УДК 004.032.26; 616.12-073.7
doi:10.31799/1684-8853-2026-5-85-95
EDN: CSPEWE

Научные статьи
Articles

Применение сверточной нейронной сети InceptionTime для классификации электрокардиограмм с множественными метками

В. М. Чубич[a], *доктор техн. наук, профессор, orcid.org/0000-0003-2006-0046, chubich@ami.nstu.ru*
Е. С. Коцелап[a], *студентка, orcid.org/0009-0008-6142-3580*
[a]*Новосибирский государственный технический университет, К. Маркса пр., 20, Новосибирск, 630073, РФ*

Введение: *проблема классификации электрокардиограмм с множественными метками возникает по причине того, что одной записи могут соответствовать несколько диагнозов. Поскольку редкие диагнозы представлены в меньшем количестве, а их признаки проявляются на участках различной длины, использование единого масштаба анализа и общего порога решения повышает риск смещения модели к частым состояниям и пропуска редких патологий.* ***Цель:*** *адаптировать и экспериментально исследовать одномерную сверточную нейронную сеть InceptionTime для классификации электрокардиограмм с множественными метками.* ***Результаты:*** *разработанная схема классификации электрокардиограмм сочетает многомасштабное извлечение признаков, взвешенную бинарную кросс-энтропию с логитами и индивидуальные пороги для пяти диагностических суперклассов. Экспериментальная проверка, проведенная на наборе десятисекундных двенадцатиканальных электрокардиограмм, включавшем записи без патологий, с инфарктом миокарда, изменениями реполяризации, нарушениями проводимости и гипертрофией, выявила, что на тестовой выборке полнота составила 0,74, а макроусредненная гармоническая мера точности и полноты — 0,72. Сравнение с рекуррентными и сверточными архитектурами показало, что модель на основе архитектуры InceptionTime обеспечила наибольшую полноту и продемонстрировала один из лучших результатов по макроусредненной гармонической мере точности и полноты.* ***Практическая значимость:*** *адаптированная модель может служить алгоритмической основой программного модуля предварительного анализа электрокардиограмм и систем поддержки врачебных решений. Независимая оценка диагностических состояний позволяет выявлять несколько патологий в одной записи, а применение отдельных порогов для каждого класса — снижать риск пропуска редких патологий и предварительно выделять записи для последующей проверки специалистом.* ***Обсуждение:*** *дальнейшие исследования целесообразно направить на повышение качества распознавания редких диагностических классов и проверку модели на других биомедицинских сигналах (анализ электроэнцефалограмм, оценка уровня стресса и эмоций человека и др.).*

Ключевые слова *— сверточная нейронная сеть, InceptionTime, глубокое обучение, бинарная кросс-энтропия, электрокардиография, дисбаланс классов, классификация с множественными метками, диагностика сердечно-сосудистых заболеваний.*

Для цитирования: Чубич В. М., Коцелап Е. С. Применение сверточной нейронной сети InceptionTime для классификации электрокардиограмм с множественными метками. *Информационно-управляющие системы*, 2026, № 5, с. 85–95. doi:10.31799/1684-8853-2026-5-85-95, EDN: CSPEWE

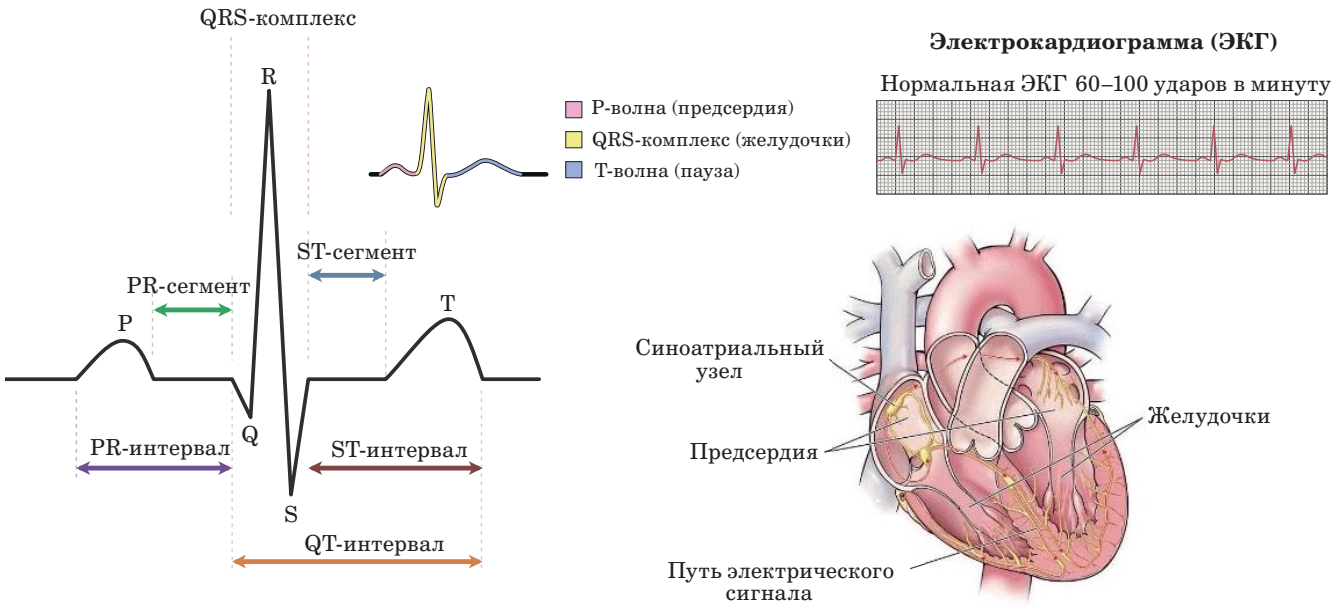
For citation: Chubich V. M., Kotselap E. S. Application of the InceptionTime convolutional neural network to multi-label classification of electrocardiograms. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2026, no. 5, pp. 85–95 (In Russian). doi:10.31799/1684-8853-2026-5-85-95, EDN: CSPEWE

Введение

По оценкам Всемирной организации здравоохранения, сердечно-сосудистые заболевания являются одной из основных причин смертности во всем мире (https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Ранняя и точная диагностика является важным фактором для своевременного медицинского вмешательства и снижения летальности. Электрокардиография традиционно выступает наиболее распространенным, неинвазивным и доступным методом оценки функционального состояния сердца.

Стандартный кардиоцикл (рис. 1) состоит из нескольких ключевых элементов (зубцов P, Q, R, S, T и соответствующих интервалов), каждый из которых отражает определенный этап электрической активности миокарда. Тем не менее ручная интерпретация записей электрокардиограмм (ЭКГ) требует высокой квалификации врачей-кардиологов, подвержена субъективности решений и является крайне трудоемким процессом, особенно в условиях постоянного роста объемов цифровых медицинских данных.

В настоящее время методы машинного и глубокого обучения успешно используются в задачах анализа биомедицинских сигналов [1–3] и являются одним из основных направлений автоматического анализа ЭКГ [4, 5]. Проведенные исследования показывают возможность применения глубоких нейронных сетей как для распознавания нарушений ритма по одноканальным амбулаторным ЭКГ [6], так и для диагностики состояний по 12-канальным записям ЭКГ [7]. В более поздних исследованиях рассматриваются задачи интерпретации ЭКГ, классификации нарушений ритма и применения нейросетевых моделей для диагностики сердечно-сосудистых заболеваний [8–10].



