

УДК 004.81

УСЛОВНАЯ ГЕНЕРАЦИЯ ЭЛЕКТРОКАРДИОГРАММ С ПОМОЩЬЮ ИЕРАРХИЧЕСКИХ ВАРИАЦИОННЫХ АВТОКОДИРОВЩИКОВ

И. А. Свиридов¹ [0009-0009-5912-1118], К. С. Егоров² [0009-0006-6991-4136]

^{1, 2}Sber AI Lab, г. Москва, Россия

¹ianatosviridov@sberbank.ru, ²Egorov.K.Ser@sberbank.ru

Аннотация

Сердечно-сосудистые заболевания являются одной из основных причин смертности. Автоматический анализ электрокардиограмм (ЭКГ) может существенно облегчить работу врачей, но его эффективность ограничена нехваткой и несбалансированностью данных. Создание синтетических ЭКГ помогает частично решить эти проблемы. Хотя чаще всего для этого применяются генеративно-состязательные сети (GAN), но последние исследования показали, что вариационные автокодировщики (VAE) могут обеспечивать сопоставимое качество.

В работе представлена модель сNVAE-ECG — модификация Nouveau VAE (NVAE), способная генерировать 12 отведений 10-секундных ЭКГ с различными патологиями. Используя компактную схему работы с каналами и встроенные представления классов для условной генерации, сNVAE-ECG улучшает результаты в задачах бинарной и multi-label классификации, обеспечивая прирост метрики AUROC до 2% по сравнению с моделями на основе GAN. Модель представлена в открытом доступе: https://github.com/univanxx/cNVAE_ECG.

Ключевые слова: ЭКГ, вариационный автокодировщик, условная генерация, GAN.

ВВЕДЕНИЕ

Сердечно-сосудистые заболевания (ССЗ) по-прежнему занимают первое место среди причин смертности во всем мире, и анализ ЭКГ играет ключевую роль в их диагностике и профилактике [1]. Современные методы глубокого обучения показывают высокий потенциал в автоматической интерпретации ЭКГ [2], однако их применение ограничено несколькими факторами: i) малым объемом доступ-

ных данных из-за требований к конфиденциальности и юридическим ограничениям [3], ii) сильным дисбалансом классов — некоторые патологии встречаются значительно реже [4], iii) наличием шумов и артефактов в записях [5].

Создание синтетических данных представляет собой перспективное направление, которое позволяет увеличить обучающие выборки, выровнять распределение классов и снизить стоимость сбора новых данных — при условии, что сгенерированные ЭКГ реалистично воспроизводят все 12 отведений и могут быть сгенерированы для конкретных патологий [6].

Хотя большинство существующих исследований основано на генеративно-состязательных сетях (GAN), вариационные автокодировщики (VAE) [7], включая архитектуру NVAE [8], ориентированную на изображения, остаются малоизученными в контексте генерации ЭКГ. Мы предлагаем **cNVAE-ECG** — условное расширение NVAE, адаптированное для генерации 12-отведенных ЭКГ. Модель доступна в открытом доступе и демонстрирует превосходство в задачах классификации патологий и обучения с переносом (transfer learning), достигая прироста AUROC до 2% по сравнению с базовыми моделями GAN.

ОБЗОР ЛИТЕРАТУРЫ

Ранние методы синтеза ЭКГ в основном опирались на GAN. Архитектуры вроде PGAN [9] и ProEGAN-MS [10] генерировали отдельные удары или одноотвенные сигналы, улучшая качество за счет продвинутых функций потерь, поэтапного обучения и физиологических ограничений [11]. Эти подходы создавали реалистичные фрагменты ЭКГ, но не могли моделировать зависимость между отведениями и плохо масштабировались на длинные записи. Позднее были предложены условные GAN, где генерация выполнялась с учетом класса заболевания [12]. Однако и здесь оставались трудности при создании длительных многоканальных сигналов. Модели WaveGAN* и Pulse2Pulse [13] стали первыми, которые сгенерировали полноценные 12-отведенные 10-секундные ЭКГ, а MLCGAN [14] показала, что условная генерация может повысить качество последующей классификации.

Альтернативой GAN стали диффузионные модели [15, 16], обладающие высокой гибкостью, но требующие значительных вычислительных ресурсов [17]. Вариационные автокодировщики (VAE), напротив, обучаются стабильнее и быстрее [18, 19]. Недавние работы [20–22] подтвердили их потенциал для генерации многоканальных ЭКГ, однако большинство таких моделей не учитывает классы патологий. В ответ на это мы разработали модель сNVAE-ECG, которая сочетает иерархическую структуру NVAE с условной генерацией по классам заболеваний.

КРАТКАЯ СПРАВКА ПО ВАРИАЦИОННЫМ АВТОКОДИРОВЩИКАМ

Для условной генерации стандартных ЭКГ потребовалось разработать архитектуру на базе NVAE, способную работать с одномерными 12-канальными сигналами и генерировать их в соответствии с заданным классом.

Вариационные автокодировщики (VAE) — это стохастические архитектуры, использующие вариационный вывод для аппроксимации распределения данных $p_\theta(x)$, параметризованного через θ , и порождения выборок из латентного представления $p_\theta(z)$. Они состоят из двух блоков: энкодера, отображающего входы в сжатое представление $p_\theta(z|x)$, и декодера, восстанавливающего $p_\theta(x|z)$ после сэмплирования из $p_\theta(z)$.

Недостаток VAE в виде более низкого качества генерации по сравнению с GAN можно смягчить продвинутыми приемами вариационного вывода, например иерархической генерацией. В этом подходе латентные переменные разделяются на группы $z = \{z_1, z_2, \dots, z_L\}$ с априорным распределением $p(z) = \prod_l p(z_l|z_{l-1})$ и аппроксимирующим апостериорным распределением $q(z|x) = \prod_l (z_l|z_{l-1}, x)$, где каждый условный фактор моделируется факторизованным нормальным распределением.

