

О.Б. ЛЕБЕДЕВ, А.Г. ШМЕЛЕВА, Н.С. ГЕЖА
**МОДЕЛЬ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ
АВТОМАТИЗИРОВАННОЙ ИНТЕРПРЕТАЦИИ
МЕДИЦИНСКИХ ЭЛЕКТРОФИЗИОЛОГИЧЕСКИХ ДАННЫХ**

Лебедев О.Б., Шмелева А.Г., Гежа Н.С. Модель глубокого обучения для автоматизированной интерпретации медицинских электрофизиологических данных.

Аннотация. Статья посвящена разработке нейросетевой модели для автоматизированного анализа медицинских данных в области электрофизиологии на основе методов глубокого обучения. Актуальность работы обусловлена растущей потребностью в повышении объективности, скорости и точности обработки сложных пространственно-временных сигналов, таких как ЭКГ или ЭЭГ. В качестве инструмента анализа были использованы сверточные нейронные сети (CNN), позволяющие извлекать пространственные особенности, и временные сверточные нейронные сети (TCN), специализирующиеся на выявлении временных зависимостей. Интеграция этих архитектур позволила выявить и обобщить сложные паттерны в сигналах и, впоследствии, классифицировать их с высокой точностью. В работе представлена архитектура предложенной гибридной нейросетевой модели, разработанной для обработки временных рядов, и приведено обоснование выбора ключевых гиперпараметров. Для оптимизации модели реализована стратегия адаптивного изменения скорости обучения на основе значения функции потерь на валидационном наборе данных. Реализован механизм ранней остановки обучения для предотвращения избыточных вычислений и осуществления контроля переобучения в случае отсутствия уменьшения значения ошибки на валидационных данных в течение заданного количества эпох. При обучении после каждой эпохи фиксировались параметры модели, включая номер эпохи, значение функции потерь и параметры оптимизатора, что позволяет восстанавливать обучение с любой эпохи, а также детально проанализировать поведение модели на различных стадиях обучения. Применение разработанной модели анализа дает возможность произвести более глубокие исследования электрофизиологических данных, что повышает объективность диагностики. Представленные в работе методы могут быть масштабированы на другие типы биомедицинских данных. Практическая значимость работы заключается в возможности интеграции предложенной модели в биомедицинские системы в качестве инструмента поддержки врачебных решений для автоматизированного выявления патологий и снижения нагрузки на персонал.

Ключевые слова: машинное обучение, глубокое обучение, нейронные сети, медицинские данные, электрофизиология.

1. Введение. Здравоохранение играет ключевую роль в формировании национальной безопасности Российской Федерации. Данный тезис подтверждает включение показателя здравоохранения в стратегические документы планирования развития Российской Федерации [1, 2]. Современные информационные технологии предоставляют широкий инструментарий анализа данных [3], позволяющий проводить обработку больших объемов информации для

решения прикладных задач из разных областей [4, 5], включая медицину. В контексте цифровой трансформации здравоохранения особую актуальность приобретают методы глубокого обучения, предоставляющие инструменты выявления скрытых закономерностей в медицинских данных, анализа биомедицинских показателей и решения задач ранней диагностики заболеваний, в том числе сердечно-сосудистых заболеваний. Следует отметить, что разработка и внедрение интеллектуальных систем анализа медицинских данных способствует обеспечению цифровой трансформации здравоохранения, определяя назначение и значимость таких систем для решения прикладных задач медицины.

Цифровизация оказывает комплексное влияние на систему здравоохранения. Во-первых, внедрение цифровых технологий в медицине способствует совершенствованию системы оказания медицинской помощи, что приводит к повышению качества диагностики, увеличению продолжительности жизни и снижению смертности. Во-вторых, технологии телемедицины и системы дистанционного мониторинга изменяют стандарты сопровождения пациентов. Кроме того, цифровизация способствует интеграции удаленных регионов в единое пространство, обеспечивая доступ к качественной медицинской помощи независимо от удаленности населенного пункта.

Решение задач автоматизированного анализа медицинских данных проводилось на основе методов статистического анализа и обработки сигналов с использованием математических моделей, таких как преобразование Фурье и вейвлет-преобразование. Современные критерии оценки эффективности цифровизации здравоохранения и показатели развития медицинских информационных систем являются предметом активного изучения научным сообществом. Накопленные массивы медицинских данных позволяют строить прогностические модели и выявлять скрытые закономерности с помощью методов машинного обучения.

Сердечно-сосудистые заболевания (ССЗ) сохраняют лидирующие позиции среди причин смертности в мире. По данным Всемирной организации здравоохранения, в 2019 году от ССЗ скончались 17,9 миллионов человек, что составило 32% от общего числа смертей в мире [6], что является серьезной проблемой для системы здравоохранения [7]. Несмотря на значительный прогресс в области кардиологии, ранняя диагностика сердечно-сосудистых патологий остается сложной и ресурсоемкой задачей, в высокой степени зависящей от квалификации врача. В условиях дефицита

медицинских кадров [8] процесс постановки правильного диагноза и назначения лечения может быть продолжительным, что подчеркивает полезность разработки автоматизированных систем анализа медицинских данных. Это позволяет решать задачи принятия обоснованных клинических решений и способствует развитию персонализированной медицины. Вышесказанное обуславливает целесообразность разработки новых подходов к анализу медицинских данных с использованием современных методов искусственного интеллекта.

2. Обзор исследований. В статье [9] рассматриваются различные аспекты и подходы применения машинного обучения к данным ЭКГ. Авторы выделяют около пяти основных задач, связанных с анализом ЭКГ с помощью алгоритмов машинного обучения: диагностика заболеваний, распознавание и интерпретация сигналов ЭКГ, анализ качества сна и определение стадий сна, разработка методов биометрической идентификации человека и биометрической аутентификации с использованием ЭКГ, а также устранение зашумлений в сигнале ЭКГ. Выделены разновидности архитектур применяемых моделей и подчеркиваются способы сбора данных об электрокардиографии.

Выявление и диагностика ССЗ являются приоритетной задачей при анализе ЭКГ. Задача осложняется наличием широкого спектра заболеваний, зачастую требующих разных подходов к работе с исходным сигналом ЭКГ. По этой причине, многие исследователи сосредотачиваются на улучшении алгоритмов выявления определенных классов заболеваний. В статье [10] рассматривают возможность применения разработанной сверточной рекуррентной модели на основе CNN-модели и модели долговременной кратковременной памяти (LSTM) для извлечения пространственных и временных характеристик ЭКГ пациентов с сердечной недостаточностью. Задействован адаптивный механизм внимания SE (Squeeze-and-Excitation), разработанный для выделения наиболее значимых каналов в сверточных нейронных сетях. Разработанная модель решает задачу многоклассовой классификации, разделяя обрабатываемые данные на четыре категории в соответствии со стандартом классификации NYHA на основе лучших сигналов сегмента ЭКГ пациентов с диагностированной сердечной недостаточностью. В качестве набора данных приводятся результаты ЭКГ, отобранные по критерию наличия разновидности сердечной недостаточности из базы данных «MIMIC-III v1.4».

Схожий подход реализуется для решения задачи классификации аритмии в работе [11]. Авторы используют данные из базы данных аритмий «MIT-BIH». В указанной статье реализуется схожая с вышеописанной работой архитектура, использующая преимущества CNN и RNN моделей. Предложенная архитектура работает с ЭКГ в формате 2D скалограмм, получаемых методом непрерывного вейвлет-преобразования (CWT) для автоматического избавления данных от шумов и артефактов.

Не менее важной задачей является распознавание и интерпретация сигналов ЭКГ методами глубокого обучения. Зачастую, эту задачу решают в контексте мониторинга сердца плода в утробе матери. В работе [12] предлагается метод извлечения комплексов QRS плода из одноканальной неинвазивной ЭКГ плода на основе многофункциональной нейронной сети слияния «ML-LSTM». Исследование, представленное в работе [13] основывается на использовании более перспективного и ресурсоемкого подхода – использование расширенной архитектуры ResNet с рядом изменений: использования 1D-сверток и удаления последнего слоя. В качестве данных рассматриваются собранные с платформы мониторинга ЭКГ «Preventice» данные об аритмиях, включая нарушения ритма, блокады сердца и задержку внутрижелудочковой проводимости с редкими желудочковыми эктопическими сокращениями (ЖЭС).

Нарушения сна ведут к серьезным проблемам со здоровьем, снижению работоспособности и концентрации, приводят к депрессивным расстройствам и способствуют ухудшению психоэмоционального состояния человека [14]. Это обуславливает важность проводимых исследований по анализу качества сна и классификации стадий сна. Так, например, в работе [15] рассматривается способ применения сверточной нейронной сети для обнаружения обструктивного апноэ сна (ОАС) с помощью извлечения пространственных признаков из ЭКГ или сигналов насыщения крови кислородом. В статье [16] предлагается использовать разработанную модель MixSleepNet для классификации стадий сна, используя при этом сигналы ЭКГ, ЭЭГ, электромиограммы (ЭМГ) и электроокулограммы (ЭОГ), что используются для выделения характеристик во временной и частотной областях. Данная архитектура сочетает в себе как сверточные нейронные сети CNN, используя при этом 3D свертки, так и графовые сверточные сети (GCN).

