



Распознавание внутреннего проговаривания фонем на основе данных электроэнцефалограммы и электромиограммы.

Студентка: Поиленкова Анна 621 группы

Научный руководитель: к.т.н., доцент Костенко В. А.

BCI

brain–computer interface



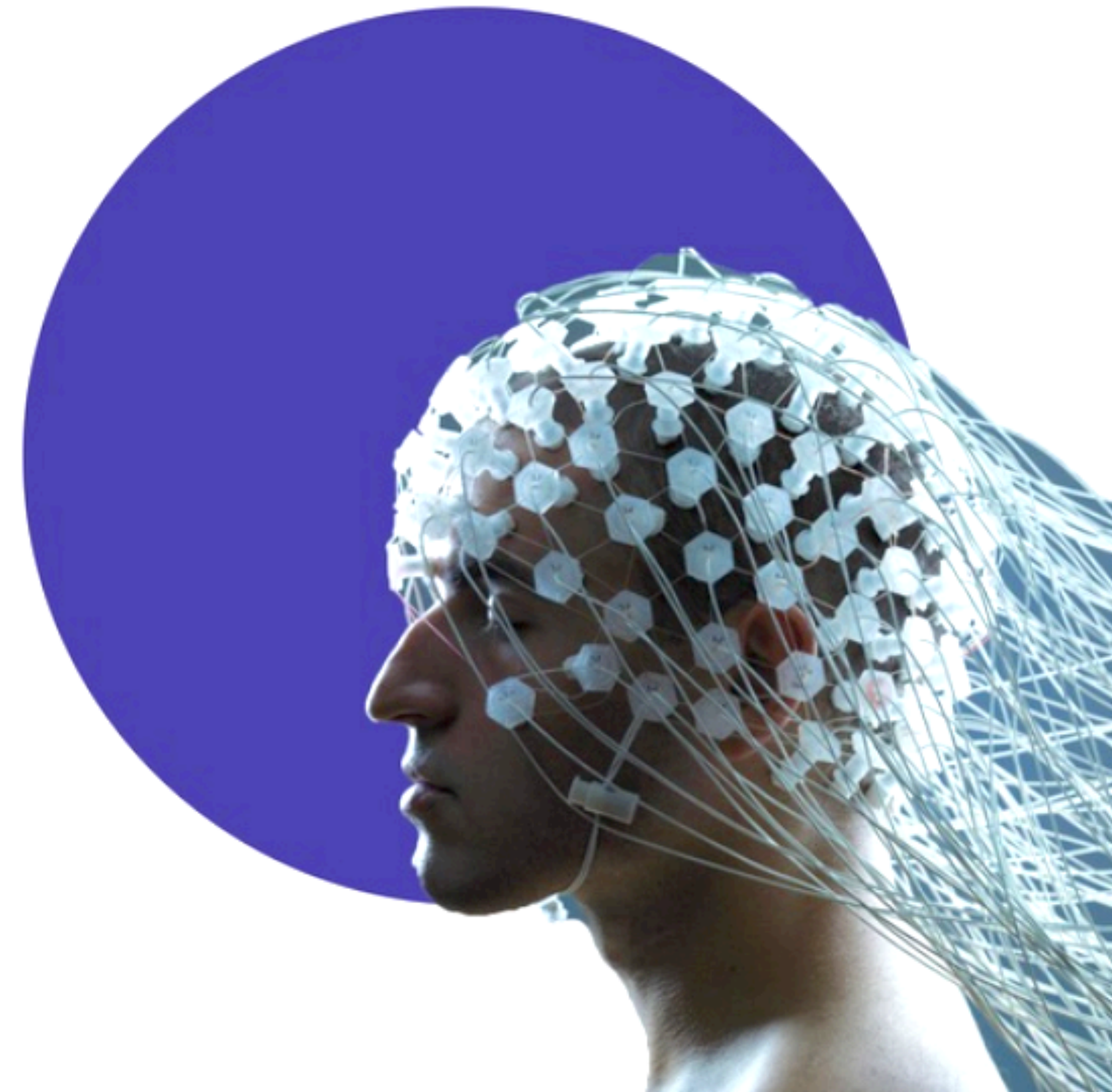
Электрофизиологические данные

ЭЭГ - Электроэнцефалография

— метод регистрации электрической активности мозга

ЭМГ - Электромиография

— метод оценки и регистрации электрической активности скелетных мышц



Цели и задачи работы



Целью работы является разработка алгоритма распознавания мысленно проговариваемых фонем русского языка на основе данных, получаемых с устройства электроэнцефалографии, и данных электромиографа, снятых в области отвечающей за речь.

