Self-Promoted Clustering-based Contrastive Learning for Brain Networks Pretraining

Junbo Ma[1,2], **Caixuan Luo**[3], **Jia Hou**[2] and **Kai Zhao**[4*]
[1]School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
[2]Lishui Institute of Hangzhou Dianzi University, Hangzhou Dianzi University, Hangzhou 310018, China
[3]Guangxi Key Lab of Multi-Source Information Mining and Security, Guangxi Normal University, Guilin 541000, China
[4]Department of Neurosurgery, First Medical Center, Chinese PLA General Hospital, Beijing 100853, China
majunbo.phd@gmail.com, luocaixuan4016@gmail.com, houjia334@sina.com, jasondna124@163.com

Abstract

Rapid advancements in neuroimaging techniques, such as magnetic resonance imaging (MRI), have facilitated the acquisition of the structural and functional characteristics of the brain. Brain network analysis is one of the essential tools for exploring brain mechanisms from MRI, providing valuable insights into the brain's organization, and stimulating the understanding of brain cognition and pathology of neurodegenerative diseases. Graph Neural Networks (GNNs) are commonly used for brain network analysis, but they are limited by the scarcity of medical data. Although Graph Contrastive Learning methods have been developed to address this, they often involve graph augmentations that distort the anatomical brain structures. To address these challenges, an augmentation-free contrastive learning method, named Self-Promoted Clustering-based Contrastive Learning(SPCCL), is proposed in this paper. Specifically, by introducing a clustering-based contrastive Learning loss and a self-promoted contrastive pairs creation scheme, the proposed SPCCL can be pre-trained from additional healthy subjects' data that are relatively easier to acquire than disorder ones. The proposed SPCCL leverages these additional data with respect to the integrity of the original brain structure, making it a promising approach for effective brain network analysis. Comprehensive experiments are conducted on an open-access schizophrenic dataset, demonstrating the effectiveness of the proposed method.

1 Introduction

The rapid development of neuroimaging techniques, such as magnetic resonance imaging (MRI), has considerably facilitated the acquisition of the brain's structural (i.e., white matter tracts) and functional (i.e., blood oxygen level-dependent signal) connections, which paves the way for more comprehensive studies into brain cognition and behavior [Hermessi *et al.*, 2021; Salama *et al.*, 2018; Zhang *et al.*, 2020]. One of the essential tools to conduct these investigations is the brain network, which represents the underlying topological organization of the brain as a graph at the macroscale [Ahmedt-Aristizabal *et al.*, 2021; Cui *et al.*, 2022; Hu *et al.*, 2021; Ma *et al.*, 2022]. In the graph, nodes correspond to anatomical brain regions, while edges illustrate the structural or functional connections between them. Graph theories and methods can then be applied to these brain networks, which allows researchers to mathematically quantify and explore the organization and interaction of the brain regions [Bullmore and Sporns, 2009; Cao *et al.*, 2023; Gordon *et al.*, 2016]. Consequently, brain network analysis has become an emerging research field with numerous downstream applications, such as brain disorder diagnosis and disease progression forecasting [Huang *et al.*, 2020; Zhu *et al.*, 2022].

Over the past few years, Graph Neural Networks (GNN) have been proven to be a promising tool for brain network analysis, which integrates graph theories with deep learning techniques, creating a paradigm that is capable of handling the complex relationships inherent in graph data [Li *et al.*, 2022; Peng *et al.*, 2022b; Zhou *et al.*, 2020]. The versatility and scalability allow GNNs to handle the complexity of brain networks, making them a powerful tool for exploring the brain's structure and function, including early diagnosis and prognosis of brain disorders, understanding the progression of various diseases, and personalizing treatments based on individual brain network structures [Ma *et al.*, 2020; Peng *et al.*, 2022a; Tanveer *et al.*, 2020; Young *et al.*, 2018].

Despite the tremendous potential, several challenges persist in applying GNNs to brain network analysis [Zhang *et al.*, 2024; Lin *et al.*, 2024]. One key challenge is that the GNNs usually demand large amounts of labeled data to train effectively [Cao *et al.*, 2023; Cui *et al.*, 2022]. However, as MRI scans are often costly, brain network data is naturally scarce, especially in the context of disease-related studies. To address this challenge, Graph Contrastive Learning (GCL) has surfaced as a promising solution [Tong *et al.*, 2021;

*Corresponding author

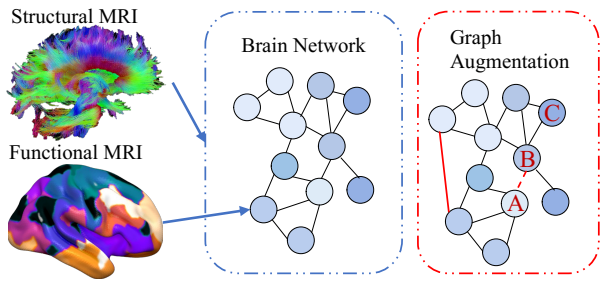

Figure 1: The Brain Networks and Graph Augmentations. Structural MRI defines the topology of the brain network reflecting the brain's anatomical structure from the white matter tracts. Functional MRI defines the node features from the blood oxygen level-dependent signal. If the graph augmentation adds a random edge (red line), it will mislead the learning model with an energy-efficient shortcut. Or, if it deletes an edge between A and B, there will be no short connection between A and C. Either will damage the topology of the real brain.