Однако иерархические VAE часто страдают от неустойчивости: исчезающие слои, «взрывающиеся» градиенты (exploding gradients). NVAE решает эти вопросы с помощью «снизу вверх» энкодера, «сверху вниз» декодера и техник стабилизации, включая дискретизированные смеси логистических распределений [23] и учет межканальных зависимостей, достигая передовых результатов на MNIST, CIFAR-10 и CelebA [24–27]. Опираясь на эти усовершенствования, мы адаптируем NVAE к одномерным многоканальным сигналам ЭКГ и вводим сNVAE-ECG для условной генерации.

ПРЕДЛОЖЕННЫЙ ПОДХОД сNVAE-ECG

Ниже описана архитектура сNVAE-ECG, представленная на рис. 1.

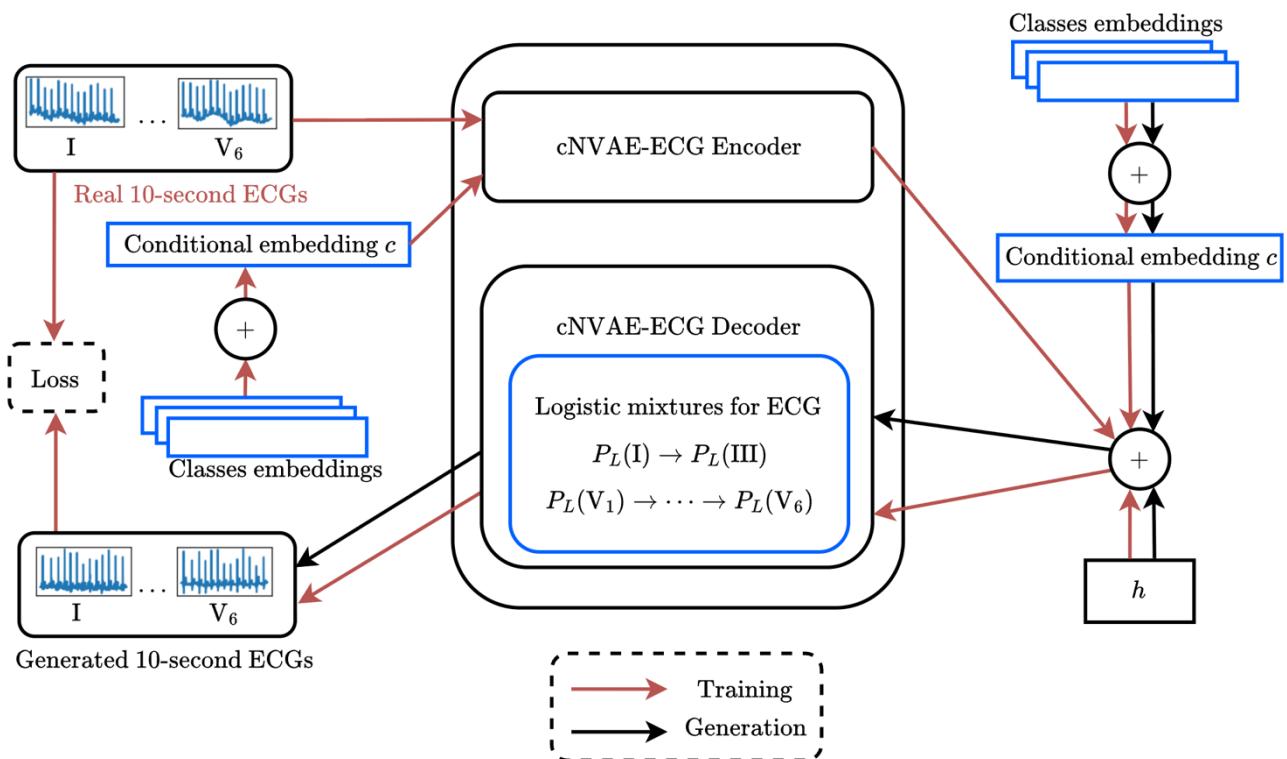


Рис. 1. Архитектура предложенного подхода сNVAE-ECG

1. Трансформация каналов (отведений)

Одной из ключевых особенностей стандартного представления через 12 основных отведений – 6 отведений конечностей (*I, II, III, aVR, aVL, aVF*) и 6 отведений грудной клетки (*V₁, ..., V₆*) – является взаимосвязь между отведениями конечностей через закон Эйнховена [28]

$$I + III = II$$

и уравнения Гольдбергера [28]

$$aVL = \frac{I - III}{2}, \quad -aVR = \frac{I + II}{2}, \quad aVF = \frac{II + III}{2}$$

Таким образом, используя эти уравнения, генерацию шести отведений конечностей *I, II, III, aVR, aVL* можно заменить генерацией только двух – *I* и *III*. В результате в процессе работы модели формируется восемь отведений, которые

затем преобразуются в полный набор из двенадцати. Для генерации этих отведений была определена процедура работы со смесями распределений и формирования каналов, описанная в предыдущем разделе.

Изначально каждое логистическое распределение $P_L(m_1), \dots, P_L(m_K)$ в смеси возвращает выходные значения методом обратного сэмплирования:

$$P_L(m_i) = \mu_{m_i} + s_{m_i} \log\left(\frac{u}{1-u}\right),$$

где μ_{m_i} и s_{m_i} являются параметрами m_i^{th} логистического распределения, $u \sim \text{Uniform}(0,1)$.

Для этого мы предлагаем алгоритм построения смеси распределений для работы с восемью одномерными ЭКГ-сигналами, точнее — переопределения способа выбора параметров распределений. Связь между каналами (отведениями) определяется следующим образом.

