatchECG

| Модель          | AFIB  |       |        | CLBBB |       |        | 1AVB  |       |        | PVC   |       |        |
|-----------------|-------|-------|--------|-------|-------|--------|-------|-------|--------|-------|-------|--------|
|                 | Se    | Sp    | G-mean | Se    | Sp    | G-mean | Se    | Sp    | G-mean | Se    | Sp    | G-mean |
| Supervised      | 0.908 | 0.946 | 0.927  | 0.916 | 0.920 | 0.918  | 0.715 | 0.865 | 0.786  | 0.926 | 0.780 | 0.850  |
| SimCLR          | 0.916 | 0.945 | 0.930  | 0.924 | 0.920 | 0.922  | 0.739 | 0.886 | 0.809  | 0.930 | 0.775 | 0.849  |
| TS-TCC          | 0.920 | 0.960 | 0.940  | 0.930 | 0.924 | 0.927  | 0.725 | 0.862 | 0.791  | 0.945 | 0.780 | 0.859  |
| LC-CL (1vs7)    | 0.920 | 0.969 | 0.944  | 0.925 | 0.954 | 0.939  | 0.770 | 0.885 | 0.825  | 0.953 | 0.787 | 0.866  |
| LC-CL (1vs1vs6) | 0.927 | 0.971 | 0.949  | 0.935 | 0.961 | 0.948  | 0.714 | 0.892 | 0.798  | 0.959 | 0.805 | 0.878  |

Результаты, представленные в таблице 13, прежде всего подтверждают ожидаемое снижение качества при переходе от 12 отведений к одному: для модели, обученной с учителем, G-mean по сравнению с 12-канальной постановкой снижается на 4.6 п.п. для AFIB ( $0.946 \rightarrow 0.900$ ), на столько же для 1AVB ( $0.889 \rightarrow 0.843$ ) и на 7.5 п.п. для PVC ( $0.949 \rightarrow 0.874$ ), тогда как по CLBBB

падение минимально – 2.4 п.п. ( $0.973 \rightarrow 0.949$ ) – морфология комплекса *QRS*, определяющая эту патологию, остаётся достаточно различимой уже по одному отведению.

На фоне этого падения контрастное предобучение оказывается заметно более полезным, чем в 12-канальной постановке (п. 3.2.4): в отличие от неё, где для AFIB прирост был близок к нулю, здесь контрастное обучение улучшает качество по сравнению с обучением с учителем по всем четырём патологиям и для всех рассмотренных методов: например, для PVC G-mean растёт с 0.874 (Supervised) до 0.937–0.970 у различных contrastive-подходов, а для 1AVB – с 0.843 до 0.862–0.890.

TS-TCC показывает лучший результат по AFIB (0.956) и PVC (0.970), тогда как предлагаемый LC-CL превосходит остальные методы по CLBBB (конфигурация 1vs7, 0.979) и 1AVB (конфигурация 1vs1vs6, 0.890). Такое распределение согласуется с общей медицинской логикой: AFIB достаточно надёжно распознаётся уже по одному отведению за счёт выраженной нерегулярности ритма, поэтому перенос межотведённых зависимостей даёт для этой патологии сравнительно небольшой дополнительный эффект, и метод, использующий временную структуру сигнала (TS-TCC), оказывается конкурентным с LC-CL или превосходит его. Для патологий, диагностика которых существенно опирается на форму комплекса *QRS* в разных проекциях (CLBBB) или на точное измерение временных интервалов (1AVB), знания о взаимосвязях между отведениями, перенесённые с помощью LC-CL, оказываются более ценными.

Сопоставление двух конфигураций LC-CL показывает, что наибольшее значение G-mean зависит от патологии: 1vs7 даёт более высокое значение по CLBBB, тогда как 1vs1vs6 — по 1AVB и, с небольшим отрывом, по AFIB. Это означает, что оптимальная схема вспомогательных энкодеров в методе LC-CL зависит от типа патологии, а не является универсальной.

Результаты на независимом тестовом наборе 1-канальных ЭКГ WatchECG (табл. 14) позволяют оценить предлагаемый LC-CL в условиях, когда целевые записи изначально зарегистрированы 1-канальным устройством, а не получены выделением первого отведения из 12-канальной ЭКГ. Обе конфигурации LC-CL в большинстве случаев превосходят модель без предварительного обучения. Для AFIB G-mean увеличивается с 0.927 до 0.944 и 0.949 для конфигураций 1vs7 и 1vs1vs6 соответственно, для CLBBB – с 0.918 до 0.939 и 0.948, а для PVC –

с 0.850 до 0.866 и 0.878. Для 1AVB наиболее заметное улучшение достигается конфигурацией 1vs7, повышающей G-mean с 0.786 до 0.825, тогда как для 1vs1vs6 значение составляет 0.798.

При этом для каждой из четырёх рассмотренных патологий наибольшее значение G-mean среди сравниваемых методов достигается одной из двух конфигураций LC-CL: 1vs1vs6 для AFIB, CLBBB и PVC и 1vs7 для 1AVB. Это дополнительно показывает, что наиболее эффективный способ организации вспомогательных энкодеров LC-CL зависит от целевой диагностической задачи.

Значения метрики на WatchECG и на первом отведении РТВ-XL при этом различаются. В частности, максимальный G-mean для 1AVB составляет 0.825 на WatchECG против 0.890 на РТВ-XL, а для PVC – 0.878 против 0.970. Прямо связывать это различие исключительно с уровнем шума некорректно, поскольку наборы одновременно различаются по способу регистрации, используемому оборудованию и составу ЭКГ-записей. Тем не менее положительный эффект LC-CL сохраняется как в контролируемой постановке с выделением первого отведения из РТВ-XL, так и при тестировании на независимых ЭКГ, изначально зарегистрированных с помощью 1-канального носимого устройства.

Таким образом, результаты данного раздела показывают, что контрастное предобучение на 12-канальных ЭКГ с использованием предложенного метода LC-CL позволяет улучшить классификацию 1-канальной ЭКГ как на редуцированных записях РТВ-XL, так и на записях, изначально зарегистрированных 1-канальными устройствами. Тем самым подтверждается возможность переноса представлений с 12-канальных ЭКГ на 1-канальные и демонстрируется практическая ценность подходов предобучения для задач, где размеченных 1-канальных данных недостаточно.

### 3.4 Выводы

В данной главе рассмотрено использование контрастного обучения для задач анализа электрокардиограмм. Показано, что применение неразмеченных данных является перспективным направлением в условиях ограниченной доступности экспертной разметки, характерной для медицинского домена. Это осо-

бенно важно для ЭКГ, где количество сырых сигналов существенно превышает объём полностью аннотированных наборов данных.

В рамках главы предложены и реализованы два метода, оба развивающие идею контрастного обучения за счёт дополнительной, специфичной для домена структуры данных. Во-первых, предложен метод контрастного предобучения с использованием метаданных пациента, который реализован следующим образом. Для каждой исходной ЭКГ формируется стандартная положительная пара с её аугментированной версией и дополнительная положительная пара с другой записью, удовлетворяющей критерию близости метаданных; в проведённых экспериментах таким критерием служила разница возраста не более двух лет. Общая функция потерь объединяет стандартную аугментационную и group-based компоненты с коэффициентом  $\alpha = 1/2$ , а батчи формируются так, чтобы дополнительные записи, удовлетворяющие тому же критерию группировки, не становились ложноотрицательными примерами. После такого предобучения полученный энкодер используется для последующего обучения классификатора.

Применение предложенного подхода к базовым методам SimCLR и TS-TCC позволило отдельно оценить вклад group-based компоненты относительно соответствующих исходных методов. Для SimCLR добавление группировки по возрасту повысило G-mean на 0.2, 1.7 и 0.8 процентных пункта для AFIB, 1AVB и PVC соответственно, а для TS-TCC – на 0.6, 1.4 и 0.6 процентных пункта. Таким образом, положительный эффект наблюдается для обоих рассмотренных базовых методов и всех трёх диагностических задач, что показывает возможность использования метаданных пациента как дополнительного источника информации непосредственно на этапе предобучения.

Во-вторых, предложен метод LC-CL для повышения качества классификации записей с меньшим числом отведений. На этапе предобучения целевое отведение обрабатывается основным энкодером, а выбранные дополнительные отведения той же записи — одним или несколькими вспомогательными энкодерами. Функция потерь одновременно поддерживает контрастное согласование исходной и аугментированной версий внутри каждой ветви и согласует представление целевого отведения с представлениями дополнительных отведений той же записи. После завершения предобучения вспомогательные энкодеры удаляются, а основной энкодер дообучается и применяется уже только к целевой записи с уменьшенным числом каналов.

Показано, что такой перенос даёт выигрыш до 7.3 п.п. G-mean по сравнению с обучением 1-канальной модели с учителем на редуцированных из РТВ-XL записях. На наборе реальных 1-канальных записей WatchECG наилучшие значения G-mean по каждой из четырёх патологий получены одной из двух рассмотренных конфигураций LC-CL. При этом абсолютные результаты на РТВ-XL и WatchECG различаются, что может быть связано с различиями между наборами по условиям регистрации, используемому оборудованию и характеристикам ЭКГ-сигналов. Полученные результаты подтверждают возможность переноса представлений с 12-канальных ЭКГ на 1-канальные записи и показывают перспективность такого подхода для сценариев, в которых объём размеченных 1-канальных данных ограничен.

