

Научная статья
УДК 004.8:616.12-073.7
<https://doi.org/10.24143/2072-9502-2026-2-15-22>
EDN VLDTPH

О генерации 12-канальных электрокардиограмм на основе гибрида диффузионной и графовой нейросетевых моделей

Евгений Юрьевич Щетинин[✉], *Анна Вячеславовна Пестрякова,*
Юлия Георгиевна Шаталова, Андрей Андреевич Шевчук

Севастопольский государственный университет,
Севастополь, Россия, riviera-molto@mail.ru[✉]

Аннотация. Представлена гибридная модель VAE-GNN-SSSD для генерации физиологически корректных 12-канальных электрокардиограмм длительностью 10 с. Предложенная архитектура объединяет три ключевых компонента: вариационный автокодировщик для выделения морфологических компонент P-QRS-T, графовую нейронную сеть с частично фиксированной матрицей смежности для обеспечения соблюдения биофизических законов Эйнтовена и Вильсона, а также диффузионную модель со структурированным пространством состояний для моделирования долгосрочных временных зависимостей. Модель позволяет генерировать сигналы с контролем по клиническим параметрам: тип аритмии, возраст, пол и частота сердечных сокращений. Экспериментальные результаты на тестовой выборке РТВ-XL показали FID = 0,052 и PRD = 10,8 %, что сопоставимо с результатами современных методов. Ключевое преимущество модели – встроенная биофизическая корректность, подтвержденная метрикой MSE по закону Эйнтовена (0,084). Практическая эффективность подтверждена в задаче классификации аритмий на MIT-BIH: аугментация синтетическими данными повысила Macro F1 с 0,84 до 0,89 (+6 %), улучшила распознавание редких желудочковых и фузурованных сокращений на 5–7 % и снизила ложный пропуск опасных аритмий на 27 %. Модель продемонстрировала хорошую обобщающую способность на независимых ICU-данных (MIMIC-IV-ECG). Результаты открывают перспективы применения для обучения диагностических систем, симуляции патологий, создания цифровых двойников сердца и подготовки медицинских специалистов при решении проблемы дефицита аннотированных данных и сохранении приватности пациентов.

Ключевые слова: генеративные модели, диффузионные модели, электрокардиография, графовые нейронные сети, трансформеры

Благодарности: работа выполнена при поддержке Севастопольского государственного университета (проект № 42-01-09/319/2025-1).

Для цитирования: Щетинин Е. Ю., Пестрякова А. В., Шаталова Ю. Г., Шевчук А. А. О генерации 12-канальных электрокардиограмм на основе гибрида диффузионной и графовой нейросетевых моделей // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. 2026. № 2. С. 15–22. <https://doi.org/10.24143/2072-9502-2026-2-15-22>. EDN VLDTPH.

Original article

On the generation of 12-channel electrocardiograms based on a hybrid of diffusion and graph neural network models

Evgeny Yu. Shchetinin[✉], *Anna V. Pestryakova, Julia G. Shatalova, Andrey A. Shevchuk*

Sevastopol State University,
Sevastopol, Russia, riviera-molto@mail.ru[✉]

Abstract. A hybrid VAE-GNN-SSSD model is presented for generating physiologically correct 12-channel electrocardiograms with a duration of 10 seconds. The proposed architecture combines three key components: a variational autoencoder for isolating the morphological components of P-QRS-T, a graph neural network with a partially fixed adjacency matrix to ensure compliance with the biophysical laws of Einthoven and Wilson, as well as a diffusion

model with a structured state space for modeling long-term time dependencies. The model allows you to generate signals controlled by clinical parameters: type of arrhythmia, age, gender, and heart rate. Experimental results on the PTB-XL test sample showed FID = 0.052 and PRD = 10.8%, which is comparable with the results of modern methods. The key advantage of the model is its built-in biophysical correctness, confirmed by the MSE metric according to Einthoven's law (0.084). The practical effectiveness was confirmed in the MIT BIH classification of arrhythmias: augmentation with synthetic data increased Macro F1 from 0.84 to 0.89 (+6%), improved the recognition of rare ventricular and fused contractions by 5-7% and reduced the false omission of dangerous arrhythmias by 27%. The model has demonstrated good generalizing ability on independent ICU data (MIMIC-IV-ECG). The results open up prospects for the use of diagnostic systems for training, pathology simulation, creation of digital heart twins and training of medical specialists in solving the problem of shortage of annotated data and maintaining patient privacy.

Keywords: generative models, diffusion models, electrocardiography, graph neural networks, transformers

Acknowledgments: The work was carried out with the support of Sevastopol State University (project no. 42-01-09/319/2025-1).