You *et al.*, 2020; Zhao *et al.*, 2023]. GCL is a representation learning technique that learns to differentiate between similar and dissimilar data instances through contrastive pairs [Lee *et al.*, 2022; Lin *et al.*, 2022; Zhang *et al.*, 2020]. The GCL model is trained to maximize the agreement, or similarity, between positive pairs while simultaneously minimizing the agreement between negative pairs. Usually, these contrastive pairs are created either from data augmentation or extracted from the training set. This pairwise organization of training data provides a valuable mechanism to alleviate the issue of data scarcity. By maximizing the utility of each data point, GCL enhances the training signal and allows for more effective learning from limited datasets.

Although GCL has achieved significant success in many fields, it still has a notable limitation when dealing with brain network data. One essential step in GCL is creating contrastive pairs, and most of the GCL methods focus on designing graph augmentation schemes [Lin *et al.*, 2022; Liu *et al.*, 2022a], which involve perturbing the graph structures, such as randomly adding or deleting nodes or edges of the original graph. As shown in Figure 1, this kind of graph augmentation scheme may not be appropriate for brain networks because the brain networks are derived from MRI images, which neurologically reflect the physical brain wiring (structural brain network) or functional brain region correspondence (functional brain network) [Bullmore and Sporns, 2009]. As such, any arbitrary change or perturbation of the brain network risks creating misleading graph representations. Learning from distorted contrastive pairs cannot help GCL models explore the brain mechanisms and could negatively impact the models' performance.

Additionally, it is important to note that brain network data is typically accompanied by experts' annotations, which is due to the fact that MRI is both a technique and cost-intensive procedure requiring the oversight of medical professionals [Chu *et al.*, 2018]. Furthermore, it is obvious that acquiring brain network data from healthy subjects is relatively easier than disorder ones in disease-related studies, as healthy individuals are the majority of the population. Thus, utilizing the extra brain network data from healthy individuals to pre-train the model is a promising way to partially alleviate the scarce issue.

In response to the aforementioned challenges and unique characteristics of brain network data, a Self-Promoted Clustering-based Contrastive Learning (SPCCL) method is proposed, aiming to help the model effectively learn instinctive graph representations from the brain network data. Specifically, extra healthy subjects are added to the training set, creating a data pool. The initial contrastive pairs are extracted from this pool. The pairs from the same group (marked by the data label) are positive pairs, whereas the pairs from different groups are negative pairs. After initializing the contrastive pairs, a Siamese Graph Convolutional Network (GCN) framework is utilized to extract graph-level representations from them [Liu *et al.*, 2022b]. Then, a self-supervised readout mechanism is proposed to refine the graph representations, ensuring they accurately reflect the underlying structural and functional characteristics of the brain networks. After that, a supervised clustering-based contrastive learning loss is adapted and integrated to ensure the clustering consistency with the data labels. Then, a self-promoted contrastive pairs creation scheme is proposed, which utilizes the cluster centroids as the prototype of each group. The samples that are far from their corresponding prototype are promoted to create contrastive pairs for the next round of training. So, in the final pre-trained contrastive representation space, the samples from the same group are tied closely together, whereas the distance between the prototypes is far away from each other. After the pre-training, the additional healthy subjects' data are removed from the training set, and the classification is performed inside this pre-trained contrastive representation space.

To test the effectiveness of the proposed SPCCL method, we conduct comprehensive experiments on an open-access schizophrenic dataset [Vohryzek *et al.*, 2020], which contains two modalities named structural and functional connectomes from 27 schizophrenic patients and 27 healthy adults. An additional dataset with 70 healthy adults is included in the pre-training phrase [Griffa *et al.*, 2019]. The details of the dataset can be found in the Experiments section.

In summary, the contributions of this paper are:

1. A Self-Promoted Clustering-based Contrastive Learning (SPCCL) method is proposed for brain network pre-training, which addresses the challenges and unique characteristics of brain network data.
2. A self-supervised readout mechanism is proposed to refine the graph representations, ensuring they accurately reflect the underlying structural and functional characteristics of the brain networks.
3. A self-promoted contrastive pairs creation scheme is proposed, which utilizes the cluster centroids as the prototype and promotes the samples far from their corresponding prototype as the candidates to form contrastive pairs.
4. Comprehensive experiments are conducted on the real schizophrenic dataset, demonstrating the effectiveness of the proposed method.

2 Related Works

2.1 Graph Contrastive Learning

Graph Contrastive Learning (GCL) is a powerful technique that leverages Graph Neural Networks (GNNs) to distinguish between similar and dissimilar instances in complex graph data [Lu *et al.*, 2021]. Over the past few years, significant advancements have been made in this field, focusing on data augmentation, pretext task designs, and contrastive objectives [Liu *et al.*, 2022a].

One of the notable works in this area is Deep Graph Infomax (DGI), which maximizes the mutual information between local and global graph features, adapting the idea from Deep InfoMax to graphs for node representation learning [Veličković *et al.*, 2018]. Similarly, InfoGraph extends this concept to learn graph-level representations by maximizing agreements between the representations of entire graphs and their various scaled substructures. Contrastive Multi-View Representation Learning on Graphs (MVGRL) is another significant contribution that maximizes mutual information from different structural views of graphs [Xu *et al.*, 2022]. Directed Graph Contrastive Learning (DiGCL) extends the contrastive paradigm to directed graphs, aiming at learning from abundant views while retaining the original structure information [Tong *et al.*, 2021]. Graph Contrastive Learning with Augmentation presents a new perspective on graph contrastive learning by introducing a novel augmentation strategy, which leverages both node-level and graph-level augmentations to enhance the performance of graph contrastive learning [You *et al.*, 2020].