Снятие ЭКГ сопровождается большим количеством шумов, в особенности из-за используемого оборудования. Для решения этой проблемы используют методы глубокого обучения, позволяющие

очистить сигнал ЭКГ от шумов и аномальных значений. В работе [17] авторы рассматривают способ применения автокодировщика с пропускными соединениями на двух наборах данных: синтетически полученные ЭКГ плода и набор из базы данных «Abdominal and Direct Fetal ECG Database». В рамках анализа применяются одномерные сверточные нейронные сети с ядрами сверток по временным измерениям ЭКГ. В свою очередь, двумерные ядра позволяют анализировать 2D-спектрограммы [11] и изображения [18]. Тем не менее, для решения более комплексных задач, затрагивающих анализ не только ЭКГ, ЭЭГ и других разделов электрофизиологии, могут применяться 3D-свертки [16]. Подход, рассматривающий применение двумерных сверточных нейронных сетей, помимо своих очевидных преимуществ, имеет ряд недостатков, самым крупным из которых является необходимость преобразования одномерных данных в двумерные. Из этого проистекает проблема роста вычислительных затрат, так как полученные тензоры изображений по большей части состоят из нулей, задействованных в процессе обработки. Этим обуславливается популярность методов, основанных на использовании 1D-CNN архитектур.

Другой разновидностью глубоких сетей, применяемых для анализа ЭКГ сигналов, являются рекуррентные нейронные сети (RNN, Recurrent Neural Networks). Основная особенность рекуррентных сетей заключается в передаче выходных сигналов предыдущего шага в качестве входных данных для следующего, что позволяет последовательно запоминать данные. Несмотря на ряд преимуществ, использование RNN сетей связано с рядом проблем, основной из которых является проблема затухающего градиента. Для решения этой проблемы была разработана наиболее часто используемая в исследованиях последних лет разновидность RNN архитектур – LSTM (Long Short-Term Memory). В работе [19] авторы применяют LSTM архитектуру для прогнозирования долгосрочных ЭКГ сигналов на наборе данных «MIT-BIH Arrhythmia».

Сверточно-рекуррентные нейронные сети (CRNN, Convolutional Recurrent Neural Networks) сочетают в себе оба описанных выше подхода, используя преимущества двух архитектур для анализа данных с пространственными и временными зависимостями [10, 20]. Возможность эффективно комбинировать анализ пространственных и временных характеристик данных делает данный подход универсальным для решения задач, требующих глубокого анализа. В ряде публикаций предпочитают реализовывать CNN-модели с большим количеством слоев, также позволяющие анализировать

временные зависимости, при этом экономя ресурсы на обучении по сравнению с аналогичными CRNN решениями [21]. Еще одним распространенным методом в анализе ЭКГ данных является использование автокодировщиков. Так, в работе [22] авторы используют маскированный автокодировщик (MAE) для самоконтролируемого обучения модели типа Vision Transformer (ViT) для улучшения ее производительности. В работе [23] авторами используется вариационный автокодировщик (VAE) для анализа данных ЭКГ с целью прогнозирования риска возникновения миокардиальной вентрикулярной аритмии (MVA) у носителей мутации гена PLN p(Arg14del). VAE-автокодировщик был предварительно обучен на 1,1 миллионе ЭКГ и использовался для сжатия ЭКГ данных и последующего сведения их к 32 объясняемым факторам. Для прогнозирования были задействованы модели на основе метода регрессии Кокса. Особенность VAE автокодировщиков заключается в возможности генерировать с их помощью новые данные на основе исходных. Это достигается благодаря обучению вероятностного распределения в латентном пространстве вместо фиксированных точек, а также использованию стохастического процесса для выборки из этого распределения. Пример такой генерации описан в работе [24], где авторы с помощью разработанного вариационного автокодировщика синтезируют ЭКГ в 12 отведениях. При этом, автокодировщик был предварительно обучен на данных с комплексного симулятора сердца на основе метода конечных элементов (FEM).

Помимо VAE автокодировщиков, для решения задачи синтеза ЭКГ сигналов также применяют генеративно-состязательные нейронные сети (GAN, Generative Adversarial Networks). Они состоят из генератора и дискриминатора, при этом каждый является отдельной нейронной сетью, обучающейся в состязательной манере с другой. В статье [23] разрабатывают GAN сеть на основе U-Net для последующего синтезирования фотоплетизмограмм из парных ЭКГ, нужных для диагностики тревожной фибрилляции предсердий.

Полносвязные слои (FC) применяются словес в более сложных архитектурах, таких как CNN или CRNN. В статье [26] представлен метод классификации сердцебиения на основе эволюционной оптимизации признаков с использованием дифференциальной эволюции (DE) и классификации с использованием вероятностной нейронной сети (PNN) для выявления нормальных и аритмичных сердечных сокращений. В ряде работ используются и гибридные подходы, как в работе [27], где используется модель глубокого обучения с тремя

подходами (CNN, LSTM и Transformer), а также традиционные методы машинного обучения, такие как метод опорных векторов (SVM), логистическую регрессию и алгоритм случайного леса.

Обобщая, отметим, что в области применения машинного обучения для анализа медицинских данных, сформировались направления, обусловленные доступностью и структурой данных. Значительное количество исследований сосредоточено на обработке ЭКГ, что отражает преобладание соответствующих наборов данных в открытом доступе. Анализ литературы демонстрирует, что машинное обучение, и в особенности глубокое обучение, успешно применяется для решения широкого спектра задач: от диагностики сердечно-сосудистых заболеваний и интерпретации сигналов до анализа сна, биометрической аутентификации и устранения шумов.

Методологически данная область характеризуется активным развитием гибридных нейросетевых архитектур, комбинирующих различные подходы для максимизации эффективности. Наиболее распространенными являются сверточные (CNN) и рекуррентные (RNN, в частности LSTM) нейронные сети, а также их комбинации (CRNN), позволяющие одновременно извлекать пространственные и временные зависимости в данных. Помимо этого, значительный исследовательский интерес вызывают генеративные модели, такие как вариационные автокодировщики (VAE) и генеративно-состязательные сети (GAN), используемые как для синтеза новых данных, так и для сжатия и выделения латентных признаков.

Таким образом, развитие направления определяется двумя взаимосвязанными факторами: спецификой решаемой прикладной задачи (диагностика, классификация, синтез) и постоянной эволюцией архитектур моделей, направленной на преодоление вычислительных сложностей, повышение точности и интерпретируемости результатов. Несмотря на явный прогресс, область применения нейронных сетей для решения прикладных задач требует дополнительных исследований, направленных на разработку более эффективных, надежных и интерпретируемых методов анализа.

Проведенный анализ существующих подходов выявил особенность, ограничивающую прямое сравнение результатов исследований с полученными в работе. Упомянутые выше научные работы сфокусированы на решении тематических задач, таких как классификация отдельных видов заболеваний (например, разные виды аритмии), используя для этого специализированные наборы данных. Такая концентрация на нескольких классах заболеваний позволяет достичь более высокой точности, но существенно сужает

область практического применения моделей. Настоящая работа ставит перед собой сложную и практикоориентированную задачу – разработку универсальной модели для мультиметочной классификации широкого спектра заболеваний. Работа с расширенным набором данных, содержащим сравнительно большее число разнородных патологий, является более комплексной задачей технологий интеллектуального анализа.

3. Анализ исходных данных. Основой для обучения модели послужили электрокардиографические (ЭКГ) данные. Ключевыми критериями выбора послужили неинвазивность метода, его повсеместная распространенность (что позволяет сформировать репрезентативный набор данных) и способность объективно фиксировать и графически отображать разность потенциалов, возникающую в процессе сокращения миокарда. В качестве исходных данных используются ЭКГ, загруженные из базы кардиологических данных «The PhysioNet / Computing in Cardiology Challenge 2021» последней редакции. Представленный в работе набор включает 88254 12-канальных записей ЭКГ (Рис. 1).

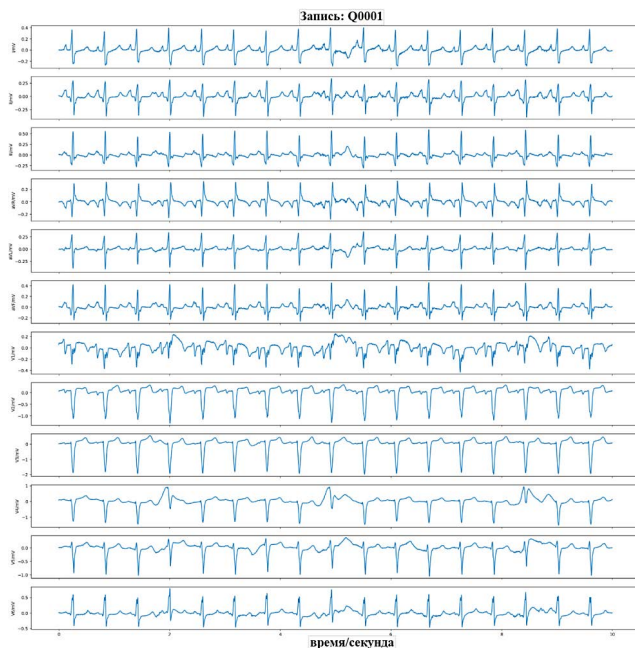


Рис. 1. Пример используемой записи ЭКГ с 12 стандартными отведениями

База данных была сформирована на основе следующих источников: CPSC и CPSC-Extra, INCART, PTB и PTB-XL, Chapman-Shaoxing и Ningbo, база данных Мичиганского университета (UMich), база данных 12-канальной ЭКГ-программы Джорджии (G12EC).