For citation: Shchetinin E. Yu., Pestryakova A. V., Shatalova J. G., Shevchuk A. A. On the generation of 12-channel electrocardiograms based on a hybrid of diffusion and graph neural network models. *Vestnik of Astrakhan State Technical University. Series: Management, computer science and informatics*. 2026,2:15-22. (In Russ.). <https://doi.org/10.24143/2072-9502-2026-2-15-22>. EDN VLDTPH.

Введение

Сердечно-сосудистые заболевания становятся причиной 30 % смертей в мире [1], а электрокардиография остается основным методом диагностики аритмий. Однако проблемой применения алгоритмов машинного обучения становится критический дефицит данных для редких, но жизнеугрожающих состояний: желудочковые аритмии составляют менее 5 % записей в публичных базах [2, 3]. Причинами являются в том числе этические ограничения на сбор данных, высокая стоимость аннотации (15–20 мин на запись) и популяционная несбалансированность датасетов [4].

Современные исследования подтвердили высокую эффективность интеграции персонифицированных математических моделей сердца с технологиями искусственного интеллекта для создания точных цифровых двойников. Генеративные модели предлагают решение через синтез искусственных ЭКГ [5, 6]. Однако существующие подходы имеют ограничения: генеративно-состязательные модели GAN подвержены коллапсу мод и склонны генерировать ограниченное разнообразие сигналов, что приводит к тому, что все синтетические сигналы одного класса становятся практически идентичными. В применении к проблеме генерации ЭКГ это означает, что, например, модель может генерировать только один «шаблон» желудочковой тахикардии, игнорируя естественное разнообразие клинических проявлений.

Существующие подходы имеют ряд недостатков. Обучение GAN нестабильно: процесс часто не сходится или сходится к локальным минимумам, что делает результаты непредсказуемыми; трудно контролировать конкретные параметры генерируемых сигналов (частоту сердечных сокращений, электрическую ось сердца и др.). Вариационные автоэнкодеры VAE генерируют «средние» сигналы, теряя

важные детали (острые пики QRS, мелкие зубцы P), что делает их клинически непригодными; кроме того, VAE плохо справляются с моделированием долгосрочных зависимостей и межканальных корреляций между отведениями.

Хотя VAE теоретически способны генерировать разнообразные сигналы, на практике они часто производят слишком однородные данные [7]. Применение методов машинного обучения в кардиологии ограничено дефицитом аннотированных записей редких аритмий. Существующие генеративные модели (GAN, VAE, DDPM) не воспроизводят закономерности между отведениями [8, 9].

Научная проблема, исследуемая в данной работе, заключается в необходимости разработки генеративной модели, способной синтезировать 12-канальные ЭКГ-сигналы, которые одновременно удовлетворяют трем критическим требованиям:

- физиологической согласованности, обеспечиваемой точным наблюдением векторкардиографических законов (включая законы Эйнтховена и Вильсона);
- способности генерировать мультицикловые записи длительностью не менее 10 секунд с сохранением дальних временных зависимостей;
- возможности контролируемой генерации по множеству клинических параметров (класс аритмии, возраст, пол, частота сердечных сокращений) при интеграции дифференцируемых биофизических ограничений на морфологию волн и интервалы.

Предлагаемая архитектура формирует физиологически согласованные 12-канальные ЭКГ-сигналы длительностью 10 с и учитывает законы Эйнтховена – Вильсона.

Архитектура гибридной модели VAE-GNN-SSSD

Модель включает 3 блока:

1) вариационный автокодировщик (VAE) выделяет морфологические компоненты P-QRS-T и формирует латентное представление;

2) графовая нейронная сеть с матрицей смежности восстанавливает структуру 12 отведений;

3) диффузионная часть состоит из DDPM (вероятностная диффузионная модель на основе де-ноизинга) + SSSD (диффузия со структурированным пространством состояний) трансформера моделирует временные зависимости длиной 5 000 отсчетов.

Условия: класс аритмии $c \in \{N, S, V, F, Q\}$, метаданные $m = [HR, age, sex] \in \mathbb{R}^3$, где HR – частота сердечных сокращений; age – возраст пациента; sex – пол (0 – женщина, 1 – мужчина). Эти метаданные преобразуются в условный вектор через двухслойный MLP ($3 \rightarrow 64 \rightarrow 128, ReLU$) $\rightarrow e \in \mathbb{R}^{128}$ и затем используются в FiLM-модуляции для масштабирования и сдвига активаций в каждом сверточном слое. Это позволяет модели персонализировать генерацию сигналов в зависимости от клинических параметров пациента.

Вариационный автокодировщик VAE отвечает за параметрическое сжатие 12-канального ЭКГ-сигнала в компактный морфологический базис $b(t) \in \mathbb{R}^{T \times 3}$ ($T = 5\,000$), интерпретируемый как ортогональные компоненты P-, QRS- и T-волн. Это обеспечивает 16-кратное сжатие по времени (с 5 000 до 312 отсчетов) при сохранении физиологической структуры.

