A Study on Graph Transformer-based Network for Brain-Computer Interface

Permana Deny, Kae Won Choi
Sungkyunkwan Univ.
denypermana@g.skku.edu, kaewonchoi@skku.edu

Graph Transformer 기반의 Brain-Computer Interface 연구

데니퍼마나, 최계원
Sungkyunkwan Univ.

Abstract

In this paper, we propose a graph transformer-based network for classifying electroencephalogram (EEG)-based motor imagery (MI) dataset. To design the classification algorithm, we represent the spatial information of the EEG channel location on a graph structure and learn through a transformer-based graph network. We evaluate the algorithm on a famous MI dataset and show significant performance improvement.

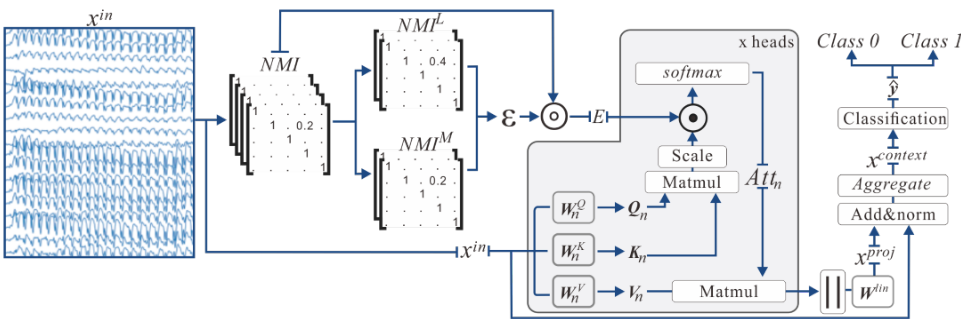


Figure 1. Design of graph representation and transformer-based network.

Ⅰ. Introduction

Currently, some studies utilize a deep learning network for designing the MI decoding algorithm. Among the study, there are two approaches that are widely explored, i.e., emphasizing either spatial or temporal feature through different deep learning network using convolutional neural network (CNN). However, CNN is known to have their own limitations on its spatial kernel. The kernel induces spatial invariance, which means that CNN all spatial locations in the same manner. Therefore, the expressive power that differs feature from different label in the channel characteristics cannot be properly represented.

In addition, some research emphasize structural-based representations such as graph-representation learning by connecting some channels through edges to farther channels as it may provide useful information propagation. This hypothesis is based on the general term of contralateral for the motoric cortex of the brain controlling the human body side.

Herein, we propose a graph representation for MI dataset using ranked-feature correlation with graph transformer-based network. The algorithm consists of two stages, i.e., graph representing stage, and graph learning stage.

Ⅱ. Methodology

A. Graph Representation

Our design for constructing graph and the classification algorithm is depicted on Fig. 1. To calculate the adjacency, we begin with calculating neighborhood mutual information (NMI) between two nodes, e.g. i and j, by measuring the feature distance of the mutual information between two different classes. The NMI with farther distance between classes can be utilized to represent the graph as it may improve the expressive power between graphs. The NMI is defined by

$$NMI_{i,j} = H(x_i) + H(x_j) - H(x_i, x_j) \tag{1}$$

The distance is then computed by splitting the train dataset into two classes, i.e., L and M, by

$$\epsilon = \begin{cases} 1, \text{if } \frac{1}{K}\sum_{l=1}^{L}\sum_{m=1}^{M} \| NMI_{i,j}^{l} - NMI_{i,j}^{m} \| \geq C_{threshold} \\ 0, \text{if } \frac{1}{K}\sum_{l=1}^{L}\sum_{m=1}^{M} \| NMI_{i,j}^{l} - NMI_{i,j}^{m} \| \leq C_{threshold} \end{cases} \tag{2}$$

Finally, the graph structure is represented by an adjacency matrix by

$$E = NMI \odot \epsilon \tag{3}$$

The next stage is designing the graph transformer-based network. Here, we propose a novel expression to calculate the self-attention weight by considering the edge feature E during the query session. The module consists of multi-head self-attention which consists of N_{head} module. Each module will calculate self-attention weight independently from D_{head} feature dimension through

$$※_n = softmax\left(\frac{Q_n \bullet K_n}{\sqrt{D_{head}}} \bullet E\right) \quad (4)$$