## Глава 4. Федеративное обучение

Четвёртая глава посвящена совместному обучению моделей на распределённых медицинских данных. В первой части главы рассматривается система федеративного обучения на основе Flower и практическая постановка ее применения для межклинического сценария. Во второй части главы предлагается метод BANT, который оценивает полезность клиентских обновлений на доверенной серверной выборке и использует полученные оценки при формировании глобальной модели. Данный результат соответствует четвёртому пункту научной новизны и раскрывает третье положение, выносимое на защиту.

### 4.1 Федеративное обучение в задачах анализа электрокардиограмм

Как отмечено в главе 1, федеративное обучение представляет собой естественную постановку для медицинских данных: исходные записи остаются в локальных информационных системах клиник, а совместное обучение выполняется за счёт обмена параметрами модели или их обновлениями. Для анализа электрокардиограмм такая схема приобретает особую значимость, поскольку данные распределены между медицинскими организациями, а их централизованное объединение часто ограничено требованиями конфиденциальности, внутренними регламентами и технической разнородностью инфраструктуры.

Практическая мотивация федеративного обучения в данной работе состоит в следующем. Отдельная клиника может располагать ограниченным числом записей, особенно для редких патологий. Как следствие, модель, обученная только на локальных данных, оказывается чувствительной к особенностям конкретного учреждения: типу кардиографа, составу пациентов, качеству разметки и частоте отдельных диагнозов. Если же несколько клиник участвуют в общем обучении, каждая из них сохраняет контроль над своими данными, однако итоговая модель обучается с использованием более разнообразных примеров. При этом федеративная постановка требует работающей инфраструктуры, соединяющей территориально разнесённые клиники.

Соответственно, в данной главе рассматриваются два аспекта федеративного обучения. В первой части (раздел 4.2) описывается практическая система федеративного обучения, предназначенная для развёртывания в реальном меж-клиническом сценарии. Рассматривается её архитектура, а также проводится эксперимент, демонстрирующий функциональные возможности такой системы на данных нескольких больниц. Во второй части (раздел 4.3) исследуется постановка устойчивости федеративного обучения к некорректным клиентским обновлениям, которые в медицинской сети могут возникать не только вследствие злонамеренных действий, но и из-за ошибок в разметке.

## 4.2 Практическая система федеративного обучения для анализа ЭКГ

### 4.2.1 Архитектура сервиса

Для проверки применимости федеративного обучения к реальному меж-клиническому сценарию в рамках работы была реализована и развёрнута система, воспроизводящая условия функционирования в распределённой медицинской инфраструктуре: клиники физически разделены, обладают собственной вычислительной инфраструктурой и не предполагают передачу исходных записей ЭКГ за пределы своего контура. Система построена на основе открытого фреймворка Flower [98], который разделяет федеративное приложение на две независимые части: инфраструктуру, отвечающую за оркестрацию обучения, и прикладной код, специфичный для конкретной задачи.

Инфраструктурный слой состоит из двух типов процессов. *SuperLink* – центральный координатор, который запускается на стороне, ответственной за сборку глобальной модели, и предоставляет три сетевых интерфейса: API для взаимодействия с серверным приложением (передача стратегии агрегации и конфигурации раунда), API для запуска обучения (используется командой `flwr run` для постановки федерации в очередь на выполнение) и API для взаимодействия с клиентскими узлами (передача заданий и приём обновлений). *SuperNode* – агент, который разворачивается непосредственно в контуре клиники и регистри-

руется на SuperLink через клиентский API; именно SuperNode гарантирует, что исходные данные не покидают локальную инфраструктуру.

Прикладной слой оформлен в виде двух контейнеризованных Docker-образов – *ServerApp* и *ClientApp*, – которые содержат код, специфичный для решаемой задачи. *ServerApp* определяет серверную логику федеративного обучения, включая выбор и настройку стратегии агрегации, управление раундами и, при необходимости, оценку глобальной модели. *ClientApp* содержит локальный цикл обучения, процедуру подготовки данных и формирования обновлений модели. Образы собираются на основе официальных базовых образов `flwr/serverapp` и `flwr/clientapp`, в которые устанавливаются необходимые для обучения и тестирования зависимости. При запуске SuperNode поднимает соответствующий *ClientApp* и монтирует в него локальный набор данных клиники в режиме «только чтение»: за пределы локального контура исходные записи не передаются.

В общем случае раунд федеративного обучения организован следующим образом. SuperLink через *ServerApp* рассылает выбранным SuperNode текущие параметры модели и конфигурацию раунда. На каждом клиенте *ClientApp* выполняет предусмотренную выбранным алгоритмом процедуру локального обучения и возвращает сформированное обновление модели. *ServerApp* получает обновления участников и обрабатывает их в соответствии с заданной стратегией агрегации. Таким образом, конкретный алгоритм агрегации не является частью инфраструктурного слоя и может быть заменён без изменения механизмов взаимодействия между SuperLink и SuperNode. Дополнительно *ServerApp* может оценивать качество полученной модели на отдельной тестовой выборке, чтобы отслеживать сходимость обучения независимо от локальных данных клиентов.

Особое внимание уделено защите канала связи, поскольку в клинической практике по нему передаются хотя и не сами записи ЭКГ, но параметры модели, а также сведения об участии конкретной клиники в обучении. Во-первых, весь трафик между SuperLink и SuperNode шифруется по TLS: генерируется собственный корневой удостоверяющий центр, на его основе выпускается сертификат сервера с указанием разрешённых адресов подключения, после чего оба типа процессов запускаются с флагами, задающими сертификат центра и ключевую пару сервера. Во-вторых, используется аутентификация клиентских узлов: для каждой клиники заранее генерируется собственная пара ключей (приватный ключ хранится на стороне клиники, публичный – заносится в список разрешён-

ных участников на сервере), и SuperLink принимает подключение только от узлов, чей публичный ключ присутствует в этом списке. Такая схема исключает подключение неавторизованных узлов к федерации и обеспечивает конфиденциальность канала передачи обновлений.

Работоспособность описанной архитектуры проверена на разнесённых физических серверах: SuperLink и один SuperNode были развёрнуты на сервере с видеокартами NVIDIA A100, а ещё два SuperNode – на отдельном сервере с видеокартами T4, соединённом с первым по сети. Все процессы взаимодействовали через реальные сетевые адреса с включённым TLS и авторизацией узлов, что подтверждает пригодность системы к развёртыванию в межклиническом сценарии, а не только к воспроизведению федеративного обучения на одной машине. Важное эксплуатационное свойство системы – возможность динамически подключать и отключать клиентов во время работающей федерации: новая клиника регистрирует свой SuperNode с указанием собственного идентификатора раздела данных без необходимости остановки федерации, а выход клиники из обучения сводится к остановке соответствующего процесса. Поскольку прикладная логика задачи полностью заключена в отдельно собираемых образах ServerApp/ClientApp, а инфраструктурный слой от неё не зависит, добавление новой медицинской постановки – другой патологии, другого типа сигнала или иного диагностического домена – сводится к замене прикладного кода задачи без изменения системы оркестрации, аутентификации и шифрования.

#### **4.2.2 Постановка эксперимента на данных ISP**

Цель эксперимента заключается в проверке практической необходимости федеративного обучения в задаче классификации ЭКГ: продемонстрировать, что федеративный сценарий с участием нескольких клиник позволяет достичь качества, существенно превосходящего локальное обучение отдельного участника и приближающегося к качеству модели, обученной централизованно на объединённых данных. Такая постановка воспроизводит реальный сценарий, в котором отдельный медицинский центр не передаёт свои ЭКГ-записи во внешнее храни-

лице, но может участвовать в совместном обучении модели через описанную выше систему.

В эксперименте используется тот же источник данных, что и в главе 2 – закрытый клинический набор ISP (см. таблицу 2), собранный в нескольких медицинских учреждениях. Для эксперимента был сформирован фиксированный подкорпус ISP50 из данных 50 медицинских учреждений, содержащий 144 650 размеченных ЭКГ-записей. В данных присутствует идентификатор `hospital_id`, что позволяет формировать клиентов в соответствии с реальным источником записей, а не на основе искусственного разбиения: каждый клиент в эксперименте представляет собой отдельную больницу из состава ISP, а варьирование числа клиентов (1, 10, 25, 50) соответствует включению в обучение первых  $n$  таких больниц.

Рассматривается multi-label классификация ЭКГ-патологий по четырём патологиям, используемым на протяжении всей работы, – AFIB, 1AVB, CLBBB и PVC. В качестве модели используется одномерная свёрточная архитектура ResNet1d18. Обучение проводится с использованием бинарной кросс-энтропии; основной метрикой качества, как и в остальной работе, является G-mean. В качестве стратегии федеративного обучения в данном эксперименте применяется алгоритм FedAvg: на каждом коммуникационном раунде клиенты выполняют локальное обучение, после чего полученные обновления агрегируются на сервере взвешенным усреднением пропорционально объёму локальных выборок.

Эксперименты с варьированием числа клиентов проводились с использованием того же прикладного кода ServerApp/ClientApp, что и в описанной архитектуре, но в режиме симуляции нескольких клиентов на общих вычислительных ресурсах, а не на физически разнесённых узлах. При этом важно, что сами клиенты в симуляции – это не синтетическое разбиение, а реальные разделы данных ISP50 по `hospital_id`, поэтому статистическая неоднородность между клиентами отражает подлинные межклинические различия.

### 4.2.3 Разбиение данных между клиентами

В федеративной постановке данные распределялись между клиентами, соответствующими клиникам – источникам записей ISP. Это существенно, поскольку в медицинской практике клиенты редко являются статистически одинаковыми: различаются численность пациентов, спектр патологий, качество сигналов и правила разметки. Даже при относительно большом общем числе записей локальная выборка отдельной клиники может содержать мало положительных примеров некоторых патологий.