Вариационный автокодировщик состоит из энкодера, латентного пространства и декодера [3].

Энкодер имеет следующую архитектуру:

– три одномерных сверточных слоя Conv1D с FiLM-модуляцией по вектору условия $e \in \mathbb{R}^{128}$;

– ядра 5, 3, 3; фильтры $64 \rightarrow 32 \rightarrow 16$; шаг 2 (апсемплинг $\times 8$ по времени);

– после каждого слоя следует блок $ReLU + FiLM$ $\gamma_i \odot h_i + \beta_i$, где $\gamma_i, \beta_i = Linear(128 \rightarrow C_i)(e)$. Далее слой max-pooling (шаг 2) и полносвязный слой, выводящий параметры латентного распределения $\mu_z, \log \sigma_z^2 \in \mathbb{R}^{312 \times 48}$ (по 16 переменных на каждую из трех компонент).

Латентное пространство состоит из векторов z , которые сэмпляются через репараметризацию.

$$z = \mu_z + \sigma_z \odot \varepsilon, \varepsilon \sim N(0, I).$$

Архитектура декодера имеет следующий вид: симметричная архитектура с тремя слоями ConvTranspose1D (ядра 3, 3, 5; фильтры $16 \rightarrow 32 \rightarrow 64$; шаг 2) и остаточными связями после каждого слоя. Выход нормализуется по SVD-компонентам реальных ЭКГ:

$$b(t) = \frac{\hat{b}(t) - \mu_b}{\sigma_b}.$$

Функция потерь имеет следующий вид:

$$L_{VAE} = MSE(b_{recon}, b_{real}) + \beta \cdot KL(q(z|x) \parallel,$$

где β линейно растет от 0 до 0,1 за первые 20 эпох (отжиг для предотвращения постериорного коллапса). VAE обеспечивает управляемое сжатие в физиологически осмысленный базис, закладывая основу для последующего восстановления 12 отведений через GNN и генерации через DDPM-SSSD.

Морфологический базис из VAE-декодера проецируется на 12 узлов графовой нейронной сети (по одному на каждое стандартное отведение) через независимые линейные слои. Затем применяются два слоя GraphSAGE с mean-агрегацией: эмбединги соседей усредняются, конкатенируются с эмбедингом узла и преобразуются MLP (128 нейронов, ReLU, dropout 0.1). Финальный слой формирует выходной сигнал $x_0^{base} \in \mathbb{R}^{T \times 12}$.

Матрица смежности $A \in \mathbb{R}^{12 \times 12}$ содержит 70 % фиксированных элементов, инициализированных аналитически в соответствии с законами Эйнтховена – Гольдбергера и Вильсона ($I + II + III = 0$, $aVR = -(I + II) / 2$, $aVL = I - II / 2$, $aVF = II - I / 2$, сумма грудных отведений $V1-V6 \approx 0$ и т. д.). Эти элементы не участвуют в градиентном спуске. Остальные 30 % обучаемых элементов матрицы A (преимущественно связи внутри группы $V1-V6$ и между грудными и конечностными отведениями) позволяют адаптироваться к индивидуальным особенностям пациента (пол, возраст, конституция). После каждого обновления обучаемых элементов матрица симметризуется:

$$A \leftarrow \frac{A + A^T}{2}.$$

Функция потерь GNN имеет вид

$$L_{GNN} = MSE(x_0^{base}, x_{real}) + L_{bio},$$

где L_{bio} – среднеквадратичная ошибка по векторкардиографическим соотношениям (коэффициент 1.0).

SSSD-трансформер (Structured State Space Diffusion Transformer) – центральный компонент обратного диффузионного процесса, отвечающий за предсказание шума $\hat{\varepsilon}_0(x_t, t, e)$ и восстановление чистого сигнала x_0 из зашумленной версии x_t . Его архитектура сочетает механизмы трансформера с моделями пространства состояний (SSM) [11], обеспечивая линейную сложность $O(T)$ и эффективное моделирование дальних зависимостей в 12-ка-

нальных последовательностях длиной 5 000 отсчетов. Прямой процесс задан как

$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \varepsilon, \varepsilon \sim N(0, I),$$

где $\bar{\alpha}_t = \prod_{s=1}^t (1 - \beta_s)$, а β_t следует cosine-расписанию

в диапазоне от 10^{-4} до 0,020.

Обратный процесс использует предсказание шума через SSSD-трансформер (4 блока, ~5 млн параметров, скрытая размерность 768). Вход – конкатенация зашумленного сигнала x_t , эмбединга шага t и условия $e \in \mathbb{R}^{128}$. Каждый блок включает:

- Multi-Head Self-Attention (8 голов);
- гибридный SSM-слой (S4/HiPPO + Mamba с входозависимыми параметрами);
- Cross-Attention;
- Feed-Forward с GELU.

