

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 [4], 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) [11]. 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 [8] 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 [9] 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 [7].

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 [10]

2 Постановка задачи

2.1 Построение матрицы связности

Исходный сигнал ЭЭГ задан в виде тензор $\mathbf{X} = [\mathbf{X}_m]_{m=1}^M$, $\mathbf{X}_m \in \mathbb{R}^{E \times N}$, где N соответствует числу отсчетов времени при измерении сигнала, а E — число электродов, считываемых сигнал, M — число испытаний. Также дана матрица координат электродов $\mathbf{Z} \in \mathbb{R}^{E \times 3}$, определяемая выбранным при снятии электроэнцефалограммы стандартом размещения. В данной работе предлагается рассматривать сигнал в качестве неориентированного динамического графа: $\mathcal{G}(m, t) = (\mathcal{V}(m, t), \mathcal{E}(m, t), \mathbf{A}_{\mathbf{X}, \mathbf{Z}}(m, t))$, для решения проблемы моделирования пространственной взаимосвязи между электродами на голове испытуемого. В качестве множества вершин $\mathcal{V}(m, t)$ мы рассматриваем электроды, а значениями в вершинах будут значения сигнала в момент t . В множество ребер $\mathcal{E}(m, t)$ задается матрицей связности графа $\mathbf{A}_{\mathbf{X}, \mathbf{Z}}(m, t)$.

2.2 Основные определения

Дана выборка $\mathcal{D} = (\mathbf{X}, \mathbf{Z}, \mathbf{y})$ активности головного мозга, где:

- $\mathbf{X} = [\mathbf{X}_m]_{m=1}^M$ — набор сигналов
- $\mathbf{X}_m = [\mathbf{x}_t]_{t \in T}$ — сигнал, полученный в m -ом испытании
- $\mathbf{x}_t \in \mathbb{R}^E$ — наблюдения сигнала в момент времени t
- $\mathbf{Z} = [\mathbf{z}_k]_{k=1}^E$, $\mathbf{z}_k \in \mathbb{R}^3$ — координаты электродов
- $\mathbf{y} = [y_m]_{m=1}^M$ — целевая переменная
- $y_m \in \{1, \dots, C\}$ — метка класса
- $T = \{t_n\}_{n=1}^N$ — набор временных отсчетов
- $E = 62$ — число электродов
- N — число наблюдений в одном отрезке сигнала

Для решения задачи декодирования рассматривается модель из класса графовых рекуррентных диффузионных нейронных сетей:

$$h_\theta : (\mathbf{X}, \Delta_{\mathbf{X}, \mathbf{Z}}^*) \rightarrow \mathbf{y}. \quad (1)$$

В качестве функции ошибки выбрана кросс-энтропия:

$$\mathcal{L} = -\frac{1}{M} \sum_{m=1}^M \left[\sum_{c=1}^C \mathbf{I}(y_m = c) \log(p_m^c) \right], \text{ где} \quad (2)$$

$p_m^c = h_\theta(\mathbf{X}_m, \Delta_{\mathbf{X}, \mathbf{Z}}^*(m))$ — вероятность класса c для \mathbf{X}_m с матрицей $\Delta_{\mathbf{X}, \mathbf{Z}}^*(m)$.

Задача поиска оптимальных параметров имеет следующий вид:

$$\hat{\theta} = \arg \max_{\theta} \mathcal{L}(\theta, \mathbf{X}, \Delta_{\mathbf{X}, \mathbf{Z}}^*). \quad (3)$$

3 Линейная корреляция Пирсона

В этом разделе описываются методы построения матрицы смежности через оценку взаимосвязи между временными рядами соответствующие электродам. Мы рассматриваем построение матрицы смежности на основе корреляции пирсона и синхронизации фаз сигналов.

