

Восстановление функциональных групп головного мозга с помощью графовых диффузных моделей

А. И. Астахов, С. К. Панченко, В. В. Стрижов

astakhov.am@phystech.edu; panchenko.sk@phystech.edu; strijov@phystech.edu

В данной работе рассматривается задача классификации многомерного временного ряда, представляющего собой электроэнцефалограмму головного мозга человека. Стандартные подходы, использующие двухмерные свертки, не могут учесть пространственную структуру сигнала, поскольку датчики, считывающие показатели, находятся на сферической поверхности. В качестве решения предлагается использовать графовое представление функциональных групп, а для моделирования использовать нейронную диффузию.

Головной мозг, ЭЭГ, Графовые нейронные сети, диффузионные модели

1 Введение

Эмоции играют ключевую роль в человеческом восприятии, принятии решений и социальном взаимодействии. Их автоматическая классификация на основе нейрофизиологических данных, таких как электроэнцефалография (ЭЭГ), открывает новые возможности в психологии, медицине, аффективных вычислениях и человеко-машинном взаимодействии. Однако, несмотря на значительный прогресс в области машинного обучения и нейронаук, точная и надежная классификация эмоций по ЭЭГ остается сложной задачей. Это связано с высокой индивидуальной вариабельностью сигналов, нелинейной природой эмоциональных процессов, а также ограничениями существующих методов предобработки и классификации.

В данной статье рассматриваются современные подходы к распознаванию эмоций по ЭЭГ, анализируются их преимущества и недостатки, а также предлагаются пути улучшения точности классификации. Особое внимание уделяется методам обработки сигналов, выделения информативных признаков и использованию алгоритмов глубокого обучения. Результаты исследования могут быть полезны для разработки более эффективных систем аффективного взаимодействия, нейрореабилитации и психофизиологических исследований. Объектом нашего исследования является сигнал полученный путем электроэнцефалографического исследования человеческого мозга, воспринимаемый как многомерный временной ряд, где каждая размерность отвечает конкретному датчику на голове испытуемого. Технические ограничения ЭЭГ-исследований включают низкое пространственное разрешение метода и высокую чувствительность к артефактам. Как показано в работе [3], движения глаз и мышечная активность могут существенно искажать сигнал. Более того, индивидуальные различия в паттернах мозговой активности между испытуемыми значительно снижают эффективность универсальных классификаторов. Исследование [1] демонстрирует, что точность межсубъектной классификации эмоций редко превышает 60%, даже при использовании современных методов глубокого обучения.

Существующие подходы преимущественно основаны на двух теоретических моделях эмоций: локационистской (базовые эмоции) и многомерной (валентность-возбуждение-доминантность, VAD) [8]. Однако, большинство современных методов не учитывают пространственные взаимосвязи между электродами, что ограничивает их эффективность.

В исследованиях используются различные методы извлечения признаков:

Временные характеристики: В работе [5] применены шесть статистических параметров ЭЭГ с последующим отбором каналов методами PCA и ReliefF, что позволило достичь

точности 81.87% на датасете DEAP. Однако авторы не рассматривали пространственные корреляции между электродами.

Частотные характеристики: Исследование [2] демонстрирует эффективность PCA для сокращения размерности признаков с последующей классификацией методом SVM (точность 85.85% на SEED). Аналогично, в [6] сравнение различных признаков показало, что статистические характеристики в сочетании с KNN дают точность 77.54-79

Важное ограничение этих работ заключается в том, что анализ проводился для каждого электрода отдельно, без учета пространственных взаимодействий между различными областями мозга. Это особенно существенно, поскольку эмоциональные состояния, как известно, связаны с скоординированной активностью распределенных нейронных сетей [4].

