ADAPTIVE BRAIN NETWORK AUGMENTATION BASED ON GROUP-AWARE GRAPH LEARNING

Ciyuan Peng[1], **Mujie Liu**[1], **Chenxuan Meng**[2], **Shuo Yu**[3]*, **Feng Xia**[4]
[1]Federation University Australia, Ballarat, Australia [2]Zhejiang Gongshang University, Hangzhou, China
[3]Dalian University of Technology, Dalian, China [4]RMIT University, Melbourne, Australia

ABSTRACT

Brain network analysis significantly improves artificial intelligence techniques in the realm of digital health. Most existing methods uniformly construct brain networks for different groups (e.g., male and female groups, healthy and sick people groups), facing the interference of group-irrelevant noises and failing to capture group-specific features to enhance brain networks. To address this issue, this paper proposes an adaptive brain network augmentation method based on group-aware graph learning. We construct group-aware brain networks, which can adapt to distinct groups, reducing the interference of noises, and improving model robustness across various tasks and subject groups.

1 INTRODUCTION

Brain network analysis stands as a crucial research topic in digital health, aiming to comprehend the complex functional and structural organization of the nervous system and assist in the diagnosis of brain diseases (Kan et al., 2022; Cui et al., 2023; Wen et al., 2023). Most current methods adopt graph learning techniques to uniformly model brain networks, in which nodes are Regions Of Interest (ROIs) and each edge represents the correlation between two ROIs (Zhao et al., 2022; Li et al., 2023; Kim et al., 2023). In particular, many existing brain network construction approaches generally extract complex information of brain regions to simply measure either brain structure or function for different brain analysis tasks. However, since the human brain can execute a diverse range of behaviors and is variable among different groups, brain networks manifest specific traits in different conditions (Krienen et al., 2014; Cole et al., 2014; Jiang et al., 2020; Chen et al., 2022). Consequently, general brain networks built on a broad population face challenges in adapting to different groups and task scenarios, and are susceptible to noise interference.

This paper hypothesizes that distinct groups have different brain network features. To prevent noise interference, such as group-irrelevant features, we propose a novel adaptive brain network augmentation method based on group-aware graph learning. Specifically, this paper aims to adaptively construct brain networks for distinct groups, reducing noise features and enhancing group features, therefore improving model robustness when it is applied to various tasks for different subject groups. Our group-adaptive brain network augmentation strategy contains two main steps, including communal attribute dropping and group feature distillation. We first drop communal features, which commonly exist across various groups and are regarded as group-irrelevant noises. Then, we design a group feature distillation method to specifically extract the group features of each group, exploring group-aware brain networks.

2 GROUP-ADAPTIVE BRAIN NETWORK AUGMENTATION STRATEGY

The overall framework of our proposed adaptive brain network augmentation method is shown in Figure 2. To improve noise resistance, the input brain signals, extracted from functional magnetic resonance imaging (fMRI) data, first encounter a process of common attribute dropping. Building upon this, we further extract the group-specific features through the method of group feature distillation, aiming to construct group-aware brain networks. The detailed algorithms are described in Appendix A.

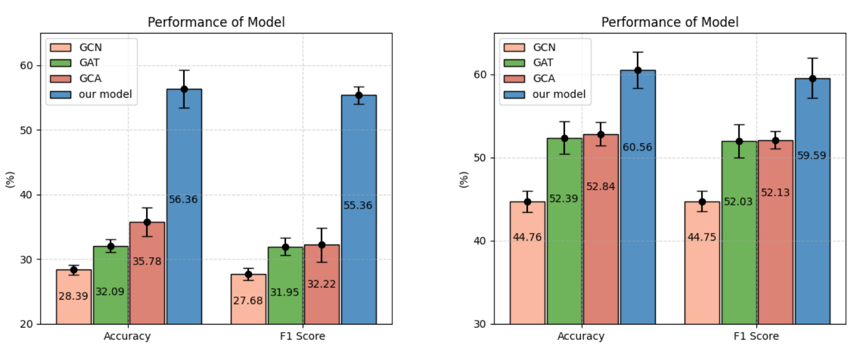
*Corresponding Author: Shuo Yu (email: shuo.yu@ieee.org).

