

Cross-attention and CCA in EEG

Eduard Vladimirov, Daniil Kazachkov, Vadim Strijov

25 ноября 2024 г.

1 Abstract

На сегодняшний день работа с мультимодальными данными набирает всё большую популярность: учет взаимосвязей между ними улучшает качество предсказания. В этой статье мы предлагаем новую архитектуру, использующую преимущества алгоритма Canonical Correlation Analysis (CCA) и механизма Attention. Ниже будет показано, что CCA - частный случай Attention, а значит, мультимодальность можно встроить внутрь фреймворка Attention. Работа полученной модели иллюстрируется на задаче классификации удара теннисного мяча по датасету Real World Table Tennis.

keywords : CCA, Attention, BCI, online-classification.

2 Introduction

Мультимодальность - мощный инструмент для улучшения качества ответов модели [What Makes Multi-modal Learning Better than Single]. Канонический корреляционный анализ (CCA) [A Survey on Canonical Correlation Analysis] является очень популярным статистическим методом, снижения размерности двух множеств данных, при котором корреляция между парными переменными в общем подпространстве взаимно максимизируется. В таких работах как [Correlational Neural Networks], [A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets] авторы показали, что он улучшает качество в задачах сопоставления событий. Однако CCA может моделировать лишь линейные зависимости.

В противоположность этому, механизм Attention находит сложные,

нелинейные зависимости. Усовершенствование cross-attention позволит лучше отсеивать информацию, снизит размерность пространства и тем самым ускорит вычисления.

Ставя перед собой цель использовать преимущества каждого метода, мы представляем модель CCT: Canonical-Correlation Transformer. Архитектура у нее следующая: из пакета PyRiemann (мб заменим на CNN EEGNet) [ссылка] мы берем энкодер и преобразуем поданный на вход ЭЭГ сигнал в скрытое пространство. Далее в качестве механизма внимания используем ... (дополнить, когда станет понятно). Выход модели - вероятность события принадлежать одному из четырех классов: попадание по мячу (1), нейтральное событие (0), промах по мячу (-1), попадание, после которого успешных ударов не было (2).

Наша задача: как по данным ЭЭГ игрока в настольный теннис в режиме реального времени классифицировать момент удара - типичный пример из области Brain-Computer Interface (BCI), когда необходимо эффективно работать с данными разных модальностей и классифицировать события в онлайн режиме. В качестве датасета мы взяли преобработанные данные из "Real World Table Tennis"[1].

2.1 Related Works

- про нейрокомпьютерный интерфейс
- про онлайн-классификацию
- мб про аналогичные оффлайн-задачи

Нейрокомпьютерный интерфейс (Brain-Computer Interface, BCI) [2] считывает сигналы поверхности кортекса головного мозга, анализирует и переводит в команды исполняющей системы. Результатом измерений является временной ряд напряжений на электродах, который используется в задаче декодирования сигнала.

3 CCT: Canonical-Correlation Transformer

3.1 Model

Мы продолжаем расширять область применимости CCT и представляем способ встраивания его в Attention.

- отдельно CCA
- отдельно Attention
- связь CCA + Attention (у Эдуарда из файла взять)

4 Experiments

[какие-то общие слова, если нужны] Например, про то, что пробо-
вали PyRiemann и трансформер из braincode.

4.1 Dataset Details

- что за датасет
- какая предобработка данных проводилась
- в каком виде данные подавались в модель

4.2 Training Details

Если возникнут какие-то проблемы или эвристики при обучении,
то пишем сюда

- На каких параметрах обучали сетку,

4.3 Experimental Results

Получилась вот такая точность и почему.

Список литературы

- [1] Amanda Studnicki and Daniel P. Ferris. Dual-layer electroencephalography data during real-world table tennis. *Data in Brief*, 52:110024, 2024.
- [2] Jerry J Shih, Dean J Krusienski, and Jonathan R Wolpaw. Brain-computer interfaces in medicine. In *Mayo Clinic Proceedings*, volume 87, pages 268–279. Elsevier, 2012.
- [3] Nick Martin and Hermine Maes. Multivariate analysis. *London, UK: Academic*, 1979.
- [4] Zhongkai Sun, Prathusha Sarma, William Sethares, and Yingyu Liang. Learning relationships between text, audio, and video via deep canonical correlation for multimodal language analysis. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 8992–8999, 2020.
- [5] Yu-Ting Lan, Wei Liu, and Bao-Liang Lu. Multimodal emotion recognition using deep generalized canonical correlation analysis with an attention mechanism. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6. IEEE, 2020.
- [6] Yu Zhang Ehsan Adeli Qingyu Zhao Kilian M. Pohl Yixin Wang, Wei Peng. Brain-cognition fingerprinting via graph-gcca with contrastive learning. 2024.
- [7] Engin Erzin Ibrahim Shoer, Berkay Kopru. Role of audio in audio-visual video summarization. In *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*. IEEE, 2023.
- [8] Bai Chenyu Pan Jiahui. Eeg-based emotion recognition via convolutional transformer with class confusion-aware attention. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 46. IEEE, 2024.