

ИСПОЛЬЗОВАНИЕ БИБЛИОТЕКИ OPENSIMILE ДЛЯ АНАЛИЗА ГОЛОСОВЫХ СИГНАЛОВ ПАЦИЕНТОВ С БАС

Михневич А.В.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Вашкевич М.И. – доктор техн. наук

Аннотация. Боковой амиотрофический склероз (БАС) – тяжелое нейродегенеративное заболевание, затрагивающее в том числе речевую функцию. Целью настоящей работы является извлечение и анализ акустических признаков речи с помощью библиотеки openSMILE, направленный на выявление различий между голосовыми сигналами здоровых лиц и пациентов с БАС. Для анализа использовались записи из базы Minsk2020_ALS. Проведен статистический анализ с коррекцией множественных сравнений (FDR), по итогам которого были выделены значимые речевые маркеры, потенциально полезные для задач автоматической диагностики.

Ключевые слова: БАС, openSMILE, речевые признаки, акустический анализ, статистика, Minsk2020, FDR-коррекция.

Введение

Боковой амиотрофический склероз (БАС, англ. ALS – Amyotrophic Lateral Sclerosis) – прогрессирующее заболевание, поражающее двигательные нейроны, включая те, что отвечают за артикуляцию речи. У значительной части пациентов наблюдается дизартрия, выражающаяся в изменении тембра, частоты, интонации и других параметров речи. Ранняя диагностика БАС исключительно важна, поскольку позволяет своевременно начать поддерживающую терапию.

Анализ акустических признаков речи в последнее десятилетие получил широкое распространение благодаря развитию алгоритмов цифровой обработки сигналов и машинного обучения. Библиотека openSMILE (Open-Source Media Interpretation by Large feature-space Extraction) является одним из наиболее мощных инструментов в данной области [1].

Целью исследования было продемонстрировать, каким образом можно автоматически извлечь и проанализировать речевые признаки, отличающие здоровых испытуемых от пациентов с БАС, используя openSMILE и данные из речевой базы Minsk2020_ALS [2].

Методы и материалы

В исследовании использовались речевые записи из базы Minsk2020_ALS, включающей записи пациентов с подтвержденным диагнозом БАС (31 диктор) и здоровых (33 диктора). Записи были стандартизированы по длительности и формату (16 кГц, WAV).

Для анализа применялась библиотека openSMILE, с использованием конфигурации ComParE_2016, которая обеспечивает извлечение 6373 акустических признаков. Программа реализована на языке Python и включает следующие ключевые этапы:

1. Извлечение признаков из каждого аудиофайла с использованием библиотеки openSMILE.
2. Сохранение и предварительная обработка данных.
3. Статистический анализ по критерию Манна – Уитни, используемый для оценки различий между двумя независимыми выборками, измеренного количественно.
4. FDR – коррекция множественного сравнения.
5. Выделение значимых признаков по скорректированным p -значениям.
6. Визуализация значимых признаков в виде диаграмм.

Коррекция множественных сравнений проводилась с использованием метода Бенжамини – Хохберга (FDR – False Discovery Rate), который контролирует долю ложноположительных результатов среди всех обнаруженных значимых различий. Это особенно важно при анализе большого количества признаков, где без коррекции высока вероятность ошибочного признания признака значимым. Пороговое значение p было установлено на уровне 0,0001, что позволило отобрать только действительно значимые признаки. Формула метода FDR приведена ниже:

$$P_{(i)} \leq \frac{i}{m} \cdot a,$$

(1)

где $P(i)$ – i -е по величине p -значение; m – общее количество тестов; a – уровень значимости.