Одношаговое восстановление:

$$\hat{x}_0 = \frac{x_t - \sqrt{1 - \bar{\alpha}_t} \hat{\varepsilon}_0}{\sqrt{\bar{\alpha}_t}}, \text{ среднее } \mu_0 \text{ вычисляется по стандартной формуле DDPM.}$$

Коррекция градиентом $\nabla_{\varepsilon} L_{\text{phys}}$ (QT, QRS, VCG) обеспечивает физиологическую согласованность. Итеративное восстановление – UniPC второго порядка (200 шагов) [12]. SSSD-трансформер гарантирует линейную сложность, устойчивость к длинным последовательностям и эффективную условную генерацию.

Обучение проводилось на массивах PTB-XL v1.0.3 (21,837 записей, 500 Гц, 10 с, folds 1–8 train, 9 val, 10 test) и MIT-BIH (48 записей, 360 Гц ресэм-

плированы до 500 Гц, протоколы межпациентский сплит DS1 train, DS2 test) [13, 14]. Предобработка: Butterworth bandpass 0,5–150 Гц, z-нормализация, клиппинг $\pm 5\sigma$. Межпациентская стратификация с балансом по возрасту/полу/классам (χ^2 test, $p > 0,05$); фаза 1 (50 эпох) – предобучение VAE+GNN (Adam, lr = 5e-4, batch 128); фаза 2 (300 эпох) – совместное обучение (AdamW, lr = 1e-4 → 1e-6 с cosine annealing, batch 64, gradient clipping norm 1.0, FP16). Аугментация обучающей выборки: гауссов шум $\sigma = 0,01$, масштабирование HR $\pm 10\%$. Процедура ранней остановки (Early stopping) на валидации по метрике FID.

Биофизика:

- QT-интервал через алгоритм Pan-Tompkins, норма 350–450 мс;
- корреляция Пирсона между I и II (ожидается $> 0,85$);
- MSE(II-I-III) (ожидается $< 0,01$).

Классификация: 1D-CNN+BiLSTM классификатор (3 Conv1D слоя, 2 LSTM слоя, функция потерь Focal Loss, 50 эпох) на MIT-BIH. Аугментация: для редких классов (V, F) генерируется синтетика до 30 % от класса (минимум 5 на пациента). Метрики точности классификации: macro F1, precision/recall, ROC AUC.

Результаты и обсуждение

Представленная гибридная модель VAE-GNN-SSSD продемонстрировала значительные улучшения в генерации физиологически корректных 12-канальных ЭКГ-сигналов по сравнению с существующими методами. Результаты приведены в табл. 1.

Таблица 1

Table 1

Сравнительный анализ результатов генерации на тестовой выборке PTB-XL test

Comparative analysis of generation results on the PTB-XL test sample

Модель	Метрики качества генерации			
	FID	PRD, %	Корреляция I–II	MSE _{bio} (II–I–III)
DiffECG [15]	0,058	14,2	0,82	0,089
BioDiffusion [16]	0,051	11,0	0,88	0,074
VAE-GNN-SSSD	0,052	10,8	0,89	0,084

Эмпирическая оценка модели LDM-ECG проводилась на тестовой выборке PTB-XL ($n = 4\ 367$ записей, 20 % от датасета, стратифицировано по пациентам). Все метрики вычислялись по 5 независимым запускам с различными инициализациями; значимость различий подтверждена парным t -тестом Уилкоксона ($p < 0,001$, поправка Бонферрони).

Сравнительный анализ результатов генерации ЭКГ показал, что предложенная модель сопоставима с передовыми моделями DiffECG [15], BioDiffusion [16] по ключевым метрикам: FID (0,052 про-

тив 0,058 и 0,051 соответственно) и PRD (10,8 % против 14,2 и 11,0 %). Ключевым достижением является интеграция биофизических ограничений на уровне архитектуры модели, что позволяет генерировать синтетические сигналы, удовлетворяющие законам Эйнтховена и другим физиологическим закономерностям. MSE_{bio} – среднеквадратичная ошибка по биофизическим (векторкардиографическим) соотношениям между отведениями ЭКГ, в первую очередь проверяющая выполнение законов Эйнтховена и Вильсона

$$MSE_{bio} = \frac{1}{T} \sum_{t=1}^T (\Pi(t) - I(t) - \text{III}(t))^2,$$

где $I(t)$, $\Pi(t)$, $\text{III}(t)$ – значения синтетической ЭКГ в отведениях I, II, III в момент времени t ; $T = 5\,000$ – длина записи (10 с при 500 Гц).