Задача нашего исследования – использовать пространственные связи между датчиками для улучшения качества классификации. Мы предлагаем рассматривать временной ряд как динамический граф, где ребра представляют взаимосвязи между датчиками в пространстве или статистически. Мы считаем, что учитывание этих факторов позволит построить более качественную и устойчивую модель классификации. Мы рассмотрим подходы построения подобных связей и посмотрим, как они влияют на результат классификации. Проводить оценку качества модели мы будем на открытом датасете SEED IV. В качестве модели мы предлагаем использовать DCGRU которая хорошо себя показала в схожей проблеме классификации эпилептических припадков по данным электроэнцефалограммы [7]

2 Общее описание проблемы

Дана выборка $D = (\mathbf{X}, \mathbf{Z}, \mathbf{y})$, где \mathbf{X} – набор сигналов, \mathbf{Z} – координаты, \mathbf{y} – целевая переменная. Цель – построить модель для декодирования активности мозга на основе иффузионной сверточной рекуррентной нейронной сети.

2.1 Постановка задачи

Данные представляют собой многомерный временной ряд m , где m – номер исследования. Таким образом все данные можно описать следующим образом:

$$\mathbf{X} = [\mathbf{X}_m]_{m=1}^M, \quad \mathbf{X}_m = [\mathbf{x}_t]_{t \in T}, \quad \mathbf{x}_t \in \mathbb{R}^E$$

Где T – количество записей в испытании, t – конкретный момент времени, E – количество электродов в приборе.

Структуру динамического графа для конкретного момента и испытания зададим следующим образом:

$$\mathcal{G}(m, t) = (\mathcal{V}(m, t), \mathcal{E}(m, t), \mathbf{A}_{\mathbf{X}, \mathbf{Z}}(m, t)),$$

Будем искать модель из параметрического семейства моделей DCGRU:

$$h_{\theta} : (\mathbf{X}, \mathbf{A}_{\mathbf{X}, \mathbf{Z}}) \rightarrow y, \quad \theta \in \Theta,$$

Необходимо найти наилучшие значения параметров θ .

Оптимизационную постановку зададим следующим образом:

$$\theta^* = \arg \min_{\theta} \mathcal{L}$$

$$\mathcal{L} = -\frac{1}{M} \sum_{m=1}^M \left[\sum_{c=1}^C \mathbb{I}(y_m = c) \log(p_m^c) \right],$$

3 Вычислительный эксперимент:

Цель вычислительного эксперимента заключатся в испытании способов построение функциональных связей между электродами. Для обучения модели мы будем использовать SEED IV датасет. Он представляет из себя испытания для пятнадцати человек поделенные на три сессии в каждой из которых участникам показывались видеозаписи длительностью около двух минут. Каждый испытуемый участвовал в двадцати четырех испытаниях в каждой сессии.

Проведем два эксперимента:

В первом в качестве признаком будем использовать показания датчиков сигнала, а в качестве связей между вершинами корреляцию Пирсона:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (1)$$

где:

- n - количество наблюдений
- X_i, Y_i - отдельные значения переменных
- \bar{X}, \bar{Y} - средние значения переменных

А во втором, в качестве признаков для модели используется дифференциальная энтропия выделенных диапазонов ритмов головного мозга: дельта (1 – 3Гц), тета (4 – 7Гц), альфа (8 – 13Гц), бета (14 – 30Гц), гамма (31 – 50Гц):

$$DE(Y) = - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \log \left(\frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \right) dy \quad (2)$$

Таким образом в момент времени t графовый сигнал, задающий значения наблюдаемой мозговой активности на вершинах имеет размерность $x_t \in \mathbb{R}^{62 \times 5}$, где 62 — число электродов, считывающих сигнал, 5 — число рассматриваемых мозговых ритмов.