Анализ результатов

Всего было проанализировано 6373 признаков. FDR – коррекции отобрано 10 значимых признаков, статистически отличающихся между группами, а именно:

- audSpec_Rfilt_sma[6]_lpc0;
- audSpec_Rfilt_sma[6]_lpc1;
- audSpec_Rfilt_sma[6]_lpc2;
- audSpec_Rfilt_sma[6]_peakRangeAbs;

- audSpec_Rfilt_sma[7]_lpc1;
- audSpec_Rfilt_sma[17]_lpc0;
- audSpec_Rfilt_sma[17]_lpc1;
- audSpec_Rfilt_sma[17]_lpc2;
- mfcc_sma[14]_range;
-

pcm_fftMag_fband1000-4000_sma_lpc1.

По результатам анализа сформированы таблицы с признаками и их ящичковые диаграммы.

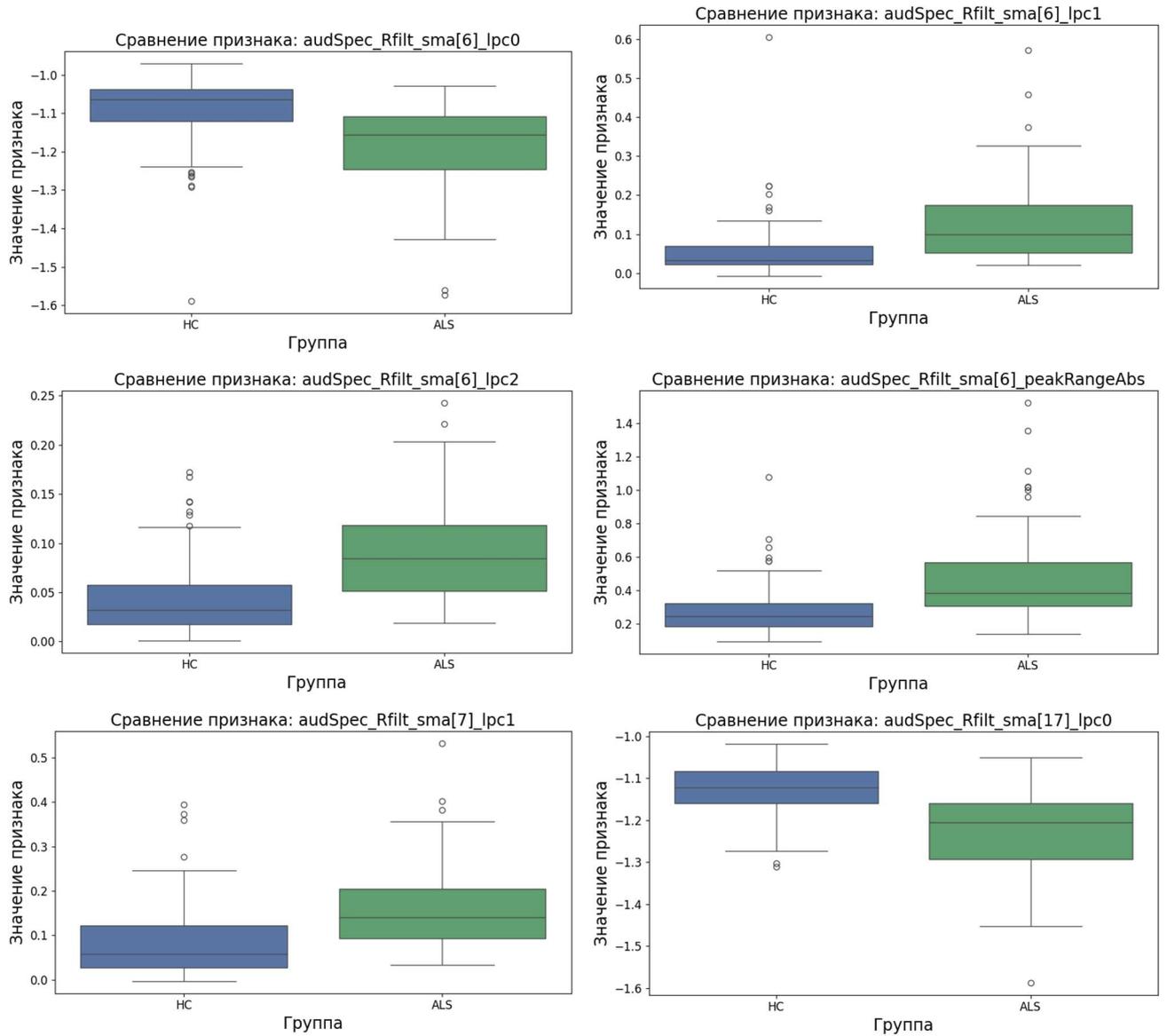
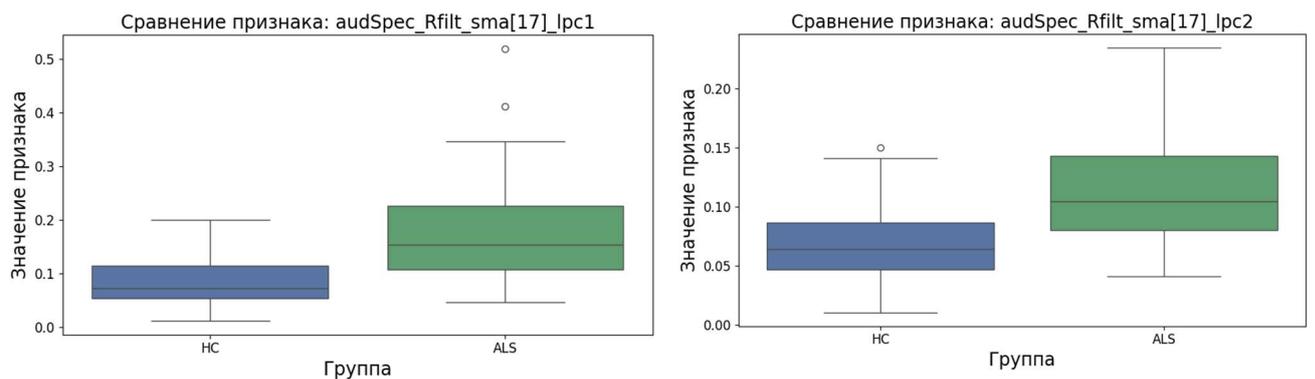