Last year, AFGRL, an augmentation-free self-supervised learning framework for graphs, was proposed to generate an alternative view of a graph by discovering nodes that share the local structural information and the global semantics with the graph [Lee *et al.*, 2022]. This method has shown superiority in various node-level tasks, including node classification, clustering, and similarity search on various real-world datasets. However, since it focuses on node-level representation, AFGRL cannot be extended to graph-level contrastive learning for brain network analysis.

These advancements in GCL have significantly expanded our understanding and capabilities in handling complex graph data, paving the way for more sophisticated applications. However, it's important to note that these methods often require careful consideration of augmentation techniques and negative sampling strategies to ensure the preservation of the underlying semantics of the graph data, which may distort the anatomical brain structure and be extremely risky when doing brain network analysis.

2.2 Clustering-based Contrastive Learning

Clustering-based Contrastive Learning (CCL) has emerged as a contrastive learning method that leverages clustering techniques to structure the learning process [Deng *et al.*, 2018; Li *et al.*, 2021; Sharma *et al.*, 2020]. In the learned representation space by CCL, the samples from the same group are pulled together, while different groups are pushed away from each other [Cheng *et al.*, 2014].

In the past few years, various CCL works have been proposed. The Supervised Contrastive (SupCon) method uses label information to pull together clusters of points belonging to the same class and push apart clusters of samples from different classes [Khosla *et al.*, 2020]. This method has shown robustness to natural corruptions and stability to hyperparameter settings, demonstrating its potential in various datasets. Prototypical Contrastive Learning (PCL) is an unsupervised representation learning method that combines contrastive learning with clustering to encode semantic structures into the learned embedding space [Li *et al.*, 2020]. PCL uses prototypes as latent variables to find the maximum-likelihood estimation of network parameters in an Expectation-Maximization framework.

For the graph data, the Graph Contrastive Clustering (GCC) method is proposed for image clustering tasks, which in-corporates category information to perform contrastive learning at both the instance and cluster levels, resulting in more discriminative and clustering-friendly features and more compact clustering assignments [Zhong *et al.*, 2021; Zhang and He, 2023]. The Multilayer Graph Contrastive Clustering Network (MGCCN) method integrates a clustering objective with graph representation learning to better capture the graph's semantic structures from multilayer graphs [Liu *et al.*, 2022a]. To address the issue of imbalanced prototype assignments that can arise from directly using K-means, the Prototypical Graph Contrastive Learning (PGCL) approach was proposed, which adds a constraint that prototype assignments must be partitioned into equally sized subsets and formulates it as an optimal transport problem [Lin *et al.*, 2022]. It also tackles the sampling bias by sampling negatives from clusters differing from the query cluster and reweighting negative samples according to their prototype distances.

Although these methods achieved great success, they still rely on graph data augmentations that are inappropriate for brain network analysis. Furthermore, because of the technique and cost-intensive nature of the brain network data, brain networks are always accompanied by expert annotations. Thus, it is not necessary to follow the unsupervised or semi-supervised setting.

3 The Proposed Method

In this paper, the brain networks are presented as graphs, with $G = \{G_i, X_i, y_i\}^N$, where N is the total number of the brain networks, G_i is the graph topology, X_i is the nodes' features, and y_i is the corresponding label of the brain network. $y_i = 1$, if the brain network is from the healthy group, whereas $y_i = 0$, if it is from the group with disease. Each graph is a set of nodes and edges, $G_i = (V_i, E_i)$, where $V_i = \left\{v_1^i, ..., v_P^i\right\}$ is the nodes set and $E_i = \left\{e_1^i, ..., e_Q^i\right\}$ is the edge set. Note that, the number of nodes is the same across all the brain networks, as the nodes represent the anatomical brain regions, and everyone should have the same anatomical brain structure. Whereas, the number of edges may vary from individual to individual, as everyone has their own brain connections that are formed by genes and/or personal experiences.

The overall goal is to train a model to classify the brain

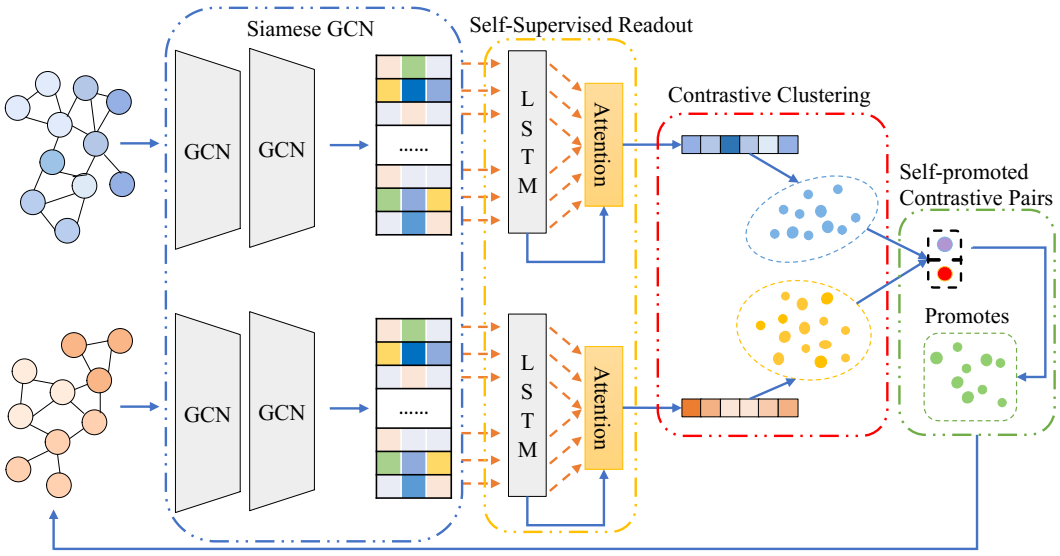

Figure 2: The overview of the proposed Self-Promoted Clustering-based Contrastive Learning (SPCCL) method's structure. The SPCCL has four modules: (a) a Siamese GCN module to process the contrastive pairs; (b) a self-supervised readout module to create graph-level representations from the nodes embeddings; (c) a contrastive clustering module to do the supervised contrastive clustering; (d) a self-promoted contrastive pair creation module to promote the contrastive pairs for the next round of training (see Figure 3) for details.