■ ***Рис. 1.*** Структура нормального ЭКГ-сигнала и путь электрического импульса в сердце
■ ***Fig. 1.*** Structure of a normal ECG signal and the path of the electrical impulse in the heart

Часто при анализе одномерных временных рядов ЭКГ в качестве базовых подходов используют рекуррентные нейронные сети (RNNs). В [11] проведено сравнение базовой RNN, Long Short-Term Memory (LSTM) и Gated Recurrent Unit (GRU) для классификации ЭКГ. Рекуррентные нейронные сети читают сигнал последовательно, из-за чего плохо улавливают признаки на длинных записях ЭКГ. В [12, 13] показано, что применение сверточных нейронных сетей (CNN) позволяет эффективно улавливать как локальные изменения в форме кардиоцикла (например, подъем сегмента ST при инфаркте), так и глобальные нарушения ритма. В [14] рассмотрена гибридная архитектура CNN-BiLSTM (такие модели часто оказываются сложнее в реализации и требуют большего числа вычислительных ресурсов).

Альтернативой выступают сверточные нейронные сети InceptionTime, которые опираются на идею параллельного извлечения признаков фильтрами разного размера [15]. В [16] описана архитектура InceptionTime, использованная для классификации временных рядов. В сравнительных экспериментах [16] InceptionTime продемонстрировала качество классификации, сопоставимое с HIVE-COTE, при существенно лучшей масштабируемости по объему и длине временных рядов. В работе [17] предложена сверточная модель для классификации ЭКГ Multireceptive Field CNN (MRF-CNN), в которой несколько размеров рецептивного поля помогают учитывать как короткие, так и более протяженные участки сигнала.

В данной работе выбрана архитектура InceptionTime, поскольку она специально разработана для классификации временных рядов, сочетает многомасштабную обработку с остаточными связями и уже применяется в InceptionTime-подобных подходах к анализу ЭКГ [18, 19]. Для ЭКГ такой подход особенно полезен, поскольку диагностические признаки могут проявляться как в коротких элементах сигнала, например в QRS-комплексе, так и в более протяженных изменениях сегмента ST и зубца T.

Для успешного обучения глубоких нейронных сетей требуются крупные и качественно размеченные медицинские наборы данных. Одним из наиболее популярных ресурсов для исследования алгоритмов автоматической диагностики является PTB-XL [20]. В сравнительном исследовании методов глубокого обучения на PTB-XL [21] ResNet- и Inception-подобные сверточные архитектуры показали одни из наиболее высоких результатов среди рассмотренных подходов к анализу ЭКГ. Данный набор включает 12-канальные записи ЭКГ длительностью 10 с и содержит диагностические метки, сформированные с использованием Standard Communications Protocol for Computerized Electrocardiography (SCP-ECG). При этом одной записи могут соответствовать несколько диагностических меток, которые дополнительно объединены в суперклассы и подклассы.

Целью работы является адаптация и экспериментальное исследование архитектуры одномерной сверточной нейронной сети InceptionTime для

автоматической классификации ЭКГ-сигналов по пяти диагностическим суперклассам набора данных PTB-XL с использованием функции взвешенной бинарной кросс-энтропии и подбором индивидуальных порогов классификации.

Постановка задачи. Материалы и методы решения

Задача автоматического анализа ЭКГ заключается в классификации записей с множественными метками. Для входной матрицы сигнала $\mathbf{X} \in \mathbb{R}^{12\times1000}$ (12 — число отведений, 1000 — длина временного ряда после приведения к частоте 100 Гц) требуется найти отображение f: $\mathbf{X}\rightarrow\mathbf{Y}$, где $\mathbf{Y} \in \{0, 1\}^5$ — бинарный вектор принадлежности записи к пяти диагностическим суперклассам. При этом, поскольку классы не являются взаимоисключающими, сумма элементов вектора $\mathbf{Y}$ может быть больше единицы.

Описание набора данных PTB-XL

В качестве базы исследования использовался открытый набор данных PTB-XL версии 1.0.3 [20, 21]. Он содержит 21 799 клинических 12-канальных записей ЭКГ длительностью 10 с, полученных от 18 869 пациентов. В данной работе использовались записи с частотой дискретизации 100 Гц. Выбор указанной частоты обусловлен результатами сравнения значений метрик качества и вычислительных затрат для представлений 100 и 500 Гц при согласованных временных масштабах сверточных ядер.

Оригинальная разметка содержит множество диагнозов, которые для решения поставленной задачи были агрегированы в пять диагностических суперклассов:

— NORM (Normal ECG) — без признаков патологий;

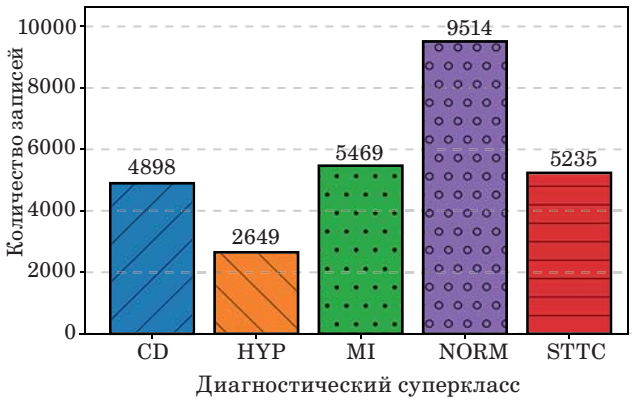
— MI (Myocardial Infarction) — инфаркт миокарда;

— STTC (ST/T Change) — изменения сегмента ST и зубца T;

— CD (Conduction Disturbance) — нарушения проводимости;

— HYP (Hypertrophy) — гипертрофия.

Используемые диагностические суперклассы представляют укрупненные категории электрокардиографических состояний. В рассматриваемой постановке классификация ЭКГ по этим суперклассам предназначена для предварительного выделения крупных групп указанных выше изменений. Результат может использоваться для предварительной сортировки ЭКГ-записей и выделения записей, требующих последующего анализа специалистом, не заменяя детальную интерпретацию ЭКГ врачом.



■ ***Рис. 2.*** Распределение количества ЭКГ-записей по пяти диагностическим суперклассам

■ ***Fig. 2.*** Distribution of ECG records across the five diagnostic superclasses

После проведения анализа были исключены записи, не содержащие диагностические суперклассы, что привело к сокращению итогового объема набора данных до 21 388 записей от 18 617 пациентов.

Распределение количества диагностических меток по суперклассам представлено на рис. 2 (подчеркнем, что одна запись может относиться сразу к нескольким классам). Видно, что набор данных несбалансирован: класс NORM представлен заметно чаще, чем класс HYP.