Communal attribute dropping. We begin with a set of brain signal data $\mathbf{M} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_D\}$, where $\mathbf{X}_d \in \mathbb{R}^{N \times U}$ denotes the feature matrix of the d-th subject, N and U are the number of signal channels (nodes) and attributes, respectively. Then, the communal attributes among the given dataset are extracted by the Singular Value Decomposition (SVD). Particularly, we adopt an SVD-based scoring function to score the group commonality of each attribute, defined as $\mathbf{S}_c = \phi(\mathbf{M})$. Afterwards, we set a threshold to sort out the scores, selecting both high-scoring and low-scoring dimensions as communal attributes. In order to eliminate the group-irrelevant noises, we drop these selected communal attributes from the original data and obtain denoising feature matrices. The updated feature matrix of the d-th subject is represented as $\hat{\mathbf{X}}_d \in \mathbb{R}^{N \times U'}$, where U' represents the number of attributes after dimensionality reduction (communal attribute dropping).

Group feature distillation. The aforementioned process eliminates group-irrelevant attributes for each subject. Group feature distillation primarily focuses on enhancing the brain network representations with group-specific features, thereby improving the ability of model to learn brain networks. Assume the input data can be divided into K groups according to subject characteristics, such as gender and brain activity states, denoted as $\hat{\mathbf{M}} = \{\hat{\mathbf{M}}_k\}_{k=1}^{K}$, where $\hat{\mathbf{M}}_k = \{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \ldots, \hat{\mathbf{X}}_{d'}\}$, representing d' subjects are classified into group k. Then, we adopt the feature scoring function on each group separately to obtain group feature scoring matrices, formulated as $\mathbf{S}_k = \phi(\hat{\mathbf{M}}_k)$. Afterwards, an adaptive brain network augmentation function $f(\hat{\mathbf{X}}_d, \mathbf{S}_k)$, where $\hat{\mathbf{X}}_d \in \hat{\mathbf{M}}_k$, is introduced to enhance the brain networks. Finally, group-aware brain networks $\{G_1, G_2, \ldots, G_D\}$ can be obtained for further analysis.

3 Experiment

We conduct experiments on the Alzheimer's Disease Neuroimaging Initiative (ADNI)[1] dataset and the Autism Brain Imaging Data Exchange (ABIDE)[2] dataset for graph classification task. The statistical information of datasets is provided in Appendix B.1. Figure 1 compares the F1 score and accuracy of our model with baselines. The detailed experimental setup and results are discussed in Appendix B.2 and B.3.



(a) Results on ADNI

(b) Results on ABIDE

Figure 1: The experimental results on two datasets.

4 Conclusion

This paper proposes a new group-adaptive brain network augmentation strategy, constructing group-specific brain networks, as well as enhancing group features and eliminating the group-irrelevant noises. The experiment results demonstrate that our method can precisely model group-aware brain networks, showing great potential in improving model robustness across various groups and tasks[3].

[1] `https://adni.loni.usc.edu/`

[2] `https://fcon_1000.projects.nitrc.org/indi/abide/`

[3] The code of this paper is provided at `https://github.com/pcyyyy/GroupBNA`.

ACKNOWLEDGEMENTS

This work is supported by Dalian Major Projects of Basic Research (grant No. 2023JJ11CG002).

URM STATEMENT

The authors acknowledge that at least one key author of this work meets the URM criteria of ICLR 2024 Tiny Papers Track.

REFERENCES

Jianzhong Chen et al. Shared and unique brain network features predict cognitive, personality, and mental health scores in the abcd study. *Nature Communications*, 13(1):2217, 2022.

Michael W Cole et al. Intrinsic and task-evoked network architectures of the human brain. *Neuron*, 83(1):238–251, 2014.

Hejie Cui et al. Braingb: A benchmark for brain network analysis with graph neural networks. *IEEE Transactions on Medical Imaging*, 42:493–506, 2023.

Rongtao Jiang et al. Task-induced brain connectivity promotes the detection of individual differences in brain-behavior relationships. *NeuroImage*, 207:116370, 2020.

Xuan Kan et al. Brain network transformer. *Advances in Neural Information Processing Systems*, 35:25586–25599, 2022.

Sein Kim et al. Heterogeneous graph learning for multi-modal medical data analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 5141–5150, 2023.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations (ICLR)*, 2017.

Fenna M Krienen et al. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1653):20130526, 2014.

Gaotang Li et al. Interpretable sparsification of brain graphs: Better practices and effective designs for graph neural networks. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1223–1234, 2023.