networks into the correct group corresponding to their labels. However, due to the scarce issue and the characteristics of the brain network data, we will pre-train the model with the proposed Self-Promoted Clustering-based Contrastive Learning (SPCCL) method without distorting the brain networks. As shown in Figure 2, the proposed SPCCL has four modules: 1) A Siamese GCN module to process the contrastive pairs; 2) A self-supervised readout module to create graph-level representations from the node embeddings; 3) A contrastive clustering module to do the supervised contrastive-based clustering, which will form the; 4) A clustering-based contrastive pair creation module to promote the contrastive pairs for the next round of training. We will introduce them one by one in the rest of this section.

3.1 Siamese GCN Module

As mentioned in the Introduction Section, the contrastive pairs are initialized from a pool incorporating the training data and additional healthy subjects. The Siamese Graph Convolutional Network (GCN) module aims to extract graph-level representations from these contrastive graph pairs. It contains two identical subnetworks with sharing weights. This is a common setting in contrastive learning. The GCN here is a standard Graph Convolutional Network with the Chebyshev polynomial [Kipf and Welling, 2016]. In this work, we use two GCN layers to process the graph data. The topology of the graph data G_i is derived from the structural MRI images, and the node features X_i for each node are extracted from the functional MRI images. Through this, two modalities are integrated into the Siamese GCN to extract node embeddings. The output node embeddings from this module are a mix of brain network structures and brain functional signals. Thus, we will need a readout module to convert the node embeddings into graph embedding [Buterez *et al.*, 2022]. This is done by our proposed self-supervised readout module in the next section.

3.2 Self-Supervised Readout Module

The proposed Self-supervised readout module exploits a Bi-directional Long Short Term Memory (LSTM) model to process the output node embeddings from the Siamese GCN module. In this case, the output node embeddings are considered as a sequence of all the nodes. Each extracted node embedding is processed step by step through the bi-directional LSTM. The Bi-directional LSTM used here is to partially eliminate the bias from the node order in the sequence. The last hidden states of the Bi-directional LSTM can be considered as a latent representation encrypting all the nodes' features and topology, which will be used as the Query of the attention mechanism, while the keys and values are the hidden states of each step. The attention score is calculated by the standard attention equation:

$$Attention = softmax(\frac{(QW_Q(KW_K)^T)}{\sqrt{(d_k)}})VW_V \quad (1)$$

where Q, K, V are Qurey, Keys and Values, and W_Q, W_K, W_V are the corresponding learnable weight matrices

The output of the attention is the weighted sum of all the hidden states. Through the attention mechanism, the output readout vector is refined by the context of node embeddings, which will comprehensively explore the underlying structural and functional brain networks and emphasize the most discriminative parts of the node embeddings into the readout vector z_i of the graph.

3.3 Contrastive Clustering Module

The readout vector z_i of each brain network will first be fed into the Supervised Contrastive $Loss$, $Loss_{sup}$, to get the loss value of each contrastive pair [Khosla *et al.*, 2020]. The $Loss_{sup}$ aims to effectively leverage label information from the contrastive pairs. Two samples coming from the same label group are called positive pairs, whereas two samples coming from different groups are called negative pairs.

$$Loss_{sup} = \sum_{i \in Pairs} -\log \frac{1}{|P|} \sum_{p \in P} \frac{exp((z_i z_p)/\tau)}{\sum_{a \in A} exp((z_i z_a)/\tau)} \tag{2}$$

where P is the positive pair set and A is the negative pair set, i is the contrastive pair from the contrastive pair set.

In this way, the label information is implicitly exploited into the contrastive loss, where the positive pairs are pulled together, and negative pairs are pushed away from each other.

After each epoch, all the readout vectors are put together to do the contrastive clustering. The number of clusters is predefined by the data labels. In our case, there are two clusters: the healthy group vs. the disorder group. The consistency loss is equipped to ensure the clustering consistency corresponding to the sample labels [Lin *et al.*, 2022].

$$Loss_{consist} = \sum_{i \in Pairs} [l(p_i, q_{i'}) + l(p_{i'}, q_i)] \tag{3}$$

where $l(p_i, q_{i'}) = \sum_{y \in Y} q(y|z_{i'}) \log p(y|z_i)$, i is the contrastive pair from the contrastive pair set, z_i and $z_{i'}$ are two readout vectors from one contrastive pair, $p(y|z_i) = softmax(z_i)$ illustrates the probability of z_i belonging to group y. Similarly, $q(y|z_{i'})$ illustrates the probability of $z_{i'}$ belonging to group y.

The final training loss can be written as:

$$Loss = Loss_{sup} + \lambda Loss_{consist}$$

where λ is a hyperparameter balancing the two terms.

3.4 Self-Promoted Contrastive Pairs Creation Module

As shown in Figure 3, the prototype of each group is chosen as the centroid of each cluster. Then, the samples of each group will be sorted by the distance to their prototype. We select the top 50% of the samples from each group to form the candidate set. The new contrastive pairs for the next training epoch will be randomly sampled from this candidate set. The total number of contrastive pairs can be predefined as a fixed number, or all the possible combinations of the candidate set. The pairs that have the same labels are considered positive pairs, whereas the pairs that have different labels are considered negative pairs.