Важность использования сгенерированных ЭКГ оценивалась в задаче классификации аритмий на тестовой выборке MIT-BIH DS2 ($n = 24$ записи, межпа-

циентский сплит). Тренировочный набор (DS1) был дополнен 20–30 % искусственно сгенерированных образцов для устранения дисбаланса классов. В качестве базового классификатора используется модель 1D-CNN + BiLSTM, предобученная на наборе PTB-XL. Оценка – по пяти независимым запускам; значимость – парный t -тест Уилкоксона с поправкой Бонферрони ($\alpha = 0,002$). Результаты представлены в табл. 2.

Таблица 2

Table 2

MIT-BIH DS2 (5 запусков, среднее \pm std)

MIT-BIH DS2 (5 runs, average \pm std)

Класс аритмии	Значения F1-меры			
	F1 (основная модель)	с аугментацией F1	$\Delta F1$, %	p -value
N	$0,94 \pm 0,02$	$0,96 \pm 0,01$	+2	<0,001
S	$0,83 \pm 0,03$	$0,91 \pm 0,02$	+6	
V	$0,74 \pm 0,04$	$0,85 \pm 0,02$	+5	
F	$0,72 \pm 0,05$	$0,81 \pm 0,03$	+7	
Q	$0,79 \pm 0,03$	$0,83 \pm 0,02$	+5	
Macro	$0,84 \pm 0,03$	$0,89 \pm 0,02$	+6	

Аугментация обучающих данных синтетическими ЭКГ, сгенерированными моделью VAE-GNN-SSSD, привела к значительному и статистически устойчивому улучшению классификации аритмий на тестовом наборе MIT-BIH DS2 (межпациентский сплит, 5 запусков). Показатель Macro F1 вырос на 6 % (с 0,84 до 0,89, $p < 0,001$).

Ключевой клинический эффект – снижение ложного пропуска желудочковых аритмий: ошибочная классификация V как N упала с 12 до 7 % (–27 %). Улучшение достигнуто без ущерба для основного класса N , где F1 вырос незначительно (+2 %), но стабильно. Результаты демонстрируют, что синтетические ЭКГ обладают высокой физиологической и диагностической релевантностью, особенно для балансировки редких паттернов, и могут эффективно

использоваться для улучшения надежности кардиологических ИИ-систем. Для строгой статистической проверки значимости улучшения метрики F1 при аугментации синтетическими данными по каждому из 5 классов аритмий (N, S, V, F, Q) в задаче классификации на MIT-BIH DS2 применялся Paired t -test с поправкой Бонферрони ($\alpha = 0,01 / 5 = 0,002$).

Проведен абляционный анализ путем последовательного удаления ключевых компонентов с последующим переобучением модели на тренировочной выборке PTB-XL (folds 1–8) и переоценкой на тестовой ($n = 2\,000$ записей). Метрики вычислялись по пяти независимым запускам; значимость оценивалась по парному t -тесту Уилкоксона ($p < 0,0001$). Результаты абляционного анализа представлены в табл. 3.

Таблица 3

Table 3

Вклад компонентов (на тестовой выборке PTB-XL, $n = 2\,000$)

Contribution of components (in the PTB-XL test sample, $n = 2\,000$)

Конфигурация	Метрики качества генерации электрокардиограмм			
	FID	PRD	MSE_{bio}	p -value
Полная	$0,052 \pm 0,004$	$10,8 \pm 0,5$	$0,084 \pm 0,003$	–
Без VAE	$0,068 \pm 0,005$	$13,2 \pm 0,8$	$0,096 \pm 0,003$	<0,0001
Без GNN	$0,152 \pm 0,007$	$14,5 \pm 0,7$	$0,115 \pm 0,005$	
Без SSSD	$0,067 \pm 0,005$	$13,7 \pm 0,6$	$0,10 \pm 0,003$	
Без priоров	$0,071 \pm 0,006$	$13,8 \pm 0,7$	$0,19 \pm 0,006$	
GNN (A обучаемая)	$0,086 \pm 0,007$	$13,2 \pm 0,6$	$0,25 \pm 0,005$	<0,001
GNN (A фиксированная)	$0,075 \pm 0,004$	$12,7 \pm 0,8$	$0,22 \pm 0,005$	<0,0001

Удаление GNN приводит к росту MSE_{bio} (с 0,084 до 0,25), что прямо подтверждает его критическую роль в соблюдении законов Эйнтховена. Замена гибридной матрицы смежности на чисто обучаемую или чисто фиксированную ухудшает все метрики, что доказывает необходимость сочетания биофизики и адаптивности. Отказ от SSSD ухудшил метрики FID и PRD, что говорит о важности моделирования долгосрочных временных зависимостей в 10-секундных записях. Потеря VAE привела к умеренному ухудшению, что согласуется с его ролью как «морфологического энкодера» – в отличие от GNN, он не отвечает напрямую за биофизику.