Для минимизации влияния дисбаланса классов в данных в условиях мультиметочной классификации, связанного с малым количеством примеров по ряду диагностических категорий, в работе рассматривается подвыборка из 29 наиболее репрезентативных заболеваний и одной нормы (Рис. 2).

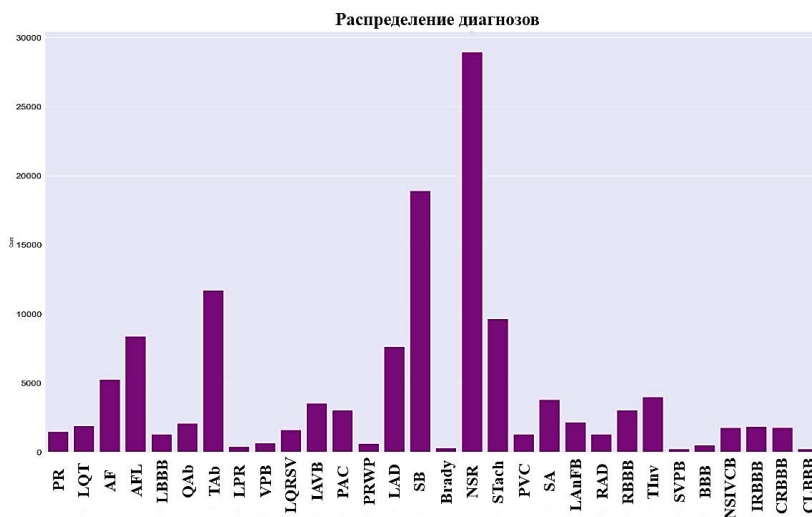


Рис. 2. Распределение выборки данных ЭКГ по диагностическим критериям

На рисунке 2 использовались аббревиатуры: PR – ритм с кардиостимулятором, LQT – удлиненный интервал QT (синдром), AF – фибрилляция предсердий, AFL – трепетание предсердий, LBVV – блокада левой ножки пучка Гиса, QAb – патологический зубец Q, Tab – патологический зубец T, LPR – удлиненный интервал PR, VPB – преждевременные сокращения желудочков, LQRSV – низкий вольтаж комплекса QRS, IABV – атриовентрикулярная блокада 1 степени, PAC – преждевременные сокращения предсердий, PRWP – слабое нарастание зубца R, LAD – отклонение электрической оси сердца влево, SB – синусовая брадикардия, Brady – брадикардия, NSR – синусовый ритм, STach – синусовая тахикардия, PVC – желудочковая экстрасистолия, SA – синусовая аритмия, LAnFB – блокада левой передней ветви пучка Гиса, RAD – отклонение электрической оси

сердца вправо, RBBB – блокада правой ножки пучка Гиса, TInv – инверсия зубца Т, SVPB – предсердная экстрасистолия, BBB – блокада ножек пучка Гиса, NSIVCB – неспецифическое нарушение внутрижелудочковой проводимости, IRBBB – неполная блокада правой ножки пучка Гиса, CRBBB – полная блокада правой ножки пучка Гиса, CLBBB – полная блокада левой ножки пучка Гиса.

В полученных данных предварительно заменяются пустые значения и выбросы, после чего сигнал стандартизируется. Для уменьшения вычислительных затрат на последующую обработку сигнала разработанной моделью, в стандартизированном сигнале находятся локальные максимумы каждые 180 временных шагов, после чего отбираются такие максимумы, вокруг которых можно извлечь сегмент длиной в 128 временных шагов. Полученные сегменты объединяются в один массив, после чего все сигналы выравниваются до 2700 временных шагов, что необходимо для стандартизации длины сигналов в каждом ЭКГ по всем отведениям, и передаются на вход модели. В модели, 90% данных отводится на обучающую выборку, а оставшиеся 10% – на валидационную.

Для оценки обобщающей способности предложенного метода на другие типы электрофизиологических данных модель была адаптирована для обработки полисомнографических сигналов и оценивалась на задаче классификации 5 стадий сна (N1, N2, N3, N4 и REM фаза быстрого сна) с использованием базы данных «Sleep-EDF Database Expanded», откуда было взято 153 записи. На рисунке 3 представлен пример данных, где цвет отражает разные стадии сна. Каждая запись представляет полисомнографию ночного сна, синхронно регистрируемую с четырёх отведений с частотой 100 Гц: два биполярных отведения энцефалограммы ЭЭГ (с электродов Fpz-Cz и Pz-Oz), горизонтальная электроокулограмма ЭОГ и электромиограмма ЭМГ подбородочной области. Все записи были разделены на тридцатисекундные отрезки (Рис. 3).

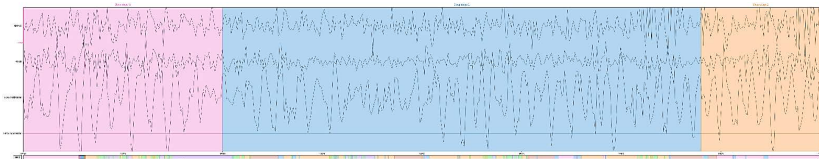


Рис. 3. Пример используемой записи ЭЭГ с 4 стандартными отведениями

Таким образом, суммарная выборка составила 86160 отрезков. В данной реализации 90% данных отводится на обучающую выборку

(2584 батча по тридцать тридцатисекундных отрезков), а оставшиеся 10% – на валидационную (288 батчей по тридцать тридцатисекундных отрезков).

4. Разработка модели. Рассмотрим архитектуру разработанной гибридной нейросетевой модели (Рис. 4).

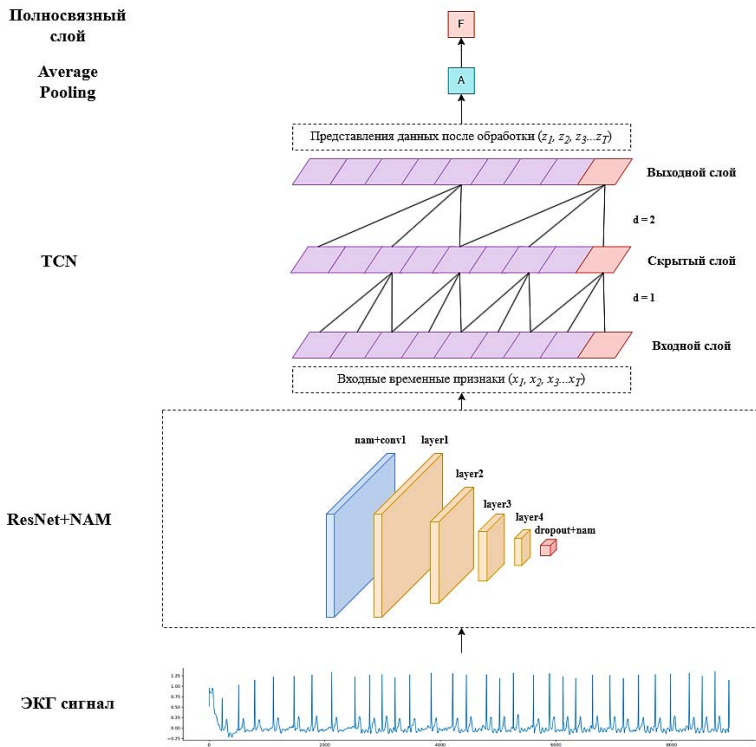


Рис. 4. Архитектура разработанной гибридной нейросетевой модели

Представленная архитектура состоит из трех частей: сверточная нейронная сеть (CNN) в виде модифицированной ResNet с механизмом внимания NAM, временная сверточная нейронная сеть TCN и выходная часть в виде среднего пуллинга и полносвязного слоя. Отчасти такую архитектуру можно назвать трансформером. Новизна заключается в интеграции пространственного и временного механизмов внимания со сверточными временными сетями: модифицированная ResNet с NAM-модулем динамически выделяет

дает возможность выстраивать нейронные сети с большим количеством слоев.

Механизм внимания NAM (normalized attention module или нормализованный механизм внимания), будучи адаптивным механизмом, используется в модели для подавления шумов и незначительных признаков в данных, усиления значимых признаков, а также для улучшения представления данных. Сам механизм состоит из двух модулей: CAM (channel attention module или каналный модуль внимания) и TAM (temporal attention module или временной модуль внимания) (Рис. 6).