Aaron van den Oord et al. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

Patrick Schober et al. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5):1763–1768, 2018.

Petar Velickovic et al. Graph attention networks. In *6th International Conference on Learning Representations (ICLR)*, 2018.

Guangqi Wen et al. Graph self-supervised learning with application to brain networks analysis. *IEEE Journal of Biomedical and Health Informatics*, 27(8):4154–4165, 2023.

Xusheng Zhao et al. Deep reinforcement learning guided graph neural networks for brain network analysis. *Neural Networks*, 154:56–67, 2022.

Yanqiao Zhu et al. Graph contrastive learning with adaptive augmentation. In *Proceedings of the Web Conference 2021*, pp. 2069–2080, 2021.

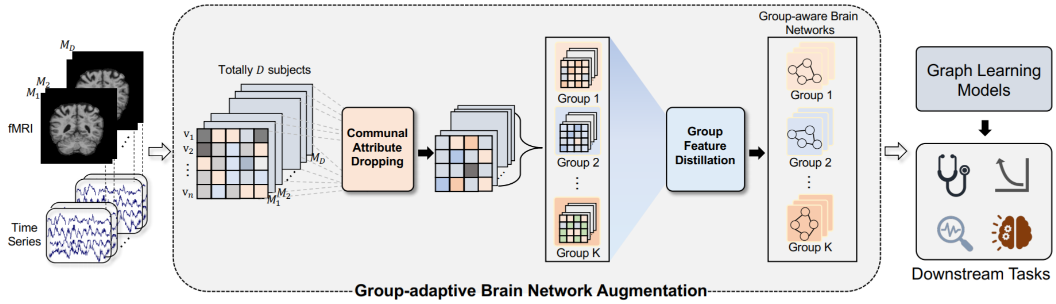


Figure 2: The framework of our proposed adaptive brain network augmentation method.

A APPENDIX-ALGORITHM DETAILS

A.1 COMMUNAL ATTRIBUTE DROPPING

To extract communal attributes across all subjects, we primarily introduce the SVD-based scoring function $\phi(\cdot)$. The first step in the scoring function is to rearrange the signal values within each channel (node) for all D subjects, resulting in a set of new signal matrices $\{\tilde{\mathbf{X}}^1, \tilde{\mathbf{X}}^2, \ldots, \tilde{\mathbf{X}}^N\}$, where $\tilde{\mathbf{X}}^n \in \mathbb{R}^{D\times U}$. Afterwards, we apply SVD to the obtained signal matrices to score the attributes on each channel. We use the first column of the left singular matrix calculated by SVD as the scoring criterion for the group commonality of each attribute, the score set of signal channel i is denoted as $\mathbf{s}^i = [s_1^i, s_2^i, \ldots, s_U^i]^\mathsf{T}$, where s_U^i is the score of U-th dimension. Finally, we can obtain a communal attribute scoring matrix $\mathbf{S}_c = [\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^N]^\mathsf{T} = \phi(\mathbf{M})$ for all signal channels (nodes). For better understanding, we provide the procedure of the SVD-based scoring function (shown in Algorithm 1). Afterwards, we differentiate communal attributes from individual attributes by setting a threshold λ, and select a subset of attributes on both high and low-scoring dimensions as communal attributes. Then, we achieve denoising and dimensionality reduction by discarding the dimensions of these communal attributes to all brain signals. For d-th subject, we denote the feature matrix as $\hat{\mathbf{X}}_d \in \mathbb{R}^{N\times U'}$, where U' can be controlled by the setting of threshold λ.

Algorithm 1: SVD-based scoring function

Input: Brain signal matrices of all subjects $\mathbf{M} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_D\},\ \mathbf{M} \in \mathbb{R}^{D\times N\times U}$, where $\mathbf{X}_d \in \mathbb{R}^{N\times U}$ is the signal matrix of the d-th subject, N is the number of channels (nodes), and U is the number of attributes

Output: Scoring matrix $\mathbf{S}_c$

Set a node-level tensor $\mathbf{O} = \{\tilde{\mathbf{X}}^1, \tilde{\mathbf{X}}^2, \ldots, \tilde{\mathbf{X}}^N\},\ \mathbf{O} \in \mathbb{R}^{N\times D\times U}$, where $\tilde{\mathbf{X}}^i \in \mathbb{R}^{D\times U}$ is the signal matrix of channel i.