Предварительная обработка данных

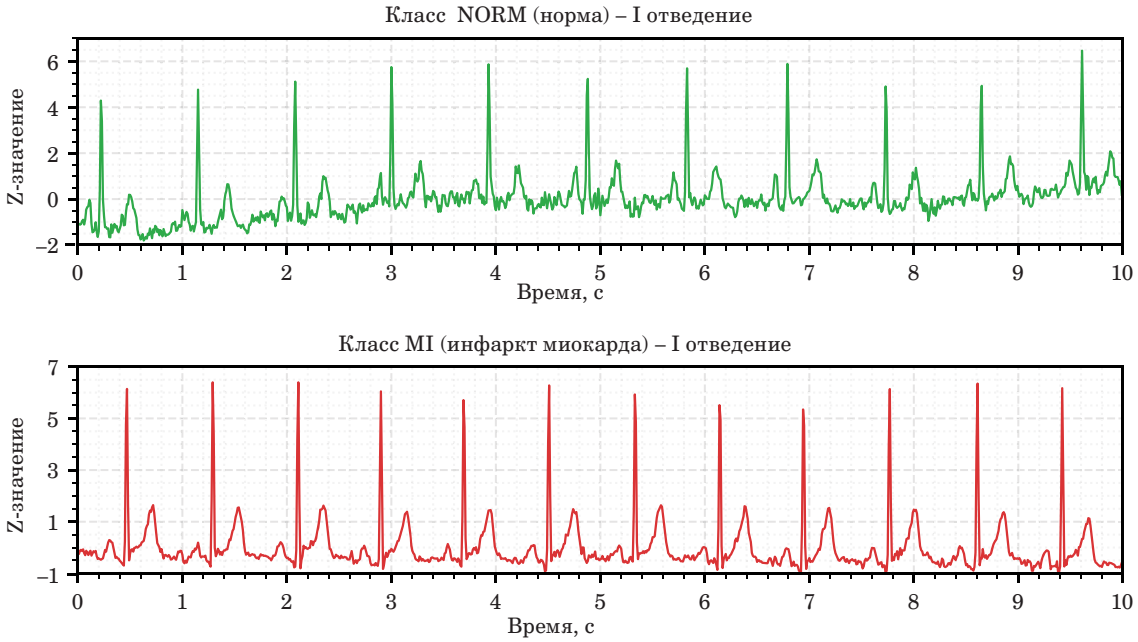
На этапе предварительной обработки выполнялась Z-score нормализация отдельно для каждой ЭКГ-записи. Среднее значение и среднеквадратическое отклонение рассчитывались совместно по всем отсчетам 12 отведений данной записи, после чего полученные общие параметры использовались для нормализации всех каналов. Такое преобразование позволяло привести ЭКГ-записи к сопоставимому численному масштабу и повысить устойчивость процесса обучения нейронной сети. При этом для всех отведений одной записи использовались общие параметры нормализации, поэтому независимое масштабирование отдельных каналов не выполнялось.

Нормализация выполнялась в соответствии с формулой

$$x_i^{'} = \frac{x_i - \overline{x}}{s},$$

где x_i — текущее значение сигнала; $\overline{x}$ — среднее значение сигнала; s — выборочное среднеквадратическое отклонение.

Примеры предварительно обработанных 10-секундных сигналов для нормальной ЭКГ и ЭКГ с инфарктом миокарда приведены на рис. 3.



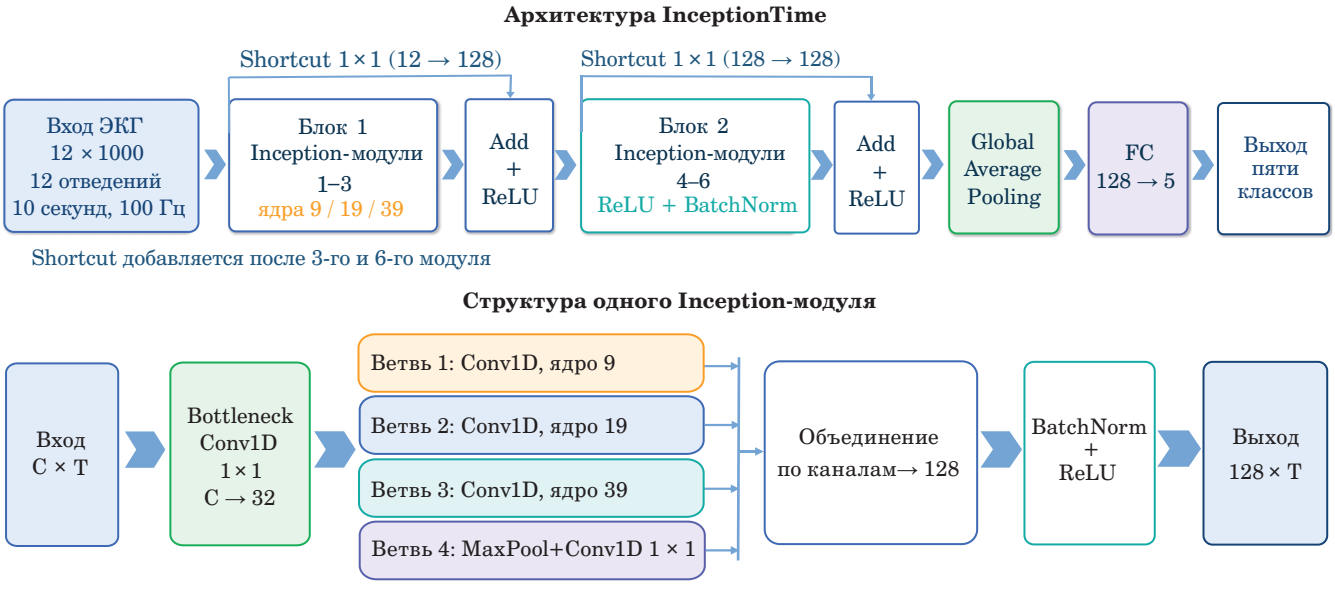
■ ***Рис. 3.*** Примеры предварительно обработанных ЭКГ-сигналов
■ ***Fig. 3.*** Examples of preprocessed ECG signals

Архитектура InceptionTime

Для решения задачи классификации многоканальных временных рядов была выбрана архитектура InceptionTime [16]. Данная архитектура развивает идею Inception-модулей, предложенную в [15], и основана на параллельном применении одномерных сверток с различными размерами ядер к одному и тому же входному сигналу. Схема разработанной модели представлена на рис. 4.

Модель состоит из шести последовательно соединенных Inception-модулей. После третьего и шестого модулей добавляются остаточные связи. Они помогают передавать информацию между блоками напрямую и делают обучение более устойчивым.

Каждый Inception-модуль содержит несколько параллельных ветвей обработки сигнала. Сначала применяется свертка с размером ядра 1 × 1, формирующая 32 признаковых канала,



■ ***Рис. 4.*** Архитектура модели InceptionTime и структура одного Inception-модуля
■ ***Fig. 4.*** Architecture of the InceptionTime model and structure of one Inception module

которые затем поступают в параллельные сверточные ветви. Затем сигнал одновременно проходит через три одномерные свертки с размерами ядер 9, 19 и 39. При частоте дискретизации 100 Гц это соответствует временным окнам около 90, 190 и 390 мс. Конфигурация размеров ядер 9, 19 и 39 была выбрана на этапе предварительной настройки модели по результатам экспериментального сравнения различных вариантов на валидационной выборке. Ядро размером 9 помогает выделять короткие элементы ЭКГ, прежде всего QRS-комплекс. Ядро размером 19 учитывает участки средней длительности, а ядро размером 39 — более протяженные изменения, связанные с сегментом ST и зубцом T. Кроме того, в модуле используется ветвь максимального объединения с последующей сверткой 1×1. Эта ветвь помогает модели быть менее чувствительной к небольшим сдвигам сигнала во времени. После этого выходы всех ветвей объединяются по каналам и проходят через пакетную нормализацию и функцию активации ReLU. На выходе одного Inception-модуля формируется 128 признаковых каналов.