Для оценки обобщающей способности модели использован независимый набор MIMIC-IV-ECG ($n = 2\ 147$ записей из отделений интенсивной терапии, длительность 10 с, частота дискретизации 500 Гц, $SNR \approx 15$ дБ) [17]. По сравнению с валидационной выборкой PTB-XL наблюдается ожидаемая деградация метрик:

- FID = $0,068 \pm 0,006$;
- PRD = $16 \pm 0,8$ %;
- SE = $2,0 \pm 0,2$;

– доля сгенерированных сигналов с физиологически нормальным QT-интервалом (350–450 мс) составила 89,4 %, что на 3,7 % ниже показателя полной модели.

Ухудшение обусловлено смещением области (ICU-сигналы существенно более зашумленные, чем амбулаторные записи PTB-XL). Тем не менее полученные результаты подтверждают хорошую обобщаемость модели на разнородные клинические данные и ее готовность к применению в реальных условиях эксплуатации.

Несмотря на достигнутые результаты, у модели есть ряд ограничений. Во-первых, она не учитывает некоторые редкие патологии, которые встречаются в клинической практике. Во-вторых, генерация длительных записей (более 10 с) может приводить к накоплению артефактов из-за ошибок в процессе диффузии. В-третьих, текущая реализация требует значительных вычислительных ресурсов для генерации (GPU с 24 ГБ памяти).

Результаты, полученные в данной работе, имеют важное практическое значение для медицинских ис-

следований. Синтетические данные, сгенерированные разработанной моделью, могут быть использованы:

- для обучения моделей диагностики редких аритмий;
- улучшения обобщающей способности существующих алгоритмов;
- снижения зависимости от дефицита аннотированных данных;
- обучения новых специалистов в кардиологии.

Заключение

В работе представлена гибридная модель VAE-GNN-SSSD для генерации физиологически корректных 12-канальных ЭКГ-сигналов длительностью 10 с. Предложенная архитектура успешно интегрирует вариационный автокодировщик для выделения морфологических компонент, графовую нейронную сеть с частично фиксированной матрицей смежности для соблюдения биофизических законов Эйнтховена и Вильсона и диффузионную модель со структурированным пространством состояний для моделирования долгосрочных временных зависимостей. На тестовой выборке PTB-XL модель достигла значений метрик FID = 0,052 и PRD = 10,8 %, продемонстрировав сопоставимое качество с передовыми аналогами при обеспечении уникального преимущества – встроенной биофизической корректности (MSE по закону Эйнтховена = 0,084).

Практическая значимость работы подтверждена успешным применением сгенерированных данных для аугментации обучающих выборок в задаче классификации аритмий на MIT-BIH. Аугментация позволила повысить Macro F1 на 6 % (с 0,84 до 0,89), значительно улучшить распознавание редких и клинически критичных классов (+5–7 % для желудочковых и фузурованных сокращений) и снизить ложный пропуск опасных аритмий на 27 %. Модель также продемонстрировала высокую обобщающую способность на независимых ICU-данных (MIMIC-IV-ECG), что подтверждает ее готовность к применению в реальных клинических условиях для решения фундаментальной проблемы дефицита аннотированных данных при сохранении приватности пациентов.

Список источников

1. World health statistics 2024 // World Health Organization. URL: <https://www.who.int/data/gho/publications/world-health-statistics> (дата обращения: 17.09.2025).
2. Щетинин Е. Ю. Современные проблемы цифровой обработки и анализа больших данных в медицине и биологии: моногр. М.: Изд. дом «Научная библиотека», 2023. 168 с.
3. Щетинин Е. Ю., Пестрякова А. В. О моделировании электрокардиограмм с корреляцией отведений методами глубокого обучения // Мягкие измерения и вычисления. 2025. № 10. Т. 95. С. 74–87. <https://www.elibrary.ru/item.asp?id=83225498>.
4. Naren Wulan, Wei Wang, Pengzhong Sun, Kuanquan Wang, Yong Xia, Henggui Zhang. Generating electrocardiogram signals by deep learning // Neurocomputing. 2020. V. 404. P. 122–136. <https://doi.org/10.1016/j.neucom.2020.04.076>.
5. Berger L., Haberbush M., Moscato F. Generative adversarial networks in electrocardiogram synthesis: Recent developments and challenges // Artif. Intell. Med. 2023. V. 143. P. 102632. DOI 10.1016/j.artmed.2023.102632.
6. Haradal S., Hayashi H., Uchida S. Biosignal Data Augmentation Based on Generative Adversarial Networks //

Ann. Int. Conf. IEEE Eng. Med. Biol. Soc. 2018. Jul. P. 368–371. DOI 10.1109/EMBC.2018.8512396.