- 1) Первое отведение I генерируется как $P_L(C_{I_i}) = P_L(\mu_{I_i}(C_{I_i}), s_{I_i}(C_{I_i}))$, где C_{I_i} — контекстный тензор, получаемый из смеси логистических распределений для отведения I .
- 2) После этого отведение III генерируется с параметрами $\mu_{III_i}(C_{III_i}, I_i) = \mu_{III_i}(C_{III_i}) + \beta \mu_{I_i}(C_{I_i})$ и $s_{III_i}(C_{III_i}, I_i) = s_{III_i}(C_{III_i}) + \beta s_{I_i}(C_{I_i})$.
- 3) Грудное отведение V_1 затем генерируется независимо от отведений конечностей с параметрами $\mu_{V_{1,i}}(C_{V_{1,i}})$ и $s_{V_{1,i}}(C_{V_{1,i}})$, так как оно отражает информацию о вертикальных плоскостях, тогда как конечностные — о горизонтальных.
- 4) Далее, отведение V_2 генерируется с параметрами $\mu_{V_{2,i}}(C_{V_{2,i}}, V_{1,i}) = \mu_{V_{2,i}}(C_{V_{2,i}}) + \alpha(V_{2,i})\mu_{V_{1,i}}(C_{V_{1,i}})$, $s_{V_{2,i}}(C_{V_{2,i}}, V_{1,i}) = s_{V_{2,i}}(C_{V_{2,i}}) + \alpha(V_{2,i}) \cdot s_{V_{1,i}}(C_{V_{1,i}})$.
- 5) Оставшиеся отведения генерируются согласно параметрам $\mu_{V_{k,i}}(C_{V_{k,i}}, V_{k,i}, V_{k-1,i}, \dots, V_{1,i}) = \mu_{V_{k,i}}(C_{V_{k,i}}) + \sum_{j=1}^{k-1} \alpha(V_{j_k,i}) \mu_{V_{j,i}}(C_{V_{j,i}})$, $s_{V_{k,i}}(C_{V_{k,i}}, V_{k,i}, V_{k-1,i}, \dots, V_{1,i}) = s_{V_{k,i}}(C_{V_{k,i}}) + \sum_{j=1}^{k-1} \alpha(V_{j_k,i}) s_{V_{j,i}}(C_{V_{j,i}})$.

2. Использование одномерных сверток

Так как каждая ЭКГ представляет собой одномерный сигнал длиной 5000 отсчетов (10 с при частоте 500 Гц), мы адаптировали NVAE для работы с одномерными свертками. Дополнительно был протестирован вариант с двухмерным представлением сигнала в виде спектrogramмы STFT [29], однако он показал худшие результаты в последующих задачах. При этом были скорректированы гиперпараметры и операции выравнивания, чтобы корректно обрабатывать сигнал полной длины.

3. Условная генерация

В задаче условной генерации одна модель обучается порождать выборки, соответствующие конкретным классам, вместо того чтобы строить отдельные архитектуры для каждого из них. Такой подход позволяет модели изучать общие закономерности, сохраняя при этом особенности, характерные для отдельных классов [12, 14, 15, 21]. Обычно это реализуется добавлением эмбеддинга класса. В NVAE на верхнем уровне иерархии используется глобальный обучаемый вектор h . В нашей модели сNVAE-ECG этот вектор обогащен эмбеддингом класса c , который в случае многометочной (multi-label) классификации представляет собой сумму эмбеддингов всех присутствующих меток. Вектор c подается как в энкодер, так и в декодер, что позволяет модели формировать глобальное представление пространства ЭКГ, одновременно учитывая признаки конкретных патологий.

ПОСТАНОВКА ЭКСПЕРИМЕНТОВ

1. Задачи и конкуренты

Мы оценивали сNVAE-ECG по решениям двух последующих задач: 1) бинарная классификация сигналов ЭКГ для выявления патологий и 2) многометочная (multi-label) классификация в режиме обучения с переносом (transfer learning) [30, 31]. Решение первой задачи проверяет, улучшает ли добавление сгенерированных сигналов метрики на тестовой выборке, а решение второй — повышает ли предобучение на синтетических ЭКГ способность модели к обобщению. В качестве классификатора использовалась XResNet1d101 [30] — модификация ResNet,

доказавшая эффективность при анализе ЭКГ и сохраняющая информативные признаки при умеренных вычислительных затратах.

В обеих задачах обучающие выборки дополнялись сгенерированными сигналами в разных пропорциях, и измерялось их влияние на качество моделей. В задаче бинарной классификации тестиировалось добавление как всех, так и только патологических сигналов, чтобы оценить влияние на баланс классов. В задаче обучения с переносом модели предварительно обучались на сбалансированных синтетических данных, затем они дообучались на целевых наборах и оценивались по метрике AUROC.

Для сравнения сNVAE-ECG сопоставлялась с условными версиями двух моделей на основе GAN — WaveGAN* и Pulse2Pulse [34]. Эти базовые модели были модифицированы для поддержки эмбеддингов классов. Все модели обучались в течение 134 ч на GPU A100, чтобы обеспечить достоверность и сопоставимость экспериментов, а также учитывать экологическую нагрузку, связанную с обучением глубоких нейросетей [32].

2. Датасеты

Мы использовали данные ЭКГ из набора PhysioNet/CinC Challenge 2021 — PTB-XL [33] (21837 записей), Georgia [34] (10344) и Ningbo [35] (34905). Все записи были стандартизированы до длины 10 с при частоте 500 Гц.

Набор PTB-XL применялся для обучения модели сNVAE-ECG, а также генерации синтетических сигналов для предобучения и выполнения задачи бинарной классификации. Для решения этой задачи мы выбрали инфаркт миокарда (MI), поскольку это один из крупнейших классов в наборе данных PTB-XL, отличающийся от синусового ритма (SR). Датасеты Georgia и Ningbo использовались для экспериментов по обучению с переносом в многометочных задачах.

Во всех наборах данных мы выбрали одни и те же семь патологий средней и редкой встречаемостей: отклонение электрической оси влево (LAD), гипертрофия левого желудочка (LVH), фибрилляция предсердий (AF), синусовая тахикардия (STach), атриовентрикулярная блокада I степени (IAVB), синусовая брадикардия (SB) и аномалии зубца Т (TAb). Все диагнозы были классифицированы по терминологии SNOMED CT [36].

Для удаления выбросов применялась перцентильная фильтрация (2.5–97.5 %). Табл. 1 суммирует обучающую выборку PTB-XL и показывает для каждого используемого класса количество реальных ЭКГ и дополнительно сгенерированных синтетических примеров, созданных сNVAE-ECG для предобучения. Это позволило оценить поведение методов в различных условиях.