Для иллюстрации локального дисбаланса в таблице 15 приведено распределение записей для одного из клиентов, использованного в режиме локального обучения. Этот клиент содержит 2885 записей, однако для отдельных целевых патологий число положительных примеров составляет лишь десятки случаев. Такая ситуация типична для клинических данных и объясняет, почему обучение только на одной локальной выборке может давать нестабильное качество.

Таблица 15 — Распределение записей одного локального клиента в эксперименте на данных ISP

| <b>Other</b> | <b>1AVB</b> | <b>PVC</b> | <b>AFIB</b> | <b>CLBBB</b> | <b>Всего</b> |
|--------------|-------------|------------|-------------|--------------|--------------|
| 2615         | 32          | 53         | 13          | 34           | 2885         |

### 4.2.4 Результаты на различном числе клиентов

В таблице 16 приведены результаты по G-mean для локального обучения, федеративного обучения при разном числе участвующих клиентов, а также централизованного обучения на объединённых данных 50 клиентов. Последний сценарий используется как ориентир, показывающий качество, достижимое при непосредственном совместном использовании данных всех рассматриваемых клиник.

Таблица 16 — Качество классификации ЭКГ при увеличении числа клиентов в федеративном обучении (G-mean)

| Режим обучения                                  | 1AVB  | PVC   | AFIB  | CLBBB |
|-------------------------------------------------|-------|-------|-------|-------|
| 1 клиент                                        | 0.482 | 0.547 | 0.779 | 0.879 |
| 10 клиентов                                     | 0.861 | 0.902 | 0.941 | 0.947 |
| 25 клиентов                                     | 0.889 | 0.907 | 0.947 | 0.961 |
| 50 клиентов                                     | 0.891 | 0.903 | 0.961 | 0.962 |
| Централизованное обучение на данных 50 клиентов | 0.896 | 0.934 | 0.958 | 0.966 |

Для одного клиента качество оказывается ограниченным: G-mean для 1AVB равен 0.482, для PVC — 0.547. При переходе к 50 клиентам значения G-mean возрастают до 0.891 для 1AVB, 0.903 для PVC, 0.961 для AFIB и 0.962 для CLBBB. Сопоставление с централизованным обучением (0.896, 0.934, 0.958 и 0.966 соответственно) показывает, что федеративная модель на 50 клиентах практически достигает уровня централизованного обучения. Таким образом, федеративное обучение позволяет компенсировать недостаточность локальных данных и получить модель, практически не уступающую по качеству модели, обученной на объединённом датасете, при том что ни одна клиника не передаёт исходные записи за пределы своего контура.

Чтобы дополнительно оценить устойчивость обученной модели, была проведена проверка на внешнем наборе PTB-XL. В этом эксперименте модель обучалась на данных ISP, а тестировалась на PTB-XL для тех же четырёх патологий.

Результаты приведены в таблице 17. Локальная модель, обученная на одном клиенте, плохо переносится на внешний домен: для 1AVB и AFIB значение G-mean равно 0.000. Напротив, федеративная модель, полученная при обучении 50 клиентов, сохраняет высокие значения G-mean по большинству классов.

Таблица 17 — Проверка переноса с ISP на PTB-XL (G-mean)

| Режим обучения               | 1AVB  | PVC   | AFIB  | CLBBB |
|------------------------------|-------|-------|-------|-------|
| Локальное обучение, 1 клиент | 0.000 | 0.488 | 0.000 | 0.682 |
| FedAvg, 50 клиентов          | 0.797 | 0.909 | 0.925 | 0.989 |

#### 4.2.5 Динамика достижения целевого качества

Отдельный практический вопрос состоит в том, как изменяется динамика достижения целевого качества при переходе от локального к федеративному обучению. Поскольку в федеративном режиме глобальная модель получает информацию от всех участников, отдельному клиенту не обязательно самостоятельно проходить весь путь локального обучения, чтобы получить модель сопоставимого качества: часть полезной информации поступает от обновлений других клиентов. Так, локальное обучение одного клиента достигает своего лучшего качества (см. таблицу 16, строка «1 клиент») после 31 локальной эпохи, тогда как в федеративном режиме с 50 клиентами сопоставимое или более высокое качество глобальной модели достигается уже примерно к 9-му коммуникационному раунду. В таблице 18 эта зависимость представлена для локального и федеративного сценариев обучения.

Таблица 18 — Динамика достижения целевого качества в локальном и федеративном сценариях обучения

| Сценарий                   | Этап достижения качества    | Результат                                     |
|----------------------------|-----------------------------|-----------------------------------------------|
| Локальное обучение         | 31 локальная эпоха          | Лучшее качество локальной модели              |
| Федеративное обучение      | ≈9-й коммуникационный раунд | Качество, сопоставимое с локальным ориентиром |
| Полный федеративный прогон | 50 коммуникационных раундов | Итоговое качество федеративной модели         |

Полученные результаты показывают, что в федеративной постановке качество, сопоставимое с лучшим результатом локального обучения отдельного клиента, достигается уже на раннем этапе обучения глобальной модели. При этом полный федеративный прогон из 50 коммуникационных раундов приводит к существенно более высокому и сбалансированному качеству по рассматриваемым патологиям (см. таблицу 16). Такая динамика объясняется тем, что при

формировании глобальной модели используются обновления всех участвующих клиник, и каждый отдельный клиент получает пользу от информации, содержащейся в данных других участников.

При этом число коммуникационных раундов нельзя непосредственно интерпретировать как объём локальных вычислений или полное время обучения. Федеративный раунд включает локальное обучение клиентов, передачу параметров модели по сети и, в синхронной постановке, ожидание завершения вычислений другими участниками. Поэтому приведённое сравнение характеризует прежде всего динамику достижения заданного уровня качества глобальной моделью, а не вычислительный выигрыш отдельного клиента. Оптимизация вычислительных и коммуникационных затрат является отдельной задачей федеративного обучения.

Совокупность приведённых результатов подтверждает практическую применимость реализованной системы: она допускает защищённое развёртывание на разнородной инфраструктуре нескольких клиник, обеспечивает качество классификации, близкое к централизованному обучению, и повышает устойчивость модели при переносе на внешние данные. Поскольку прикладная логика задачи в этой архитектуре отделена от инфраструктурного слоя, аналогичная схема развёртывания без изменения ядра системы допускает подключение новых клиник или постановку новых медицинских задач поверх уже работающей федерации. Это является основой для дальнейшего рассмотрения условий, в которых невозможно достичь абсолютного доверия к клиентским обновлениям.

### **4.3 Оценка клиентских обновлений при ошибочной разметке в федеративном обучении**

#### **4.3.1 Проблема ненадёжной разметки в медицинской федерации**

Стандартный алгоритм FedAvg, использованный в разделе 4.2, предполагает, что все клиенты корректно выполняют локальное обучение и отправляют на сервер полезные обновления модели. В классической постановке федеративного

обучения нарушение этого предположения обычно связывают с византийскими клиентами. Византийским называют участника, поведение которого не соответствует протоколу обучения: он может отправлять произвольные, зашумлённые или специально искажённые обновления, ухудшающие сходимость глобальной модели.

В медицинском приложении аналогичная ситуация может возникать без злого умысла. Врач может ошибаться при выставлении отдельных диагнозов, в локальной информационной системе могут применяться неодинаковые правила кодирования заключений, а часть записей может содержать неполную или противоречивую разметку. В результате локальная модель клиента обучается на искажённых метках и формирует обновления, которые ухудшают качество глобальной модели. С точки зрения сервера такое поведение аналогично византийскому: клиент отправляет обновление, отклоняющееся от полезного направления обучения. Однако содержательно такой клиент не обязательно является атакующим. В клиническом сценарии он может быть участником с повреждённой или непоследовательной разметкой.

Это наблюдение позволяет иначе интерпретировать методы византийской устойчивости. В классической постановке устойчивая агрегация стремится исключить или ослабить вклад подозрительных клиентов. В медицинской постановке низкое доверие к обновлениям клиента можно рассматривать не только как основание для уменьшения его веса в агрегации, но и как сигнал о возможной проблеме в локальной выборке. Если обновления некоторой больницы регулярно ухудшают качество модели на доверенной выборке, это может указывать на систематические ошибки разметки или на отличие локальных правил аннотирования от правил, принятых в общей группе больниц, участвующих в федеративном обучении.

В дальнейшем будем рассматривать следующий сценарий. Пусть каждый клиент  $k$  обладает локальным набором данных

$$\mathcal{D}_k = \left\{ \left( x_i^{(k)}, y_i^{(k)} \right) \right\}_{i=1}^{n_k}, \quad (4.1)$$

где  $x_i$  — ЭКГ-запись, а  $y_i$  — вектор диагностических меток.