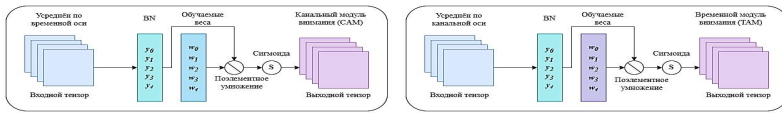


Рис. 6. Структура нормализованного механизма внимания NAM

Входной тензор $x \in R^{B \times C \times T}$, размерности (размер батча, количество каналов, длина временной последовательности) передается на вход каналного модуля внимания CAM, где первоначально усредняется по временной оси:

$$\bar{x}_{b,c} = \frac{1}{T} \sum_{t=1}^T x_{b,c,t}, \quad (1)$$

где $\bar{x}_c \in R^{B \times C}$ – усредненные значения для каждого канала. По каналным осям применяется батч-нормализация (BN), для стандартизации значения по каналам:

$$\tilde{x}_c = \text{BN}(\bar{x}_c). \quad (2)$$

Для вычисления BN предварительно для каждого канала рассчитывается среднее значение и дисперсия:

$$\begin{aligned} \mu_c &= \frac{1}{B} \sum_{b=1}^B \bar{x}_c^{(b)} \\ \sigma_c^2 &= \frac{1}{B} (\bar{x}_c^{(b)} - \mu_c)^2. \end{aligned} \quad (3)$$

Затем входной тензор нормализуется по батчу для каждого канала:

$$\hat{x}_c^{(b)} = \frac{\tilde{x}_c^{(b)} - \mu_c}{\sqrt{\sigma_c^2 + \varepsilon}}. \quad (4)$$

Далее применяется аффинное преобразование:

$$y_c^{(b)} = \gamma_c \hat{x}_c^{(b)} + \beta_c, \quad (5)$$

где γ_c и β_c – обучаемые параметры масштаба и сдвига.

Для расчета внимания по каналам нормализованные значения умножаются на обучаемые веса внимания, после чего к результату применяется сигмоидная функция активации:

$$s_c = \sigma(\tilde{x}_c \cdot \omega_{cam}), \quad (6)$$

где s_c – коэффициент внимания для канала, σ – сигмоидная функция активации, ω_{cam} – обучаемые веса внимания.

Полученный результат масштабируется для дальнейшего применения к исходному тензору:

$$x' = x \cdot s_{cam}, \quad (7)$$

где s_{cam} – веса канального внимания, расширенные вдоль временной оси.

Для расчета внимания по временной оси исходный тензор x передается на вход временному модулю внимания ТАМ, где аналогично усредняется по каналам:

$$\tilde{x}_{b,t} = \frac{1}{C} \sum_{c=1}^C x_{b,c,t}, \quad (8)$$

где $\tilde{x}_t \in R^{B \times T}$ – усредненные значения для каждого временного шага.

По временной оси применяется батч-нормализация:

$$\tilde{x}_t = \text{BN}(\tilde{x}_t). \quad (9)$$

Аналогично модулю САМ, рассчитывается внимание по временным шагам с использованием обучаемых весов:

$$s_t = \sigma(\tilde{x}_t \cdot \omega_{tam}). \quad (10)$$

Результат масштабируется и применяется к исходному тензору:

$$x' = x \cdot s_{tam}. \quad (11)$$

Итоговая формула имеет вид:

$$x' = x \cdot s_{cam} \cdot s_{tam} = x \cdot \sigma(\text{BN}(\bar{x}_c) \cdot \omega_{cam}) \cdot \sigma(\text{BN}(\bar{x}_t) \cdot \omega_{tam}). \quad (12)$$

Полученный тензор передается на вход сверточной нейронной сети, состоящей из слоя свертки и 4 слоев с разным количеством блоков с остаточными соединениями.

Каждая свертка вычисляется по следующей формуле одномерной свертки:

$$x'(c, t) = \sum_{c'} \sum_k \tilde{x}(c', S \cdot t - k) \cdot \omega(c, c', k), \quad (13)$$

где $x(c, t)$ – выходной тензор, c' – входной канал, c – выходной канал, t – временной шаг, k – индекс смещения, $\omega(c, c', k)$ – веса ядра свертки, S – страйд или шаг перемещения фильтра свертки, \tilde{x} – входной тензор, к которому применили паддинг:

$$\tilde{x}(c, t) = \begin{cases} x(c', t), & \text{если } P \leq t < T + P, \\ 0, & \text{иначе} \end{cases}, \quad (14)$$

где P – значение паддинга. Паддинг нужен для того, чтобы при применении операции свертки не терялась информация на краях данных.

Для сохранения размерности данных паддинг рассчитывается по формуле:

$$P = \frac{K-1}{2}, \quad (15)$$

где K – размер ядра свертки. Таким образом, в указанной модели при применении операции свертки с ядром 17 используется паддинг 8.

Итоговая формула применения одномерной свертки к исходному тензору в матричном виде имеет вид:

$$X' = X \cdot W, \quad (16)$$

где W – матрица свертки, X – входной тензор, развернутый в виде матрицы, Теплица.

В отличие от стандартной архитектуры ResNet (остаточная нейронная сеть), в которой используется базовый остаточный блок для всех вычислений, в представленной архитектуре используются так называемые ResStage блоки (Рис. 7).

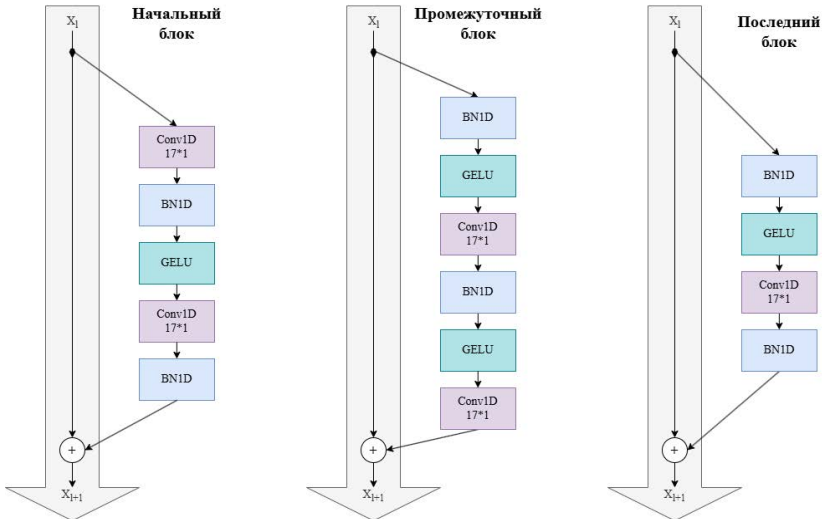


Рис. 7. Структура ResStage блоков сверточной нейронной сети

Данный подход заключается в использовании разных блоков в процессе вычисления: начального, промежуточного и заключительного. В зависимости от количества блоков в слое количество промежуточных блоков может варьироваться. На рисунке 7 кругом со знаком плюс обозначаются остаточные соединения, работающие по принципу:

$$y = F(x) + x, \quad (17)$$

где x – входной тензор блока, $F(x)$ – тензор после математических преобразований внутри блока, y – выходной тензор блока.

В представленной модели используется простое (identity shortcut) соединение. Данный механизм позволяет повышать точность обучения, решать проблему затухающих градиентов и упрощать обучение глубоких нейронных сетей. В качестве функции активации в блоках используется GELU (Гауссов линейный блок ошибки):

$$\text{GELU}(x) = x \cdot \Phi(x) = \frac{1}{2} x (1 + \text{erf}(\frac{x}{\sqrt{2}})), \quad (18)$$

где erf – функция ошибки из распределения Гаусса.

Использование данной функции активации объясняется рядом преимуществ: она более адаптивна, не имеет резкого порога (в отличие от ReLU), благодаря чему модель легче поддается оптимизации, обрабатывает малозначимые признаки, адаптивно уменьшая их влияние и тем самым предотвращая ситуацию, когда нейрон перестает активироваться и обновлять свои веса в процессе обучения. Перечисленные факторы обуславливают применение GELU в современных моделях глубокого обучения и архитектурах типа трансформер.

После вычисления всех скрытых резидуальных слоев применяется Dropout, являющийся эффективным методом регуляризации, предотвращающим совместную адаптацию нейронов. При применении Dropout к входному тензору, случайным образом обнуляются некоторые элементы входного тензора с вероятностью p на время обучения. Обнуляемые элементы выбираются независимо для каждого прямого вызова и выбираются из распределения Бернулли. Остальные элементы входного тензора масштабируются на $\frac{1}{1-p}$ на время обучения.

После преобразования полученный тензор передается на вход временной сверточной нейронной сети TCN (Рис. 8).

Основной принцип TCN сетей заключается в использовании остаточных соединений и каузальных разреженных сверток. Формула каузальной разреженной свертки имеет вид:

$$x'_t = \sum_{i=0}^{k-1} w_i \cdot x_{t-i \cdot d}, \quad (19)$$

где k – размер ядра свертки, w_i – веса сверточного фильтра, d – коэффициент разрежения (диляция), $x_{t-i \cdot d}$ – входное значение, взятое с разрежением.