Табл. 1. Распределение классов в датасете PTB-XL

Название класса	SR	MI	LAD	TAB	LVH	AF	STach	SB	IAVB
Реальные сигналы	12.243	3.300	3.276	1.556	908	849	502	456	445
Сгенерированные сигналы	0	8.943	8.967	10.687	11.335	11.394	11.741	11.787	11.798

РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

1. Количественные результаты

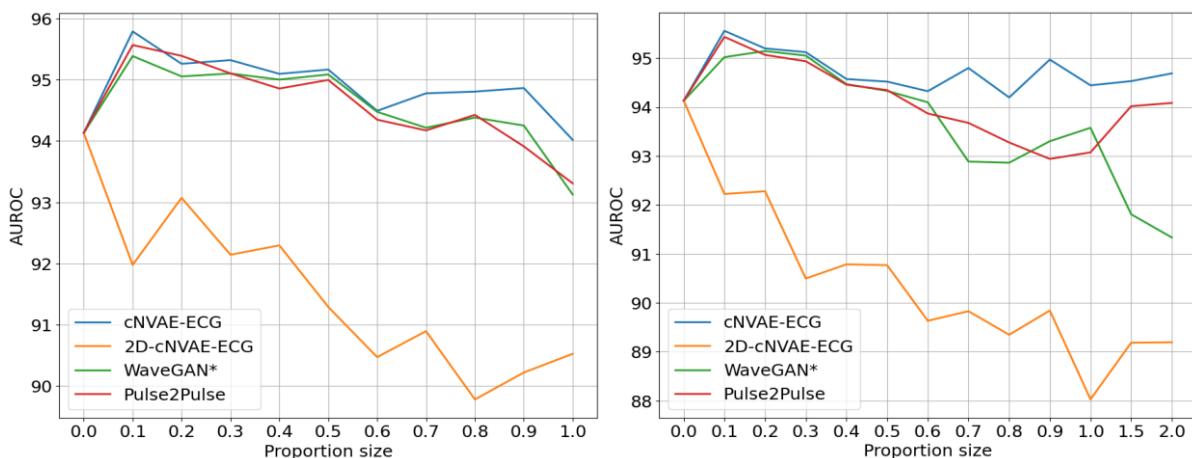


Рис. 2. Результаты экспериментов по задаче обогащения обучающей выборки PTB-XL для бинарной классификации: класса инфаркта миокарда (слева) и обоих классов, включая синусовый ритм (справа)

Левая часть рис. 2 показывает, что при добавлении только сигналов инфаркта миокарда (MI) для снижения дисбаланса классов сNVAE-ECG показывает наилучшие результаты при малых пропорциях, обеспечивая прирост AUROC

до 2%. Напротив, версия 2D-cNVAE-ECG ухудшает результаты. Кроме того, с увеличением доли синтетических данных качество падает, вероятно, из-за ограниченного разнообразия исходных примеров MI, что делает сгенерированные записи менее обобщаемыми при масштабировании.

Как показано на правой части рис. 2, добавление сигналов, сгенерированных моделью cNVAE-ECG для обоих классов, улучшает значение AUROC во всех пропорциях в задаче бинарной классификации. Модель cNVAE-ECG снова превосходит WaveGAN* и Pulse2Pulse при большинстве соотношений, достигая максимального прироста AUROC до 1,5% при долях 0.1–0.3 и 0.9. Эти улучшения показывают, что небольшое добавление синтетических сигналов действует как регуляризатор, а большее количество повышает разнообразие паттернов.

Табл. 2. Значения AUROC (%) на датасете Georgia в зависимости от стратегии предобучения (усреднено по всем пропорциям)

Имя класса	Без предобучения	Реальные данные	Предложенный метод (cNVAE-ECG)	WaveGAN*	Pulse2Pulse
LAD	93.80	94.67	95.23	95.19	93.98
Tab	89.77	92.07	92.40	91.23	88.40
LVH	92.51	97.33	97.99	96.30	94.45
AF	91.26	93.15	93.67	91.44	90.71
STach	98.46	99.43	99.39	98.99	98.37
SB	86.70	88.19	87.99	86.42	83.97
IAVB	91.18	93.81	93.17	92.26	89.73

Из табл. 2 видно, что использование предобучения для датасета Georgia в целом улучшает качество модели по всем классам. Особенно предобучение, обогащенное cNVAE-ECG, демонстрирует наилучшие результаты среди методов, основанных на GAN, превосходя базовое предобучение для четырех самых распространенных (согласно табл. 1) классов в наборе, использованном при обучении cNVAE-ECG. Это наблюдение указывает, что cNVAE-ECG уловила зависимости для

редких классов слабее, чем для более распространенных. Кроме того, в некоторых случаях модели Pulse2Pulse и WaveGAN* показывают худшие результаты, чем без предобучения вовсе.

Табл. 3. Значения AUROC (%) на датасете Ningbo в зависимости от стратегии предобучения (усреднено по всем пропорциям)

Имя класса	Без предобучения	Реальные данные	Предложенный метод (cNVAE-ECG)	WaveGAN*	Pulse2Pulse
LAD	97.79	97.84	98.02	97.73	97.62
TAb	88.86	89.55	89.65	89.59	88.55
LVH	91.58	91.63	92.08	90.73	89.67
STach	98.49	99.02	99.50	99.07	99.05
SB	99.71	99.80	99.79	99.74	99.71
IAVB	96.74	97.56	97.38	97.24	96.38

Табл. 3 показывает, что для большинства классов добавление сигналов, сгенерированных с помощью cNVAE-ECG и моделей, подобных GAN, дает приемлемый прирост качества на тестовом наборе Ningbo по сравнению с результатами без предобучения. Модель cNVAE-ECG снова демонстрирует лучшие результаты по всем классам по сравнению с другими генеративными методами. Однако для редких классов, таких как SB и IAVB, качество оказалось ниже, чем при использовании исходного предобучения. С учетом предыдущих результатов можно заключить, что для этих классов у cNVAE-ECG недостаточно данных, чтобы успешно генерировать соответствующие типы ЭКГ-сигналов.

2. Качественные результаты

На рис. 3 представлены реальные и сгенерированные сигналы ЭКГ для синусового ритма (SR) и инфаркта миокарда (MI). Несмотря на наличие отдельных артефактов, сгенерированные сигналы воспроизводят основные волновые характеристики, наблюдаемые в обучающих данных.

