

## АНАЛИЗ ЭКГ-СИГНАЛОВ НА ОСНОВЕ МЕТОДОВ КОНТРАСТИВНОГО ОБУЧЕНИЯ

*Кривохижин Даниил Николаевич*

*Магистрант*

*Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия*

*E-mail: daniil.krivohizhin@mail.ru*

*Научный руководитель — Шиликин Алексей Геннадьевич*

Сердечно-сосудистые заболевания — одна из ведущих причин смертности в мире, требующая регулярного мониторинга и точной диагностики на ранних стадиях, в том числе по данным электрокардиографии. Однако ручной анализ тысяч ЭКГ-записей кардиологами может приводить к ошибкам и снижению качества медицинской помощи, особенно в отдалённых регионах, где не хватает специалистов. Развитие вычислительных технологий и машинного обучения позволяет автоматизировать диагностику и постановку первичного диагноза, однако большинство таких подходов требует большого объёма размеченных данных, что увеличивает нагрузку на врачей. Перспективным подходом является контрастивное обучение [1], позволяющее использовать неразмеченные данные, снижая потребность в разметке и затраты на создание синтетических данных.

Одним из наиболее популярных методов в рамках этого подхода является SimCLR [1]. Суть метода заключается в обучении нейросети на неразмеченных данных в два этапа. На первом этапе модель с помощью специальной контрастивной функции потерь (1) и репрезентационной нейросети обучается распознавать сходство и различия между двумя случайно аугментированными версиями одного сигнала, образуя позитивные пары, одновременно отличая их от других сигналов, составляя негативные пары.

Пусть  $z_i$  и  $z_j$  — эмбединги пары примеров. Тогда функция потерь определяется следующим образом:

$$\ell_{i,j} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbf{1}_{[k \neq i]} \exp(\text{sim}(z_i, z_k)/\tau)}. \quad (1)$$

В отличие от классических подходов, SimCLR направлен на создание качественных внутренних представлений или эмбедингов, которые затем используются для дообучения классификатора на меньшем объёме размеченных данных.

Целью работы является изучение возможности применения мето-

да контрастивного обучения для предварительной подготовки нейросетевой модели к решению задачи классификации ЭКГ-сигналов с минимальным использованием размеченных данных.

Для исследований использовались датасеты CODE-15% [2] и Chapman–Shaoxing [3]. В них представлены 45 152 и 345 779 записей ЭКГ продолжительностью, в среднем, по 10 секунд и частотами дискретизации 500 Гц и 400 Гц соответственно. В работе реализован полный цикл подготовки данных, включающий удаление блуждающей нулевой линии, с помощью вейвлет-преобразования Daubechies db6, формирование R-пиковых сегментов с использованием инструментов NeuroKit2 и построение обучающих примеров на основе усреднённых кардиоциклов (УКЦ) по сегментам каждой записи. Также проведен сравнительный анализ следующих подходов:

- Контрастивное предобучение с последующей классификацией на основе архитектуры ResNet1D, работающей напрямую с одномерными сигналами;
- Контрастивное предобучение с последующей классификацией на основе рекуррентной архитектуры LSTM, предназначенной для моделирования временной динамики;
- Классический подход без контрастивного этапа: градиентный бустинг, обучаемый только на размеченной части данных.

Представленное сравнение позволяет лучше оценить вклад контрастивного обучения в итоговую точность классификации на целевой задаче.

Таким образом, контрастивное обучение позволяет эффективно использовать большие объёмы неразмеченных данных, снижая нагрузку на медицинских специалистов и затраты на подготовку данных. Подобный подход открывает широкие перспективы для решения конкретных медицинских задач, таких как классификация сердечных патологий по данным ЭКГ.

### Литература

1. Chen T., Kornblith S., Norouzi M., Hinton G. A Simple Framework for Contrastive Learning of Visual Representations.
2. Ribeiro A. H. et al. CODE-15% dataset. Zenodo.  
DOI: 10.5281/zenodo.4916206.
3. Zheng J., Zhang J. et al. Chapman–Shaoxing ECG database for arrhythmia research. Scientific Data, 2020, vol. 7, art. 48.  
DOI: 10.1038/s41597-020-0386-x.