Для достижения указанной цели необходимо решить следующие задачи:

1. Изучить метод проведения эксперимента по сбору данных, а также методы их предварительной обработки,
2. Разработать математическую постановку задачи распознавания фонем,
3. Провести обзор существующих алгоритмов для задачи распознавания мысленно проговариваемых фонем,
4. На основе обзора разработать и реализовать выбранный метод распознавания мысленно проговариваемых фонем,
5. Провести экспериментальное исследование реализованного алгоритма на реальных данных.

Внутренняя речь



Внутренняя речь - беззвучная, мысленная речь, которая возникает в тот момент, когда мы думаем о чем-либо или по-другому, когда произносим что-то про себя

во время мысленной речи возникает отчетливая речедвигательная импульсная активности либо в форме повышения общего тонуса речевой мускулатуры, либо в форме кратковременных вспышек.

Задача классификации мысленно произносимых проговариваемых фонем



Сбор данных

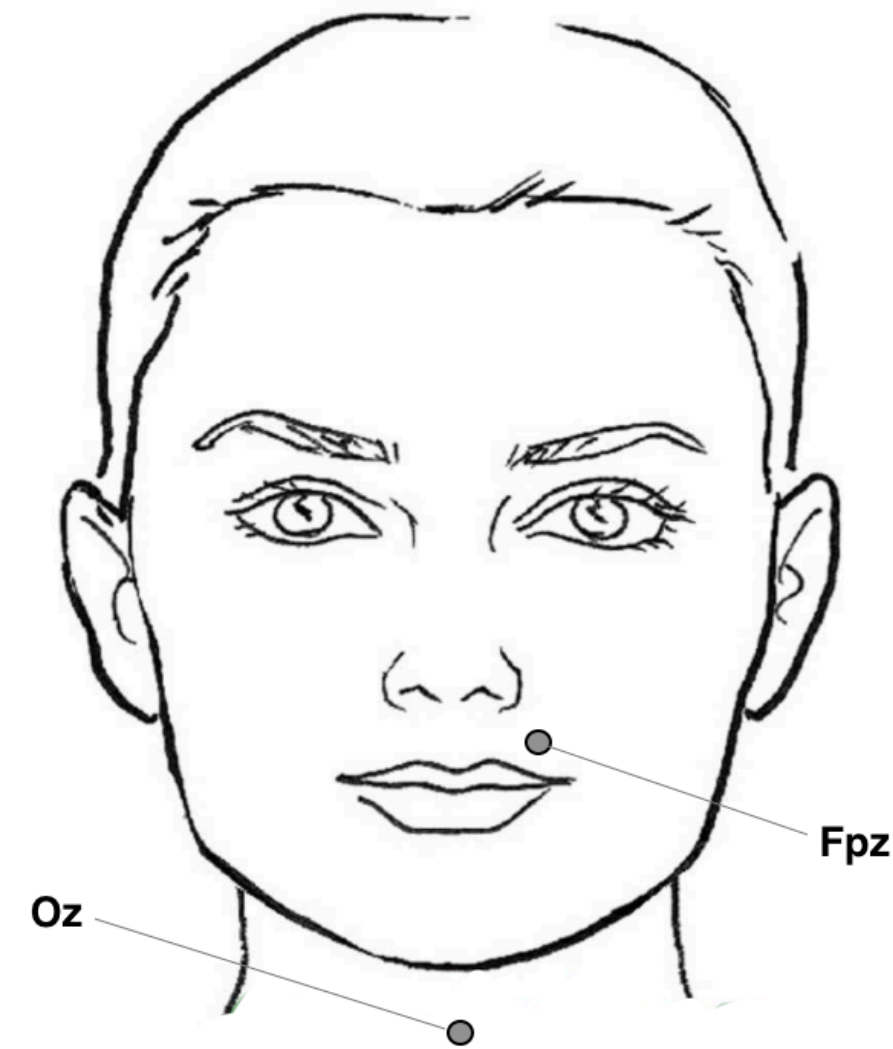


Схема расположения электродов электромиографа

Зона Брока и Вернике

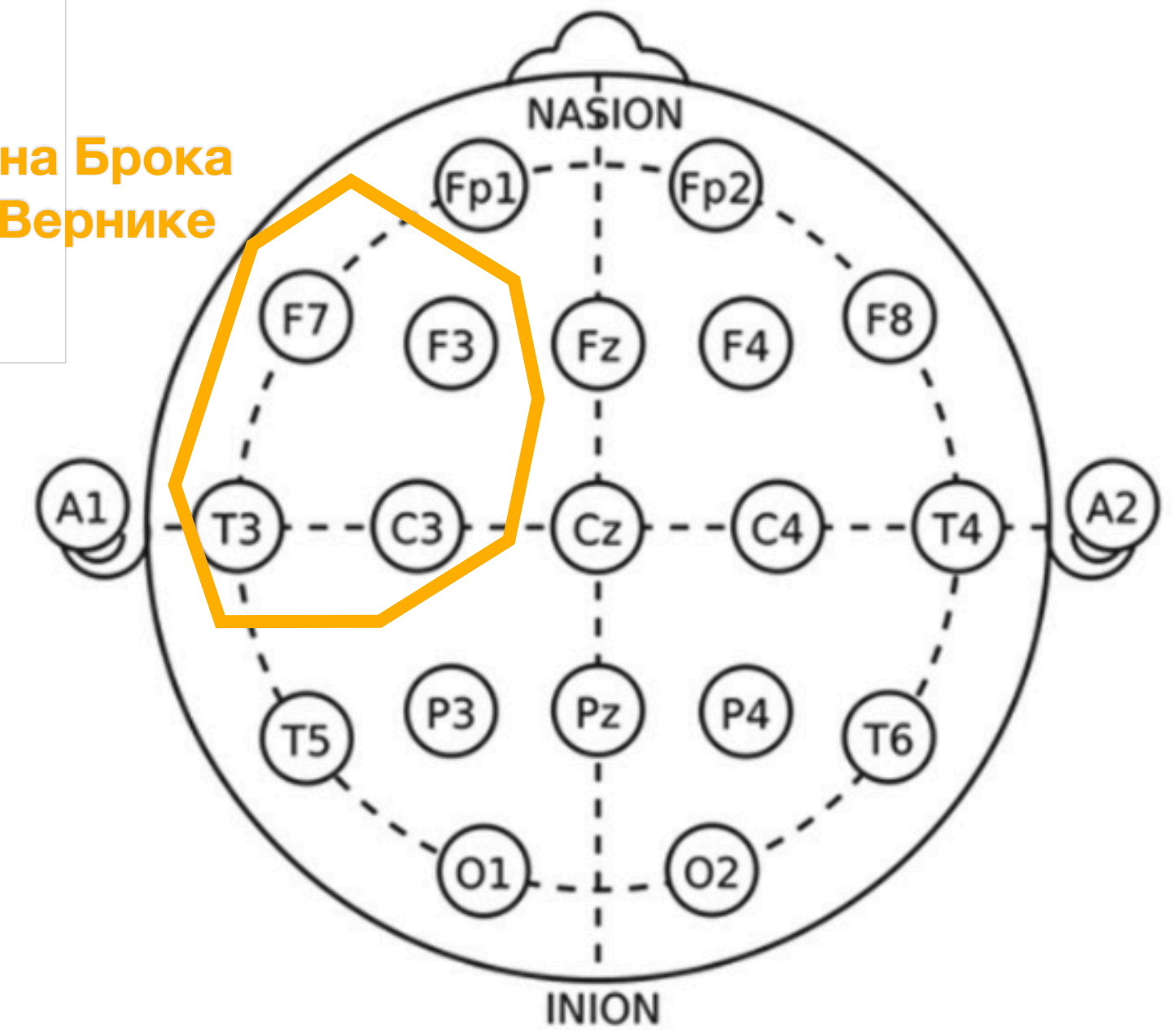


Схема расположения электродов электроэнцефалографа

- визуальный

время реакции

Команда



N - уникальный номер одной из фонемы русского языка: 1 : А – [а], 2 : Б – [б], 3 : Ф – [ф], 4 : Г – [г], 5 : М – [м], 6 : Р – [р], 7 : У – [у]

Задача распознавания мысленного произношения



Дано:

- Данные ЭЭГ: $E^{(i)} = [e_1^{(i)}, e_2^{(i)}, \dots, e_j^{(i)}, \dots, e_{K_i}^{(i)}]$

, где $e_j^{(i)} = \langle e_{j,C3}^{(i)}, e_{j,T3}^{(i)}, e_{j,F3}^{(i)}, e_{j,F7}^{(i)} \rangle$

- Данные ЭМГ: $M^{(i)} = [m_1^{(i)}, m_2^{(i)}, \dots, m_j^{(i)}, \dots, m_{K_i}^{(i)}]$

, где $m_j^{(i)} = \langle m_{j,Oz}^{(i)}, m_{j,Fp}^{(i)} \rangle$

- Обозначим $W^{(i)}$ за множество, полученное объединением элементов множеств $E^{(i)}$ и $M^{(i)}$.