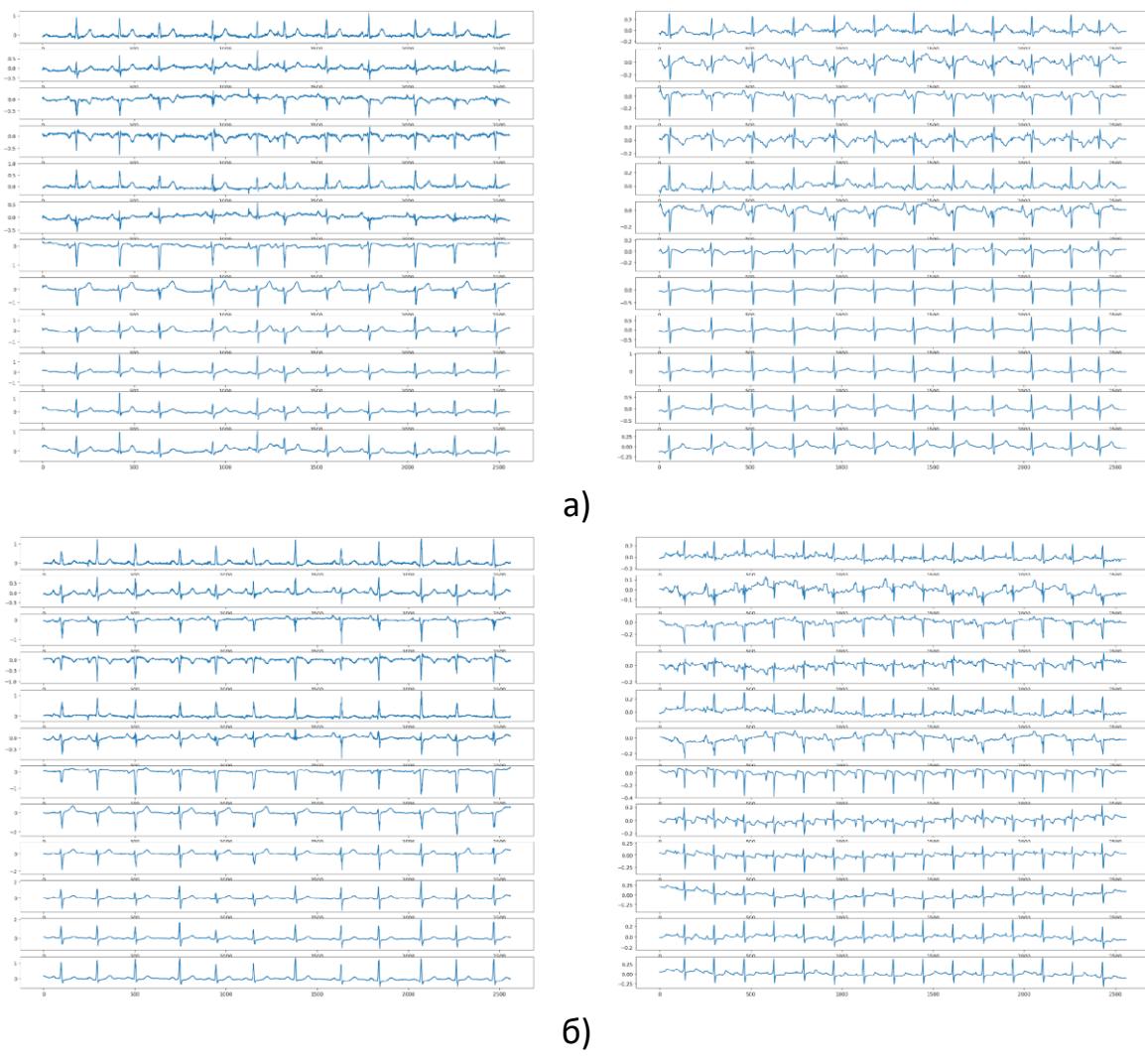


Рис. 3. Результаты генерации ЭКГ-сигналов (в правой части) по сравнению с оригинальными сигналами (в левой части) для классов: а) синусовый ритм, б) инфаркта миокарда

Мы также провели более детальное сравнение одного сердечного цикла для отведения *I* синусового ритма (см. на рис. 4).

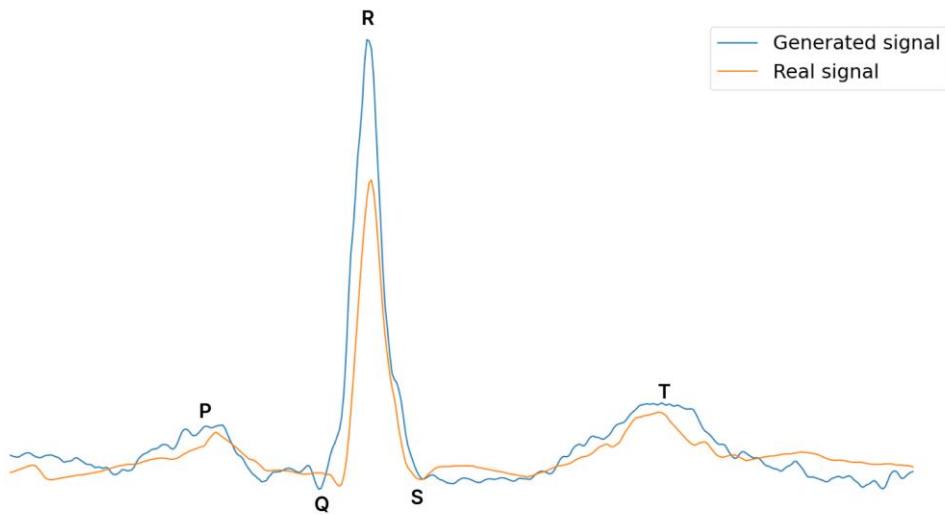


Рис. 4. Сравнение реального и сгенерированного одного сердечного цикла для отведения I в классе синусового ритма

Это сравнение показывает, что модель сNVAE-ECG успешно воспроизводит основную структуру и характерные особенности реального ЭКГ-сигнала, такие как зубцы P, Q, R, S и T [37].