7. Kuznetsov V. V., Moskalenko V. A., Zolotykh N. Y. Electrocardiogram Generation and Feature Extraction Using a Variational Autoencoder // arXiv preprint arXiv. 2020. V. 2002. P. 00240.

8. Berger L., Habermusch M., Moscato F. Generative adversarial networks in electrocardiogram synthesis: Recent developments and challenges // *Artificial Intelligence in Medicine*. 2023. V. 143. P. 102632. <https://doi.org/10.1016/j.artmed.2023.102632>.

9. Jiang Y. Research on Denoising Diffusion Probabilistic Models // *Highlights in Science, Engineering and Technology*. 2024. V. 107. P. 560–572. doi.org/10.54097/sxd49274.

10. Hamilton W. L., Ying R., Leskovec J. Inductive Representation Learning on Large Graphs // *Advances in Neural Information Processing Systems*. 2017. V. 30. P. 1024–1034. DOI 10.48550/arXiv.1706.02216.

11. Gu A., Goel K., Ré C. Efficiently Modeling Long Sequences with Structured State Spaces // *International Conference on Learning Representations (ICLR)*. 2022. URL: <https://openreview.net/pdf?id=uYLFoz1vIAC> (дата обращения: 09.09.2025).

12. Zhao W., Bai L., Rao Y., Zhou J., Lu J. UniPC: A Unified Predictor-Corrector Framework for Fast Sampling of Diffusion Models // *Advances in Neural Information Pro-*

cessing Systems. 2023. URL: <https://arxiv.org/abs/2302.04867> (дата обращения: 09.09.2025).

13. Wagner P., Strodthoff N., Bousseljot R. D., Kreiseler D., Lunze F. I., Samek W., Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset // *Sci Data*. 2020. May 25. V. 7 (1). P. 154. DOI 10.1038/s41597-020-0495-6.

14. Moody G. B., Mark R. G. The impact of the MIT-BIH Arrhythmia Database // *IEEE Eng. Med. Biol. Mag.* 2001. V. 20 (3). P. 45–50.

15. Neifar N., Ben-Hamadou A., Mdhaffar A., Jmaiel M. DiffECG: A Versatile Probabilistic Diffusion Model for ECG Signals Synthesis // *2024 IEEE/ACIS 22nd International Conference on Software Engineering Research, Management and Applications (SERA)*. 2024. P. 182–188. DOI 10.1109/SERA61261.2024.10685651.

16. Li X., Sakevych M., Atkinson G., Metsis V. Bio-Diffusion: A Versatile Diffusion Model for Biomedical Signal Synthesis // *Bioengineering*. 2024. V. 11. P. 299. <https://doi.org/10.3390/bioengineering11040299>.

17. Johnson A. E. W., Bulgarelli L., Shen L., Gayles A., Shammout A., Horng S., Pollard T. J., Hao S., Moody B., Gow B., Lehman L.-W. H., Celi L. A., Mark R. G. MIMIC-IV, a freely accessible electronic health record dataset // *Scientific Data*. 2023. V. 10. N. 1. Art. 1. DOI 10.1038/s41597-022-01899-x.

References

1. *World health statistics 2024*. World Health Organization. Available at: <https://www.who.int/data/gho/publications/world-health-statistics> (accessed: 17.09.2025).

2. Shchetinin E. Yu. *Sovremennyye problemy cifrovoj obrabotki i analiza bol'shikh dannykh v medicine i biologii: monografiya* [Modern problems of digital processing and analysis of big data in medicine and biology: monograph]. Moscow, Izd. dom «Nauchnaya biblioteka», 2023. 168 p.

3. Shchetinin E. Yu., Pestryakova A. V. O modelirovaniy elektrokardiogramm s korrelyatsionnyimi metodami glubokogo obucheniya [On modeling electrocardiograms with lead correlation using deep learning methods]. *Myagkie izmereniya i vychisleniya*, 2025, no. 10, vol. 95, pp. 74-87. <https://www.elibrary.ru/item.asp?id=83225498>.

4. Naren Wulan, Wei Wang, Pengzhong Sun, Kuanquan Wang, Yong Xia, Henggui Zhang. Generating electrocardiogram signals by deep learning. *Neurocomputing*, 2020, vol. 404, pp. 122-136. <https://doi.org/10.1016/j.neucom.2020.04.076>.

5. Berger L., Habermusch M., Moscato F. Generative adversarial networks in electrocardiogram synthesis: Recent developments and challenges. *Artif. Intell. Med.*, 2023, vol. 143, p. 102632. DOI 10.1016/j.artmed.2023.102632.

6. Haradal S., Hayashi H., Uchida S. Biosignal Data Augmentation Based on Generative Adversarial Networks. *Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2018. Jul. Pp. 368-371. DOI 10.1109/EMBC.2018.8512396.