Связи будем считать как синхронизацию фаз между сигналами датчиков:

Фазовая синхронизация представляет собой подход к анализу возможных нелинейных взаимозависимостей и фокусируется на фазах сигналов. Предполагается, что две динамические системы могут иметь синхронизацию фаз, даже если их амплитуды независимы. Обозначим $x(t)$, $y(t)$ динамические системы, соответствующие наблюдениям сигнала, x_{mi} и x_{mj} в отрезке времени $[t_n - T_w, t_n]$ в m -ом испытании. Синхронизация фаз понимается как:

$$|\varphi_x(t) - \varphi_y(t)| = \text{const.} \quad (3)$$

Для оценки фазы вычисляется аналитическое представление сигнала с использованием преобразования Гильберта:

$$H(t) = x(t) + i\dot{x}(t), \quad \text{где} \quad (4)$$

$$\dot{x}(t) = \frac{1}{\pi} \text{v.p.} \int_{-\infty}^{\infty} \frac{x(t')}{t-t'} dt' - \text{преобразование Гильберта сигнала } x(t), \quad (5)$$

где v.p. — главное значение интеграла по Коши.

Фаза аналитического сигнала определяется как:

$$\varphi(t) = \arctan \left(\frac{\dot{x}(t)}{x(t)} \right). \quad (6)$$

Таким образом, для двух сигналов $x(t)$, $y(t)$ равной продолжительности T_w с фазами $\varphi_x(t)$, $\varphi_y(t)$ значение синхронизации фаз (phase locking value) [?] задаётся уравнением:

$$p_{ij}(m, t_n) = \left| \frac{1}{T_w} \sum_{k=1}^{T_w} \exp(s(\varphi_x(k\Delta t) - \varphi_y(k\Delta t))) \right|, \quad \text{где} \quad (7)$$

$$\Delta t = \text{шаг по времени}, \quad s = \sqrt{-1}. \quad (8)$$

Итоговая матрица синхронизации:

$$A_{\Sigma, Z}^*(m, t) = [a_{ij}(m, t)] \in \mathbb{R}_+^{E \times E}, \quad a_{ij}(m, t) = \begin{cases} p_{ij}(m, t), & \text{если } p_{ij}(m, t) \geq \rho(p) \\ 0, & \text{иначе.} \end{cases} \quad (9)$$

Результаты получились следующие:



Рис. 1 Для первого подхода



Рис. 2 Для второго подхода

4 Теоретическая часть

В данном разделе представлено решение задачи классификации многомерных временных рядов ЭЭГ с использованием графовых диффузионных моделей. Предлагаемый метод сочетает пространственное моделирование функциональных связей мозга с анализом временной динамики сигналов.

4.1 Предлагаемое решение

Решение задачи включает следующие этапы:

1. Построение графа функциональных связей между электродами
2. Применение диффузионных сверток для пространственного моделирования
3. Использование рекуррентных слоев для анализа временных зависимостей
4. Обучение модели с учителем для классификации паттернов ЭЭГ

4.2 Математическая модель

Графовая структура мозга представляется как $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$, где:

- \mathcal{V} — множество вершин (электродов), $|\mathcal{V}| = E$
- \mathcal{E} — множество ребер (функциональных связей)
- $\mathbf{A} \in \mathbb{R}^{E \times E}$ — взвешенная матрица смежности

Матрица смежности строится на основе корреляции сигналов между электродами:

$$A_{ij} = \rho(x_i, x_j), \quad (10)$$

где $\rho(x_i, x_j)$ — коэффициент корреляции Пирсона между сигналами электродов i и j .

Для улучшения устойчивости модели можно применить пороговую обработку:

$$A_{ij} = \begin{cases} \rho(x_i, x_j), & \text{если } \rho(x_i, x_j) > \tau \\ 0, & \text{иначе} \end{cases} \quad (11)$$

где τ — пороговое значение корреляции.