In the end, we derive the context tensors that are the learning result from all module that represents the expressive power as in Fig. 1. The expressiveness represents the two graphs are different when both are truly different. The graph from the same label will tend to present similar expression. And the different label will generate different expression. In order to summarize the expressiveness, we use an aggregation with summation over the last dimension, as

$$\hat{y} = softmax(Aggregate(x^{trans})W^{cls}) \quad . \quad (5)$$

Ⅲ. Result and Discussion

A. Dataset

In this study, we utilize the dataset from the moabb, which is an open source project to build benchmarks for popular BCI algorithms applied to freely available EEG datasets. Among the various datasets, we chose the Lee2019 [1] datasets. The dataset has 62 nodes with 54 subjects and two class labels.

B. Result

The overall result is shown in Table 1, and it shows that our algorithm is superior amongst the non-graph and graph representation learning. We believe the improvement proves and supports our hypothesis. Representing the MI dataset as a graph representation with our strategy is relevant to the concept of neuroscience. As the result in dataset Lee2019 with all previous study obtain significantly lower performance than our result.

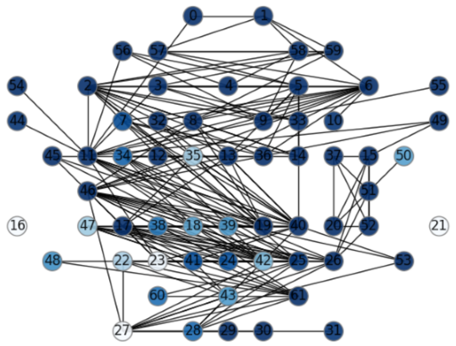
Table 1. Classification accuracy of the proposed method compared to previous studies.

Dataset	Method	Subjects	Acc.
Lee2019	Deep-CNN [1]	54	0.713
	PSPD [2]	54	0.732
	FBCNet [3]	54	0.747
	Ours	54	**0.768**

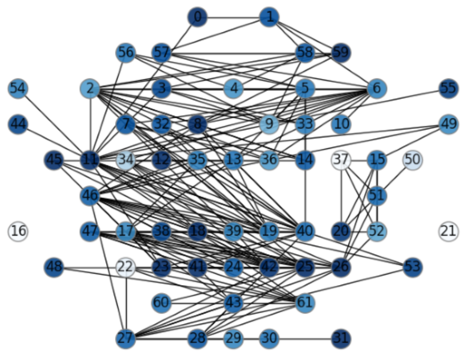
B. Graph Expressive Power

The expressiveness can be captured through a gradient class activation function (Grad-CAM) [4]. Our visualization for the interpretability is depicted in Fig. 2. The figure includes selected edges connecting the nodes and node weights. Lets observe the node edge that varies among the training dataset. Although the edge does not provide interpretability, it is interesting to see that the node that located in the middle of y-axis has higher node degree. It correlates to the position of brain motoric cortex that regulates the motoric activity. It indicates that our algorithm form graph representation that retain high distinguishable feature and remove unimportant feature such feature from the nodes that do not cover or close to motoric cortex area.

On the other hand, we depict the node weight in color scale, where the lower intensity indicates lower weight. Generally, the trends show that two classes depict the reverse pattern, where class 0 has higher weighted nodes activated on the left side and class 1 on the right side. These trends underpin the contralateral of the brain in controlling the human body.



(a)



(b)

Figure 2. Node weights with selected edges from Lee2019 dataset (a) class label 0, (b) class label 1. Color intensity represents the weight from class activation mapping.

Acknowledgement

This work was supported by the BK21 FOUR Project.

Reference

[1] O.-Y. Kwon et al, "Subject-independent Brain - Computer Interfaces based on deep convolutional neural networks," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 10, pp. 3839 - 3852, Nov. 2019.

[2] Daemen, J., and Rijmen, V. "AES Proposal: Rijndael, Version2.," Submission to NIST, March 1999.

[3] R. Mane, et al., "FBCNet: A multi-view convolutional neural network for Brain-Computer Interface," 2021. [Online]. Available: arXiv:2104.01233

[4] R. R. Selvaraju, et al., "Grad-cam: Visual explanations from deep networks via gradient-based localization", Proc. ICCV'17, Venice, Italy, October 2017, pp. 618 - 626.