

**ИСПОЛЬЗОВАНИЕ МЕТОДОВ ЛОКАЛИЗАЦИИ
ИСТОЧНИКОВ АКТИВНОСТИ МОЗГА ДЛЯ РЕШЕНИЯ
ЗАДАЧИ РАСПОЗНАВАНИЯ МЫСЛЕННО
ПРОИЗНОСИМЫХ ФОНЕМ ПО ДАННЫМ
ЭЛЕКТРОЭНЦЕФАЛОГРАФИИ**

Порывай Максим Викторович

Студент

Филиал МГУ им. М. В. Ломоносова в г. Севастополе, Севастополь, Россия

E-mail: mvporyvaj@asvk.cs.msu.ru

Научный руководитель — Костенко Валерий Алексеевич

В данной работе рассматривается задача классификации мысленно произносимых фонем русского языка. Это задача является подзадачей более общей задачи распознавания внутренней речи человека, которая в свою очередь является одной из перспективных для развития технологии «Интерфейс Мозг–Компьютер» (англ. Brain Computer Interface, или BCI).

Фонема — минимальная смысловозначительная единица языка. Частными случаями фонем являются отдельные звуки.

Задача распознавания мысленно произносимых фонем хорошо исследована для некоторых других языков отличных от русского, например, английского [1]. Однако перенос существующих наработок является отдельной задачей, часто требующей создания собственного решения с нуля, которое только частично опирается на существующие.

Проведенный аналитический обзор литературы показал, что подходы на основе глубокого обучения хорошо себя зарекомендовали в подобных исследованиях [2]. В данной работе задача распознавания мысленно произносимых фонем была сведена к задаче классификации изображений, поэтому использовался подход на основе сверточных нейронных сетей.

Решение задачи классификации мысленно произносимых фонем состоит из следующих этапов: извлечение, сбор и предобработка данных; применение алгоритма локализации; формирование обучающей и тестовой выборки; разработка модели классификации фонем.

В качестве данных электрической активности головного мозга в данном исследовании используются данные электроэнцефалографии (ЭЭГ). ЭЭГ, являясь неинвазивным методом, имеет низкое пространственное разрешение, которое можно повысить за счет исполь-

зования алгоритмов локализации источников активности мозга. Повышение разрешения потенциально может позволить повысить качество классификации.

В качестве фонем были выбраны 7 фонем из различных групп классификации Плоткина: А, Б, Ф, Г, М, Р, У. Для сбора данных схема, представленная на рис. 1 (слева). Для дальнейшего исследования были отобраны только каналы $F7$, $F3$, $T3$, $C3$ (зоны Брока и Вернике), так как именно в них наблюдается наибольшая мозговая активность во время мысленного проговаривания. Детальное описание методики сбора данных и алгоритма их предобработки приведено в работе [3].

Самыми популярными алгоритмами локализации источников активности мозга являются методы семейства LORETA [4]. Но такие алгоритмы применимы в большей степени к локализации высокочастотного гамма-ритма ЭЭГ (выше 32 Гц), в то время как внутренняя речь относится к ритмам ЭЭГ меньшей частоты. Поэтому в данной работе был выбран метод, основанный на идеи добавления «виртуально вживленных электродов» [5], который лишен данного недостатка.

Данные ЭЭГ представляют собой временные ряды. После применения к ним алгоритма локализации, добавляется некоторое количество временных рядов от «виртуально вживленных электродов». Количество «виртуально вживленных электродов» контролируется параметром алгоритма локализации. На рис. 1 в качестве примера показано расположение 4 реальных электродов ($F7$, $F3$, $T3$, $C3$, слева) и 9 «виртуальных» электродов (справа). Временные ряды, полученные с электродов, переводятся из амплитудно-временной в частотно-временную область с помощью вейвлет-преобразования Морле [5]. Это позволяет перейти к задаче классификации изображений, а именно: спектрограммам. Пример спектрограмм приведен на рис. 2.

В качестве функционала качества разработанного алгоритма классификации используется доля верно распознанных фонем. На данный момент среднее значение функционала качества по всем испытуемым равняется 68%. Дальнейшие исследования направлены на улучшение качества модели засчет направленного перебора гиперпараметров. Также планируется разработать способ локализации источников активности мозга, основанный на методах машинного или глубокого обучения, что потенциально позволит увеличить точность локализации, а, следовательно, и точность классификации.

Иллюстрации

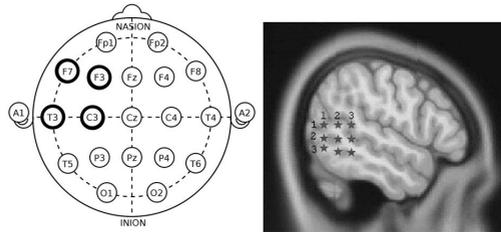


Рис. 1. Слева: Международная система размещения электродов «10 – 20» (выделены электроды, соответствующие областям Брока и Вернике).
Справа: Расположение «виртуально вживленных» электродов

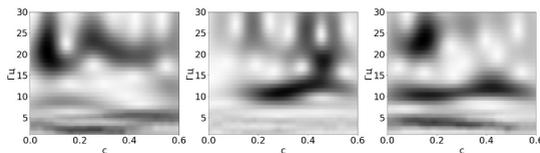


Рис. 2. Примеры спектрограмм для канала $F7$ и двух локализованных каналов при внутреннем проговаривании испытуемым фонемы Φ

Литература

1. Xuemin C., Hagedorn J. B., Schoonover D., D'Zmura M. EEG-based discrimination of imagined speech phonemes // International Journal of Bioelectromagnetism. 2011. № 4. P. 201–206.
2. Santos E., San Martin R., Fraga F. J. Comparison of LORETA and CSP for Brain-Computer interface Applications // In proceedings of the 18th International Multi-Conference on Systems, Signals & Devices (SSD'21), Monastir, Tunisia, 2021, P. 817–822.
3. Gavrilenko Y., Saada D., Ilyushin E., Vartanov A. V., Shevchenko A. The Electroencephalogram Based Classification of Internally Pronounced Phonemes // Brain-Inspired Cognitive Architectures for Artificial Intelligence: BICA*AI 2020. P. 97–105.
4. LORETA localization tool: <http://www.uzh.ch/keyinst/loreta>
5. Vartanov A. V. A new method of localizing brain activity using the scalp EEG data // The 13th Annual Meeting of the BICA Society, Guadalajara, Mexico, 2022, P. 41–48.