После последнего Inception-модуля значения каждого признакового канала усредняются по всей длине сигнала. В результате для каждого канала остается одно итоговое значение, а полученный вектор поступает на полносвязный слой с пятью выходами — по числу диагностических суперклассов. Для каждого из них затем применяется сигмоидная функция. Она позволяет отдельно оценивать вероятность каждого класса и не требует, чтобы сумма вероятностей была равна единице. Это важно, поскольку одна ЭКГ-запись может одновременно относиться к нескольким диагностическим категориям.

Обучение модели и подбор порогов классификации

Программная реализация модели и процедуры обучения была выполнена на языке Python. Архитектура нейронной сети реализована средствами PyTorch [22]. Для организации процесса обучения, работы с временными рядами, применения преобразований данных и расчета метрик качества использовались библиотеки fastai [23] и tsai [24].

Вычислительные эксперименты проводились в облачной среде Kaggle с использованием графического ускорителя NVIDIA Tesla T4. Для сопоставимой оценки вычислительных затрат время обучения представлений 100 и 500 Гц измерялось на фиксированном интервале в 25 эпох. При 100 Гц обучение заняло 11,88 мин, при 500 Гц с ядрами (9; 19; 39) — 55,63 мин, а в варианте 500 Гц с ядрами (45; 95; 195) — 178,92 мин. Пиковое использование памяти составило 0,88 ГБ для 100 Гц и 3,76 ГБ для варианта 500 Гц с увеличенными ядрами. Поскольку переход от 100 Гц к 500 Гц не обеспечил улучшения Macro-F1, было сохранено представление ЭКГ с частотой дискретизации 100 Гц.

Обучение проводилось мини-пакетами по 64 записи. Использовалось рекомендованное в [20] разбиение набора данных: обучающая выборка — блоки 1–8, валидационная выборка — блок 9, тестовая выборка — блок 10. После исключения записей, не содержащих рассматриваемых диагностических суперклассов, обучающая выборка включала 17 084 записи от 14 823 пациентов, валидационная — 2 146 записей от 1 917 пациентов, тестовая — 2 158 записей от 1 877 пациентов. Проверка идентификаторов пациентов показала отсутствие пересечений между обучающей, валидационной и тестовой выборками, поэтому различные ЭКГ-записи одного пациента не попадали в разные выборки. Валидационная выборка применялась для настройки параметров модели и порогов классификации, а тестовая — только для финальной оценки качества.

Для оптимизации параметров модели использовался алгоритм AdamW [25] с коэффициентом затухания весов 0,1. Скорость обучения изменялась по «политике одного цикла» [26], при этом максимальное значение скорости обучения составляло 10^{-4}. Значения максимальной скорости обучения и коэффициента затухания весов дополнительно проверялись в сравнительных экспериментах на валидационной выборке. Максимальное число эпох было равно 100. Для снижения риска переобучения применялась ранняя остановка: если значение функции потерь на валидационной выборке не улучшалось в течение восьми эпох, обучение прекращалось автоматически. Для итоговой оценки использовалось состояние модели, соответствующее минимальному значению функции потерь на валидационной выборке.

Поскольку в работе решается задача классификации с множественными метками, для обучения использовалась бинарная кросс-энтропия с логитами L_{BCE}. Эта функция рассчитывает ошибку независимо для каждого диагностического суперкласса и не предполагает, что классы являются взаимоисключающими. Для учета дисбаланса классов применялась взвешенная модификация функции потерь, которая вычислялась по формуле

$$L_{\mathrm{BCE}}(\mathbf{y}, \hat{\mathbf{y}}) =$$

$$= -\frac{1}{5}\sum_{i=1}^{5}[w_i \cdot y_i \log(\sigma(\hat{y}_i)) + (1 - y_i)\log\big(1 - \sigma(\hat{y}_i)\big)],$$

где для i-го класса: w_i — весовой коэффициент, компенсирующий дисбаланс путем штрафа за ложноотрицательные предсказания; $y_i \in \{0, 1\}$ — истинная метка (1 — класс присутствует, 0 — класс отсутствует); $\hat{y}_i$ — логит («сырой» выход сети

до функции активации); $\sigma(\hat{y}_i) = \frac{1}{1+e^{-\hat{y}_i}}$ — сигмоидная функция активации, преобразующая логит в вероятность.

При проведении вычислительных экспериментов вектор весов $\mathbf{W} = (w_1; w_2; w_3; w_4; w_5) = (w_{CD}; w_{HYP}; w_{MI}; w_{NORM}; w_{STTC})$ выбирался равным (1,0; 2,0; 2,0; 1,0; 1,5). Повышенные коэффициенты для классов HYP, MI и STTC увеличивали вклад положительных примеров этих классов в функцию потерь и использовались как умеренная мера учета дисбаланса классов. Выбор коэффициентов был проверен экспериментальным сравнением нескольких схем взвешивания.

После обучения выполнялся подбор индивидуальных порогов классификации. Учитывая, что стандартный порог 0,5 не всегда оптимален для несбалансированных медицинских данных (для редких классов модель может выдавать более низкие значения вероятности), целесообразно для каждого диагностического суперкласса на валидационной выборке перебирать значения порога от 0,10 до 0,90 с шагом 0,05. При этом оптимальным считался порог, при котором F1-мера для данного класса достигала наибольшего значения. Найдены индивидуальные оптимальные пороги классификации: NORM — 0,45; MI — 0,50; STTC — 0,50; CD — 0,40; HYP — 0,45.

Результаты экспериментов и их обсуждение

Оценка качества модели осуществлялась на тестовой выборке. Для анализа результатов использовались общая точность (Accuracy), точность положительных предсказаний (Precision), полнота (Recall), F1-мера и макроусредненная F1-мера (Macro-F1), которые рассчитывались по следующим формулам:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN};$$

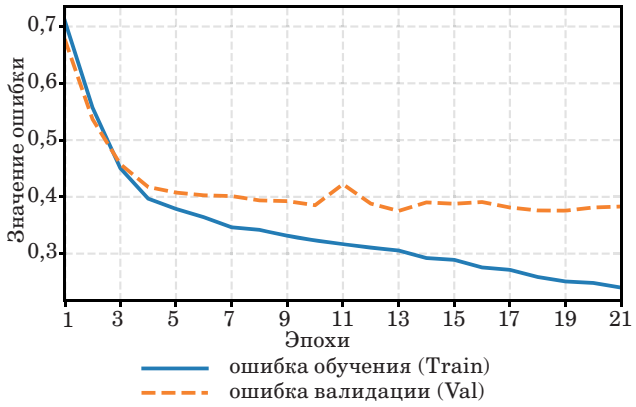
$$Precision = \frac{TP}{TP + FP};$$

$$Recall = \frac{TP}{TP + FN};$$

$$F1 = \frac{2 \times TP}{2 \times TP + FN + FP};$$

$$Macro\text{-}F1 = \frac{1}{5}\sum_{i=1}^{5} \mathrm{F1}_i,$$

где TP, TN, FP и FN — соответственно истинно положительные, истинно отрицательные, ложноположительные и ложноотрицательные решения.



■ ***Рис. 5.*** Динамика изменения функции потерь L_{BCE} при обучении

■ ***Fig. 5.*** Dynamics of the loss function during training

Подчеркнем, что для несбалансированных медицинских данных особое значение имеет полнота, поскольку она отражает долю реально присутствующих патологий.

Графики изменения функции потерь в зависимости от эпохи на обучающей и валидационной выборках представлены на рис. 5. Отметим, что минимальное значение ошибки на валидационной выборке было достигнуто на 13-й эпохе, после чего дальнейшее обучение приводило к переобучению.