Для части клиентов метки могут быть искажены некоторым правилом

$$\tilde{y}_i^{(k)} = f_k \left( y_i^{(k)} \right). \quad (4.2)$$

В данной работе такая ситуация моделируется синтетическим искажением разметки типа *label flip*. В простейшем случае для выбранной доли локальных примеров метка одного или нескольких классов заменяется на противоположную:

$$\tilde{y}_{i,c}^{(k)} = \begin{cases} 1 - y_{i,c}^{(k)}, & i \in \mathcal{I}_k, c \in \mathcal{C}_{\text{flip}}, \\ y_{i,c}^{(k)}, & \text{иначе,} \end{cases} \quad (4.3)$$

где  $\mathcal{I}_k$  — подмножество примеров клиента  $k$  с повреждённой разметкой, а  $\mathcal{C}_{\text{flip}}$  — набор классов, для которых выполняется подмена. В классификационной задаче такая ошибка может соответствовать как случайному изменению отдельных диагностических меток, так и систематической путанице между конкретными классами. В медицинской интерпретации такое искажение используется как модель непреднамеренной ошибки разметки, а не как описание намеренной атаки. Если удаётся обнаружить клиента с такой ошибкой и частично восстановить характер искажения, то клиент может улучшить локальную выборку и продолжить участие в федеративном обучении без ухудшения глобальной модели.

#### 4.3.2 Доверенная выборка для повышения устойчивости

В условиях неоднородных медицинских данных методы, основанные исключительно на геометрии клиентских обновлений, могут работать неидеально. Обновление клиента может отличаться от остальных не вследствие ошибки, а потому что его локальная выборка содержит другую популяцию пациентов или иной набор редких патологий. С другой стороны, клиент с небольшой, но систематической ошибкой разметки может не выглядеть как грубый выброс, хотя его обновления будут постепенно ухудшать качество модели. При этом методы, уже использующие доверенную выборку на сервере, сами по себе не свободны от проблем: их надёжность существенно зависит от того, насколько эта выборка репрезентативна. Если доверенная выборка мала или её распределение классов заметно отличается от распределения по всей федерации — например, содержит непропорционально мало редких патологий, — оценка полезности клиентского обновления по такой выборке может быть смещена: обновление, полезное для

редких классов, но слабо влияющее на доминирующий класс доверенной выборки, может получить заниженную оценку, даже если оно отражает не ошибку разметки, а полезную информацию о недопредставленной патологии. Поэтому в рассматриваемой медицинской постановке доверенную выборку целесообразно использовать как дополнительный источник информации при оценке клиентских обновлений, учитывая при этом требования к её составу и репрезентативности. Формирование такой доверенной выборки непосредственно из данных участвующих медицинских организаций может быть затруднено необходимостью дополнительной экспертной верификации и организационными ограничениями на передачу данных. Одним из практических вариантов является использование подходящего публичного набора данных с проверенной разметкой, если его распределение и набор целевых классов достаточно хорошо соответствуют решаемой задаче.

В данной работе для повышения устойчивости федеративного обучения в условиях недоверия к обновлениям клиентов используется новый метод BANT (Byzantine Antidote via Trial Function and Trust Scores). Его ключевая идея состоит в оценке полезности каждого клиентского обновления непосредственно по изменению качества глобальной модели на небольшой доверенной выборке. В медицинском сценарии такая выборка может представлять собой малый набор ЭКГ с проверенной разметкой, сформированный заранее и не участвующий в обычном локальном обучении клиентов.

Пусть на раунде  $t$  сервер имеет текущие параметры глобальной модели  $w^t$ , а клиент  $k$  возвращает локальное обновление  $\Delta_k^t$ . На сервере определяется функция потерь  $L_{\text{trust}}$ , соответствующая функции потерь на доверенной выборке  $\mathcal{D}_{\text{trust}}$ . Для оценки обновления клиента сервер временно применяет его к текущей модели и вычисляет изменение  $L_{\text{trust}}$ :

$$\delta_k^t = L_{\text{trust}}(w^t) - L_{\text{trust}}(w^t + \Delta_k^t). \quad (4.4)$$

Таким образом,  $\delta_k^t > 0$  означает, что обновление клиента уменьшает значение функции потерь на доверенной выборке, тогда как при  $\delta_k^t \leq 0$  такое улучшение отсутствует.

На основе полученных значений формируется нормированная текущая оценка клиента:

$$s_k^t = \frac{[\delta_k^t]_+}{\sum_{j=1}^K [\delta_j^t]_+}, \quad [\delta]_+ = \max\{\delta, 0\}, \quad (4.5)$$

если хотя бы для одного из участвующих клиентов  $\delta_j^t > 0$ . Следовательно, обновления, уменьшающие значение функции потерь на доверенной выборке, получают положительную оценку, пропорциональную достигнутому уменьшению ошибки, тогда как для обновлений с  $\delta_k^t \leq 0$  текущая оценка равна нулю.

В BANT итоговый доверительный вес клиента учитывает не только его оценку на текущем раунде, но и значение, накопленное на предыдущих раундах. Для этого используется экспоненциальное сглаживание с параметром  $\beta \in (0, 1]$ :

$$\alpha_k^t = (1 - \beta)\alpha_k^{t-1} + \beta s_k^t. \quad (4.6)$$

В начальный момент всем клиентам назначаются одинаковые веса  $\alpha_k^0 = 1/K$ . Использование предыдущего значения позволяет сгладить случайные колебания доверительной оценки: единичное неудачное локальное обновление не приводит к немедленной потере накопленного клиентом доверительного веса.

Отдельно определяется вырожденный случай, когда ни одно клиентское обновление на текущем раунде не уменьшает значение функции потерь на доверенной выборке, то есть  $[\delta_j^t]_+ = 0$  для всех  $j$ . В этом случае текущая составляющая доверительной оценки принимается равномерной,  $s_k^t = 1/K$ , и вес обновляется по правилу

$$\alpha_k^t = (1 - \beta)\alpha_k^{t-1} + \frac{\beta}{K}. \quad (4.7)$$

Таким образом, правило обновления накопленных доверительных весов остаётся однозначным и при отсутствии положительно оценённых обновлений. После вычисления доверительных весов сервер формирует новое состояние глобальной модели:

$$w^{t+1} = w^t + \sum_{k=1}^K \mathbb{I}[\delta_k^t > 0] \alpha_k^t \Delta_k^t. \quad (4.8)$$

Индикатор  $\mathbb{I}[\delta_k^t > 0]$  не позволяет клиенту повлиять на глобальную модель обновлением, которое на текущем раунде не уменьшает ошибку на доверенной выборке, даже если за счёт предыдущих раундов у этого клиента сохраняется

ненулевой доверительный вес. При этом накопленный вес отклонённого клиента не обнуляется и может использоваться на последующих раундах.

BANT в рассматриваемой постановке позволяет интерпретировать доверительные веса как дополнительный источник информации о качестве клиентских обновлений. Низкое значение текущей оценки  $s_k^t$  означает, что пробное применение обновления клиента не приводит к существенному уменьшению функции потерь на доверенной выборке. При этом для формирования обратной связи целесообразно учитывать не единичную оценку, а сглаженный доверительный вес  $\alpha_k^t$ , поскольку используемый в BANT механизм сглаживания учитывает оценки клиента на предыдущих раундах и уменьшает влияние случайных колебаний. Устойчиво низкий доверительный вес может служить основанием для сообщения о необходимости проверки локальных данных, однако сам по себе не позволяет делать вывод о наличии ошибки в разметке.

### 4.3.3 Эксперименты

В данной работе проводится сравнение метода BANT с существующими методами федеративного обучения и устойчивой агрегации: FedAvg [76], FLTrust [89], Zeno [88], SafeGuard [99], Central Clip [91]. В качестве базового метода используется FedAvg, не содержащий специальной защиты и усредняющий обновления клиентов пропорционально размеру локальных выборок. Этот метод позволяет оценить, насколько существенно повреждённая разметка влияет на стандартное федеративное обучение.

Метод FLTrust использует доверенную выборку на стороне сервера и сопоставляет клиентские обновления с направлением обновления, полученным на доверенных данных. Клиенты, направление обновлений которых плохо согласуется с доверенным обновлением, получают меньший вес. Метод Zeno также использует проверочную выборку, однако рассматривает клиентские обновления с точки зрения их влияния на значение функции потерь и отбирает более полезные обновления.

Метод SafeGuard относится к подходам византийски-устойчивой стохастической оптимизации и направлен на сохранение сходимости при наличии некор-

ректных участников. Метод Central Clip ограничивает влияние отклоняющихся обновлений за счёт клиппирования относительно центрального направления агрегации. Такие методы эффективны для подавления грубых выбросов, но не всегда позволяют отличить ошибочную разметку от статистической неоднородности медицинских клиентов.

Дополнительно рассматривается FedAvg на данных без искусственного повреждения меток. Этот сценарий используется как референсный для оценки величины деградации, возникающей при распределённой ошибочной разметке.

#### 4.3.4 Эксперимент с распределённой ошибочной разметкой

Для проверки предложенной постановки был проведён эксперимент на подвыборке набора данных ISFed. Эксперимент направлен на моделирование более реалистичной ситуации, в которой большинство клиентов не являются злонамеренными атакующими, но содержат небольшую долю ошибочно размеченных примеров. Такая постановка ближе к медицинской практике, чем сценарий с малым числом клиентов, полностью инвертирующих метки.

Рассматривалась задача классификации ЭКГ по классам, использованным в предыдущих главах: атриовентрикулярная блокада первой степени (1AVB), желудочковая экстрасистолия (PVC), фибрилляция предсердий (AFIB) и полная блокада левой ножки пучка Гиса (CLBBB).

Для данного эксперимента использовался отдельный от ISP50 протокол формирования данных: подвыборка ISFed была распределена между 100 клиентами, соответствующими больницам. Один из клиентов сохранял исходную разметку без внесения искусственных ошибок, тогда как в данные остальных 99 клиентов вносилось контролируемое повреждение меток. Отдельно на стороне сервера формировалась доверенная выборка ISTrust с проверенной разметкой. Одна и та же доверенная выборка одинакового объёма использовалась всеми сравниваемыми методами, для работы которых требуются доверенные серверные данные, включая BANT, FLTrust и Zeno.