В таком случае: $W^{(i)} = [w_1^{(i)}, w_2^{(i)}, \dots, w_j^{(i)}, \dots, w_{K_i}^{(i)}]$

$$w_j^{(i)} = \langle e_{j,C3}^{(i)}, e_{j,T3}^{(i)}, e_{j,F3}^{(i)}, e_{j,F7}^{(i)}, m_{j,Oz}^{(i)}, m_{j,Fp}^{(i)} \rangle$$

i - номер испытуемого

K_i - число сегментов испытуемого i

$m_{j,Oz}^{(i)}$ - временной ряд канала Oz

Требуется разработать:

- Алгоритм построения классификатора для испытуемого i .

$$classifier : M \rightarrow P$$

, где M - множество всех возможных временных рядов длины T , P - множество $\{1, 2, 3, 4, 5, 6, 7\}$.

$$accuracy = \frac{\sum_{k=1}^{K_i} w_k^{(i)} \in W^{(i)}, [classifier_1^{(i)}\{w_k^{(i)}\} = L(i, k)]}{|\{w_k^{(i)} : w_k^{(i)} \in W^{(i)}\}|}$$

$$accuracy \rightarrow max$$

Обзор



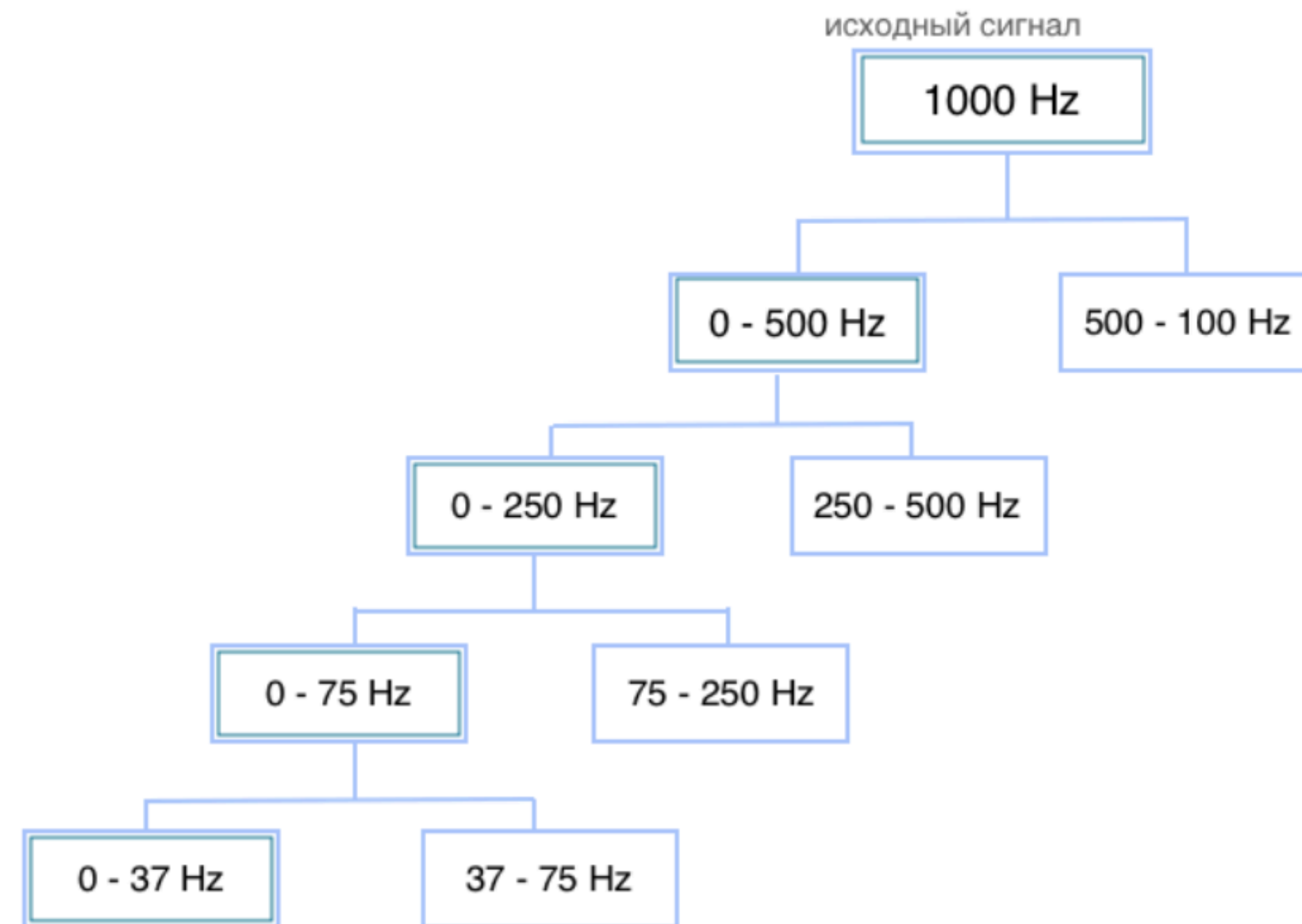
Статья		Точность распознавания
[2]	2 хинди фонемы	SVM - 75–80%
[3]	11 (фонемы + слова)	SVM - 20%
[4]	9 слов	CNN - 85 %
[5]	4 слова (вверх, вниз, налево, направо)	SVM - 26.2%, XGBoos - 27,9%, BiLSTM - 36,1%
[6]	4 слова	SVM - 28,95% до 30,25%