Поскольку распределение диагностических классов в используемом наборе данных является несбалансированным, Accuracy в данной работе рассматривается как вспомогательная метрика, характеризующая динамику процесса обучения, и не используется в качестве основного критерия качества модели. Основные выводы о качестве классификации основываются прежде всего на Recall и Macro-F1. График изменения общей точности в зависимости от эпохи на валидационной выборке представлен на рис. 6. Рост значения метрики в первые эпохи и последующая стабилизация свидетельствуют о выходе процесса обучения на устойчивый уровень без дальнейшего существенного изменения данной метрики на валидационной выборке.

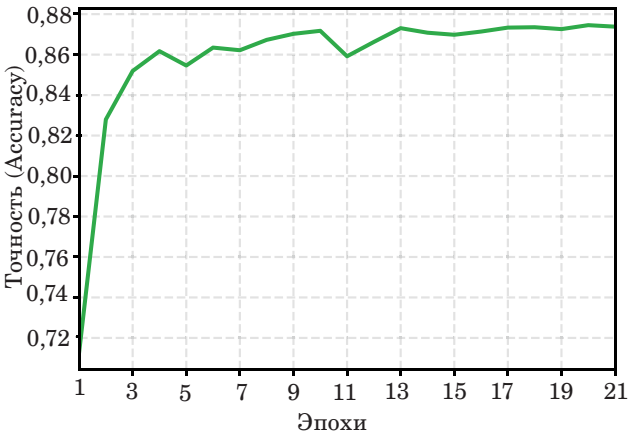
Для проверки эффективности предложенного подхода значения метрик качества модели InceptionTime сравнивались с соответствующими метриками других моделей, применявшихся для анализа ЭКГ на базе PTB-XL (таблица).

Анализ таблицы показывает, что предложенная модель InceptionTime обеспечивает наибольшее значение метрики полноты 0,74 по сравнению с рекуррентными и сверточными моделями. Для задачи предварительного анализа ЭКГ этот результат особенно важен, поскольку Recall отражает способность модели находить реально присутствующие патологические состояния.

■ Сравнение значений метрик качества классификации моделей по пяти диагностическим суперклассам PTB-XL

■ Comparison of classification performance metrics for models across the five PTB-XL diagnostic superclasses

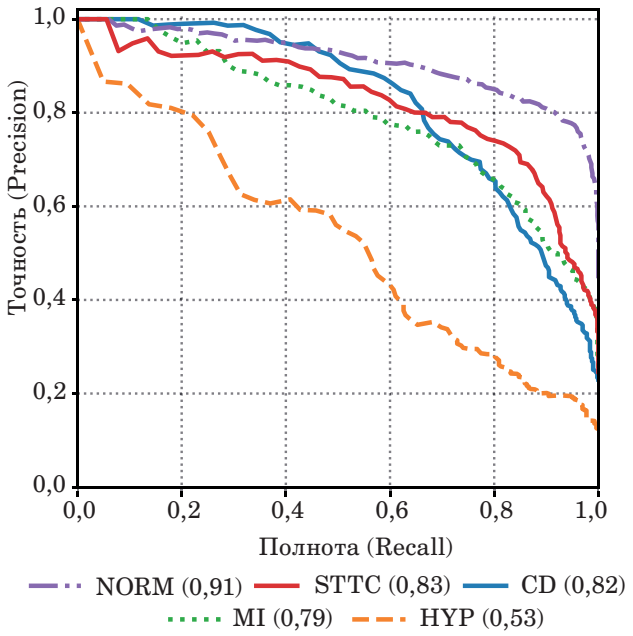
Модель	Accuracy	Precision	Recall	Macro-F1	Источник
RNN	0,76	0,01	0,07	0,09	[11]
GRU	0,86	0,82	0,52	0,62	[11]
LSTM	0,87	**0,82**	0,57	0,66	[11]
CNN	0,72	0,64	0,60	0,61	[12]
SincNet	0,73	0,67	0,59	0,60	[12]
CNN + entropy features	0,77	0,71	0,66	0,68	[12]
MRF-CNN	**0,90**	0,73	0,71	0,72	[17]
Inception-Time	0,87	0,70	**0,74**	**0,72**	Данная работа



■ ***Рис. 6.*** Динамика изменения точности модели на валидационной выборке

■ ***Fig. 6.*** Dynamics of model accuracy on the validation set

По значению метрики Macro-F1 модель InceptionTime превосходит все представленные в таблице модели за исключением MRF-CNN, для которой значения совпадают. Отметим, что модель MRF-CNN демонстрирует более высокую Accuracy, хотя для несбалансированных медицинских данных она не является основной метрикой, поскольку может быть завышена за счет преобладания часто встречающихся классов. Небольшое снижение метрики Precision в модели InceptionTime по сравнению с MRF-CNN связано с повышением чувствительности модели. Такой компромисс допустим для систем предварительного анализа ЭКГ: ложноположительное срабатывание может



■ ***Рис. 7.*** Кривые точности-полноты для пяти диагностических суперклассов

■ ***Fig. 7.*** Precision-recall curves for the five diagnostic superclasses

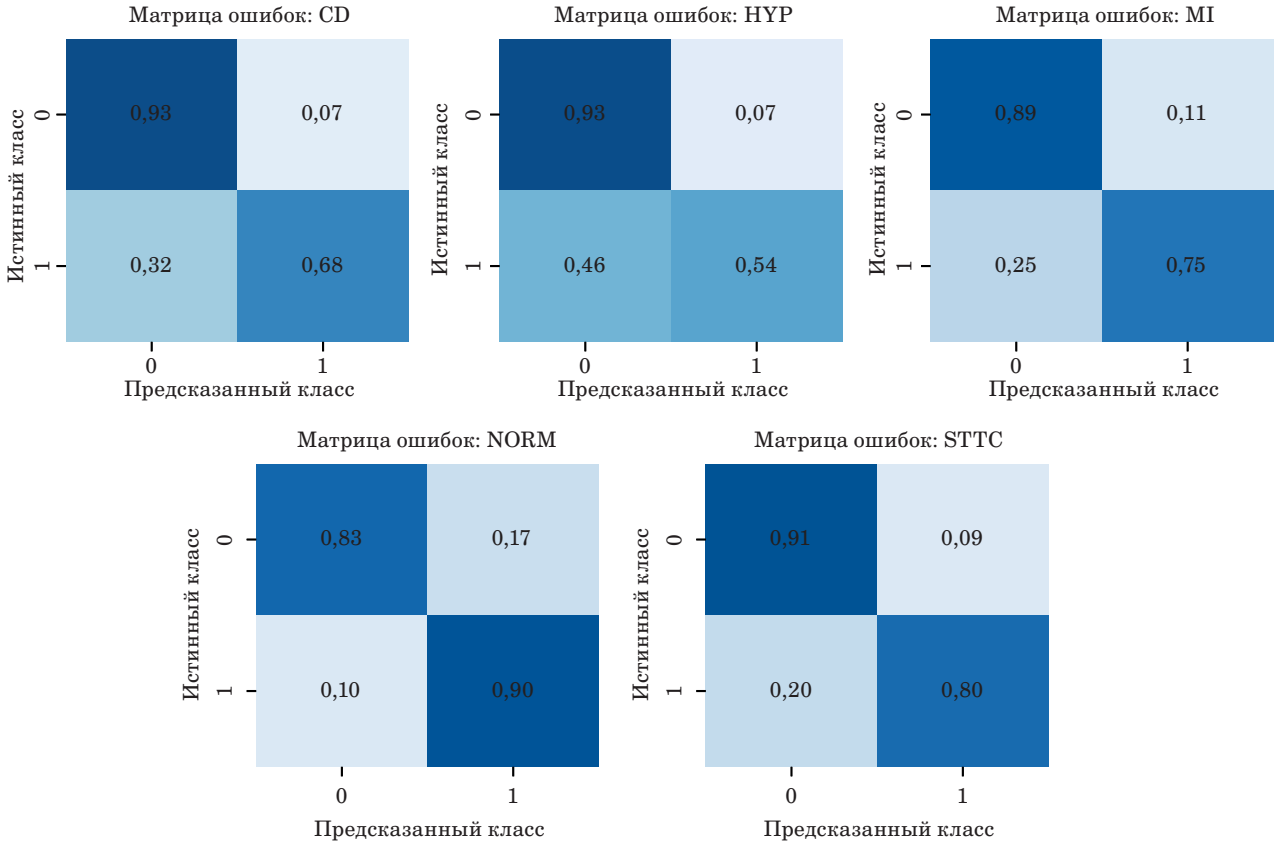
быть дополнительно проверено врачом, в то время как пропуск патологического состояния несет больший риск. В данной работе основной акцент сделан на метрики Recall и Macro-F1, которые лучше отражают качество распознавания патологий в условиях дисбаланса данных.