99 клиентов с поврежденными метками были разделены на девять групп по 11 клиентов. В этих группах доля примеров с изменённой меткой составляла

соответственно 1%, 3%, 6%, 8%, 10%, 13%, 15%, 17% и 19%. Средний уровень повреждения разметки составлял примерно 10%. Повреждение разметки генерировалось случайно. Для каждого клиента формировалась бинарная матрица меток размера  $N_k \times C$ , где  $N_k$  – число локальных записей клиента, а  $C$  – число рассматриваемых классов. В соответствии с заданным для группы уровнем повреждения случайно выбиралась доля элементов этой матрицы, значения которых инвертировались ( $0 \rightarrow 1$  и  $1 \rightarrow 0$ ). Маска повреждения формировалась один раз перед началом обучения и оставалась неизменной на протяжении всех коммуникационных раундов. В отличие от экстремальной византийской постановки, где атакующий клиент может полностью менять направление обучения, здесь каждый клиент содержит лишь небольшую долю ошибочных меток. Поэтому такой эксперимент можно интерпретировать как модель медицинской сети, в которой ошибки разметки распределены по многим участникам.

В таблице 19 приведено распределение классов в основной федеративной выборке ISPfed и доверенной выборке ISPtrust. Доверенная выборка отличается от основной по распределению классов: в ней выше доля патологических записей, что делает её более информативной для оценки качества обновлений клиентов.

Таблица 19 — Распределение классов в основной и доверенной выборках

| Выборка  | Other | 1AVB  | PVC   | AFIB  | CLBBB |
|----------|-------|-------|-------|-------|-------|
| ISPfed   | 0.91  | 0.009 | 0.020 | 0.004 | 0.010 |
| ISPtrust | 0.77  | 0.036 | 0.068 | 0.025 | 0.024 |

Результаты экспериментов для методов FedAvg, FLTrust, Zeno, SafeGuard, Central Clip, BANT на данных с поврежденными метками и для метода FedAvg на данных с чистыми метками приведены в таблице 20.

Таблица 20 — Результаты обучения при распределённой ошибочной разметке

| Метод                | G-mean | Sensitivity | Specificity |
|----------------------|--------|-------------|-------------|
| FedAvg               | 0.711  | 0.820       | 0.617       |
| FLTrust              | 0.581  | 0.573       | 0.589       |
| Zeno                 | 0.761  | 0.640       | 0.905       |
| SafeGuard            | 0.474  | 0.958       | 0.235       |
| Central Clip         | 0.647  | 0.725       | 0.577       |
| BANT                 | 0.869  | 0.834       | 0.905       |
| FedAvg, чистые метки | 0.911  | 0.892       | 0.930       |

Прежде всего видна разница, которую повреждённая разметка создаёт для незащищённого FedAvg: его G-mean на повреждённых данных (0.711) заметно ниже, чем у того же FedAvg на чистых метках (0.911). Наилучший результат среди рассмотренных методов защиты показывает BANT (G-mean 0.869), наиболее близко приближаясь к сценарию обучения на чистых метках. Zeno также улучшает результат относительно незащищённого FedAvg и достигает G-mean 0.761. В то же время FLTrust, SafeGuard и Central Clip в данном эксперименте демонстрируют более низкие значения G-mean, чем FedAvg на повреждённых данных.

Особенностью рассматриваемого сценария является распределённый характер ошибок: небольшая доля повреждённых меток присутствует у 99 из 100 клиентов, тогда как полностью искажённые обновления отдельных клиентов не моделируются. Полученные результаты характеризуют устойчивость сравниваемых методов в такой постановке.

На уровне отдельных метрик наиболее заметное отличие BANT от незащищённого FedAvg наблюдается по специфичности: на повреждённой разметке Specificity FedAvg составляет 0.617, тогда как BANT достигает 0.905, приближаясь к результату FedAvg на чистых метках (0.930). Таким образом, в рассматриваемом эксперименте применение BANT позволяет существенно снизить число ложноположительных решений по сравнению с незащищённым FedAvg.

Эти результаты подтверждают практическую значимость рассмотренного сценария: в отличие от искусственных атак, где отдельные византийские клиенты могут формировать грубо искажённые обновления, здесь небольшое повреждение разметки распределено между большинством участников и моделиру-

ет возможные ошибки в медицинских данных. В рассматриваемой постановке BANT показывает более высокое качество по сравнению как с незащищённым FedAvg, так и с другими рассмотренными методами устойчивой агрегации. Кроме того, BANT формирует доверительные оценки клиентов, которые могут быть использованы для последующей обратной связи.

#### **4.3.5 Возможное использование доверительных оценок для контроля качества локальных данных**

Результаты эксперимента показывают, что доверительная выборка позволяет оценивать полезность клиентских обновлений даже тогда, когда повреждение меток у каждого клиента относительно невелико. Следующий естественный шаг состоит в том, чтобы использовать доверительные оценки не только для агрегации, но и для улучшения локальных данных.

Если клиент имеет стабильно низкую доверительную оценку, сервер может сообщить ему, что его локальная разметка статистически плохо согласуется с доверенной выборкой. После получения глобальной модели клиент может выполнять локальный пересмотр меток: для каждой записи сравниваются исходная метка врача и предсказание глобальной модели. Записи с наибольшим расхождением могут образовывать список кандидатов на повторную проверку. В простейшем автоматическом варианте клиент может выполнить переразметку этих записей с помощью глобальной модели; в более клинически корректном варианте эти записи передаются врачу для ручной верификации.

Более сложный вариант состоит в восстановлении правила ошибки. Если предполагается, что клиент искажает метки по некоторому правилу  $f_k(y)$ , то задача обратной связи может быть сформулирована как оценка этого правила по наблюдаемым расхождениям. Например, если у клиента систематически происходит замена одной диагностической метки на другую, сервер может обнаружить характерный паттерн: обновления клиента ухудшают качество именно по соответствующим классам, а локальные предсказания глобальной модели регулярно расходятся с исходными метками в одном и том же направлении. После

этого клиенту можно рекомендовать преобразование, приближённое к  $f_k^{-1}$ , то есть исправление меток в обратную сторону.

Важно, что такая процедура должна рассматриваться как система поддержки качества данных, а не как автоматическая замена медицинского решения. Глобальная модель может ошибаться, поэтому её предсказания не должны безусловно заменять врачебную разметку. Более безопасный сценарий состоит в том, чтобы использовать модель для ранжирования подозрительных записей: система показывает врачу ограниченный список ЭКГ, которые сильнее всего расходятся с общим федеративным знанием, а окончательное решение о корректировке остаётся за медицинским экспертом.

Таким образом, подход BANT в данной работе состоит в оценке полезности клиентских обновлений и использовании этих оценок при агрегации глобальной модели. Применение накопленных доверительных весов для организации проверки локальной разметки может рассматриваться как перспективное продолжение развития метода, требующее отдельной процедуры анализа записей на стороне клиента.

#### 4.4 Выводы

В данной главе рассмотрено применение федеративного обучения к задачам анализа электрокардиограмм в двух взаимодополняющих аспектах.

Во-первых, построена и развёрнута практическая система федеративного обучения на основе фреймворка Flower, включающая шифрование транспортного канала, аутентификацию клиентских узлов и разделение инфраструктурного и прикладного слоёв. Работоспособность системы подтверждена на разнесённых физических серверах. Проведенные эксперименты продемонстрировали, что обучение с участием нескольких медицинских организаций существенно превосходит обучение одного локального клиента. При участии 50 клиентов качество по G-mean для всех четырёх рассматриваемых патологий становится близким к результатам централизованного обучения, при котором данные тех же 50 клиентов непосредственно объединяются в единую обучающую выборку. Проверка на PTB-XL дополнительно показывает, что модель, обученная с ис-

пользованием данных множества клиентов, значительно лучше переносится на внешний набор данных, чем модель одного клиента.

Во-вторых, рассмотрена задача федеративного обучения в условиях неполного доверия к клиентским обновлениям. Такая постановка связана с византийской устойчивостью, однако отличается от классической модели злонамеренной атаки: в клиническом сценарии повреждённые метки могут возникать непреднамеренно из-за ошибок врача или различий в правилах кодирования диагнозов.

Для этой задачи предложен метод BANT, оценивающий полезность клиентских обновлений на доверенной выборке. На каждом раунде сервер для каждого клиентского обновления временно применяет его к текущей глобальной модели и вычисляет изменение функции потерь на доверенной выборке. Положительное уменьшение потерь интерпретируется как полезность обновления; текущие оценки нормируются и объединяются с накопленными доверительными весами посредством экспоненциального сглаживания. Обновления, не уменьшающие ошибку на доверенной выборке, исключаются из текущей агрегации, а остальные входят в глобальный шаг с весами, определяемыми их накопленной полезностью.

Эксперимент с распределённой подменой меток на данных ISP показал, что BANT заметно повышает качество по сравнению с FedAvg на повреждённых данных и приближает результат к обучению на чистых метках.

Таким образом, разработанная система обеспечивает техническую основу для совместного межклинического обучения без передачи исходных данных, а предложенный механизм оценки полезности клиентских обновлений уменьшает снижение качества глобальной модели при рассмотренном сценарии распределённых ошибок разметки.