In this way, the difficult samples of each label group are promoted into the candidate set. Our model will then be trained on the contrastive pairs created from this difficult sample set in the next epoch. For each epoch, the candidate set is updated by the clustering results. With the help of $Loss_{sup}$ and $Loss_{consist}$, the learned readout vector will be updated accordingly, which will make the learned graph-level representations more discriminative.

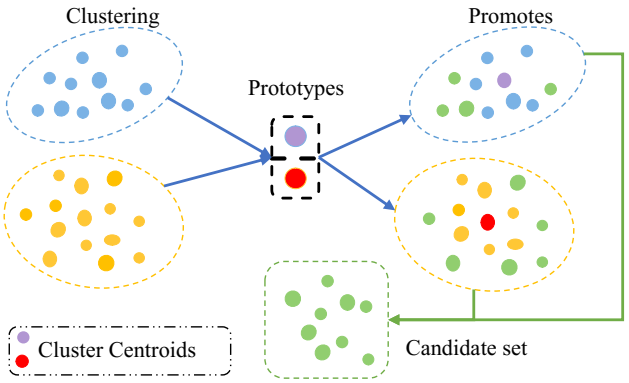

Figure 3: The Self-Promoted Contrastive pairs creation procedure, in which the difficult samples are promoted into the candidate set to create contrastive pairs for the next training epoch. In our case, there are two clusters: the healthy group vs. the disorder group.

Algorithm 1 SPCCL pre-training

Data pool: Mix training set and extra healthy subjects.
Input: Random select M contrastive pairs.
Output: Graph representations

```
Let t = 0.
while t < T or not converge do
    for pairs in M do
        do Siamese GCN
        do Self-supervised readout
    end for
    do Contrastive clustering
    backpropagate with Loss = Loss_sup + λ Loss_consist
    promote M contrastive pairs with the centroids.
end while
return Graph representations of data
```

3.5 Classification

The pre-training process can be found in Algorithm 1. It will loop through the 4 modules until it converges or reaches a given number of epochs. After pre-tained by the proposed SPCCL model, we remove all the auxiliary modules and only use the readout vectors as the graph-level representation of brain networks. These representations then train a simple multilayer perceptron network to do the final classification. The detailed experimental settings and results are reported in the next section.

4 Experiments

4.1 Dataset

The schizophrenic dataset utilized in this study is from an open-access dataset that includes two types of brain network modalities: structural and functional connectomes. It consists of MRI data acquired from 27 schizophrenic patients, and 27 matched healthy adults [Vohryzek *et al.*, 2020]. An additional 70 healthy adults' MRI data are used to do the pre-training [Griffa *et al.*, 2019], which reflects the characteristic of the brain network data that it is relatively easier to acquire healthy samples than the samples with diseases. All the data

Method	KNN	Logistic Reg.	K-Means	SVC	GCBN	SPCCL
ACC	59.2±5.2	71.3±3.8	52.6±6.1	70.9±3.4	73.1 ± 4.6	85.2±3.8
SENS	69.1±8.3	70.4±7.3	50.3±6.5	72.5±7.7	78.2 ± 5.8	88.1±5.8
SPEC	56.2±8.2	73.2±4.6	56.7±8.9	70.7±6.3	68.1 ± 8.1	83.7±3.2
PPV	62.4±6.2	71.6±7.2	55.7±4.6	73.4±5.1	71.5 ± 5.3	84.3±4.6
NPV	67.2±8.7	72.7±5.1	52.4±8.5	74.6±5.2	74.5 ± 5.1	83.2±7.0
F1	0.619	0.717	0.533	0.716	0.728	0.858

Table 1: Comparison of classification results.

used in this study was collected at the Service of General Psychiatry at the Lausanne University Hospital. The diagnosis of schizoaffective disorders for the patients was established based on their fulfillment of the DSM-IV criteria. The control subjects (healthy adults), on the other hand, had no previous history of neurological disorders, ensuring a comparison between healthy individuals and those with schizophrenia. All subjects underwent scanning using a 3 Tesla Siemens Trio scanner equipped with a 32-channel head coil. The data collection protocol involved three distinct imaging modalities. Firstly, a magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence was employed to capture white/gray matter contrast, with a high resolution of 1-mm in-plane resolution and 1.2-mm slice thickness. Secondly, a Diffusion Spectrum Imaging (DSI) sequence was utilized, comprising 128 diffusion-weighted volumes and a single b0 volume. This sequence utilized a maximum b-value of 8,000 s/mm² and had a voxel size of 2.2x2.2x3.0 mm. Finally, a gradient echo EPI sequence sensitive to BOLD contrast was employed, resulting in 280 images per participant. This sequence had an in-plane resolution and slice thickness of 3.3-mm with a 0.3-mm gap, TE 30 ms, and TR 1,920 ms. For the structural connectivity analysis, deterministic streamline tractography was performed on the reconstructed DTI data. The tractography pipeline initiated 32 streamline propagations per diffusion direction per white matter voxel. Fiber density, representing the number of streamlines normalized by the average streamline length and surface area of the regions, was computed to measure the structural connectivity between pairs of regions of interest (ROIs). To estimate functional connectomes, the blood oxygen level-dependent (BOLD) time series from fMRI data was utilized. Specifically, the absolute value of the Pearson correlation was computed between the time courses of individual brain regions, enabling the assessment of functional connectivity across the brain.

