

Задачи классификации мозговой активности при помощи синоптических сетей

Власенко Даниил Владимирович

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

Научный руководитель: к.ф.-м.н., доцент Шпилев П.В.
Рецензент: руководитель отдела алгоритмической структурной
биоинформатики АО "Биокад" Зенкова Н.В.

Санкт-Петербург, 2023

Функциональная магнитно-резонансная томография

Технология, позволяющая измерять изменения в токе крови головного мозга, вызванные нейронной активностью.

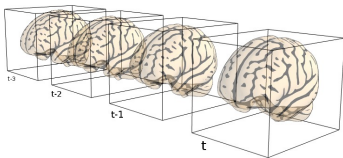


Рис. 1: Последовательных трехмерных изображений.

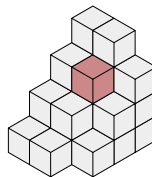


Рис. 2: Воксели.

Воксель

Элемент объёмного растрового изображения, содержащий значение конкретной его области, аналог пикселя в двухмерном изображении.

Данные: представление фМРТ данных в памяти компьютера

Пусть Ω — множество фМРТ, а $\Sigma = \{I, II\}$ — множество режимом мозговой активности. $(\tilde{\Omega}, \tilde{\Sigma}) = \{(\omega_n, \sigma_n)\}_n$ — конечная выборка из (Ω, Σ) .

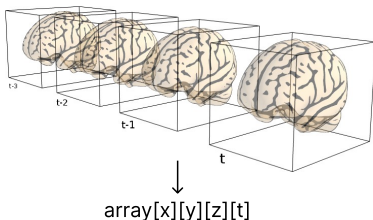


Рис. 3: Конвертирование данных в массив.

$\omega \in \Omega$ конвертируется массив a . $a_{xyzt} \in \mathbb{R}_+$ — значение вокселя с индексами x, y, z в момент времени t , а a_{xyz} — все значения вокселя с индексами x, y, z .

- Сравнение качества классификации режимов мозговой активности с и без данных из конкретной области мозга позволяет оценить важность этой области для данных режимов (Li X. et al. 2019).
- Доказана эффективность применения методов анализа графов к данным о мозге (Li X. et al. 2019; Saueressig C. et al. 2021).
- Доказана эффективность применения синолитических сетей в анализе сложных многомерных данных (Nazarenko T. et al. 2021).

Цель работы

Предложить метод представления фМРТ данных в форме графов, которые будут отражать в себе полезную информацию о взаимосвязях областей мозга для последующей классификации.

Метод будет протестирован с помощью классификация на основе характеристик графов.

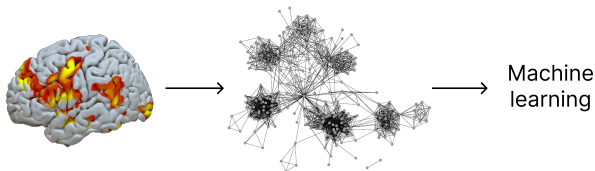


Рис. 4: Представление мозга в форме графа, классификация на данных графа.

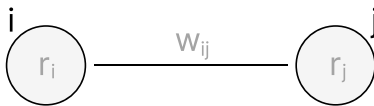


Рис. 5: Вершины i , j и ребро ij .

С помощью статистики $T : \{a_{xyz}\}_{xyz} \rightarrow \mathbb{R}$ вычисляются значения вершин r_i, r_j как образы от временных рядов соответствующих вокселей.

Лучший результат метода наблюдался, когда T — медианное значение.

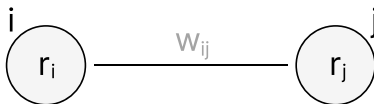


Рис. 6: Вершины i , j и ребро ij .

Вес ребра w_{ij} вычисляется по следующей формуле:

$$w_{ij} = P(II|r_i, r_j) - P(I|r_i, r_j).$$

Для вычисления вероятностей используется вероятностный классификатор $Cl_{ij} : \{y|(r_i, r_j), \{(r_i^n, r_j^n)\}_n, \{y_n\}_n\} \rightarrow [0, 1]$, обученный на выборке $(\tilde{\Omega}, \tilde{\Sigma})$.

В качестве Cl_{ij} использовался метод опорных векторов.

Соседние воксели

Два вокселя называются соседними, если у них есть общая сторона, ребро или угол.

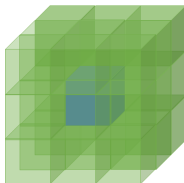


Рис. 7: Соседние воксели к центральному.