## Заключение

В диссертации выполнен систематический анализ современных методов автоматизированного анализа электрокардиограмм. На основе обзора существующих подходов выявлены ограничения, наиболее существенные для практического применения моделей анализа ЭКГ: зависимость качества от состава и источника обучающих данных, ограниченность экспертной разметки для сегментации, недостаточное использование неразмеченных записей и дополнительной структуры медицинских данных, потерю информации при уменьшении числа доступных отведений, а также ограничения централизованного объединения клинических выборок и влияние ошибочных клиентских обновлений при распределённом обучении. Указанные ограничения определили структуру дальнейшего исследования: во второй главе рассматриваются задачи классификации, междатасетного переноса и сегментации при ограниченной разметке, в третьей — методы самообучения с использованием метаданных пациентов и многоканальной структуры сигнала, а в четвёртой — практическое федеративное обучение и снижение влияния аномальных клиентских обновлений.

Во второй главе на крупном закрытом клиническом наборе ISP (более 1,5 млн записей) проведено прямое экспериментальное сравнение foundation-моделей (ECGFounder, ECG-JEPA, ST-MEM, HuBERT-ECG) с моделями, обученными с учителем (resnet1d50, densenet1d121). Показано, что заявленный статус foundation-модели и масштаб предобучения сами по себе не гарантируют превосходства: компактная архитектура densenet1d121 продемонстрировала качество, превосходящее или сопоставимое с foundation-моделями по всем четырём рассмотренным патологиям, и была выбрана в качестве базовой архитектуры для дальнейших экспериментов. Исследование зависимости качества от объёма обучающей выборки показало, что эффект от увеличения выборки различается при оценке внутри домена обучения и при переносе на внешний набор: на данных исходного домена прирост для большинства рассмотренных патологий становится незначительным уже при сравнительно ограниченной доле выборки, тогда как при тестировании на внешнем наборе увеличение объёма обучающих данных в проведённых экспериментах давало более выраженный эффект. Эксперименты по дообучению подтвердили практическую целесообразность transfer

learning: модель, предварительно обученная на большом гетерогенном клиническом наборе ISP и дообученная на целевом наборе PTB-XL, превзошла как модель, обученную только на целевых данных, так и модель без адаптации к целевому домену.

Результатом второй главы является разработка архитектуры ECG-CODE и метода сегментации цифровых ЭКГ с использованием self-training. Базовая модель обучается на экспертно размеченных данных, после чего применяется для формирования псевдоразметки дополнительных записей. Наиболее уверенные псевдометки отбираются отдельно для зубцов  $P$ , комплекса QRS и зубца  $T$ . Затем новая модель обучается на сформированном псевдоразмеченном наборе и дообучается на экспертной выборке. Использование псевдоразметки для предложенной нейросетевой модели позволило повысить  $F_1$ -меру по всем шести рассматриваемым характерным точкам и получить результаты, превосходящие или сопоставимые с результатами сравниваемых современных методов сегментации. Этот результат соответствует первому пункту научной новизны и раскрывает первое положение, выносимое на защиту.

В третьей главе предложен метод контрастного предобучения нейросетевых моделей с использованием метаданных пациента при формировании дополнительных положительных пар. В проведенных экспериментах в качестве критерия группировки использовалась близость возраста, а group-based составляющая функции потерь объединялась со стандартной контрастной компонентой. Добавление предложенной group-based компоненты повышало значение G-mean относительно соответствующего базового метода: для SimCLR прирост составил от 0,2 до 1,7 процентного пункта, а для TS-TCC — от 0,6 до 1,4 процентного пункта в зависимости от патологии. Тем самым продемонстрирована возможность использовать индивидуальные характеристики пациентов в качестве дополнительного источника обучающего сигнала непосредственно на этапе предобучения.

Второй предложенный в третьей главе метод, LC-CL (*Lead Combinations with Contrastive Learning*), использует взаимосвязи между различными группами отведений одной записи как дополнительный контрастный сигнал. На этапе предобучения представление целевого отведения согласуется с представлениями дополнительных отведений той же ЭКГ, формируемыми вспомогательными энкодерами; после завершения предобучения вспомогательные энкодеры удаляются, а основной энкодер применяется к записи с уменьшенным числом каналов.

На первом отведении РТВ-XL применение LC-CL обеспечило прирост G-mean относительно одноканальной модели, обученной с учителем, до 7,3 процентного пункта. На независимом наборе WatchECG применение LC-CL повысило G-mean относительно модели, обученной с учителем, до 3,9 процентного пункта. Два предложенных в третьей главе метода соответствуют второму и третьему пунктам научной новизны и раскрывают второе положение, выносимое на защиту.

В четвёртой главе исследовано применение федеративного обучения к задачам анализа ЭКГ в межклиническом сценарии. Реализована и развёрнута на физически разнесённых серверах практическая система на основе фреймворка Flower с защищённым каналом связи и аутентификацией клиентских узлов. Эксперименты на данных ISP показали, что в рассмотренной конфигурации федеративное обучение с участием нескольких клиентов превосходит локальное обучение отдельного участника по качеству и устойчивости к переносу на внешний набор данных, приближаясь к качеству централизованного обучения без передачи исходных записей между участниками. Для сценария, в котором недоверие к обновлениям, получаемым от клиентов, обусловлено распределёнными ошибками медицинской разметки, предложен метод BANT, основанный на оценивании полезности клиентских обновлений на доверенной серверной выборке. На каждом раунде обновление временно применяется к текущей глобальной модели, его полезность определяется по изменению функции потерь на доверенной выборке, после чего накопленные доверительные оценки используются при формировании глобального шага. В проведённом эксперименте BANT достиг значения G-mean 0,869 против 0,711 для FedAvg на данных с повреждёнными метками и приблизил результат к FedAvg на чистых метках (G-mean 0,911). Этот результат соответствует четвёртому пункту научной новизны и раскрывает третье положение, выносимое на защиту.

По совокупности полученных результатов могут быть сформулированы следующие рекомендации по практическому применению разработанных методов. При построении систем автоматической классификации ЭКГ выбор базовой архитектуры нейросетевой модели целесообразно обосновывать непосредственным сравнением моделей на данных целевого клинического домена: большой объём предобучающих данных сам по себе не гарантирует превосходства модели после переноса. При переносе моделей между медицинскими учреждениями с

различающимися распределениями данных целесообразно использовать предварительное обучение на большом и разнородном наборе записей с последующим дообучением на данных целевого учреждения.

Для задач сегментации ЭКГ при ограниченном объёме экспертной разметки целесообразно применять подходы с привлечением неразмеченных записей и псевдоразметки, что позволяет увеличить эффективный объём обучающей выборки без дополнительной ручной аннотации. При организации контрастного обучения моделей для медицинских временных рядов следует, при наличии такой возможности, использовать доступные метаданные пациентов и многоканальную структуру записи в качестве дополнительных источников обучающего сигнала. Результаты group-based контрастного обучения и LC-CL показывают, что такая информация может повышать качество формируемых представлений без необходимости получения дополнительной диагностической разметки.

При развёртывании федеративного обучения в сети медицинских учреждений выбор механизма агрегации следует связывать с предполагаемым типом ошибок, которые могут возникать на стороне клиента. В рассмотренном сценарии распределённых ошибок разметки целесообразно использовать оценку полезности клиентских обновлений на небольшой доверенной выборке, как в предложенном методе BANT. При этом репрезентативность доверенной выборки относительно целевого распределения данных является существенным условием корректной оценки клиентских обновлений.

Полученные результаты открывают несколько направлений дальнейшего развития предложенных методов. Для LC-CL представляет интерес расширение предобучения на более широкий набор комбинаций отведений и исследование переноса на другие конфигурации ЭКГ. Для метода сегментации перспективным направлением является дополнительная регуляризация границ выделяемых комплексов, позволяющая уменьшить локальные ошибки определения характерных точек. Для федеративного обучения дальнейшая работа может быть направлена на расширение метода на более широкий спектр неоднородности клиентских данных и разработку способов адаптивного формирования доверенной выборки.

Таким образом, поставленная цель диссертационной работы — разработка и реализация методов и программных средств автоматического анализа медицинских временных рядов на примере анализа цифровых электрокардиограмм — достигнута. Разработанные методы охватывают анализ ЭКГ в условиях ограни-

ченного объёма экспертной разметки, использования больших массивов данных без целевой разметки, уменьшения числа доступных каналов и распределённого хранения клинических данных, а полученные экспериментальные результаты подтверждают применимость предложенных решений в рассмотренных постановках.

### Список литературы

1. Avetisyan A. et al. Deep neural networks generalization and fine-tuning for 12-lead ECG classification // Biomedical Signal Processing and Control. – 2024. – Т. 93. – С. 106160.
2. Ananov V. V. et al. Assessment of the impact of non-architectural changes in the predictive model on the quality of ECG classification // Proceedings of the Institute for System Programming of the RAS (Proceedings of ISP RAS). – 2021. – Т. 33. – № 4. – С. 87–98.
3. Skorik S. et al. Transferring Knowledge from 12-lead to 1-lead ECGs via Contrastive Learning // 2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC). – IEEE, 2024. – С. 1825–1830.
4. Avetisyan A. et al. Self-trained model for ECG complex delineation // ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – IEEE, 2025. – С. 1–5.
5. Skorik S., Avetisyan A. SSSD-ECG-nle: New Label Embeddings with Structured State-Space Models for ECG generation // 2024 Ivannikov Memorial Workshop (IVMEM). – IEEE, 2024. – С. 55–61.
6. Molodtsov G. et al. Bant: Byzantine antidote via trial function and trust scores // Proceedings of the AAAI Conference on Artificial Intelligence. – 2026. – Т. 40. – № 29. – С. 24431–24440.
7. Дятлинко Е. С. и др. Групповое контрастивное обучение для анализа медицинских данных // Doklady Mathematics. – 2026. – Т. 531. – № 1. – С. 96–107.
8. Молодцов Г. Л. и др. Защита от византийских атак через доверительное взвешивание устройств // Успехи математических наук. – 2025. – Т. 80. – № 6 (486). – С. 191–194.
9. Аветисян А. А. и др. Инструмент для развертывания нейросетевой модели предсказания патологий по данным 12-канальных электрокардиограмм – 2024. – Свидетельство о государственной регистрации программы для ЭВМ. Российская Федерация.