for $i = 1$ **to** D **do**
    for $j = 1$ **to** N **do**
        for $z = 1$ **to** U **do**
            $\mathbf{O}[j, i, z] = \mathbf{M}[i, j, z]$;

for $i = 1$ **to** N **do**
    $\mathbf{S}'_c = \mathrm{SVD}(\mathbf{O}[i, :, :])$;
    for $j = 1$ **to** U **do**
        $\mathbf{S}_c[j, i] = \mathbf{S}'_c[j, 1]$;

return $\mathbf{S}_c$

A.2 GROUP FEATURE DISTILLATION

Based on communal attribute dropping, we can obtain a set of denoising signal matrices $\{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \ldots, \hat{\mathbf{X}}_D\}$, where $\hat{\mathbf{X}}_d \in \mathbb{R}^{N\times U'}$. In this step, we aim to enhance the group-specific features of each brain signal matrix from different groups. These denoising signal matrices are classified into K groups, represented as $\hat{\mathbf{M}} = \{\hat{\mathbf{M}}_k\}_{k=1}^{K}$. In order to extract the features within each group, we apply SVD-based scoring function group by group and conclude the scoring matrix of each group, respectively. For group k, the group feature scoring matrix is $\mathbf{S}_k = \phi(\hat{\mathbf{M}}_k)$. According to the scores assigned to each attribute, our aim is to enhance the values of high-scoring attributes while reducing those of low-scoring attributes, achieving intra-group enhancement of signal features. Thereby, we introduce a feature augmentation function, formulated as:

$$\begin{aligned} f(\hat{\mathbf{X}}_d, \mathbf{S}_k) &= \hat{\mathbf{X}}_d \odot \hat{\mathbf{S}}_k,\ \hat{\mathbf{X}}_d \in \hat{\mathbf{M}}_k, \\ \hat{\mathbf{S}}_k &= \left[\alpha \Big/ \min \left(\frac{\ln s_j^{max} - \ln s_j^{u}}{\ln s_j^{max} - \ln s_j^{\mu}}, s_l \right) \right]_{u=1,j=1}^{u=U',j=N}, \end{aligned} \tag{1}$$

where $\odot$ is element-wise multiplication, α is a hyperparameter to adjust the magnitude of feature augmentation, $\mathbf{s}_j^{max}$ and $\mathbf{s}_j^{\mu}$ is the maximum and the average scores of all dimensions in channel j, and s_l is the limitation of the lowest value. Afterwards, for all the subjects, we can obtain the augmented brain signal matrices. Finally, we construct brain networks based on the augmented signal data by using Pearson's rule (Schober et al., 2018), denoted as $\{G_1, G_2, \ldots, G_D\}$. We consider the signal channels as nodes, and the relations between them are regarded as the edges. The procedure of our group-adaptive brain network augmentation strategy is shown in Algorithm 2.

Afterwards, the constructed brain networks are input into a graph learning model for downstream tasks. We first use a two-layer graph convolution network to encode the input data. Then, we adopt a contrastive loss function, InfoNCE loss (Oord et al., 2018), as the training objective for graph classification task.

Algorithm 2: Group-adaptive brain network augmentation strategy

Input: Brain signal matrices of all subjects $\mathbf{M} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_D\},\ \mathbf{X}_d \in \mathbb{R}^{N \times U}$
Output: Brain networks $\{G_1, G_2, \ldots, G_D\}$
Set threshold value λ and hyperparameter α
Calculate the commual attribute scoring matrix $\mathbf{S}_c = \phi\{\mathbf{M}\}$
Drop the communal attributes obtained by comparing $\mathbf{S}_c[i, j]$ with λ and form the reduced matrices $\{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \ldots, \hat{\mathbf{X}}_D\}$
Classify $\{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \ldots, \hat{\mathbf{X}}_D\}$ into K groups, $\hat{\mathbf{M}} = \{\hat{\mathbf{M}}_k\}_{k=1}^{K}$
Calculate group feature scoring matrix for each group $\mathbf{S}_k = \phi\{\hat{\mathbf{M}}_k\}$
Obtain augmented signal matrix for each subject based on $\mathbf{S}_k$ (Eq. (1))
Construct group-aware brain networks $\{G_1, G_2, \ldots, G_D\}$ based on Pearson's rule
return $\{G_1, G_2, \ldots, G_D\}$