В каузальных свертках индексы смещены так, чтобы при обработке временной последовательности не заглядывать в будущее, в отличие от обычных сверток в которых используется симметричное окно. Это позволяет модели принимать решения на основе текущих и прошлых данных, что достигается за счет использования следующего паддинга:

$$P = (K - 1) \cdot d, \quad (20)$$

где K – размер ядра свертки, d – коэффициент разрежения (дилатции).

Тем не менее, в работе используются некаузальные разреженные свертки, позволяющие анализировать данные более глобально, но хуже справляющиеся с анализом в реальном времени. Использование дилатции позволяет экспоненциально увеличивать область восприятия без роста числа слоев, тем самым становится возможным захватывать долгосрочные зависимости без увеличения вычислительных затрат. В качестве нормализации в архитектуре резидуального блока TCN используется WeightNorm (нормализация весов):

$$W = g \frac{v}{\|v\|}, \tag{21}$$

где g – скалярный параметр масштаба (обучаемый параметр, контролирующий обучение весов), v – вектор параметров (обучаемый вектор направления весов), W – нормализованные веса, $\|v\|$ – L2-норма вектора v , вычисляемая по формуле:

$$\|v\| = \sqrt{\sum_i v_i^2}, \tag{22}$$

где v_i – компонента вектора v .

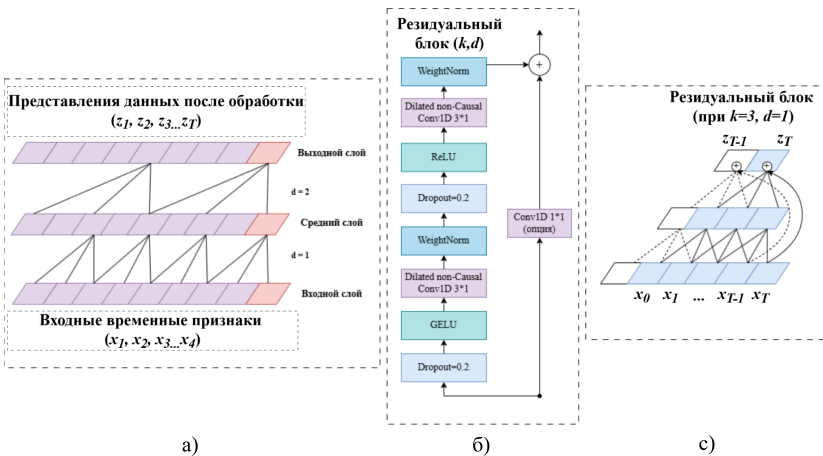


Рис. 8. Архитектура временной сверточной нейронной сети: а) схема TCN сети; б) архитектура резидуального блока TCN сети; в) схема работы резидуального блока при использовании каузальной свертки с разрежением с ядром 3 и дилатцией 1

Нормализованные веса в виде единичного вектора умножаются на нормализованный вектор g , что позволяет контролировать размер весов. Использование нормализации весов в TCN архитектурах позволяет стабилизировать обучение за счет решения проблемы с резким изменением градиентов и уменьшения зависимости от начальной инициализации весов, ускорять сходимость обучения и решать проблему переобучения, благодаря более равномерной распределенности весов. Помимо функции активации GELU в резидуальном блоке TCN сети также используется стандартная функция ReLU:

$$\text{ReLU}(x) = \max(0, x), \quad (23)$$

где x – входное значение.

Функция возвращает x если он положительный и 0 если отрицательный. Подход с использованием двух функций активации продемонстрировал более высокий результат при тестировании.

Выбор TCN-архитектур в современных задачах машинного обучения обуславливается возможностью параллельных вычислений на графических ускорителях благодаря использованию сверток (в то время как скрытое состояние в рекуррентных сетях, таких как LSTM, передается последовательно), отсутствием проблемы затухающих градиентов при обработке длинных последовательностей, возможностью обрабатывать последовательности разной длины, а также отсутствием утечки будущей информации за счет использования специализированного паддинга, в отличие от сверточных нейронных сетей. Благодаря вышеописанным факторам TCN-архитектуры хорошо зарекомендовали себя в текущих задачах глубокого обучения и демонстрируют высокий уровень точности и эффективности при обработке последовательностей данных. После TCN модели тензор передается слой адаптивного усредняющего пулинга:

$$x_{b,c,1} = \frac{1}{T} \sum_{t=1}^T x_{b,c,t}, \quad (24)$$

где T – исходная длина последовательности.

Адаптивный усредняющий пулинг вычисляет среднее значение по всей временной оси для каждого канала в тензоре, после чего масштабирует исходный тензор. Он применяется для уменьшения размерности и агрегации глобальных временных признаков перед полносвязным слоем, что необходимо для сохранения временных признаков. После применения пулинга канальная и временная

размерности тензора разворачиваются в одномерный вектор, тем самым получая тензор размерностью $x_{b,c}$, который затем передается на полносвязный слой:

$$x' = xW^T + b, \quad (25)$$

где x – входной тензор, W^T – транспонированная матрица весов, b – вектор смещения.

На выходе имеем тензор с выходными значениями для каждого класса, которые затем необходимо преобразовать в бинарные значения, соответствующие классифицированным заболеваниям, с помощью пороговых значений.

В модели применяется байесовская оптимизация для нахождения приближенных оптимальных пороговых значений и локальная минимизация методом Пауэлла – для их уточнения. Выбранная метрика для оптимизации – точность, которая будет максимизироваться. Для этого, оптимизатор строит аппроксимацию выбранной целевой функции ее нахождения с помощью гауссова процесса:

$$p(f(t)|D) = N(\mu(t), \sigma^2(t)), \quad (26)$$

где $\mu(t)$ – среднее (оценка целевой функции в точке t), $\sigma^2(t)$ – дисперсия в точке t , $D = \{(t_i, f(t_i))\}$ – наблюдаемые точки.

Таким образом, находятся те пороги, в которых целевая функция может иметь лучшее значение. Для дальнейшего поиска среди выбранных порогов применяется функция полезности $EI(t)$:

$$EI(t) = (\mu(t) - f^* - \xi)\Phi(Z) + \sigma(t)\phi(Z), \quad (27)$$

где f^* – наилучшее значение целевой функции, ξ – коэффициент исследования для баланса между выбором и проверкой порогов, $\Phi(Z)$ – кумулятивная функция распределения, $\phi(Z)$ – плотность нормального распределения. В данном случае параметр Z рассчитывается по формуле:

$$Z = \frac{\mu(t) - f^* - \xi}{\sigma(t)}. \quad (28)$$

В случае функции полезности $EI(t)$, если $\mu(t)$ намного больше значения функции f^* , порог будет проверяться. Если $\sigma(t)$ имеет большое значение, алгоритм не уверен в ответе и данный порог также

будет проверяться на возможность дальнейшего улучшения. При $\mu(t) \approx f^*$ и низком значении $\sigma(t)$ порог исключается. Найденные пороги передаются в оптимизатор для построения новой аппроксимации целевой функции.

После нахождения оптимальных порогов с помощью байесовской оптимизации точки передаются в локальный минимизатор, уточняющий значения порогов методом Пауэлла, являющимся безградиентным методом оптимизации на основе принципа координатного спуска, последовательно улучшающим параметры вдоль разных направлений. Для этого первоначально выбирается минимизируемая функция, в данном случае функция точности (accuracy), взятая со знаком минус. Таким образом, решается задача $\min_T f(T)$, где $f(T)$:

$$f(T) = -\text{accuracy_score}(y_{\text{true}}, (y_{\text{pred}} \geq T)). \quad (29)$$

Далее выполняется одномерная оптимизация по каждому направлению для каждого порога:

$$\alpha_i = \min_{\alpha} f(T^{(i-1)} + \alpha d_i), \quad (30)$$

где d_i – текущее направление (изначально берутся базисные единичные векторы), α – шаг в выбранном направлении.

После оптимизации по направлению точка обновляется:

$$T^{(i)} = T^{(i-1)} + \alpha_i d_i. \quad (31)$$

После прохода по всем направлениям добавляется новое направление:

$$d_{\text{new}} = T^{(n)} - T^{(0)}, \quad (32)$$

где $T^{(0)}$ – начальная точка, полученная после байесовской оптимизации, $T^{(n)}$ – точка, полученная после прохода по всем направлениям.

Алгоритм выполняется снова, после чего проверяется сходимость с новой найденной точкой. Алгоритм завершает работу если достигает предельного количества итераций или если выполняется условие сходимости для всех пороговых значений:

$$\sqrt{\sum_{i=1}^m (T_i^{(new)} - T_i^{(n)})^2} < \varepsilon, \quad (33)$$

где ε – некоторое пороговое значение, определяющее сходимость.

Полученные пороги применяются к выходам модели, преобразовывая их в бинарные значения.