10. Аветисян А. А. и др. Программа классификации синдромов для 12-канальных цифровых ЭКГ на основе нейросетевой модели – 2024. – Свидетельство о государственной регистрации программы для ЭВМ. Российская Федерация.
11. Kligfield P. et al. Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society endorsed by the International Society for Computerized Electrocardiology // *Circulation*. – 2007. – Т. 115. – № 10. – С. 1306–1324.
12. Sörnmo L., Laguna P. Bioelectrical signal processing in cardiac and neurological applications. – Academic press, 2007.
13. Hannun A. Y. et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network // *Nature medicine*. – 2019. – Т. 25. – № 1. – С. 65–69.
14. Wagner P. et al. PTB-XL, a large publicly available electrocardiography dataset // *Scientific data*. – 2020. – Т. 7. – № 1. – С. 154.
15. Kalyakulina A. I. et al. LUDB: a new open-access validation tool for electrocardiogram delineation algorithms // *IEEE access*. – 2020. – Т. 8. – С. 186181–186190.
16. Pan J., Tompkins W. J. A real-time QRS detection algorithm // *IEEE transactions on biomedical engineering*. – 1985. – № 3. – С. 230–236.
17. Hamilton P. S., Tompkins W. J. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database // *IEEE transactions on biomedical engineering*. – 1986. – № 12. – С. 1157–1165.
18. Christov I. I. Real time electrocardiogram QRS detection using combined adaptive threshold // *Biomedical engineering online*. – 2004. – Т. 3. – № 1. – С. 28.
19. Kohler B. U., Hennig C., Orglmeister R. The principles of software QRS detection // *IEEE Engineering in Medicine and biology Magazine*. – 2002. – Т. 21. – № 1. – С. 42–57.

20. Elgendi M. Fast QRS detection with an optimized knowledge-based method: Evaluation on 11 standard ECG databases // PloS one. – 2013. – T. 8. – № 9. – C. e73557.
21. Trahanias P. E. An approach to QRS complex detection using mathematical morphology // IEEE transactions on biomedical engineering. – 1993. – T. 40. – № 2. – C. 201–205.
22. Li C., Zheng C., Tai C. Detection of ECG characteristic points using wavelet transforms // IEEE Transactions on biomedical Engineering. – 1995. – T. 42. – № 1. – C. 21–28.
23. Addison P. S. Wavelet transforms and the ECG: a review // Physiological measurement. – 2005. – T. 26. – № 5. – C. R155-R199.
24. Martínez J. P. et al. A wavelet-based ECG delineator: evaluation on standard databases // IEEE Transactions on biomedical engineering. – 2004. – T. 51. – № 4. – C. 570–581.
25. de Lannoy G. et al. Supervised ECG delineation using the wavelet transform and hidden Markov models // 4th European Conference of the International Federation for Medical and Biological Engineering: ECIFMBE 2008 23–27 November 2008 Antwerp, Belgium. – Berlin, Heidelberg : Springer Berlin Heidelberg, 2009. – C. 22–25.
26. Fasano A., Villani V. Baseline wander removal in ECG and AHA recommendations // Computing in Cardiology 2013. – IEEE, 2013. – C. 1171–1174.
27. Romero F. P. et al. Baseline wander removal methods for ECG signals: A comparative study // arXiv preprint arXiv:1807.11359. – 2018.
28. Hu Y. H., Palreddy S., Tompkins W. J. A patient-adaptable ECG beat classifier using a mixture of experts approach // IEEE transactions on biomedical engineering. – 1997. – T. 44. – № 9. – C. 891–900.
29. Osowski S., Hoai L. T., Markiewicz T. Support vector machine-based expert system for reliable heartbeat recognition // IEEE transactions on biomedical engineering. – 2004. – T. 51. – № 4. – C. 582–589.

30. De Chazal P., O'Dwyer M., Reilly R. B. Automatic classification of heartbeats using ECG morphology and heartbeat interval features // IEEE transactions on biomedical engineering. – 2004. – T. 51. – № 7. – C. 1196–1206.
31. Ye C., Kumar B. V. K. V., Coimbra M. T. Heartbeat classification using morphological and dynamic features of ECG signals // IEEE Transactions on Biomedical Engineering. – 2012. – T. 59. – № 10. – C. 2930–2941.
32. Goldberger A. L. et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals // circulation. – 2000. – T. 101. – № 23. – C. e215-e220.
33. Moody G. B., Mark R. G. The impact of the MIT-BIH arrhythmia database // IEEE engineering in medicine and biology magazine. – 2001. – T. 20. – № 3. – C. 45–50.
34. Ribeiro A. H. et al. Automatic diagnosis of the 12-lead ECG using a deep neural network // Nature communications. – 2020. – T. 11. – № 1. – C. 1760.
35. Perez Alday E. A. et al. Classification of 12-lead ecgs: the physionet/computing in cardiology challenge 2020 // Physiological measurement. – 2020. – T. 41. – № 12. – C. 124003.
36. Strodthoff N. et al. Deep learning for ECG analysis: Benchmarks and insights from PTB-XL // IEEE journal of biomedical and health informatics. – 2020. – T. 25. – № 5. – C. 1519–1528.
37. He K. et al. Deep residual learning for image recognition // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – C. 770–778.
38. Szegedy C. et al. Going deeper with convolutions // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2015. – C. 1–9.
39. Ismail Fawaz H. et al. Inceptiontime: Finding alexnet for time series classification // Data mining and knowledge discovery. – 2020. – T. 34. – № 6. – C. 1936-1962.
40. Warrick P., Homsy M. N. Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks // 2017 Computing in Cardiology (CinC). – IEEE, 2017. – C. 1–4.

41. Natarajan A. et al. A wide and deep transformer neural network for 12-lead ECG classification // 2020 Computing in Cardiology. – IEEE, 2020. – C. 1–4.
42. Cheng R. et al. Msw-transformer: Multi-scale shifted windows transformer networks for 12-lead ecg classification // arXiv preprint arXiv:2306.12098. – 2023.
43. Fu L. et al. Hybrid network with attention mechanism for detection and location of myocardial infarction based on 12-lead electrocardiogram signals // Sensors. – 2020. – T. 20. – № 4. – C. 1020.
44. Zhou F., Chen L. Leadwise clustering multi-branch network for multi-label ECG classification // Medical Engineering & Physics. – 2024. – T. 130. – № 1. – C. 104196.
45. Zhang D. et al. Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram // Iscience. – 2021. – T. 24. – № 4.
46. Raghunath S. et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network // Nature medicine. – 2020. – T. 26. – № 6. – C. 886–891.
47. Attia Z. I. et al. Age and sex estimation using artificial intelligence from standard 12-lead ECGs // Circulation: Arrhythmia and Electrophysiology. – 2019. – T. 12. – № 9. – C. e007284.
48. Li J. et al. An electrocardiogram foundation model built on over 10 million recordings // Nejm ai. – 2025. – T. 2. – № 7. – C. AIoa2401033.
49. Kim S. Learning general representation of 12-lead electrocardiogram with a joint-embedding predictive architecture // arXiv preprint arXiv:2410.08559. – 2024.
50. Na Y. et al. Guiding masked representation learning to capture spatio-temporal relationship of electrocardiogram // International conference on learning representations. – 2024. – T. 2024. – C. 15012–15035.
51. Coppola E. et al. HuBERT-ECG as a self-supervised foundation model for broad and scalable cardiac applications // medRxiv. – 2024. – C. 2024.11.14.24317328.

52. Hsu W. N. et al. Hubert: Self-supervised speech representation learning by masked prediction of hidden units // IEEE/ACM transactions on audio, speech, and language processing. – 2021. – T. 29. – C. 3451–3460.
53. Liu C. et al. Zero-shot ecg classification with multimodal learning and test-time clinical knowledge enhancement // arXiv preprint arXiv:2403.06659. – 2024.
54. Laguna P. et al. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG // Computers in cardiology 1997. – IEEE, 1997. – C. 673–676.
55. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical image computing and computer-assisted intervention. – Cham : Springer international publishing, 2015. – C. 234–241.
56. Jimenez-Perez G., Alcaine A., Camara O. Delineation of the electrocardiogram with a mixed-quality-annotations dataset using convolutional neural networks // Scientific reports. – 2021. – T. 11. – № 1. – C. 863.
57. Moskalenko V., Zolotykh N., Osipov G. Deep learning for ECG segmentation // International conference on neuroinformatics. – Cham : Springer International Publishing, 2019. – C. 246–254.
58. Duraj K. et al. Semantic segmentation of 12-lead ECG using 1D residual U-net with squeeze-excitation blocks // Applied sciences. – 2022. – T. 12. – № 7. – C. 3332.
59. Chen Z. et al. Post-processing refined ECG delineation based on 1D-UNet // Biomedical Signal Processing and Control. – 2023. – T. 79. – C. 104106.
60. Chen T. et al. A simple framework for contrastive learning of visual representations // International conference on machine learning. – PmLR, 2020. – C. 1597–1607.
61. Grill J. B. et al. Bootstrap your own latent-a new approach to self-supervised learning // Advances in neural information processing systems. – 2020. – T. 33. – C. 21271–21284.
62. Oord A., Li Y., Vinyals O. Representation learning with contrastive predictive coding // arXiv preprint arXiv:1807.03748. – 2018.