SVM

Support Vector Machine



Извлечение признаков



Четырехуровневое дискретное вейвлет-преобразование для исходного сигнала

Классификация

SVM

AR + MAD + STD

AR - коэффициенты авторегрессии
MAD - среднее абсолютное отклонение
STD - среднеквадратичное отклонение

SVM

Support Vector Machine



	Испытуемый					
Кол-во коэфф. авторегрессии	1	2	3	4	5	6
Данные ЭЭГ						
AR4	26.8%	13.6%	22%	23.4%	15.3%	22%
AR6	24.3%	24.1%	32.5%	11.1%	27.8%	23.6%
AR8	17.5%	26.8%	24.3%	14.7%	21.7%	19.2%
Объединённые данных ЭЭГ и ЭМГ						
AR4	21.6%	12.7%	20%	23.1%	13.4%	10.4%
AR6	23.5%	18.2%	20.3%	16.9%	18.8%	15.1%
AR8	15%	15.4%	18.4%	17.2%	18.3%	13.4%

Усредненные значения точности. AR4, AR6, AR8 — коэффициенты авторегрессионных моделей 4-го, 6-го и 8-го порядка соответственно.

Полученные результаты

- по данным ЭЭГ - 23.9% в среднем по всем испытуемым
- при объединении данных ЭЭГ + ЭМГ — 18.8%

Литература



1. Шевченко А. О., Вартанов А. В. ВЫЗВАННЫЙ АРТИКУЛЯТОРНЫЙ ОТВЕТ ПРИ ВНУТРЕННЕМ И ВНЕШНЕМ ПРОГОВОРЕНИИ // Современная наука: актуальные проблемы теории и практики. Серия: Познание. -2023. -№01. -С. 108-111 DOI 10.37882/2500-3682.2023.01.20
2. Khan, Munna. Classification of myoelectric signal for sub-vocal Hindi phoneme speech recognition / Munna Khan, Mosarrat Jahan // Journal of Intelligent & Fuzzy Systems. — 2018. — Vol. 35, no. 5. — Pp. 5585–5592.
3. Cooney, Ciaran. Mel frequency cepstral coefficients enhance imagined speech decoding accuracy from EEG / Ciaran Cooney, Rafaella Folli, Damien Coyle // 2018 29th Irish Signals and Systems Conference (ISSC) / IEEE. — 2018. — Pp. 1–7.
4. Silent eeg-speech recognition using convolutional and recurrent neural network with 85% accuracy of 9 words classification / Darya Vorontsova, Ivan Menshikov, Aleksandr Zubov et al. // Sensors. — 2021. — Vol. 21, no. 20. — P. 6744.
5. Gasparini F., Cazzaniga E., Saibene A. Inner speech recognition through electroencephalographic signals //arXiv preprint arXiv:2210.06472. – 2022.
6. Cooney C. et al. Evaluation of hyperparameter optimization in machine and deep learning methods for decoding imagined speech EEG //Sensors. – 2020. – T. 20. – No. 16. – C. 4629.