4.2 Experiment Setup

To ensure unbiased performance evaluation, we employ a 6-fold cross-validation strategy during the training process. This involves randomly dividing the dataset into three equal parts, where one-third of the samples from each class are selected as the testing set, while the remaining two-thirds serve as the training set. To evaluate the classification performance, we employ six distinct metrics: accuracy (ACC), sensitivity (SENS), specificity (SPEC), positive predictive values (PPV), negative predictive values (NPV), and F1 score.

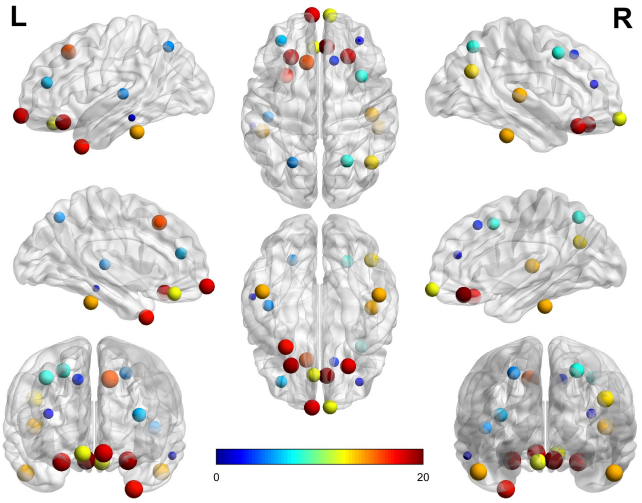

Figure 4: The view of the top 20 important nodes on the brain. The size and color indicate the importance of the node. The bigger size and warmer color mean the more important the node is.

By employing these metrics, we gain comprehensive insights into the performance of our classification method. Furthermore, we compare the classification performance of our approach with three commonly used learning-based methods: K-nearest neighbor (KNN), logistic regression, and K-means clustering. By conducting these comparisons, we can gauge the effectiveness and superiority of our proposed method. Lastly, we benchmark our method against the state-of-the-art approach for schizophrenia classification reported in the literature [Ghosh *et al.*, 2023], called Graph Convolution on Brain Network (GCBN), which utilized the Graph Convolutional Networks to do the schizophrenia classification task. In GCBN, a data augmentation process is proposed to create new artificial data by inducing artificial perturbation to the available data set, which may downgrade the model performance as the perturbation may distort the brain's anatomical structures. These experimental settings enable us to ascertain the advancements and contributions of our proposed method in relation to existing methodologies.

4.3 Discussion

In order to evaluate the performance of our proposed SPCCL method, we compared it against baseline and competing approaches using six distinct quality metrics, with higher values signifying superior classification performance. The summary

of our experimental results can be seen in Table 1, where we report the mean and standard deviation for each metric over the 6-fold cross-validation. The results unequivocally demonstrate that our proposed Self-Promoted Clustering-based Contrastive Learning (SPCCL) method outperforms the other methods in terms of classification accuracy. Notably, a higher F1 score is achieved by our method, indicating its superior capabilities. The improvements were especially marked when compared to the GCBN method tailored for schizophrenia analysis, underscoring the value of the SPCCL approach for brain network analysis. As shown in Figure 4, the brain regions that have the highest attention scores throughout the patients' data are identified, such as the temporal gyrus, frontal gyrus, insula, and amygdala. These identified regions are consistent with the findings reported by neuroscience papers [Henkel *et al.*, 2022; Huckins *et al.*, 2019; Rubinov and Bullmore, 2013].

5 Conclusion

In the field of cognitive and behavioral neuroscience, employing Graph Neural Networks (GNN) for brain network analysis is often challenging due to the scarcity of labeled data from both technique and cost intensive MRI scans. This paper proposed a Self-Promoted Cluster-ing-based Contrastive Learning (SPCCL) method to address these limitations with respect to the unique characteristics of brain network data that its topology should be strictly aligned with the brain's anatomical brain structure. The proposed SPCCL employs a Siamese Graph Convolutional Network (GCN) framework for extracting graph-level representations, a self-supervised readout mechanism to refine the representations, a supervised clustering-based contrastive learning module for clustering consistency, and then a self-promoted contrastive pair creation scheme to fully explore the dataset without distort the brain networks. It overcomes the limitations of existing Graph Contrastive Learning methods and enhances model robustness. Tested on an open-access schizophrenia dataset, SPCCL demonstrates significant potential in advancing the diagnosis and prognosis of brain disorders.

Ethical Statement

There are no ethical issues.

References

[Ahmedt-Aristizabal *et al.*, 2021] David Ahmedt-Aristizabal, Mohammad Ali Armin, Simon Denman, Clinton Fookes, and Lars Petersson. Graph-based deep learning for medical diagnosis and analysis: past, present and future. *Sensors*, 21(14):4758, 2021.

[Bullmore and Sporns, 2009] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews neuroscience*, 10(3):186–198, 2009.

[Buterez *et al.*, 2022] David Buterez, Jon Paul Janet, Steven J Kiddle, Dino Oglic, and Pietro Liò. Graph neural networks with adaptive readouts. *Advances in Neural Information Processing Systems*, 35:19746–19758, 2022.

[Cao *et al.*, 2023] Peng Cao, Guangqi Wen, Wenju Yang, Xiaoli Liu, Jinzhu Yang, and Osmar Zaiane. A unified framework of graph structure learning, graph generation and classification for brain network analysis. *Applied Intelligence*, 53(6):6978–6991, 2023.