Для анализа зависимости между точностью Precision и полнотой Recall при различных значениях порога классификации для каждого из пяти диагностических суперклассов были построены кривые точности-полноты (рис. 7); в скобках приведены соответствующие значения площадей под кривыми Precision–Recall (PR-AUC).

Наибольшее значение PR-AUC получено для суперкласса NORM — 0,91. Наименьшее значение PR-AUC, равное 0,53, соответствует суперклассу HYP, который остается наиболее сложным для распознавания среди рассматриваемых диагностических категорий.

Нормализованные матрицы ошибок для диагностических суперклассов (рис. 8) показывают долю верных и ошибочных решений.

Матрицы ошибок подтверждают, что модель уверенно распознает нормальные записи и демонстрирует высокую чувствительность к инфаркту миокарда и изменениям сегмента ST/T. Для суперклассов CD и HYP задача остается более сложной, что связано с меньшим числом записей и большей вариативностью проявлений этих состояний. Подбор индивидуальных порогов позволяет уменьшить число ложноотрицательных решений для таких классов.



■ ***Рис. 8.*** Нормализованные матрицы ошибок для пяти диагностических суперклассов
■ ***Fig. 8.*** Normalized confusion matrices for the five diagnostic superclasses

Заключение

Предложен и экспериментально исследован подход к многометочной классификации 12-канальных ЭКГ-сигналов по пяти диагностическим суперклассам PTB-XL на основе одномерной сверточной нейронной сети InceptionTime.

Основной результат работы состоит в том, что архитектура была адаптирована не только для обработки временных рядов, но и как часть полной схемы классификации несбалансированных медицинских данных, включающей в себя нормализацию ЭКГ-записей, многомасштабное извлечение признаков, взвешенную бинарную кросс-энтропию и индивидуальные пороги классификации для каждого диагностического суперкласса.

Экспериментальная проверка показала, что для предложенной модели InceptionTime на тестовой выборке Macro-F1 = 0,72 и Recall = 0,74. При этом по сравнению с другими моделями InceptionTime представила наибольшее значение Recall и один из лучших результатов по Macro-F1. Отметим, что высокое значение Recall существенно именно для медицинских задач, поскольку оно отражает способность модели выявлять реально присутствующие патологические состояния.

Полученные результаты демонстрируют применимость архитектуры InceptionTime для задачи предварительного многометочного анализа ЭКГ. Разработанная модель обеспечивает одновременное выявление нескольких диагностических категорий и показывает конкурентоспособные результаты по полноте и Macro-F1 в сравнении с рассмотренными подходами. Предложенная схема может служить основой для дальнейшего развития программных средств предварительного анализа ЭКГ и систем поддержки врачебных решений. При этом модель рассматривается именно как инструмент предварительного анализа, дополняющий интерпретацию специалиста, а не как автономная система постановки диагноза.

Дальнейшие исследования в рамках рассмотренной задачи могут быть направлены на повышение качества распознавания наиболее сложных и менее представленных диагностических классов. Представляется целесообразной адаптация сверточной нейронной сети InceptionTime к другим видам биомедицинских сигналов (анализ электроэнцефалограмм, оценка уровня стресса и эмоций человека и др.).