ЗАКЛЮЧЕНИЕ

Мы представили сNVAE-ECG — модификацию NVAE для условной генерации 12-отведенных 10-секундных ЭКГ. Модель продемонстрировала стабильное улучшение показателя AUROC по сравнению с базовыми моделями на основе GAN (WaveGAN*, Pulse2Pulse) как в задаче бинарной классификации, так и в задаче многометочного обучения с переносом. Предобучение с использованием сNVAE-ECG улучшило перенос на датасеты Georgia и Ningbo, особенно для классов средней частоты, хотя редкие патологии остаются сложными для генерации. Сгенерированные ЭКГ в целом сохраняют ключевые волновые характеристики при наличии лишь незначительных артефактов. В дальнейшем планируется провести клиническую валидацию и реализовать модель в федеративной среде для повышения ее обобщающей способности и устойчивости [38].

СПИСОК ЛИТЕРАТУРЫ

1. *Tsao C.W., Aday A.W., Almarzooq Z.I., et al.* Heart Disease and Stroke Statistics—2023 Update: A Report from the American Heart Association // *Circulation*. 2023. Vol. 147, No. 8. P. e93–e621.
2. *Liu X., Wang H., Li Z., Qin L.* Deep learning in ECG diagnosis: A review // *Knowledge-Based Systems*. 2021. Vol. 227. P. 107187.
3. *Gerke S., Minssen T., Cohen G.* Ethical and legal challenges of artificial intelligence-driven healthcare // *Artificial Intelligence in Healthcare* / Eds. A. Bohr, K. Memarzadeh. Academic Press, 2020. P. 295–336.
4. *Reyna M.A., Sadr N., Alday E.A.P., et al.* Will two do? Varying dimensions in electrocardiography: The PhysioNet/Computing in Cardiology Challenge 2021 // *Computing in Cardiology*. 2021. Vol. 48. P. 1–4.
5. *Friesen G., Jannett T., Jadallah M., Yates S., Quint S., Nagle H.* A comparison of the noise sensitivity of nine QRS detection algorithms // *IEEE Transactions on Biomedical Engineering*. 1990. Vol. 37, No. 1. P. 85–98.
6. *Maron B.J., Friedman R.A., Kligfield P., et al.* Assessment of the 12-lead ECG as a screening test for detection of cardiovascular disease in healthy general populations of young people (12–25 years of age) // *Circulation*. 2014. Vol. 130, No. 15. P. 1303–1334.
7. *Kingma D.P., Welling M.* Auto-encoding variational Bayes. 2022.
8. *Vahdat A., Kautz J.* NVAE: A deep hierarchical variational autoencoder. 2020.
9. *Golany T., Radinsky K.* PGANs: Personalized generative adversarial networks for ECG synthesis to improve patient-specific deep ECG classification // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019. Vol. 33, No. 1. P. 557–564.
10. *Yang H., Liu J., Zhang L., Li Y., Zhang H.* ProEGAN-MS: A progressive growing generative adversarial networks for electrocardiogram generation // *IEEE Access*. 2021. Vol. 9. P. 52089–52100.
11. *Golany T., Radinsky K., Freedman D.* SimGANs: Simulator-based generative adversarial networks for ECG synthesis to improve deep ECG classification // *Proceedings of the 37th International Conference on Machine Learning (PMLR)*. 2020. Vol. 119. P. 3597–3606.

12. *Nankani D., Baruah R.D.* Investigating deep convolution conditional GANs for electrocardiogram generation // 2020 International Joint Conference on Neural Networks (IJCNN). 2020. P. 1–8.
 13. *Thambawita V., Isaksen J.L., Hicks S.A., et al.* DeepFake electrocardiograms using generative adversarial networks are the beginning of the end for privacy issues in medicine // *Scientific Reports*. 2021. Vol. 11. P. 21896.
 14. *Wu J., Wang L., Pan H., Wang B.* MLCGAN: Multi-lead ECG synthesis with multi label conditional generative adversarial network // ICASSP 2023 – IEEE International Conference on Acoustics, Speech and Signal Processing. 2023. P. 1–5.
 15. *Alcaraz J.M.L., Strodthoff N.* Diffusion-based conditional ECG generation with structured state space models. 2023.
 16. *Ho J., Jain A., Abbeel P.* Denoising diffusion probabilistic models // *CoRR*. 2020. abs/2006.11239.
 17. *Dhariwal P., Nichol A.* Diffusion models beat GANs on image synthesis. 2021.
 18. *Xia Y., Wang W., Wang K.* ECG signal generation based on conditional generative models // *Biomedical Signal Processing and Control*. 2023. Vol. 82. P. 104587.
 19. *El-Kaddoury M., Mahmoudi A., Himmi M.M.* Deep generative models for image generation: A practical comparison between variational autoencoders and generative adversarial networks // *Mobile, Secure, and Programmable Networking*. Cham: Springer, 2019. P. 1–8.
 20. *Kuznetsov V., Moskalenko V., Gribanov D., Zolotykh N.* Interpretable feature generation in ECG using a variational autoencoder // *Frontiers in Genetics*. 2021. Vol. 12. P. 638191.
 21. *Sang Y., Beetz M., Grau V.* Generation of 12-lead electrocardiogram with subject-specific, image-derived characteristics using a conditional variational autoencoder // 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). 2022. P. 1–5.
 22. *Beetz M., Banerjee A., Sang Y., Grau V.* Combined generation of electrocardiogram and cardiac anatomy models using multi-modal variational autoencoders // 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). 2022. P. 1–4.
 23. *Salimans T., Karpathy A., Chen X., Kingma D.P.* PixelCNN++: Improving the PixelCNN with discretized logistic mixture likelihood and other modifications. 2017.
-