В качестве функции потерь в работе используется бинарная кросс-энтропия с логитами (BCEWithLogitsLoss), применяемая в задачах бинарной многоклассовой и мультиметочной классификации:

$$L = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C -[y_{ij} \log \sigma(z_{ij}) + (1 - y_{ij}) \log(1 - \sigma(z_{ij}))], \quad (34)$$

где N – количество примеров в батче, C – количество каналов, z_{ij} – логиты модели (непреобразованные выходы модели), y_{ij} – истинные бинарные метки, σ – сигмоидная функция активации.

Оптимизатором модели был выбран стандартный Adam (оценка адаптивного метода). При его использовании веса модели обновляются согласно представленному ниже алгоритму. В первую очередь, вычисляется градиент функции потерь:

$$g_t = \nabla_{\theta} L(\theta_t), \quad (35)$$

где θ_t – параметры модели на шаге t , ∇_{θ} – оператор градиента, $L(\theta_t)$ – функция потерь.

Затем рассчитывается экспоненциальное скользящее среднее первого момента (или среднее значение градиента):

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad (36)$$

где β_1 – коэффициент сглаживания.

Затем, вычисляется экспоненциальное скользящее среднее второго момента (или среднее значение квадрата градиента):

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \quad (37)$$

где β_2 – также коэффициент сглаживания.

На первом шаге выполняется коррекция смещения:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \hat{v}_t = \frac{v_t}{1 - \beta_2^t}. \quad (38)$$

Наконец, параметры модели обновляются:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \varepsilon}} \hat{m}_t, \quad (39)$$

где η – скорость обучения, ε – некоторое малое число для предотвращения деления на ноль.

Преимущества использования оптимизатора Adam заключаются в том, что он автоматически адаптирует скорость обучения для каждого параметра модели, сочетает в себе разные подходы, что делает его более универсальным и стабилизированным, Adam не требует сложной настройки гиперпараметров, а также демонстрирует быструю и стабильную сходимость в сложных задачах.

Вся реализация модели выполнена с использованием фреймворка PyTorch. Обучение модели осуществлялось с использованием объектно-ориентированного подхода, в котором основной класс инкапсулирует всю логику тренировочного и валидационного циклов. Для повышения устойчивости обучения была реализована стратегия адаптивного изменения скорости обучения на основе значения функции потерь на валидационном наборе данных. Для контроля переобучения и предотвращения избыточных вычислений был реализован механизм ранней остановки, срабатывающий при отсутствии улучшения ошибки на валидационных данных в течение заданного количества эпох. После каждой эпохи состояние модели, включая параметры оптимизатора, номер текущей эпохи и значение функции потерь, сохранялось на диск в виде контрольной точки. Это позволяло при необходимости восстановить обучение с нужной эпохи, а также проанализировать поведение модели на различных стадиях оптимизации.

5. Верификация модели. Для оценки качества модели использовались несколько метрик, каждая из которых отражает разные стороны работы модели. Полученные результаты оценки модели на валидационной выборке представлены в таблице 1.

Таблица 1. Результаты оценки модели на валидационной выборке для 30 классов

| Метрика качества | Результат |
|------------------|-----------|
| Accuracy | 47% |
| F1-score | 59% |
| Roc-Auc | 64% |
| Precision | 67% |
| Recall | 59% |

Отметим, что для оценки моделей анализа медицинских данных необходимо применять несколько метрик, так как обучение моделей производится на несбалансированных выборках, где присутствует дисбаланс классов, поэтому, несмотря на низкий показатель глобальной точности (Accuracy = 47%), что является закономерным следствием экстремального дисбаланса при 30 классах, модель демонстрирует эффективность. Ключевые метрики, устойчивые к дисбалансу, подтверждают её работоспособность: высокое значение Precision (67%) указывает на точность предсказаний для редких категорий, а F1-score (59%) – на баланс между точностью и полнотой их обнаружения. ROC-AUC (64%) подтверждает общую способность модели к содержательной дискриминации всех классов. Таким образом, представленные метрики свидетельствуют о практической применимости модели для задач классификации в условиях сильного дисбаланса, где критически важным является корректное выявление редких, но значимых событий.

Дополнительно разработанная модель была обучена на задаче классификации 12 наиболее репрезентативных заболеваний. Полученные результаты оценки модели на валидационной выборке представлены в таблице 2.

Таблица 2. Результаты оценки модели на валидационной выборке для 12 классов

| Метрика качества | Результат |
|------------------|-----------|
| Accuracy | 53% |
| F1-score | 64% |
| Roc-Auc | 72% |
| Precision | 65% |
| Recall | 60% |

Также, разработанная модель была адаптирована и обучена для решения задачи многоклассовой классификации 5 стадий сна по данным ЭЭГ. Полученные результаты оценки модели на валидационной выборке представлены в таблице 3.

Таблица 3. Результаты оценки модели на валидационной выборке при решении задачи классификации 5 стадий сна

| Метрика качества | Результат |
|------------------|-----------|
| Accuracy | 78% |
| F1-score | 78% |
| Precision | 79% |
| Recall | 78% |

Представленная верификация подтверждает гипотезу о высокой практической применимости предложенного в работе подхода для анализа сложных медицинских данных. Разработанная гибридная архитектура, основанная на комбинации сверточных и временных сверточных нейронных сетей с механизмом внимания, продемонстрировала высокую результативность в задачах автоматизированной диагностики на разнородных физиологических сигналах, в частности, для выявления патологий по данным ЭКГ и классификации стадий сна на основе многоканальных записей ЭЭГ с включением ЭОГ и ЭМГ. Универсальность и эффективность данных методов открывает перспективы для их адаптации при проектировании других специализированных архитектур для анализа медицинских данных.

6. Заключение. В рамках данной статьи показана принципиальная возможность и обоснована целесообразность применения алгоритмов машинного обучения для анализа многомерных медицинских данных. Использование методов интеллектуальной обработки, в частности, глубоких сверточных сетей, позволило достичь высоких показателей точности в задачах автоматизированной диагностики и многоклассовой классификации патологий. Разработанная архитектура гибридной нейросетевой модели реализует полный конвейер обработки данных, включающий этапы предобработки сигналов, селекции информативных признаков, а также их последующего анализа с помощью современных адаптивных моделей. Применение данного подхода продемонстрировало высокую эффективность, что подтверждается результатами оценки модели.

Полученные результаты обуславливают перспективность использования предложенных методов для создания систем поддержки принятия решений в медицине. Разработанная архитектура обладает значительным потенциалом для дальнейшего совершенствования и развития. Пути повышения эффективности включают в себя увеличение объема и качества размеченных данных для обучения, применение методов аугментации, а также тонкую настройку гиперпараметров моделей и использование более сложных ансамблевых алгоритмов. Научная и практическая значимость полученных результатов подтверждает целесообразность дальнейших исследований в области создания автоматизированных информационных систем обработки и анализа медицинских данных.

Литература

1. Президент России. Указ Президента РФ от 02.07.2021 № 400 «О Стратегии национальной безопасности Российской Федерации». URL: <http://kremlin.ru/acts/bank/47046> (дата обращения: 20.02.2025).
2. Президент России. Указ Президента РФ от 13.05.2017 № 208 «О Стратегии экономической безопасности Российской Федерации на период до 2030 года». URL: <http://www.kremlin.ru/acts/bank/41921> (дата обращения: 20.02.2025).
3. Шевцов Е.С., Шамин Р.В. Логическая интеграция информационных систем на основе экспертных систем // Российский технологический журнал. 2025. Т. 13. №2. С. 27–35. DOI: 10.32362/2500-316X-2025-13-2-27-35.
4. Mityakov E.S., Lapaev D.N., Beskhmelnov M.I., Ladynin A.I., et al. Knowledge-intensive Enterprises Multi-criteria Decision-making Energy Security Management Information System // Proceedings of the 2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). 2022. pp. 1711–1713. DOI: 10.1109/EIConRus54750.2022.9755655.
5. Красноперова А.С., Твердохлебов А.С., Карташов А.А., Вебер В.И., Куприц В.Ю. Исследование эффективности применения моделей нейронных сетей YOLO для распознавания объектов на радиолокационных изображениях // Российский технологический журнал. 2025. Т. 13. №4. С. 25–36. DOI: 10.32362/2500-316X-2025-13-4-25-36.
6. Всемирная организация здравоохранения. Сердечно-сосудистые заболевания (ССЗ). URL: <https://www.who.int/ru/news-room/fact-sheets/detail/cardiovascular-diseases-cvds> (дата обращения: 25.02.2025).
7. Министерство здравоохранения Российской Федерации. Федеральный проект «Борьба с сердечно-сосудистыми заболеваниями». URL: <https://minzdrav.gov.ru/poleznye-resursy/natsproektzdravoohranenie/bssz> (дата обращения: 25.02.2025).
8. Шишкин С.В., Шейман И.М., Алмазов А.А., Бирюкова А.И., Власов В.В., Потапчик Е.Г., Сажина С.В., Салахутдинова С.К. Российское здравоохранение: перспективы развития. Доклад НИУ ВШЭ / Под ред. Шишкина С.В. // М.: Изд. дом ВШЭ. 2024. 60 с.
9. Hong S., Zhou Y., Shang J., Xiao C., Sun J. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review // Computers in Biology and Medicine. 2020. vol. 122. pp. 1–16. DOI: 10.1016/j.compbimed.2020.103801.
10. Zhang C.-J., Yuan L., Tang F.-Q., Cai H.-P., Qian Y.-F., Wang C. Heart failure classification using deep learning to extract spatiotemporal features from ECG // BMC Medical Informatics and Decision Making. 2024. vol. 24. no. 1. pp. 1–19. DOI: 10.1186/s12911-024-02415-4.
11. Madan P., Singh V., Singh D.P., Diwakar M., Pant B., Kishor A. A Hybrid Deep Learning Approach for ECG-based Arrhythmia Classification // Bioengineering. 2022. vol. 9. no. 4. DOI: 10.3390/bioengineering9040152.
12. Huang Z., Yu J., Shan Y., Wang X. A Non-Invasive Fetal QRS Complex Detection Method Based on a Multi-Feature Fusion Neural Network // Applied Sciences. 2024. vol. 14. no. 19.
13. Teplitzky B.A., McRoberts M., Ghanbari H. Deep learning for comprehensive ECG annotation // Heart Rhythm. 2020. vol. 17. no. 5. pp. 881–888. DOI: 10.1016/j.hrthm.2020.02.015.
14. Riemann D., Krone L.B., Wulff K., Nissen C. Sleep, insomnia, and depression // Neuropsychopharmacology. 2020. vol. 45. no. 1. pp. 74–89. DOI: 10.1038/s41386-019-0411-y.