Литература

1. **Ayano Y. M., Schwenker F., Dufera B. D., Debelee T. G.** Interpretable machine learning techniques in ECG-based heart disease classification: A systematic review. *Diagnostics,* 2023, vol. 13, no. 1, art. 111. doi:10.3390/diagnostics13010111
2. **Marwa S. R., Shakir M. A.** Improving ECG signals classification by using deep learning techniques: A review. *ITM Web of Conferences*, 2024, vol. 64, art. 01023. doi:10.1051/itmconf/20246401023
3. **Wang L., Bi T., Hao J., Zhou T. H.** Heart diseases recognition model based on HRV feature extraction over 12-lead ECG signals. *Sensors*, 2024, vol. 24, no. 16, art. 5296. doi:10.3390/s24165296
4. **Petmezas G., Stefanopoulos L., Kilintzis V., Tzavelis A., Rogers J. A., Katsaggelos A. K., Maglaveras N.** State-of-the-art deep learning methods on electrocardiogram data: Systematic review. *JMIR Medical Informatics*, 2022, vol. 10, no. 8, e38454. doi:10.2196/38454
5. **Martínez-Sellés M., Marina-Breysse M.** Current and future use of artificial intelligence in electrocardiography. *Journal of Cardiovascular Development and Disease*, 2023, vol. 10, no. 4, art. 175. doi:10.3390/jcdd10040175
6. **Hannun A. Y., Rajpurkar P., Haghpanahi M., Tison G. H., Bourn C., Turakhia M. P., Ng A. Y.** Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 2019, vol. 25, no. 1, pp. 65–69. doi:10.1038/s41591-018-0268-3
7. **Ribeiro A. H., Ribeiro M. H., Paixão G. M. M., Oliveira D. M., Gomes P. R., Canazart J. A., Ferreira M. P. S., Andersson C. R., Macfarlane P. W., Meira W. Jr., Schön T. B., Ribeiro A. L. P.** Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nature Communications*, 2020, vol. 11, art. 1760. doi:10.1038/s41467-020-15432-4
8. **Ose B., Sattar Z., Gupta A., Toquica C., Harvey C., Noheria A.** Artificial intelligence interpretation of the electrocardiogram: A state-of-the-art review. *Current Cardiology Reports*, 2024, vol. 26, no. 6, pp. 561–580. doi:10.1007/s11886-024-02062-1
9. **Wu Z., Guo C.** Deep learning and electrocardiography: Systematic review of current techniques in cardiovascular disease diagnosis and management. *BioMedical Engineering OnLine*, 2025, vol. 24, art. 23. doi:10.1186/s12938-025-01349-w
10. **Gustafsson S., Ribeiro A. H., Gedon D., Abreu P. E. O. G. B., Pielawski N., Paixão G. M. M., Gutierrez M. A., Krieger J. E., Dias F. M., Ribeiro A. L. P., Lindholm D., Schön T. B., Sundström J.** A deep learning ECG model for identification and localization of occlusion myocardial infarction. *Nature Communications*, 2026, vol. 17, art. 4336. doi:10.1038/s41467-026-73023-1
11. **Sæther S.** *Comparing Recurrent Neural Networks for ECG Analysis*: Master's thesis. Oslo, Oslo Metropolitan University, 2023. 61 p. https://home.simula.no/~paalh/students/SanderSaether-OsloMet-2023.pdf (дата обращения: 12.05.2026).
12. **Śmigiel S., Pałczyński K., Ledziński D.** ECG signal classification using deep learning techniques based on the PTB-XL dataset. *Entropy*, 2021, vol. 23, no. 9, art. 1121. doi:10.3390/e23091121
13. **Lekkas G., Vrochidou E., Papakostas G. A.** Time-frequency transformations for enhanced biomedical signal classification with convolutional neural networks. *BioMedInformatics*, 2025, vol. 5, no. 1, art. 7. doi:10.3390/biomedinformatics5010007
14. **Jaiswal G. K., Mitra S.** Hybrid CNN–BiLSTM Model for ECG signal classification and accurate arrhythmia detection (MIT-BIH and PTB-XL Validation). *International Journal of Computational and Experimental Science and Engineering*, 2025, vol. 11, no. 4, pp. 10058–10067. doi:10.22399/ijcesen.4575
15. **Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A.** Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9. doi:10.1109/CVPR.2015.7298594
16. **Ismail Fawaz H., Lucas B., Forestier G., Pelletier C., Schmidt D. F., Weber J., Webb G. I., Idoumghar L., Muller P.-A., Petitjean F.** InceptionTime: Finding AlexNet for time series classification. *Data Mining and Knowledge Discovery*, 2020, vol. 34, no. 6, pp. 1936–1962. doi:10.1007/s10618-020-00710-y
17. **Feyisa D. W., Debelee T. G., Ayano Y. M., Kebede S. R., Assore T. F.** Lightweight multireceptive field CNN for 12-lead ECG signal classification. *Computational Intelligence and Neuroscience*, 2022, vol. 2022, art. 8413294. doi:10.1155/2022/8413294
18. **Crocker H. J., Costall A. W.** An inceptiontime-inspired convolutional neural network to detect cardiac abnormalities in reduced-lead ECG data. *Computing in Cardiology*, 2021, vol. 48, pp. 1–4. doi:10.23919/CinC53138.2021.9662678
19. **Borra D., Andalò A., Severi S., Corsi C.** On the application of convolutional neural networks for 12-lead ECG multi-label classification using datasets from multiple centers. *Computing in Cardiology*, 2020, vol. 47, pp. 1–4. doi:10.22489/CinC.2020.349
20. **Wagner P., Strodthoff N., Bousseljot R.-D., Kreiseler D., Lunze F. I., Samek W., Schaeffter T.** PTB-XL, a large publicly available electrocardiography dataset. *Scientific Data*, 2020, vol. 7, art. 154. doi:10.1038/s41597-020-0495-6
21. **Strodthoff N., Wagner P., Schaeffter T., Samek W.** Deep learning for ECG analysis: Benchmarks and Insights from PTB-XL. *IEEE Journal of Biomedical and Health Informatics*, 2021, vol. 25, no. 5, pp. 1519–1528. doi:10.1109/JBHI.2020.3022989

22. **Paszke A., Gross S., Massa F., Lerer A., Bradbury J., Chanan G., Killeen T., Lin Z., Gimelshein N., Antiga L., Desmaison A., Köpf A., Yang E., DeVito Z., Raison M., Tejani A., Chilamkurthy S., Steiner B., Fang L., Bai J., Chintala S.** PyTorch: An imperative style, high-performance deep learning library. arXiv preprint, 2019. doi:10.48550/arXiv.1912.01703
23. **Howard J., Gugger S.** Fastai: A layered API for deep learning. *Information*, 2020, vol. 11, no. 2, art. 108. doi:10.3390/info11020108
24. **Oguiza I., Rodriguez-Fernandez V., Neoh D., filipj8, J-M., Kainkaryam R., Mistry D., Yang Z., Williams D., Cho R., geoHeil, Versek C., samlloydig, Mistry D., imilas, Golinski A., Muhr D., H J., MichaelDee, Sachdeva K., Capelle T.** timeseriesAI/tsai: v0.3.9 [Software]. *Zenodo*, 2024. doi:10.5281/zenodo.10647659
25. **Loshchilov I., Hutter F.** Decoupled weight decay regularization. *International Conference on Learning Representations*, New Orleans, LA, USA, 2019. https://openreview.net/forum?id=Bkg6RiCqY7 (дата обращения: 12.05.2026).
26. **Smith L. N., Topin N.** Super-convergence: Very fast training of neural networks using large learning rates. *Proceedings of SPIE*, 2019, vol. 11006, art. 1100612. doi:10.1117/12.2520589

UDC 004.032.26; 616.12-073.7
doi:10.31799/1684-8853-2026-5-85-95
EDN: CSPEWE

Application of the InceptionTime convolutional neural network to multi-label classification of electrocardiograms

V. M. Chubich[a], Dr. Sc., Tech., Professor, orcid.org/0000-0003-2006-0046, chubich@ami.nstu.ru
E. S. Kotselap[a], Student, orcid.org/0009-0008-6142-3580
[a]Novosibirsk State Technical University, 20, K. Marks Pr., 630073, Novosibirsk, Russian Federation

Introduction: The challenge of multi-label electrocardiogram classification arises due to several diagnoses corresponding to a single record. Rare diagnoses have fewer examples, while their features appear in signal segments of different durations. Consequently, a single analysis scale and a common decision threshold increase the risk of bias toward frequent conditions and the risk of missing rare pathologies. **Purpose:** To adapt and experimentally evaluate a one-dimensional InceptionTime convolutional neural network for multi-label electrocardiogram classification. **Results:** The study develops a classification scheme combining multiscale feature extraction, weighted binary cross-entropy with logits, and individual thresholds for five diagnostic superclasses. We evaluate the model on ten-second, twelve-lead electrocardiograms representing records without pathology, myocardial infarction, repolarization abnormalities, conduction disturbances, and hypertrophy. On the test set, recall reaches 0.74, while the macro-averaged harmonic mean of precision and recall reaches 0.72. Comparison with recurrent and convolutional architectures shows that the InceptionTime-based model has achieved the highest recall and one of the highest macro-averaged harmonic means of precision and recall. **Practical relevance:** The adapted model can underpin a software module for preliminary electrocardiogram analysis and clinical decision support systems. Independent assessment of diagnostic conditions enables the detection of several pathologies in one record, while class-specific thresholds help reduce the risk of missing rare pathologies and identify records for specialist review. **Discussion:** The study identifies improving the recognition of rare classes and evaluating the model on other biomedical signals, including electroencephalograms and signals related to stress and emotions, as priorities for further research.

Keywords — convolutional neural network, InceptionTime, deep learning, binary cross-entropy, electrocardiography, class imbalance, multi-label classification, cardiovascular diagnostics.