B APPENDIX-EXPERIMENTAL ANALYSIS

B.1 DATASET

We collect fMRI data from ADNI and ABIDE datasets. The detailed information of collected ADNI dataset is summarized in Table 1. The dataset used in this study is categorized into 3 groups, including Normal Control (NC), Alzheimer's Disease (AD), and Mild Cognitive Impairment (MCI). Mini-Mental State Examination (MMSE) scores range from 0 to 30, with higher scores indicating better cognitive function. Clinical Dementia Rating (CDR) is commonly used to assess the severity of dementia. For ABIDE dataset, we collect the fMRI data of 175 subjects, including 75 autism patients and 100 normal controls. Therefore, the collected ABIDE dataset can be categorized into 2 groups.

Table 1: Clinical and demographic information of subjects in ADNI dataset.

Group	Number	Age Range	Age Statistics	MMSE	CDR
NC	211	57-93	72.8±8.3	28.9±1.7	0.2±0.8
MCI	195	49-96	72.8±7.9	27.6±2.2	1.6±1.2
AD	54	55-89	75.5±7.0	22.4±2.8	4.7±2.0

B.2 Experimental Setup

B.2.1 Implementation Details

Our model is implemented using PyTorch Geometric v2.4.0 and PyTorch v1.9.1. Model training is performed on an NVIDIA A40 GPU with 24GB of memory. The model parameters are trained using the Adam SGD optimizer. We train our model for 200 epochs, with a learning rate of 1e-3. A weight decay of 1e-5 is applied for regularization. Data is randomly split, with 40% used for training, 40% for validation, and 20% for testing.

B.2.2 Baselines

We compare our model with several representative baseline methods, including graph convolution network (GCN) (Kipf & Welling, 2017), graph attention network (GAT) (Velickovic et al., 2018) and graph contrastive learning with adaptive augmentation (GCA) (Zhu et al., 2021). GCN leverages graph convolutional layers to capture and propagate information through graph nodes, effectively modeling relations between nodes. GAT uses attention mechanisms to dynamically weigh the contributions of neighboring nodes during the aggregation, allowing the model to selectively focus on relevant information in the graph. GCA employs adaptive data augmentation strategies to create informative positive and negative node pairs for contrastive learning, promoting better graph embeddings through self-supervised learning.

B.3 Experimental Results

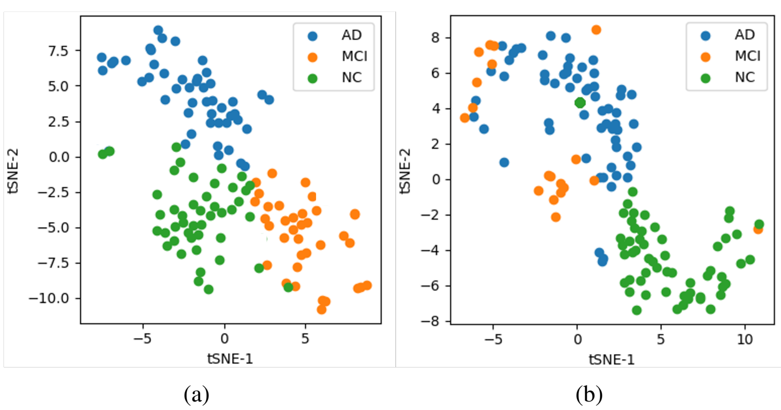


Figure 3: (a) visualizes the brain network feature distribution obtained by our model, while (b) represents the result of GCA model.

As illustrated in Figure 1, our model exhibits significant performance improvements compared to the baselines. Specifically, on the ADNI dataset, the accuracy of our model increases from 20.58% to 27.97%, and on the ABIDE dataset, it improves from 7.72% to 15.8%. Furthermore, our model demonstrates higher F1 scores, showing an improvement of about 23.14% to 27.68% on the ADNI dataset and 7.46% to 14.84% on the ABIDE dataset, in comparison to the baselines. In general, our method shows high adaptability when being applied to various groups and high scalability when handling different tasks, as well as effectively overcoming noise interference.

We also use the visualization tool t-SNE to visualize the brain network feature distribution of ADNI dataset. The visualization results are illustrated in Figure 3. In comparison to GCA, which lacks the group-adaptive brain network augmentation strategy, our model distinctly separates brain networks from different groups, forming 3 distinct clusters. These findings further validate the outstanding performance of our model in constructing group-aware brain networks, providing robust support for downstream tasks.