4.3 Диффузионная свертка

Для пространственного моделирования используется диффузионная свертка:

$$\mathbf{H}^{(l+1)} = \sum_{k=0}^K \mathbf{P}^k \mathbf{H}^{(l)} \mathbf{W}_k,$$

где:

- $\mathbf{P} = \mathbf{D}^{-1} \mathbf{A}$ — нормированная матрица смежности
- $\mathbf{H}^{(l)}$ — скрытые представления на слое l
- \mathbf{W}_k — обучаемые параметры
- K — количество шагов диффузии

4.4 Временное моделирование

Для анализа временных зависимостей применяется GRU-слой:

$$\mathbf{z}_t = \sigma(\mathbf{W}_z[\mathbf{h}_{t-1}, \mathbf{x}_t]),$$

$$\mathbf{r}_t = \sigma(\mathbf{W}_r[\mathbf{h}_{t-1}, \mathbf{x}_t]),$$

$$\begin{aligned}\tilde{\mathbf{h}}_t &= \tanh(\mathbf{W}[\mathbf{r}_t \odot \mathbf{h}_{t-1}, \mathbf{x}_t]), \\ \mathbf{h}_t &= (1 - \mathbf{z}_t) \odot \mathbf{h}_{t-1} + \mathbf{z}_t \odot \tilde{\mathbf{h}}_t,\end{aligned}$$

где $\mathbf{z}_t, \mathbf{r}_t$ — gates, \mathbf{h}_t — скрытое состояние.

4.5 Алгоритм обучения

Алгоритм 1 Обучение DCRNN модели

Вход: Множество данных $D = (\mathbf{X}, \mathbf{Z}, \mathbf{y})$, параметры модели

Выход: Обученная модель h_θ

- 1: Инициализировать параметры θ случайным образом
 - 2: Построить матрицы смежности $\mathbf{A}_{\mathbf{X}, \mathbf{Z}}$ для всех примеров
 - 3: **для** эпоха = 1 до MaxEpochs
 - 4: **для** пакет (\mathbf{X}_m, y_m) в D
 - 5: Вычислить предсказание: $\hat{y}_m = h_\theta(\mathbf{X}_m, \mathbf{A}_{\mathbf{X}, \mathbf{Z}}(m))$
 - 6: Вычислить потерю: $\mathcal{L} = -\log p_m^{y_m}$
 - 7: Обновить параметры: $\theta \leftarrow \theta - \alpha \nabla_\theta \mathcal{L}$
 - 8: **return** θ^*
-

4.6 Свойства решения

Предлагаемое решение обладает следующими свойствами:

- Учитывает неевклидову природу расположения электродов
- Моделирует как пространственные, так и временные зависимости
- Позволяет интерпретировать функциональные связи через матрицу смежности
- Устойчиво к шумам благодаря диффузионному процессу

Литература

- [1] Salma Alhagry, Aly Aly, and Reda A. Emotion recognition based on eeg using lstm recurrent neural network. *International Journal of Advanced Computer Science and Applications*, 8(10), 2017.
- [2] Md. Asadur Rahman, Md. Foisal Hossain, Mazhar Hossain, and Rasel Ahmmed. Employing pca and t-statistical approach for feature extraction and classification of emotion from multichannel eeg signal. *Egyptian Informatics Journal*, 21(1):23–35, March 2020.
- [3] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, March 2004.
- [4] Kristen A. Lindquist, Tor D. Wager, Hedy Kober, Eliza Bliss-Moreau, and Lisa Feldman Barrett. The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 35(3):121–143, May 2012.
- [5] Yishu Liu and Guifang Fu. Emotion recognition by deeply learned multi-channel textual and eeg features. *Future Generation Computer Systems*, 119:1–6, June 2021.
- [6] Rab Nawaz, Kit Hwa Cheah, Humaira Nisar, and Vooi Voon Yap. Comparison of different feature extraction methods for eeg-based emotion recognition. *Biocybernetics and Biomedical Engineering*, 40(3):910–926, July 2020.
- [7] Siyi Tang, Jared A. Dunnmon, Khaled Saab, Xuan Zhang, Qianying Huang, Florian Dubost, Daniel L. Rubin, and Christopher Lee-Messer. Self-supervised graph neural networks for improved electroencephalographic seizure analysis. 2021.

- [8] Mirosław Wyczesany and Tomasz S. Ligeza. Towards a constructionist approach to emotions: verification of the three-dimensional model of affect with eeg-independent component analysis. *Experimental Brain Research*, 233(3):723–733, November 2014.