For citation: Chubich V. M., Kotselap E. S. Application of the InceptionTime convolutional neural network to multi-label classification of electrocardiograms. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2026, no. 5, pp. 85–95 (In Russian). doi:10.31799/1684-8853-2026-5-85-95, EDN: CSPEWE

References

1. Ayano Y. M., Schwenker F., Dufera B. D., Debelee T. G. Interpretable machine learning techniques in ECG-based heart disease classification: A systematic review. *Diagnostics*, 2023, vol. 13, no. 1, art. 111. doi:10.3390/diagnostics13010111
2. Marwa S. R., Shakir M. A. Improving ECG signals classification by using deep learning techniques: A review. *ITM Web of Conferences*, 2024, vol. 64, art. 01023. doi:10.1051/itmconf/20246401023
3. Wang L., Bi T., Hao J., Zhou T. H. Heart diseases recognition model based on HRV feature extraction over 12-lead ECG signals. *Sensors*, 2024, vol. 24, no. 16, art. 5296. doi:10.3390/s24165296
4. Petmezas G., Stefanopoulos L., Kilintzis V., Tzavelis A., Rogers J. A., Katsaggelos A. K., Maglaveras N. State-of-the-art deep learning methods on electrocardiogram data: Systematic review. *JMIR Medical Informatics*, 2022, vol. 10, no. 8, e38454. doi:10.2196/38454
5. Martínez-Sellés M., Marina-Breysse M. Current and future use of artificial intelligence in electrocardiography. *Journal of Cardiovascular Development and Disease*, 2023, vol. 10, no. 4, art. 175. doi:10.3390/jcdd10040175
6. Hannun A. Y., Rajpurkar P., Haghpanahi M., Tison G. H., Bourn C., Turakhia M. P., Ng A. Y. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 2019, vol. 25, no. 1, pp. 65–69. doi:10.1038/s41591-018-0268-3
7. Ribeiro A. H., Ribeiro M. H., Paixão G. M. M., Oliveira D. M., Gomes P. R., Canazart J. A., Ferreira M. P. S., Andersson C. R., Macfarlane P. W., Meira W. Jr., Schön T. B., Ribeiro A. L. P. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nature Communications*, 2020, vol. 11, art. 1760. doi:10.1038/s41467-020-15432-4
8. Ose B., Sattar Z., Gupta A., Toquica C., Harvey C., Noheria A. Artificial intelligence interpretation of the electrocardio-

gram: A state-of-the-art review. *Current Cardiology Reports*, 2024, vol. 26, no. 6, pp. 561–580. doi:10.1007/s11886-024-02062-1
9. Wu Z., Guo C. Deep learning and electrocardiography: Systematic review of current techniques in cardiovascular disease diagnosis and management. *BioMedical Engineering OnLine*, 2025, vol. 24, art. 23. doi:10.1186/s12938-025-01349-w
10. Gustafsson S., Ribeiro A. H., Gedon D., Abreu P. E. O. G. B., Pielawski N., Paixão G. M. M., Gutierrez M. A., Krieger J. E., Dias F. M., Ribeiro A. L. P., Lindholm D., Schön T. B., Sundström J. A deep learning ECG model for identification and localization of occlusion myocardial infarction. *Nature Communications*, 2026, vol. 17, art. 4336. doi:10.1038/s41467-026-73023-1
11. Sæther S. *Comparing Recurrent Neural Networks for ECG Analysis*: Master's thesis. Oslo, Oslo Metropolitan University, 2023. 61 p. Available at: https://home.simula.no/~paalh/students/SanderSaether-OsloMet-2023.pdf (accessed 12 May 2026).
12. Śmigiel S., Pałczyński K., Ledziński D. ECG signal classification using deep learning techniques based on the PTB-XL dataset. *Entropy*, 2021, vol. 23, no. 9, art. 1121. doi:10.3390/e23091121
13. Lekkas G., Vrochidou E., Papakostas G. A. Time-frequency transformations for enhanced biomedical signal classification with convolutional neural networks. *BioMedInformatics*, 2025, vol. 5, no. 1, art. 7. doi:10.3390/biomedinformatics5010007
14. Jaiswal G. K., Mitra S. Hybrid CNN–BiLSTM Model for ECG signal classification and accurate arrhythmia detection (MIT-BIH and PTB-XL Validation). *International Journal of Computational and Experimental Science and Engineering*, 2025, vol. 11, no. 4, pp. 10058–10067. doi:10.22399/ijcesen.4575
15. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9. doi:10.1109/CVPR.2015.7298594
16. Ismail Fawaz H., Lucas B., Forestier G., Pelletier C., Schmidt D. F., Weber J., Webb G. I., Idoumghar L., Muller P.-A., Petitjean F. InceptionTime: Finding AlexNet for time series classification. *Data Mining and Knowledge Discovery*, 2020, vol. 34, no. 6, pp. 1936–1962. doi:10.1007/s10618-020-00710-y
17. Feyisa D. W., Debelee T. G., Ayano Y. M., Kebede S. R., Assore T. F. Lightweight multireceptive field CNN for 12-lead ECG signal classification. *Computational Intelligence and Neuroscience*, 2022, vol. 2022, art. 8413294. doi:10.1155/2022/8413294
18. Crocker H. J., Costall A. W. An inceptiontime-inspired convolutional neural network to detect cardiac abnormalities in reduced-lead ECG data. *Computing in Cardiology*, 2021, vol. 48, pp. 1–4. doi:10.23919/CinC53138.2021.9662678
19. Borra D., Andalò A., Severi S., Corsi C. On the application of convolutional neural networks for 12-lead ECG multi-label classification using datasets from multiple centers. *Computing in Cardiology*, 2020, vol. 47, pp. 1–4. doi:10.22489/CinC.2020.349
20. Wagner P., Strodthoff N., Bousseljot R.-D., Kreiseler D., Lunze F. I., Samek W., Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset. *Scientific Data*, 2020, vol. 7, art. 154. doi:10.1038/s41597-020-0495-6
21. Strodthoff N., Wagner P., Schaeffter T., Samek W. Deep learning for ECG analysis: Benchmarks and Insights from PTB-XL. *IEEE Journal of Biomedical and Health Informatics*, 2021, vol. 25, no. 5, pp. 1519–1528. doi:10.1109/JBHI.2020.3022989
22. Paszke A., Gross S., Massa F., Lerer A., Bradbury J., Chanan G., Killeen T., Lin Z., Gimelshein N., Antiga L., Desmaison A., Köpf A., Yang E., DeVito Z., Raison M., Tejani A., Chilamkurthy S., Steiner B., Fang L., Bai J., Chintala S. PyTorch: An imperative style, high-performance deep learning library. arXiv preprint, 2019. doi:10.48550/arXiv.1912.01703
23. Howard J., Gugger S. Fastai: A layered API for deep learning. *Information*, 2020, vol. 11, no. 2, art. 108. doi:10.3390/info11020108
24. Oguiza I., Rodriguez-Fernandez V., Neoh D., filipj8, J-M., Kainkaryam R., Mistry D., Yang Z., Williams D., Cho R., geoHeil, Versek C., samlloydig, Mistry D., imilas, Golinski A., Muhr D., H J., MichaelDee, Sachdeva K., Capelle T. timeseriesAI/tsai: v0.3.9 [Software]. *Zenodo*, 2024. doi:10.5281/zenodo.10647659
25. Loshchilov I., Hutter F. Decoupled weight decay regularization. *International Conference on Learning Representations*, New Orleans, LA, USA, 2019. Available at: https://openreview.net/forum?id=Bkg6RiCqY7 (accessed 12 May 2026).
26. Smith L. N., Topin N. Super-convergence: Very fast training of neural networks using large learning rates. *Proceedings of SPIE*, 2019, vol. 11006, art. 1100612. doi:10.1117/12.2520589