15. Paul T., Hassan O., Islam S.K., Mosa A.S.M. Real-time obstructive sleep apnea detection from raw ECG and SpO₂ signal using convolutional neural network // AMIA Joint Summits on Translational Science Proceedings. 2024. pp. 662–669.
16. Ji X., Li Y., Wen P., Barua P., Acharya U.R. MixSleepNet: A Multi-Type Convolution Combined Sleep Stage Classification Model // Computer Methods and Programs in Biomedicine. 2024. vol. 244. 107992 p. DOI: 10.1016/j.cmpb.2023.107992.
17. Mvuh F.L., Ebode Ko'a C.O.V., Bodo B. Multichannel high noise level ECG denoising based on adversarial deep learning // Scientific Reports. 2024. vol. 14. no. 1. DOI: 10.1038/s41598-023-50334-7.
18. Aversano L., Bernardi M.L., Cimitile M., Montano D., Pecori R. Characterization of Heart Diseases per Single Lead Using ECG Images and CNN-2D // Sensors. 2024. vol. 24. no. 11. DOI: 10.3390/s24113485.
19. Zacarias H., Marques J.A.L., Felizardo V., Pourvahab M., Garcia N.M. ECG Forecasting System Based on Long Short-Term Memory // Bioengineering. 2024. vol. 11. no. 1. DOI: 10.3390/bioengineering11010089.
20. Kamanditya B., Fuadah Y.N., Mahardika T.N.Q., Lim K.M. Continuous blood pressure prediction system using Conv-LSTM network on hybrid latent features of photoplethysmogram (PPG) and electrocardiogram (ECG) signals // Scientific Reports. 2024. vol. 14. no. 1.
21. Khan F., Yu X., Yuan Z., Rehman A.U. ECG classification using 1-D convolutional deep residual neural network // PLOS One. 2023. vol. 18. no. 4. e0284791 p. DOI: 10.1371/journal.pone.0284791.
22. Sawano S., Kodera S., Setoguchi N., et al. Applying masked autoencoder-based self-supervised learning for high-capability vision transformers of electrocardiographies // PLOS One. 2024. vol. 19. no. 8. pp. 1–17. DOI: 10.1371/journal.pone.0307978.
23. van de Leur R.R., de Brouwer R., Bleijendaal H., et al. ECG-only explainable deep learning algorithm predicts the risk for malignant ventricular arrhythmia in phospholamban cardiomyopathy // Heart Rhythm. 2024. vol. 21. no. 7. pp. 1102–1112. DOI: 10.1016/j.hrthm.2024.02.038.
24. Nishikimi R., Nakano M., Kashino K., Tsukada S. Variational autoencoder-based neural electrocardiogram synthesis trained by FEM-based heart simulator // Cardiovascular Digital Health Journal. 2023. vol. 5. no. 1. pp. 19–28. DOI: 10.1016/j.cvdhj.2023.12.002.
25. Sohn J., Shin H., Lee J., Kim H.C. Validation of Electrocardiogram Based Photoplethysmogram Generated Using U-Net Based Generative Adversarial Networks // Journal of Healthcare Informatics Research. 2024. vol. 8. no. 1. pp. 140–157. DOI: 10.1007/s41666-023-00156-z.
26. Nasim A., Kim Y.S. DE-PNN: Differential Evolution-Based Feature Optimization with Probabilistic Neural Network for Imbalanced Arrhythmia Classification // Sensors. 2022. vol. 22. no. 12. 4450 p. DOI: 10.3390/s22124450.
27. Din S., Qaraqe M., Mourad O., Qaraqe K., Serpedin E. ECG-based cardiac arrhythmias detection through ensemble learning and fusion of deep spatial-temporal and long-range dependency features // Artificial Intelligence in Medicine. 2024. vol. 150. 102818 p.

Лебедев Олег Борисович — д-р техн. наук, доцент, профессор, кафедра информатики, Федеральное государственное бюджетное образовательное учреждение высшего образования "МИРЭА – Российский технологический университет" (РТУ МИРЭА). Область научных интересов: искусственный интеллект, машинное обучение, алгоритмы роевой оптимизации, системы поддержки принятия решений, разработка и применение

математических и программных инструментов. Число научных публикаций — 270. lebedev_o@mirea.ru; пр. Вернадского, 78, 119454, Москва, Россия; р.т.: +7(499)215-6565.

Шмелева Анна Геннадьевна — канд. физ.-мат. наук, доцент, доцент, кафедра информатики, Федеральное государственное бюджетное образовательное учреждение высшего образования "МИРЭА – Российский технологический университет" (РТУ МИРЭА). Область научных интересов: искусственный интеллект, машинное обучение, системы поддержки принятия решений. Число научных публикаций — 107. ann_shmeleva@mail.ru; пр. Вернадского, 78, 119454, Москва, Россия; р.т.: +7(499)215-6565.

Гежа Николай Сергеевич — студент, кафедра информатики, Федеральное государственное бюджетное образовательное учреждение высшего образования "МИРЭА – Российский технологический университет" (РТУ МИРЭА). Область научных интересов: машинное обучение, искусственный интеллект, анализ биомедицинских данных, автоматизация поддержки принятия решений. Число научных публикаций — 1. n.s.gezha@gmail.com; пр. Вернадского, 78, 119454, Москва, Россия; р.т.: +7(499)215-6565.

O. LEBEDEV, A. SHMELEVA, N. GEZHA
**MEDICAL ELECTROPHYSIOLOGICAL DATA AUTOMATED
INTERPRETATION DEEP LEARNING MODEL**

Lebedev O., Shmeleva A., Gezha N. **Medical Electrophysiological Data Automated Interpretation Deep Learning Model.**

Abstract. The article is devoted to the development of a neural network model for automated analysis of medical data in the field of electrophysiology based on deep learning methods. The presented study is of high relevance due to the growing need for improvement of complex spatiotemporal signal processing. Such signals include electrocardiograms and electroencephalograms in clinical practice. Convolutional neural networks (CNNs), which extract spatial features, and temporal convolutional neural networks (TCNs), which specialize in identifying long-term time dependencies, were used as analysis tool. The integration of such architectures made it possible to identify and generalize complex patterns in signals and, subsequently, classify them with high accuracy. The paper presents a hybrid neural network model architecture developed for processing one-dimensional time series and provides a rationale for choosing key hyperparameters. To optimize the model, a strategy for adaptively changing the learning rate based on the value of the loss function on the validation dataset was implemented. An early training stopping mechanism was implemented to prevent redundant computations and to control overfitting in the event of no decrease in the error value on the validation data over a specified number of epochs. During training, model parameters were recorded after each epoch, including the epoch number, loss function value, and optimizer parameters. This allows for the reconstruction of training from any epoch and a detailed analysis of the model's behavior at various stages of training. The application of the developed analysis model enables more in-depth analysis of electrophysiological data, improving the objectivity of diagnostics. The methods presented in this paper can also be scaled to different biomedical data. The practical significance of the paper lies in the integration of the developed model into biomedical systems as a tool for supporting medical decisions, utilizing automated detection of pathologies and reducing the burden on doctors.

Keywords: machine learning, deep learning, neural networks, medical data, electrophysiology.