Рисунок 1 – Отобранные признаки после FDR – коррекции



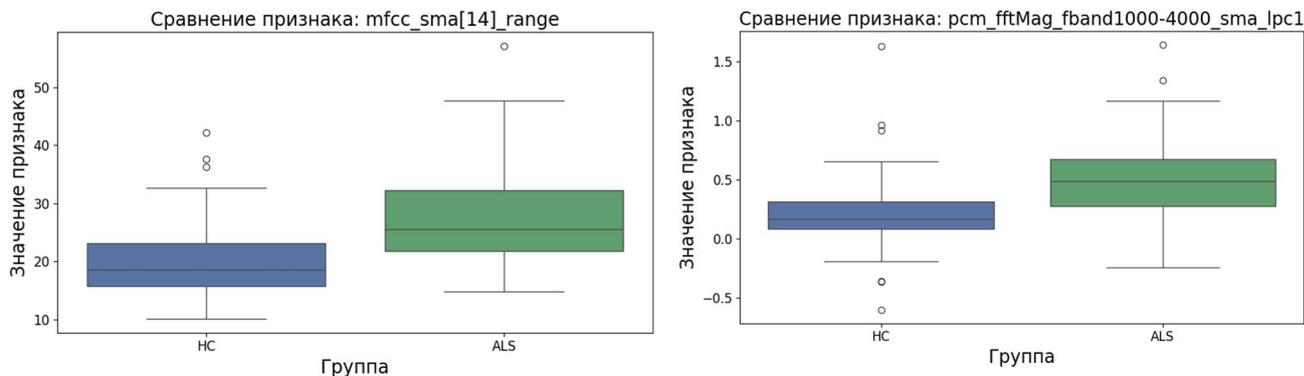


Рисунок 2 – Отобранные признаки после FDR – коррекции

На построенных ящичковых диаграммах (рисунки 1 и 2) представлено распределение значений отобранных голосовых признаков в двух группах: здоровые пациенты (НС) и пациенты с БАС (ALS). Большинство признаков демонстрируют статистически значимые различия между группами, что подтверждено FDR-коррекцией ($p < 0.0001$). Признаки `mfcc_sma[14]_range` и `pcm_fftMag_fband1000-4000_sma_lpc1` могут служить наиболее чувствительными биомаркерами речевых нарушений при БАС. Другие, несмотря на менее выраженные различия, могут быть полезны как дополнительные стабильные признаки при построении модели классификации или мониторинга болезни.

В данной работе использовался автоматизированный подход к анализу голосовых данных пациентов с БАС и здоровых пациентов. Для оценки полученных результатов было проведено сопоставление с методологией и выводами работы [3], в которой также рассматривалась задача автоматического детектирования БАС по акустическим характеристикам речи.

В исследовании [3] также использовалась библиотека OpenSMILE с конфигурацией ComParE13, и отбор признаков осуществлялся с использованием методов машинного обучения – таких как рекурсивное исключение признаков (RFE), машины опорных векторов (SVM) и метод случайного леса (Random Forest). Сравнение с результатами данного исследования показывает, что, несмотря на различие в конкретных названиях признаков, имеется содержательное пересечение по группам параметров. Таким образом, наиболее выраженное пересечение наблюдается в области Мел-кепстральных коэффициентах частот (MFCC) и частотных полос, что подтверждает стабильность этих признаков в разных выборках и методологиях.

Несмотря на различия в подходах и версиях конфигурации ComParE (2013 и 2016), оба исследования подчеркивают значимость спектральных и частотно – временных признаков речи в диагностике БАС. Частичное совпадение в содержательных категориях признаков подтверждает обоснованность нашего подхода и подчеркивает потенциал библиотеки OpenSMILE как универсального инструмента в акустической биомедицинской аналитике.

Заключение

В ходе исследования показано, что автоматическое извлечение речевых признаков с использованием openSMILE позволяет эффективно выявлять различия между здоровыми испытуемыми и пациентами с БАС. Полученные признаки могут быть использованы как основа для построения алгоритмов ранней диагностики или встраивания в системы поддержки клинических решений.

Список использованных источников:

1. OpenSMILE 3.0 [Электронный ресурс] / audEERING - Режим доступа: <https://www.audeering.com/research/opensmile/> – Дата доступа: 18.04.2025.
2. Minsk2020_ALS_database [Электронный ресурс] / GitHub - Режим доступа: https://github.com/Mak-Sim/Minsk2020_ALS_database – Дата доступа: 17.04.2025.
3. Norel R. et al. Detection of amyotrophic lateral sclerosis (ALS) via acoustic analysis // Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. – 2018. – P. 377-381.

ANALYSIS AND EXTRACTION OF SPEECH FEATURES IN HEALTHY INDIVIDUALS AND PATIENTS WITH ALS USING THE OPENSMILE LIBRARY

Mikhnevich A.V.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Vashkevich M.I. – PhD

Annotation. Amyotrophic lateral sclerosis (ALS) is a severe neurodegenerative disease that also affects speech function. The aim of this work is to extract and analyze acoustic features of speech using the openSMILE library, aimed at identifying differences between speech signals of healthy subjects and patients with ALS. Records from the Minsk2020 database were used for the analysis. Statistical analysis with multiple comparison correction (FDR) was performed, based on the results of which significant speech markers were identified that are potentially useful for automatic diagnostics.

Keywords: ALS, openSMILE, speech features, acoustic analysis, statistics, Minsk2020, FDR correction.