3.1 Корреляция пирсона

Обозначим $\mathbf{x} = \mathbf{x}_{mi}$ и $\mathbf{y} = \mathbf{x}_{mj}$ строки матрицы \mathbf{X}_m , соответствующие сигналам в отрезке времени $[t_n - T_w, t_n]$ в m -ом испытании для i и j электрода. Коэффициент корреляции задается следующим образом

$$\tilde{r}_{ij}(m, t_n) = \frac{\sum_{k=t_n-T_w}^{t_n} (x_k - \bar{\mathbf{x}})(y_k - \bar{\mathbf{y}})}{\sqrt{s_x^2 s_y^2}},$$

где $x_k = (\mathbf{x})_k$, $y_k = (\mathbf{y})_k$,

$\bar{\mathbf{x}}, \bar{\mathbf{y}}, s_x^2, s_y^2$ — выборочное среднее и дисперсия сигналов на i и j электроде соответственно. Таким образом получаем:

$$\tilde{r}_{ij}(m, t_n) = \frac{\sum_{k=t_n-T_w}^{t_n} (x_k - \bar{\mathbf{x}})(y_k - \bar{\mathbf{y}})}{\sqrt{\sum_{k=t_n-T_w}^{t_n} (x_k - \bar{\mathbf{x}})^2 \sum_{k=t_n-T_w}^{t_n} (y_k - \bar{\mathbf{y}})^2}}$$

Матрица связности определяется как:

$$A_{\mathbf{X}, \mathbf{Z}}^*(m, t) = [a_{ij}(m, t)] \in \mathbb{R}_+^{E \times E}, \quad a_{ij}(m, t) = \begin{cases} r_{ij}(m, t), & \text{если } r_{ij}(m, t) \geq \rho(p) \\ 0, & \text{иначе,} \end{cases} \quad (4)$$

где $r_{ij}(m, t) = |\tilde{r}_{ij}(m, t)|$.

3.2 Синхронизация фаз сигналов

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, \mathbf{x}_{mi} и \mathbf{x}_{mj} в отрезке времени $[t_n - T_w, t_n]$ в m -ом испытании. Синхронизация фаз понимается как:

$$|\varphi_x(t) - \varphi_y(t)| = \text{const.} \quad (5)$$

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

$$H(t) = x(t) + i\dot{x}(t), \quad \text{где} \quad (6)$$

$$\dot{x}(t) = \frac{1}{\pi} \text{v.p.} \int_{-\infty}^{\infty} \frac{x(t')}{t - t'} dt' - \text{преобразование Гильберта сигнала } x(t), \quad (7)$$

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

Фаза аналитического сигнала определяется как:

$$\varphi(t) = \arctan \left(\frac{\dot{x}(t)}{x(t)} \right). \quad (8)$$

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) [3] задается уравнением:

$$p_{ij}(m, t_n) = \left| \frac{1}{T_w} \sum_{k=1}^{T_w} \exp(s(\varphi_x(k\Delta t) - \varphi_y(k\Delta t))) \right|, \quad (9)$$

где Δt — шаг по времени, $s = \sqrt{-1}$.

Матрица связности определяется как:

$$\mathbf{A}_{\mathbf{x}, \mathbf{z}}^*(m, t) = [a_{ij}(m, t)] \in \mathbb{R}_+^{E \times E}, \quad a_{ij}(m, t) = \begin{cases} p_{ij}(m, t), & \text{если } p_{ij}(m, t) \geq \rho(p) \\ 0, & \text{иначе.} \end{cases} \quad (10)$$

4 Модель классификации

Для решения проблем классификации мы предлагаем использовать модель DCGRU [6] хорошо показавшую себя в решении задачи классификации эпилептический припадков по ЭЭГ в работе [10]. Мы считаем, что использование диффузии позволит учитывать отдаленные вершины графа, что улучшить качество классификации, а так же сделает модель более устойчивой к шуму, что важно поскольку ЭЭГ данные крайне индивидуальны.

Для моделирования диффузии на графе используется спектральная свертка:

Спектральная свертка на графе определяется как:

$$X_{:,p} \star_{\mathcal{G}} f_{\theta} = \Phi F(\theta) \Phi^T X_{:,p}$$

где:

- $L = \Phi \Lambda \Phi^T$ — спектральное разложение лапласиана графа
- $F(\theta) = \sum_{k=0}^{K-1} \theta_k \Lambda^k$ — полиномиальный фильтр
- p — индекс признака вершины

Данная свёртка эквивалентна диффузионной свёртке на графе с точностью до преобразования подобия, когда граф \mathcal{G} является неориентированным [6].