24. *Deng L.* The MNIST database of handwritten digit images for machine learning research [Best of the Web] // IEEE Signal Processing Magazine. 2012. Vol. 29, No. 6. P. 141–142.
 25. *Krizhevsky A.* Learning multiple layers of features from tiny images. 2009.
 26. *Larsen A.B.L., Sønderby S.K., Larochelle H., Winther O.* Autoencoding beyond pixels using a learned similarity metric. 2016.
 27. *Karras T., Aila T., Laine S., Lehtinen J.* Progressive growing of GANs for improved quality, stability, and variation. 2018.
 28. *Malmivuo J., Plonsey R.* Bioelectromagnetism. 15. 12-Lead ECG System. 1975. P. 277–289.
 29. *Griffin D., Lim J.* Signal estimation from modified short-time Fourier transform // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1984. Vol. 32, No. 2. P. 236–243.
 30. *Strothoff 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. P. 1519–1528.
 31. *Donahue C., McAuley J., Puckette M.* Adversarial audio synthesis. 2019.
 32. *Wu C.-J., Raghavendra R., Gupta U., et al.* Sustainable AI: Environmental implications, challenges and opportunities // ArXiv. 2021. abs/2111.00364.
 33. *Wagner P., Strothoff N., Bousseljot R.-D., et al.* PTB-XL, a large publicly available electrocardiography dataset // Scientific Data. 2020. Vol. 7. P. 154.
 34. *Goldberger A., Amaral L., Glass L., et al.* PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals // Circulation. 2000. Vol. 101. P. e215–e220.
 35. *Zheng J., Chu H., Struppa D., et al.* Optimal multi-stage arrhythmia classification approach // Scientific Reports. 2020. Vol. 10.
 36. *El-Sappagh S., Franda F., Ali F., Kwak K.-S.* SNOMED CT standard ontology based on the ontology for general medical science // BMC Medical Informatics and Decision Making. 2018. Vol. 18, No. 1. P. 76.
 37. *Berkaya S.K., Uysal A.K., Gunal E.S., et al.* A survey on ECG analysis // Biomedical Signal Processing and Control. 2018. Vol. 43. P. 216–235.
 38. *Zhang M., Wang Y., Luo T.* Federated learning for arrhythmia detection of non-IID ECG // 2020 IEEE 6th International Conference on Computer and Communications (ICCC). 2020. P. 1176–1180.
-

CONDITIONAL ELECTROCARDIOGRAM GENERATION USING HIERARCHICAL VARIATIONAL AUTOENCODERS

I. A. Sviridov¹ [0009-0009-5912-1118], K. S. Egorov² [0009-0006-6991-4136]

^{1, 2}*Sber AI Lab, Moscow, Russia*

¹ianatosviridov@sberbank.ru, ²Egorov.K.Ser@sberbank.ru

Abstract

Cardiovascular diseases remain the leading cause of mortality, and automated electrocardiogram (ECG) analysis can ease clinical workloads but is limited by scarce and imbalanced data. Synthetic ECG can mitigate these issues, and while most methods use Generative Adversarial Networks (GANs), recent work show variational autoencoders (VAEs) perform comparably. We introduce **cNVAE-ECG**, a conditional Nouveau VAE (NVAE) that generates high-resolution, 12-lead, 10-second ECGs with multiple pathologies. Leveraging a compact channel-generation scheme and class embeddings for multi-label conditioning, cNVAE-ECG improves downstream binary and multi-label classification, achieving up to a 2% AUROC gain in transfer learning over GAN-based models.

Keywords: *ECG, variational autoencoder, conditional generation, GAN.*

REFERENCES

1. *Tsao C.W., Aday A.W., Almarzooq Z.I., et al.* Heart Disease and Stroke Statistics—2023 Update: A Report from the American Heart Association // *Circulation*. 2023. Vol. 147, No. 8. P. e93–e621.
2. *Liu X., Wang H., Li Z., Qin L.* Deep learning in ECG diagnosis: A review // *Knowledge-Based Systems*. 2021. Vol. 227. P. 107187.
3. *Gerke S., Minssen T., Cohen G.* Ethical and legal challenges of artificial intelligence-driven healthcare // *Artificial Intelligence in Healthcare* / Eds. A. Bohr, K. Memarzadeh. Academic Press, 2020. P. 295–336.
4. *Reyna M.A., Sadr N., Alday E.A.P., et al.* Will two do? Varying dimensions in electrocardiography: The PhysioNet/Computing in Cardiology Challenge 2021 // *Computing in Cardiology*. 2021. Vol. 48. P. 1–4.
5. *Friesen G., Jannett T., Jadallah M., Yates S., Quint S., Nagle H.* A comparison of the noise sensitivity of nine QRS detection algorithms // *IEEE Transactions on Biomedical Engineering*. 1990. Vol. 37, No. 1. P. 85–98.
6. *Maron B.J., Friedman R.A., Kligfield P., et al.* Assessment of the 12-lead ECG as a screening test for detection of cardiovascular disease in healthy general populations of young people (12–25 years of age) // *Circulation*. 2014. Vol. 130, No. 15. P. 1303–1334.
7. *Kingma D.P., Welling M.* Auto-encoding variational Bayes. 2022.
8. *Vahdat A., Kautz J.* NVAE: A deep hierarchical variational autoencoder. 2020.
9. *Golany T., Radinsky K.* PGANs: Personalized generative adversarial networks for ECG synthesis to improve patient-specific deep ECG classification // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019. Vol. 33, No. 1. P. 557–564.
10. *Yang H., Liu J., Zhang L., Li Y., Zhang H.* ProEGAN-MS: A progressive growing generative adversarial networks for electrocardiogram generation // *IEEE Access*. 2021. Vol. 9. P. 52089–52100.
11. *Golany T., Radinsky K., Freedman D.* SimGANs: Simulator-based generative adversarial networks for ECG synthesis to improve deep ECG classification // *Proceedings of the 37th International Conference on Machine Learning (PMLR)*. 2020. Vol. 119. P. 3597–3606.