References

1. Ukaz Prezidenta RF of 02.07.2021 № 400 [On the National Security Strategy of the Russian Federation]. Available at: <http://kremlin.ru/acts/bank/47046> (accessed 20.02.2025). (In Russ.).
2. Ukaz Prezidenta RF of 13.05.2017 № 208 [On the Strategy of Economic Security of the Russian Federation for the period up to 2030]. Available at: <http://www.kremlin.ru/acts/bank/41921> (accessed 20.02.2025). (In Russ.).
3. Shevtsov E.S., Shamin R.V. [Logical integration of information systems based on expert systems]. *Rossiyskiy tekhnologicheskii zhurnal – Russian Journal of Technology*. 2025. vol. 13. no. 2. pp. 27–35. DOI: 10.32362/2500-316X-2025-13-2-27-35. (In Russ.).
4. Mityakov E.S., Lapaev D.N., Beskhmelov M.I., Ladynin A.I., et al. Knowledge-intensive Enterprises Multi-criteria Decision-making Energy Security Management Information System. Proceedings of the 2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). 2022. pp. 1711–1713. DOI: 10.1109/EIConRus54750.2022.9755655.

5. Krasnoperova A.S., Tverdokhlebov A.S., Kartashov A.A., Weber V.I., Kuprits V.Yu. [Investigation of the effectiveness of YOLO neural network models for object recognition in radar images]. *Rossiyskiy tekhnologicheskiy zhurnal – Russian Journal of Technology*. 2025. vol. 13. no. 4. pp. 25–36. DOI: 10.32362/2500-316X-2025-13-4-25-36. (In Russ.).
6. Vsemirnaya organizatsiya zdavookhraneniya. [Cardiovascular diseases (CVD)]. Available at: <https://www.who.int/ru/news-room/fact-sheets/detail/cardiovascular-diseases-cvds> (accessed 25.02.2025). (In Russ.).
7. Ministerstvo zdavookhraneniya Rossiyskoy Federatsii. [Federal project «Fight against cardiovascular diseases»]. Available at: <https://minzdrav.gov.ru/poleznyerekursy/natsproektzdravookhraneniya/bssz> (accessed 25.02.2025). (In Russ.).
8. Shishkin S.V., Sheiman I.M., Almazov A.A., Biryukova A.I., Vlasov V.V., Potapchik E.G., Sazhina S.V., Salakhutdinova S.K. Rossiyskoe zdavookhranenie: perspektivy razvitiya. Doklad NIU VShE [Russian healthcare: development prospects. Report of the Higher School of Economics] (Pod red. Shishkina S.V.) M.: Izd. dom VShE. 2024. 60 p. (In Russ.).
9. Hong S., Zhou Y., Shang J., Xiao C., Sun J. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. *Computers in Biology and Medicine*. 2020. vol. 122. pp. 1–16. DOI: 10.1016/j.cmpbiomed.2020.103801.
10. Zhang C.-J., Yuan L., Tang F.-Q., Cai H.-P., Qian Y.-F., Wang C. Heart failure classification using deep learning to extract spatiotemporal features from ECG. *BMC Medical Informatics and Decision Making*. 2024. vol. 24. no. 1. pp. 1–19. DOI: 10.1186/s12911-024-02415-4.
11. Madan P., Singh V., Singh D.P., Diwakar M., Pant B., Kishor A. A Hybrid Deep Learning Approach for ECG-based Arrhythmia Classification. *Bioengineering*. 2022. vol. 9. no. 4. DOI: 10.3390/bioengineering9040152.
12. Huang Z., Yu J., Shan Y., Wang X. A Non-Invasive Fetal QRS Complex Detection Method Based on a Multi-Feature Fusion Neural Network. *Applied Sciences*. 2024. vol. 14. no. 19.
13. Teplitzky B.A., McRoberts M., Ghanbari H. Deep learning for comprehensive ECG annotation. *Heart Rhythm*. 2020. vol. 17. no. 5. pp. 881–888. DOI: 10.1016/j.hrthm.2020.02.015.
14. Riemann D., Krone L.B., Wulff K., Nissen C. Sleep, insomnia, and depression. *Neuropsychopharmacology*. 2020. vol. 45. no. 1. pp. 74–89. DOI: 10.1038/s41386-019-0411-y.
15. Paul T., Hassan O., Islam S.K., Mosa A.S.M. Real-time obstructive sleep apnea detection from raw ECG and SpO₂ signal using convolutional neural network. *AMIA Joint Summits on Translational Science Proceedings*. 2024. pp. 662–669.
16. Ji X., Li Y., Wen P., Barua P., Acharya U.R. MixSleepNet: A Multi-Type Convolution Combined Sleep Stage Classification Model. *Computer Methods and Programs in Biomedicine*. 2024. vol. 244. 107992 p. DOI: 10.1016/j.cmpb.2023.107992.
17. Mvuh F.L., Ebode Ko'a C.O.V., Bodo B. Multichannel high noise level ECG denoising based on adversarial deep learning. *Scientific Reports*. 2024. vol. 14. no. 1. DOI: 10.1038/s41598-023-50334-7.
18. Aversano L., Bernardi M.L., Cimitile M., Montano D., Pecori R. Characterization of Heart Diseases per Single Lead Using ECG Images and CNN-2D. *Sensors*. 2024. vol. 24. no. 11. DOI: 10.3390/s24113485.
19. Zacarias H., Marques J.A.L., Felizardo V., Pourvahab M., Garcia N.M. ECG Forecasting System Based on Long Short-Term Memory. *Bioengineering*. 2024. vol. 11. no. 1. DOI: 10.3390/bioengineering11010089.

20. Kamanditya B., Fuadah Y.N., Mahardika T.N.Q., Lim K.M. Continuous blood pressure prediction system using Conv-LSTM network on hybrid latent features of photoplethysmogram (PPG) and electrocardiogram (ECG) signals. *Scientific Reports*. 2024. vol. 14. no. 1.
21. Khan F., Yu X., Yuan Z., Rehman A.U. ECG classification using 1-D convolutional deep residual neural network. *PLOS One*. 2023. vol. 18. no. 4. e0284791 p. DOI: 10.1371/journal.pone.0284791.
22. Sawano S., Kodera S., Setoguchi N., et al. Applying masked autoencoder-based self-supervised learning for high-capability vision transformers of electrocardiographies. *PLOS One*. 2024. vol. 19. no. 8. pp. 1–17. DOI: 10.1371/journal.pone.0307978.
23. van de Leur R.R., de Brouwer R., Bleijendaal H., et al. ECG-only explainable deep learning algorithm predicts the risk for malignant ventricular arrhythmia in phospholamban cardiomyopathy. *Heart Rhythm*. 2024. vol. 21. no. 7. pp. 1102–1112. DOI: 10.1016/j.hrthm.2024.02.038.
24. Nishikimi R., Nakano M., Kashino K., Tsukada S. Variational autoencoder-based neural electrocardiogram synthesis trained by FEM-based heart simulator. *Cardiovascular Digital Health Journal*. 2023. vol. 5. no. 1. pp. 19–28. DOI: 10.1016/j.cvdhj.2023.12.002.
25. Sohn J., Shin H., Lee J., Kim H.C. Validation of Electrocardiogram Based Photoplethysmogram Generated Using U-Net Based Generative Adversarial Networks. *Journal of Healthcare Informatics Research*. 2024. vol. 8. no. 1. pp. 140–157. DOI: 10.1007/s41666-023-00156-z.
26. Nasim A., Kim Y.S. DE-PNN: Differential Evolution-Based Feature Optimization with Probabilistic Neural Network for Imbalanced Arrhythmia Classification. *Sensors*. 2022. vol. 22. no. 12. 4450 p. DOI: 10.3390/s22124450.
27. Din S., Qaraqe M., Mourad O., Qaraqe K., Serpedin E. ECG-based cardiac arrhythmias detection through ensemble learning and fusion of deep spatial-temporal and long-range dependency features. *Artificial Intelligence in Medicine*. 2024. vol. 150. 102818 p.

Lebedev Oleg — Ph.D., Dr. Sci., Associate professor, Professor, Department of Computer Science, Federal State Budgetary Educational Institution of Higher Education "MIREA – Russian Technological University" (RTU MIREA). Research interests: artificial intelligence, machine learning, swarm optimization algorithms, decision support systems, development and application of mathematical and software tools. The number of publications — 270. lebedev_o@mirea.ru; 78, Vernadsky Ave., 119454, Moscow, Russia; office phone: +7(499)215-6565.

Shmeleva Anna — Ph.D., Associate professor, Department of Computer Science, Federal State Budgetary Educational Institution of Higher Education "MIREA – Russian Technological University" (RTU MIREA). Research interests: artificial intelligence, machine learning, decision support systems. The number of publications — 107. ann_shmeleva@mail.ru; 78, Vernadsky Ave., 119454, Moscow, Russia; office phone: +7(499)215-6565.

Gezha Nikolay — Student, Department of Computer Science, Federal State Budgetary Educational Institution of Higher Education "MIREA – Russian Technological University" (RTU MIREA). Research interests: machine learning, artificial intelligence, biomedical data analysis, and decision support automation. The number of publications — 1. n.s.gezha@gmail.com; 78, Vernadsky Ave., 119454, Moscow, Russia; office phone: +7(499)215-6565.