$$\begin{aligned} r^{(t)} &= \sigma(\Theta_r \star_g [X^{(t)}, H^{(t-1)}] + b_r) \\ u^{(t)} &= \sigma(\Theta_u \star_g [X^{(t)}, H^{(t-1)}] + b_u) \\ C^{(t)} &= \tanh(\Theta_C \star_g [X^{(t)}, (r^{(t)} \odot H^{(t-1)})] + b_c) \\ H^{(t)} &= u^{(t)} \odot H^{(t-1)} + (1 - u^{(t)}) \odot C^{(t)} \end{aligned}$$

где:

- $X^{(t)}, H^{(t)}$ — вход и выход на временном шаге t
- $r^{(t)}, u^{(t)}$ — вентиль сброса и вентиль обновления на шаге t соответственно
- \star_g — оператор диффузионной свёртки
- $\Theta_r, \Theta_u, \Theta_C$ — параметры соответствующих фильтров

5 Признаковое описание

В качестве признаков для модели используется дифференциальная энтропия выделенных диапазонов ритмов головного мозга:

- дельта (1–3 Гц)
- тета (4–7 Гц)
- альфа (8–13 Гц)
- бета (14–30 Гц)
- гамма (31–50 Гц)

Формула дифференциальной энтропии:

$$DE(Y) = - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \log \left(\frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \right) dy \quad (11)$$

где $Y \in \mathcal{N}(\mu, \sigma^2)$ — временной ряд.

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

6 План вычислительного эксперимента:

Гипотеза. Учет пространственной и функциональной структуры сигнала, а также использование диффузионных методов повышает качество решения задачи классификации человеческих эмоций.

Цели эксперимента.

- 1) Построить матрицы связей электродов исследуемыми методами.
- 2) Оценить качество работы предложенной пространственно- временной модели на основе полученных оценок матрицы.

В исследовании использовался набор данных [5] для изучения аффективных состояний человека. В эксперименте приняли участие 15 испытуемых, соответствующих необходимым медицинским требованиям. Все участники подписали информированное согласие и были ознакомлены с протоколом исследования.

В качестве визуальных стимулов применялись видеоролики, отобранные из 6 различных категорий. Критерии выбора видеофрагментов включали:

- ограничение общей продолжительности эксперимента для минимизации утомления участников;
- понятность содержания без дополнительных объяснений;
- способность вызывать чётко определённую целевую эмоцию.

Длительность каждого видеофрагмента составляла приблизительно 4 минуты. Материалы были специально отредактированы для создания последовательного эмоционального воздействия с максимальной выразительностью.

Эксперимент проводился в 3 серии по 15 попыток в каждой. Порядок демонстрации видео был организован таким образом, чтобы фрагменты, направленные на одну и ту же эмоцию, не следовали друг за другом. Для сбора обратной связи участники заполняли опросные листы с описанием пережитых эмоций.

Регистрация ЭЭГ проводилась с использованием 62 электродов, размещённых согласно международной системе 10–20. Исходная частота дискретизации сигнала составляла 1 кГц.

На этапе предварительной обработки данных применялись следующие процедуры:

- фильтрация сигнала полосовым фильтром (диапазон 0.3–50 Гц) для устранения шумов и артефактов
- понижение частоты дискретизации до 200 Гц

7 Результаты



Рис. 1 Построение матрицы на основе корреляции Пирсона, признаки значения сигнала



Рис. 2 Построение связи на основе синхронизации фаз, признаки дифференциальная энтропия фаз

Как видно, значение потерь во втором подходе падает значительно быстрее, что говорит об его эффективности по сравнению с первым.

8 Дальнейшие планы

В дальнейшем планируется рассмотреть еще способы построения графа, а также протестировать модель

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] Sergul Aydore, Dimitrios Pantazis, and Richard M. Leahy. A note on the phase locking value and its properties. *NeuroImage*, 74:231–244, July 2013.
- [4] 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.
- [5] Ruo-Nan Duan, Jia-Yi Zhu, and Bao-Liang Lu. Differential entropy feature for eeg-based emotion classification. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, page 81–84. IEEE, November 2013.
- [6] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting, 2017.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.