В графе g ребра соединяют те вершины, которые ассоциированы с соседними вокселями. Из g удаляются ребра $\{ij : r_i < r | r_j < r | |w_{ij}| < w\}_{ij}$.

Тестирование: классификация на основе характеристик графа

Вычисляются характеристики графа $\{f_u\}_u = \{F_u(g)\}_u$:

- медианный вес ребра,
- квантиль весов ребер уровня 10%,
- квантиль весов ребер уровня 90%,
- стандартное отклонение весов ребер.

С помощью классификатора $Cl : \{\{f_u\}_u | \{\{f_u^n\}_u\}_n, \{y_n\}_n\} \rightarrow \{0, 1\}$, обученного на выборке $(\tilde{\Omega}, \tilde{\Sigma})$, происходит итоговая классификация фМРТ данных ω .

В качестве Cl использовался метод опорных векторов.

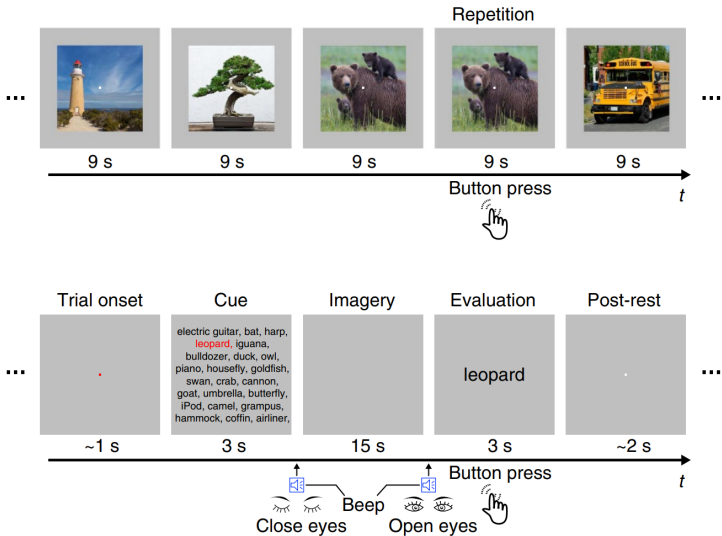


Рис. 8: Наблюдение или воображение объекта.

Внутригрупповое тестирование: разделение выборки

| | seen | | imagined | | |
|--------|----------|------|----------|------|-----|
| | training | test | training | test | |
| sub-01 | 17 | 7 | 14 | 6 | 44 |
| sub-02 | 17 | 7 | 14 | 6 | 44 |
| sub-03 | 17 | 7 | 14 | 6 | 44 |
| sub-04 | 17 | 7 | 14 | 6 | 44 |
| sub-05 | 16 | 8 | 14 | 6 | 44 |
| | 84 | 36 | 70 | 30 | 220 |
| | 120 | | 100 | | |

Таблица 1: Разделение выборки на тренировочную и тестовую части.

Данные не независимые, однако позволяют проверить, возможно ли предсказать поведение людей, на которых метод обучался.

Внутригрупповое тестирование: характеристики графов

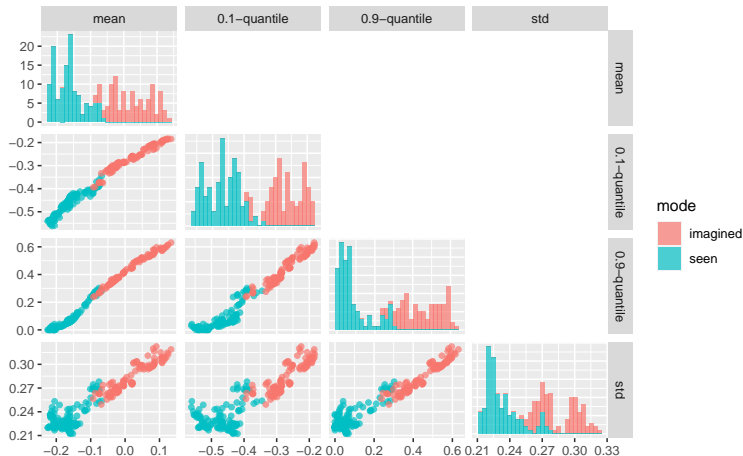


Рис. 9: Распределения характеристик графов для всей выборки фМРТ, когда T — медианное значение вокселя, $r = 1$, $w = 0$, где imagined mode — режим визуализации на основе памяти, seen mode — режим зрительного восприятия.

| | seen | imagined |
|----------|------|----------|
| seen | 36 | 0 |
| imagined | 1 | 29 |

Таблица 2: Матрица классификации когда T — медианное значение вокселя, $r = 1$, $w = 0.2$, где imagined — режим визуализации на основе памяти, seen — режим зрительного восприятия

Для сравнения результатов использовался метод класса Decoder из библиотеки nilearn (Abraham A. 2014).

| | Точность |
|--------------------------|----------|
| Synolitic networks + SVM | 98.5% |
| nilearn.Decoder | 100% |

Таблица 3: Точность классификации синолитического метода и метода nilearn.Decoder.