7. Kuznetsov V. V., Moskalenko V. A., Zolotykh N. Y. Electrocardiogram Generation and Feature Extraction Using a Variational Autoencoder. *arXiv preprint arXiv*, 2020, vol. 2002, p. 00240.

8. Berger L., Habermusch M., Moscato F. Generative adversarial networks in electrocardiogram synthesis: Recent developments and challenges. *Artificial Intelligence in Medicine*, 2023, vol. 143, p. 102632. <https://doi.org/10.1016/j.artmed.2023.102632>.

9. Jiang Y. Research on Denoising Diffusion Probabilistic Models. *Highlights in Science, Engineering and Technology*, 2024, vol. 107, pp. 560-572. doi.org/10.54097/sxd49274.

10. Hamilton W. L., Ying R., Leskovec J. Inductive Representation Learning on Large Graphs. *Advances in Neural Information Processing Systems*, 2017, vol. 30, pp. 1024-1034. DOI 10.48550/arXiv.1706.02216.

11. Gu A., Goel K., Ré C. Efficiently Modeling Long Sequences with Structured State Spaces. *International Conference on Learning Representations (ICLR)*, 2022. Available at: <https://openreview.net/pdf?id=uYLFoz1vIAC> (accessed: 09.09.2025).

12. Zhao W., Bai L., Rao Y., Zhou J., Lu J. UniPC: A Unified Predictor-Corrector Framework for Fast Sampling of Diffusion Models. *Advances in Neural Information Processing Systems*, 2023. Available at: <https://arxiv.org/abs/2302.04867> (accessed: 09.09.2025).

13. Wagner P., Strodthoff N., Bousseljot R. D., Kreiseler D., Lunze F. I., Samek W., Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset. *Sci Data*, 2020, May 25, vol. 7 (1), p. 154. DOI 10.1038/s41597-020-0495-6.

14. Moody G. B., Mark R. G. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng. Med. Biol. Mag.*, 2001, vol. 20 (3), pp. 45-50.

15. Neifar N., Ben-Hamadou A., Mdhaffar A., Jmaiel M. DiffECG: A Versatile Probabilistic Diffusion Model for ECG Signals Synthesis. *2024 IEEE/ACIS 22nd International Conference on Software Engineering Research, Management and Applications (SERA)*, 2024. Pp. 182-188. DOI 10.1109/SERA61261.2024.10685651.

16. Li X., Sakevych M., Atkinson G., Metsis V. Bio-Diffusion: A Versatile Diffusion Model for Biomedical Signal Synthesis. *Bioengineering*, 2024, vol. 11, p. 299. <https://doi.org/10.3390/bioengineering11040299>.

17. Johnson A. E. W., Bulgarelli L., Shen L., Gayles A., Shammout A., Horng S., Pollard T. J., Hao S., Moody B., Gow B., Lehman L.-W. H., Celi L. A., Mark R. G. MIMIC-IV, a freely accessible electronic health record dataset. *Scientific Data*, 2023, vol. 10, no. 1, art. 1. DOI 10.1038/s41597-022-01899-x.

Статья поступила в редакцию 19.11.2025; одобрена после рецензирования 25.12.2025; принята к публикации 31.03.2026
The article was submitted 19.11.2025; approved after reviewing 25.12.2025; accepted for publication 31.03.2026

Информация об авторах / Information about the authors

Евгений Юрьевич Щетинин – доктор физико-математических наук, доцент; профессор кафедры информационных технологий и систем; Севастопольский государственный университет; riviera-molto@mail.ru

Evgeny Yu. Shchetinin – Doctor of Physico-Mathematical Sciences, Assistant Professor; Professor of the Department of Information Technologies and Systems; Sevastopol State University; riviera-molto@mail.ru

Анна Вячеславовна Пестрякова – ассистент кафедры информационных технологий и систем; Севастопольский государственный университет; pestryakova@mail.sevsu.ru

Anna V. Pestryakova – Lecturer of the Department of Information Technologies and Systems; Sevastopol State University; pestryakova@mail.sevsu.ru

Юлия Георгиевна Шаталова – кандидат технических наук, доцент; доцент кафедры информационных технологий и систем; Севастопольский государственный университет; bastion2417@mail.ru

Julia G. Shatalova – Candidate of Technical Sciences, Assistant Professor; Assistant Professor of the Department of Information Technologies and Systems; Sevastopol State University; bastion2417@mail.ru

Андрей Андреевич Шевчук – аспирант кафедры информационных технологий и систем; Севастопольский государственный университет; andreiluck11@yandex.ru

Andrey A. Shevchuk – Postgraduate Student of the Department of Information Technologies and Systems; Sevastopol State University; andreiluck11@yandex.ru