12. *Nankani D., Baruah R.D.* Investigating deep convolution conditional GANs for electrocardiogram generation // 2020 International Joint Conference on Neural Networks (IJCNN). 2020. P. 1–8.
 13. *Thambawita V., Isaksen J.L., Hicks S.A., et al.* DeepFake electrocardiograms using generative adversarial networks are the beginning of the end for privacy issues in medicine // *Scientific Reports*. 2021. Vol. 11. P. 21896.
 14. *Wu J., Wang L., Pan H., Wang B.* MLCGAN: Multi-lead ECG synthesis with multi label conditional generative adversarial network // ICASSP 2023 – IEEE International Conference on Acoustics, Speech and Signal Processing. 2023. P. 1–5.
 15. *Alcaraz J.M.L., Strodthoff N.* Diffusion-based conditional ECG generation with structured state space models. 2023.
 16. *Ho J., Jain A., Abbeel P.* Denoising diffusion probabilistic models // *CoRR*. 2020. abs/2006.11239.
 17. *Dhariwal P., Nichol A.* Diffusion models beat GANs on image synthesis. 2021.
 18. *Xia Y., Wang W., Wang K.* ECG signal generation based on conditional generative models // *Biomedical Signal Processing and Control*. 2023. Vol. 82. P. 104587.
 19. *El-Kaddoury M., Mahmoudi A., Himm M.M.* Deep generative models for image generation: A practical comparison between variational autoencoders and generative adversarial networks // *Mobile, Secure, and Programmable Networking*. Cham: Springer, 2019. P. 1–8.
 20. *Kuznetsov V., Moskalenko V., Gribanov D., Zolotykh N.* Interpretable feature generation in ECG using a variational autoencoder // *Frontiers in Genetics*. 2021. Vol. 12. P. 638191.
 21. *Sang Y., Beetz M., Grau V.* Generation of 12-lead electrocardiogram with subject-specific, image-derived characteristics using a conditional variational autoencoder // 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). 2022. P. 1–5.
 22. *Beetz M., Banerjee A., Sang Y., Grau V.* Combined generation of electrocardiogram and cardiac anatomy models using multi-modal variational autoencoders // 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). 2022. P. 1–4.
 23. *Salimans T., Karpathy A., Chen X., Kingma D.P.* PixelCNN++: Improving the PixelCNN with discretized logistic mixture likelihood and other modifications. 2017.
-

24. *Deng L.* The MNIST database of handwritten digit images for machine learning research [Best of the Web] // IEEE Signal Processing Magazine. 2012. Vol. 29, No. 6. P. 141–142.
 25. *Krizhevsky A.* Learning multiple layers of features from tiny images. 2009.
 26. *Larsen A.B.L., Sønderby S.K., Larochelle H., Winther O.* Autoencoding beyond pixels using a learned similarity metric. 2016.
 27. *Karras T., Aila T., Laine S., Lehtinen J.* Progressive growing of GANs for improved quality, stability, and variation. 2018.
 28. *Malmivuo J., Plonsey R.* Bioelectromagnetism. 15. 12-Lead ECG System. 1975. P. 277–289.
 29. *Griffin D., Lim J.* Signal estimation from modified short-time Fourier transform // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1984. Vol. 32, No. 2. P. 236–243.
 30. *Strothoff 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. P. 1519–1528.
 31. *Donahue C., McAuley J., Puckette M.* Adversarial audio synthesis. 2019.
 32. *Wu C.-J., Raghavendra R., Gupta U., et al.* Sustainable AI: Environmental implications, challenges and opportunities // ArXiv. 2021. abs/2111.00364.
 33. *Wagner P., Strothoff N., Bousseljot R.-D., et al.* PTB-XL, a large publicly available electrocardiography dataset // Scientific Data. 2020. Vol. 7. P. 154.
 34. *Goldberger A., Amaral L., Glass L., et al.* PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals // Circulation. 2000. Vol. 101. P. e215–e220.
 35. *Zheng J., Chu H., Struppa D., et al.* Optimal multi-stage arrhythmia classification approach // Scientific Reports. 2020. Vol. 10.
 36. *El-Sappagh S., Franda F., Ali F., Kwak K.-S.* SNOMED CT standard ontology based on the ontology for general medical science // BMC Medical Informatics and Decision Making. 2018. Vol. 18, No. 1. P. 76.
 37. *Berkaya S.K., Uysal A.K., Gunal E.S., et al.* A survey on ECG analysis // Biomedical Signal Processing and Control. 2018. Vol. 43. P. 216–235.
 38. *Zhang M., Wang Y., Luo T.* Federated learning for arrhythmia detection of non-IID ECG // 2020 IEEE 6th International Conference on Computer and Communications (ICCC). 2020. P. 1176–1180.
-

СВЕДЕНИЯ ОБ АВТОРАХ



СВИРИДОВ Иван Анатольевич – получил степень бакалавра по прикладной математике и информатике в МГТУ им. Н.Э. Баумана (Москва, Россия) в 2021 году и степень магистра по компьютерным наукам в Высшей школе экономики (Москва, Россия) в 2023 году. С 2023 года по настоящее время работает исследователем в Центре практического искусственного интеллекта Sber AI Lab (Москва, Россия), где занимается исследованиями в области искусственного интеллекта в медицине. Области его научных интересов: большие языковые модели, многоагентные системы, генеративное моделирование, биосигналы.

Ivan Anatolevich SVIRIDOV – received a B.S. degree in applied mathematics and informatics from Bauman Moscow State Technical University, Moscow, Russia, in 2021 and an M.S. degree in computer science from the Higher School of Economics, Moscow, Russia, in 2023. From 2023 to the present, he has been a researcher at Sber AI Lab in Moscow, Russia, mainly focused on AI in medicine. His research interests are: large language models, multi-agent systems, generative modeling, and biosignals.

email: ianatosviridov@sberbank.ru

ORCID 0009-0009-5912-1118



ЕГОРОВ Константин Сергеевич – получил степень магистра в Московском государственном техническом университете им. Н.Э. Баумана (Москва, Россия). Свою карьеру начал инженером-электронщиком, специализируясь на слаботочных системах, с 2010 по 2018 год. С 2018 года работает исследователем в Центре практического искусственного интеллекта Sber AI Lab, где основное внимание уделяет медицинским сигналам, таким как ЭКГ, ЭЭГ и PPG.

Konstantin Sergeevich EGOROV – received a master's degree from Bauman Moscow State Technical University, Moscow, Russia. He began his career as an electronics engineer, specializing in low-current systems from 2010 to 2018. Since 2018, he has been a researcher at Sber AI Lab, where his primary focus is on medical signals such as ECG, EEG, and PPG.

email: Egorov.K.Ser@sberbank.ru

ORCID 0009-0006-6991-4136

Материал поступил в редакцию 12 октября 2025 года