Было проведено 5 запусков метода так, что при каждом запуске тестовая выборка состояла из всех фМРТ одного испытуемого, остальные фМРТ попадали в обучающую выборку.

| | Средняя точность | Дисперсия |
|--------------------------|------------------|-----------|
| Synolitic networks + SVM | 51.6% | 0.46% |
| nilearn.Decoder | 39.54% | 2.45% |

Таблица 4: Средняя точность и дисперсия классификации для синолитического метода и метода nilearn.Decoder.

- ❶ На языке Python был реализован метод представления данных фМРТ в форме графов, в основе которого лежит идея синоптических сетей.
- ❷ По результатам тестирования метод оказался внутригрупповым и неэффективным при межгрупповом применении.
- ❸ Было проверено, что синоптические сети применимы в сфере изучения функционирования мозга, но требуются дополнительное развитие данной области.

Код ВКР и его документация выложены в открытый доступ, DOI: 10.5281/zenodo.7927929.

Есть с.в. $\omega = (\omega_1, \dots, \omega_k)$. На основе ω будет строится граф $g = (V, E, R, W)$, где

- $V = \{i\}_i$ — множество вершин,
- $R = \{r_i\}_i$ — множество значений вершин,
- $E = \{ij\}_{ij}$ — множество неориентированных ребер,
- $W = \{w_{ij}\}_{ij}$ — множество весов ребер.

Синолитическая сеть

Представление с.в. ω в форме графа g , в котором вершины V отражают признаки $(\omega_1, \dots, \omega_k)$, а ребра E отражают взаимосвязи между признаками $(\omega_1, \dots, \omega_k)$.

При этом вес любого ребра w_{ij} вычисляется на основе значений инцидентных ребру вершин r_i, r_j таким образом, чтобы нести полезную для классификации информацию.

| | seen | | imagined | | |
|--------|----------|------|----------|------|-----|
| | training | test | training | test | |
| sub-01 | 0 | 24 | 0 | 20 | 44 |
| sub-02 | 24 | 0 | 20 | 0 | 44 |
| sub-03 | 24 | 0 | 20 | 0 | 44 |
| sub-04 | 24 | 0 | 20 | 0 | 44 |
| sub-05 | 24 | 0 | 20 | 0 | 44 |
| | 96 | 24 | 80 | 20 | 220 |
| | 120 | | 100 | | |

Таблица 5: Разделение выборки на тренировочную и тестовую части.

Межгрупповое тестирование: характеристики графов

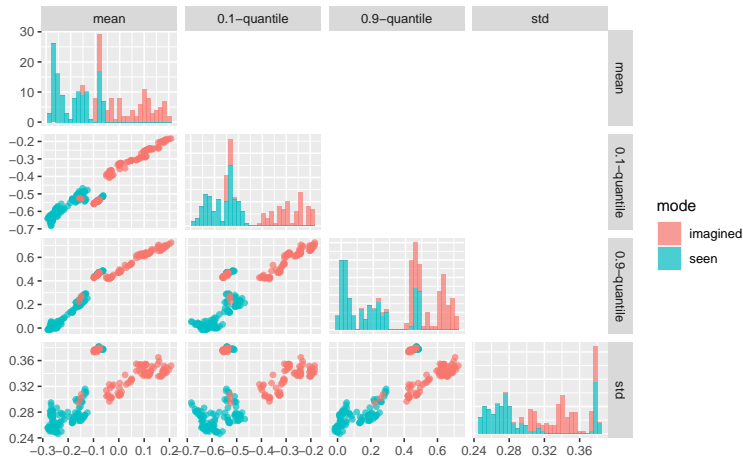


Рис. 10: Распределения характеристик графов для всей выборки фМРТ, когда T — медианное значение вокселя, $r = 1$, $w = 0$, где imagined mode — режим визуализации на основе памяти, seen mode — режим зрительного восприятия.