63. Um T. T. et al. Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks // Proceedings of the 19th ACM international conference on multimodal interaction. – 2017. – C. 216–220.
64. Iwana B. K., Uchida S. An empirical survey of data augmentation for time series classification with neural networks // Plos one. – 2021. – T. 16. – № 7. – C. e0254841.
65. Kiyasseh D., Zhu T., Clifton D. A. Clocs: Contrastive learning of cardiac signals across space, time, and patients // International conference on machine learning. – PmLR, 2021. – C. 5606–5615.
66. Gopal B. et al. 3kg: Contrastive learning of 12-lead electrocardiograms using physiologically-inspired augmentations // Machine learning for health. – PMLR, 2021. – C. 156–167.
67. Eldele E. et al. Self-supervised contrastive representation learning for semi-supervised time-series classification // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2023. – T. 45. – № 12. – C. 15604–15618.
68. Zhang X. et al. Self-supervised contrastive pre-training for time series via time-frequency consistency // Advances in neural information processing systems. – 2022. – T. 35. – C. 3988–4003.
69. He K. et al. Masked autoencoders are scalable vision learners // 2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR). – IEEE, 2022. – C. 15979–15988.
70. Mehari T., Strodthoff N. Self-supervised representation learning from 12-lead ECG data // Computers in biology and medicine. – 2022. – T. 141. – C. 105114.
71. Lee D. H. et al. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks // Workshop on challenges in representation learning, ICML. – 2013. – T. 3. – № 2. – C. 896.
72. Xie Q. et al. Self-training with noisy student improves imagenet classification // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – IEEE, 2020. – C. 10684–10695.

73. Tarvainen A., Valpola H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results // *Advances in neural information processing systems*. – 2017. – T. 30.
74. Sohn K. et al. Fixmatch: Simplifying semi-supervised learning with consistency and confidence // *Advances in neural information processing systems*. – 2020. – T. 33. – C. 596–608.
75. Konečný J. et al. Federated optimization: Distributed machine learning for on-device intelligence // *arXiv preprint arXiv:1610.02527*. – 2016.
76. McMahan B. et al. Communication-efficient learning of deep networks from decentralized data // *Artificial intelligence and statistics*. – Pmlr, 2017. – C. 1273–1282.
77. Kairouz P., McMahan H. B. Advances and open problems in federated learning // *Foundations and trends in machine learning*. – 2021. – T. 14. – № 1–2. – C. 1–210.
78. Rieke N. et al. The future of digital health with federated learning // *NPJ digital medicine*. – 2020. – T. 3. – № 1. – C. 119.
79. Sheller M. J. et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data // *Scientific reports*. – 2020. – T. 10. – № 1. – C. 12598.
80. Nguyen D. C. et al. Federated learning for smart healthcare: A survey // *ACM Computing Surveys (Csur)*. – 2022. – T. 55. – № 3. – C. 1–37.
81. Teo Z. L. et al. Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture // *Cell Reports Medicine*. – 2024. – T. 5. – № 2.
82. Jimenez Gutierrez D. M. et al. Application of federated learning techniques for arrhythmia classification using 12-lead ecg signals // *International Symposium on Algorithmic Aspects of Cloud Computing*. – Cham : Springer Nature Switzerland, 2023. – C. 38–65.
83. Agrawal V. et al. Federated learning and differential privacy techniques on multi-hospital population-scale electrocardiogram data // *Proceedings of the 2024 8th international conference on medical and health informatics*. – 2024. – C. 143–152.

84. Li T. et al. Federated optimization in heterogeneous networks // Proceedings of Machine learning and systems. – 2020. – T. 2. – C. 429–450.
85. Blanchard P. et al. Machine learning with adversaries: Byzantine tolerant gradient descent // Advances in neural information processing systems. – 2017. – T. 30.
86. Yin D. et al. Byzantine-robust distributed learning: Towards optimal statistical rates // International conference on machine learning. – Pmlr, 2018. – C. 5650–5659.
87. El-Mhamdi E. M., Guerraoui R., Rouault S. The hidden vulnerability of distributed learning in byzantium // International conference on machine learning. – PMLR, 2018. – C. 3521–3530.
88. Xie C., Koyejo S., Gupta I. Zeno: Distributed stochastic gradient descent with suspicion-based fault-tolerance // International conference on machine learning. – PMLR, 2019. – C. 6893–6901.
89. Cao X. et al. Fltrust: Byzantine-robust federated learning via trust bootstrapping // arXiv preprint arXiv:2012.13995. – 2020.
90. Yan H. et al. Recess vaccine for federated learning: Proactive defense against model poisoning attacks // Advances in Neural Information Processing Systems. – 2023. – T. 36. – C. 8702–8713.
91. Karimireddy S. P., He L., Jaggi M. Learning from history for byzantine robust optimization // International conference on machine learning. – PMLR, 2021. – C. 5311–5319.
92. Huang G. et al. Densely connected convolutional networks // 2017 IEEE conference on computer vision and pattern recognition (CVPR). – Ieee, 2017. – C. 2261–2269.
93. Hochreiter S., Schmidhuber J. Long short-term memory // Neural computation. – 1997. – T. 9. – № 8. – C. 1735–1780.
94. Cho K. et al. Learning phrase representations using RNN encoder–decoder for statistical machine translation // Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). – 2014. – C. 1724–1734.
95. Vaswani A. et al. Attention is all you need // Advances in neural information processing systems. – 2017. – T. 30.

96. Howard A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications // arXiv preprint arXiv:1704.04861. – 2017.
97. Liang X. et al. ECG\_SegNet: An ECG delineation model based on the encoder-decoder structure // Computers in biology and medicine. – 2022. – T. 145. – C. 105445.
98. Beutel D. J. et al. Flower: A friendly federated learning research framework // arXiv preprint arXiv:2007.14390. – 2020.
99. Allen-Zhu Z. et al. Byzantine-resilient non-convex stochastic gradient descent // arXiv preprint arXiv:2012.14368. – 2020.

## Список рисунков

|     |                                                                                                                         |    |
|-----|-------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | Элементы кардиоцикла . . . . .                                                                                          | 13 |
| 1.2 | Треугольник Эйнтховена и направления стандартных отведений<br>на конечностях . . . . .                                  | 14 |
| 2.1 | Архитектура модели ECG-CODE . . . . .                                                                                   | 63 |
| 2.2 | Примеры сегментации ЭКГ моделью Pseudolabeled ECG-CODE . .                                                              | 69 |
| 3.1 | Схема LC-CL предобучения для переноса представлений с<br>12-канальной ЭКГ на модель, использующую одно отведение. . . . | 90 |

## Список таблиц

|    |                                                                                                                                  |     |
|----|----------------------------------------------------------------------------------------------------------------------------------|-----|
| 1  | Наборы данных, используемые в работе . . . . .                                                                                   | 17  |
| 2  | Сравнение наборов данных PTB-XL и ISP, используемых в задаче классификации ЭКГ . . . . .                                         | 50  |
| 3  | Сравнение архитектур нейросетевых моделей на тестовой части набора данных ISP . . . . .                                          | 53  |
| 4  | Зависимость качества классификации от размера обучающей выборки для модели <code>densenet1d121</code> (Тест на ISP) . . . . .    | 55  |
| 5  | Зависимость качества классификации от размера обучающей выборки для модели <code>densenet1d121</code> (Тест на PTB-XL) . . . . . | 55  |
| 6  | Оценка междатасетной устойчивости модели <code>densenet1d121</code> . .                                                          | 59  |
| 7  | Влияние дообучения на качество классификации на целевом наборе PTB-XL . . . . .                                                  | 59  |
| 8  | Характеристики наборов данных с ручной разметкой, использованных для сегментации . . . . .                                       | 62  |
| 9  | Результаты сегментации модели ECG-CODE при обучении на LUDB и ISPdeli . . . . .                                                  | 68  |
| 10 | Сравнение качества сегментации на ISPdeli с современными методами . . . . .                                                      | 69  |
| 11 | Результаты классификации 12-канальных ЭКГ с использованием различных методов <code>contrastive learning</code> . . . . .         | 83  |
| 12 | Наборы данных, используемые для переноса представлений с 12-канальных ЭКГ на 1-канальные . . . . .                               | 87  |
| 13 | Результаты моделей классификации на первом отведении PTB-XL .                                                                    | 92  |
| 14 | Результаты моделей классификации 1-канальных ЭКГ на данных WatchECG . . . . .                                                    | 92  |
| 15 | Распределение записей одного локального клиента в эксперименте на данных ISP . . . . .                                           | 102 |
| 16 | Качество классификации ЭКГ при увеличении числа клиентов в федеративном обучении (G-mean) . . . . .                              | 103 |
| 17 | Проверка переноса с ISP на PTB-XL (G-mean) . . . . .                                                                             | 103 |

|    |                                                                                                  |     |
|----|--------------------------------------------------------------------------------------------------|-----|
| 18 | Динамика достижения целевого качества в локальном и<br>федеративном сценариях обучения . . . . . | 104 |
| 19 | Распределение классов в основной и доверенной выборках . . . . .                                 | 112 |
| 20 | Результаты обучения при распределённой ошибочной разметке . .                                    | 113 |