[Cheng *et al.*, 2014] Debo Cheng, Shichao Zhang, Zhenyun Deng, Yonghua Zhu, and Ming Zong. k nn algorithm with data-driven k value. In *Advanced Data Mining and Applications: 10th International Conference, ADMA 2014, Guilin, China, December 19-21, 2014. Proceedings 10*, pages 499–512. Springer, 2014.

[Chu *et al.*, 2018] Shu-Hsien Chu, Keshab K Parhi, and Christophe Lenglet. Function-specific and enhanced brain structural connectivity mapping via joint modeling of diffusion and functional mri. *Scientific reports*, 8(1):4741, 2018.

[Cui *et al.*, 2022] Hejie Cui, Wei Dai, Yanqiao Zhu, Xuan Kan, Antonio Aodong Chen Gu, Joshua Lukemire, Liang Zhan, Lifang He, Ying Guo, and Carl Yang. Braingb: A benchmark for brain network analysis with graph neural networks. *IEEE transactions on medical imaging*, 42(2):493–506, 2022.

[Deng *et al.*, 2018] Zhenyun Deng, Shichao Zhang, Lifeng Yang, Ming Zong, and Debo Cheng. Sparse sample self-representation for subspace clustering. *Neural Computing and Applications*, 29:43–49, 2018.

[Ghosh *et al.*, 2023] Sanjay Ghosh, Eshan Bhargava, Chieh-Te Lin, and Srikantan S Nagarajan. Graph convolutional learning of multimodal brain connectome data for schizophrenia classification. *bioRxiv*, pages 2023–01, 2023.

[Gordon *et al.*, 2016] Evan M Gordon, Timothy O Laumann, Babatunde Adeyemo, Jeremy F Huckins, William M Kelley, and Steven E Petersen. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebral cortex*, 26(1):288–303, 2016.

[Griffa *et al.*, 2019] Alessandra Griffa, Yasser Alemán-Gómez, and Patric Hagmann. Structural and functional connectome from 70 young healthy adults. *Zenodo*, May 2019.

[Henkel *et al.*, 2022] Nicholas D Henkel, Xiajoun Wu, Sinead M O'Donovan, Emily A Devine, Jessica M Jiron, Laura M Rowland, Zoltan Sarnyai, Amy J Ramsey, Zhexing Wen, Margaret K Hahn, et al. Schizophrenia: a disorder of broken brain bioenergetics. *Molecular Psychiatry*, 27(5):2393–2404, 2022.

[Hermessi *et al.*, 2021] Haithem Hermessi, Olfa Mourali, and Ezzeddine Zagrouba. Multimodal medical image fusion review: Theoretical background and recent advances. *Signal Processing*, 183:108036, 2021.

[Hu *et al.*, 2021] Rongyao Hu, Ziwen Peng, Xiaofeng Zhu, Jiangzhang Gan, Yonghua Zhu, Junbo Ma, and Guorong Wu. Multi-band brain network analysis for functional neuroimaging biomarker identification. *IEEE transactions on medical imaging*, 40(12):3843–3855, 2021.

[Huang *et al.*, 2020] Jiashuang Huang, Luping Zhou, Lei Wang, and Daoqiang Zhang. Attention-diffusion-bilinear neural network for brain network analysis. *IEEE transactions on medical imaging*, 39(7):2541–2552, 2020.

[Huckins *et al.*, 2019] Laura M Huckins, Amanda Dobbyn, Douglas M Ruderfer, Gabriel Hoffman, Weiqing Wang, Antonio F Pardinas, Veera M Rajagopal, Thomas D Als, Hoang T. Nguyen, Kiran Girdhar, et al. Gene expression imputation across multiple brain regions provides insights into schizophrenia risk. *Nature genetics*, 51(4):659–674, 2019.

[Khosla *et al.*, 2020] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. *Advances in neural information processing systems*, 33:18661–18673, 2020.

[Kipf and Welling, 2016] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2016.

[Lee *et al.*, 2022] Namkyeong Lee, Junseok Lee, and Chanyoung Park. Augmentation-free self-supervised learning on graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 7372–7380, 2022.

[Li *et al.*, 2020] Junnan Li, Pan Zhou, Caiming Xiong, and Steven Hoi. Prototypical contrastive learning of unsupervised representations. In *International Conference on Learning Representations*, 2020.

[Li *et al.*, 2021] Yunfan Li, Peng Hu, Zitao Liu, Dezhong Peng, Joey Tianyi Zhou, and Xi Peng. Contrastive clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 8547–8555, 2021.

[Li *et al.*, 2022] Michelle M Li, Kexin Huang, and Marinka Zitnik. Graph representation learning in biomedicine and healthcare. *Nature Biomedical Engineering*, 6(12):1353–1369, 2022.

[Lin *et al.*, 2022] Shuai Lin, Chen Liu, Pan Zhou, Zi-Yuan Hu, Shuojia Wang, Ruihui Zhao, Yefeng Zheng, Liang Lin, Eric Xing, and Xiaodan Liang. Prototypical graph contrastive learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[Lin *et al.*, 2024] Han Lin, Yingjian Li, Zheng Zhang, Lei Zhu, and Yong Xu. Learning with noisy labels by semantic and feature space collaboration. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.

[Liu *et al.*, 2022a] Liang Liu, Zhao Kang, Jiajia Ruan, and Xixu He. Multilayer graph contrastive clustering network. *Information Sciences*, 613:256–267, 2022.

[Liu *et al.*, 2022b] Yixin Liu, Ming Jin, Shirui Pan, Chuan Zhou, Yu Zheng, Feng Xia, and S Yu Philip. Graph self-supervised learning: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 35(6):5879–5900, 2022.

[Lu *et al.*, 2021] Yuanfu Lu, Xunqiang Jiang, Yuan Fang, and Chuan Shi. Learning to pre-train graph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 4276–4284, 2021.

[Ma *et al.*, 2020] Junbo Ma, Xiaofeng Zhu, Defu Yang, Jiazhou Chen, and Guorong Wu. Attention-guided deep graph neural network for longitudinal alzheimer's disease analysis. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23*, pages 387–396. Springer, 2020.

[Ma *et al.*, 2022] Junbo Ma, Jilian Zhang, and Zeyu Wang. Multimodality alzheimer's disease analysis in deep riemannian manifold. *Information Processing & Management*, 59(4):102965, 2022.

[Peng *et al.*, 2022a] Liang Peng, Rongyao Hu, Fei Kong, Jiangzhang Gan, Yujie Mo, Xiaoshuang Shi, and Xiaofeng Zhu. Reverse graph learning for graph neural network. *IEEE transactions on neural networks and learning systems*, 2022.

[Peng *et al.*, 2022b] Liang Peng, Nan Wang, Nicha Dvornek, Xiaofeng Zhu, and Xiaoxiao Li. Fedni: Federated graph learning with network inpainting for population-based disease prediction. *IEEE Transactions on Medical Imaging*, 2022.

[Rubinov and Bullmore, 2013] Mikail Rubinov and Ed Bullmore. Schizophrenia and abnormal brain network hubs. *Dialogues in clinical neuroscience*, 15(3):339–349, 2013.

[Salama *et al.*, 2018] Gayle R Salama, Linda A Heier, Praneil Patel, Rohan Ramakrishna, Rajiv Magge, and Apostolos John Tsiouris. Diffusion weighted/tensor imaging, functional mri and perfusion weighted imaging in glioblastoma—foundations and future. *Frontiers in neurology*, 8:660, 2018.

[Sharma *et al.*, 2020] Vivek Sharma, Makarand Tapaswi, M Saquib Sarfraz, and Rainer Stiefelhagen. Clustering based contrastive learning for improving face representations. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pages 109–116. IEEE, 2020.

[Tanveer *et al.*, 2020] Muhammad Tanveer, Bharat Richhariya, Riyaj Uddin Khan, Ashraf Haroon Rashid, Pritee Khanna, Mukesh Prasad, and CT Lin. Machine learning techniques for the diagnosis of alzheimer's disease: A review. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 16(1s):1–35, 2020.

[Tong *et al.*, 2021] Zekun Tong, Yuxuan Liang, Henghui Ding, Yongxing Dai, Xinke Li, and Changhu Wang. Directed graph contrastive learning. *Advances in neural information processing systems*, 34:19580–19593, 2021.

[Veličković *et al.*, 2018] Petar Veličković, William Fedus, William L Hamilton, Pietro Liò, Yoshua Bengio, and R Devon Hjelm. Deep graph infomax. In *International Conference on Learning Representations*, 2018.

[Vohryzek *et al.*, 2020] Jakub Vohryzek, Yasser Aleman-Gomez, Alessandra Griffa, Jeni Raoul, Martine Cleusix, Philipp S. Baumann, Philippe Conus, Kim Do Cuenod, and Patric Hagmann. Structural and functional connectomes from 27 schizophrenic patients and 27 matched healthy adults. *Zenodo*, April 2020.

[Xu *et al.*, 2022] Jie Xu, Yazhou Ren, Huayi Tang, Zhimeng Yang, Lili Pan, Yang Yang, Xiaorong Pu, S Yu Philip, and Lifang He. Self-supervised discriminative feature learning for deep multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 2022.

[You *et al.*, 2020] Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, and Yang Shen. Graph contrastive learning with augmentations. *Advances in neural information processing systems*, 33:5812–5823, 2020.

[Young *et al.*, 2018] Alexandra L Young, Razvan V Marinescu, Neil P Oxtoby, Martina Bocchetta, Keir Yong, Nicholas C Firth, David M Cash, David L Thomas, Katrina M Dick, Jorge Cardoso, et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with subtype and stage inference. *Nature communications*, 9(1):4273, 2018.

[Zhang and He, 2023] Zheng Zhang and Wen-Jue He. Tensorized topological graph learning for generalized incomplete multi-view clustering. *Information Fusion*, 100:101914, 2023.

[Zhang *et al.*, 2020] Li Zhang, Mingliang Wang, Mingxia Liu, and Daoqiang Zhang. A survey on deep learning for neuroimaging-based brain disorder analysis. *Frontiers in neuroscience*, 14:779, 2020.

[Zhang *et al.*, 2024] Zheng Zhang, Xu Yuan, Lei Zhu, Jingkuan Song, and Liqiang Nie. Badcm: Invisible backdoor attack against cross-modal learning. *IEEE Transactions on Image Processing*, 2024.

[Zhao *et al.*, 2023] Peiyao Zhao, Yuangang Pan, Xin Li, Xu Chen, Ivor W Tsang, and Lejian Liao. Coarse-to-fine contrastive learning on graphs. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[Zhong *et al.*, 2021] Huasong Zhong, Jianlong Wu, Chong Chen, Jianqiang Huang, Minghua Deng, Liqiang Nie, Zhouchen Lin, and Xian-Sheng Hua. Graph contrastive clustering. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9224–9233, 2021.

[Zhou *et al.*, 2020] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI open*, 1:57–81, 2020.

[Zhu *et al.*, 2022] Yonghua Zhu, Junbo Ma, Changan Yuan, and Xiaofeng Zhu. Interpretable learning based dynamic graph convolutional networks for alzheimer's disease analysis. *Information Fusion*, 77:53–61, 2022.