参赛学生姓名：范宇桓

中学：北京师范大学附属实验中学

省份：北京市

国家/地区：中国

指导老师 1 姓名：高跃

指导老师 1 单位：清华大学

指导老师 2 姓名：马静

指导老师 2 单位：北京师范大学附属实验中学

论文题目：Diagnosing Autism Spectrum Disorder via Brain-Population Graph-in-Graph Neural Networks

Diagnosing Autism Spectrum Disorder via Brain-Population Graph-in-Graph Neural Networks

Yuhuan Fan
The Experimental High School Attached to Beijing Normal University
fyh1029384756@163.com

Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that significantly impacts brain development and function with increasing prevalence, making early and accurate diagnosis critical for effective intervention. While existing diagnostic methods typically focus on either internal brain connectivity or external population relations, effectively integrating these two crucial sources of information remains an open problem. To address this challenge, we propose a novel Brain-Population Graph-in-Graph Neural Network (BP-GiGNN) that enhances ASD diagnosis by effectively combining brain-level and population-level data. The BP-GiGNN consists of two parts: an internal brain-GNN and an external population-GNN. The internal brain-GNN captures the intricate neural interactions within the brain, and the external population-GNN models the relationships between individuals using both brain-level embeddings and non-imaging phenotypic data. Our method is evaluated on two public datasets, ABIDE-NYU and ABIDE-UCLA, and consistently outperforms state-of-the-art methods, demonstrating its effectiveness in ASD diagnosis tasks.

Keywords: Autism Spectrum Disorder, Graph Neural Networks, rs-fMRI

Contents

Preprint. Under review.

1 Introduction

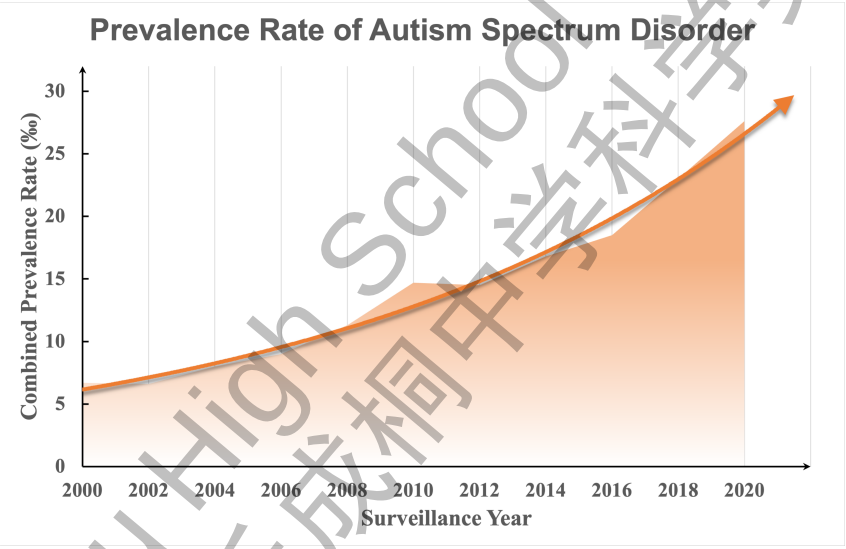
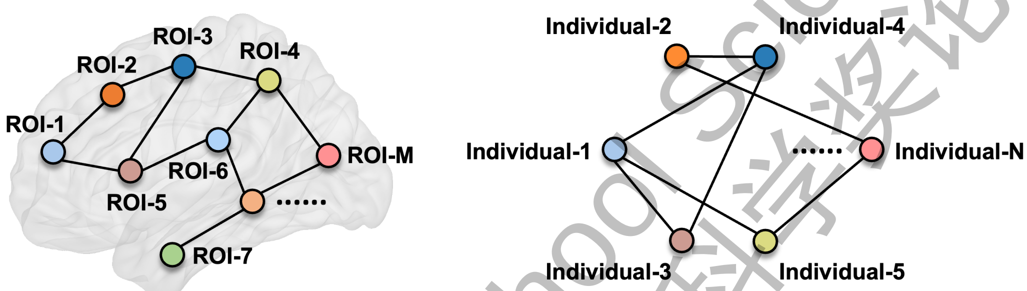


Figure 1: Prevalence rate in the past 20 years (2000-2020).

Autism Spectrum Disorder (ASD) is a representative type of neurodevelopmental disorder, which causes severe impact on the development and function of the human brain in childhood [1]. Research has shown that ASD can affect a child's ability to develop typical social skills, leading to challenges in forming social relationships and communicating effectively with others [2]. By 2015, the number of ASD patients in China has exceeded 10 million [3]. As shown in Fig. 1, the prevalence rate of ASD in America has increased from 6.7‰ in 2000 to 27.6‰ in 2020, more than quadrupling in just twenty years [4]. The severe impact on children and rapid increase in prevalence calls for greater attention to the diagnosis and treatment of ASD.

Accurate diagnosis and timely intervention are crucial in mitigating the long-term effects of ASD [5]. Research indicates that children who receive intensive early interventions are more likely to make substantial gains in cognitive and language abilities, as well as in adaptive functioning [6, 7]. However, the current diagnostic process for ASD is fraught with challenges that contribute to high variability and the potential for misdiagnosis [8]. Diagnosis typically relies on behavioral assessments and clinical observations which often involves structured interviews, parent-reported questionnaires, and standardized behavioral assessments [9]. Due to the high variability and easy misdiagnosis of the current diagnostic process, there is an urgent need to develop an effective method for automatic ASD diagnosis with more accuracy and robustness.

Fortunately, recent advances in neuroimaging and machine learning offer promising avenues for effective automatic ASD diagnosis, potentially transforming the way ASD is diagnosed in the future. In neuroimaging, **resting-state functional Magnetic Resonance Imaging (rs-fMRI)** is a useful source to assess brain activity by tracking hemodynamic fluctuations. By capturing the dynamic interactions between brain regions, rs-fMRI provides insights into the functional connectivity of the brain and reveals how different regions interact during various cognitive tasks [10]. This makes rs-fMRI an invaluable tool in neurodevelopmental research, offering a window into the characteristics of neurodevelopmental disorders such as ASD [11]. In machine learning, by analyzing complex and high-dimensional brain rs-fMRI data, techniques such as support vector machine [12], random forests [13], and deep learning models [14, 15], have been applied to classify individuals with ASD. Especially, **Graph Neural Networks (GNNs)** have gained significant attention for their ability to effectively model complex, non-Euclidean data such as brain networks derived from rs-fMRI. By capturing structural information within the brain graphs, researchers have adopted GNNs in disorder diagnosis and proposed many GNN-based ASD prediction methods [16, 17].



(a) Brain Connectivity Graph (b) Population Relation Graph

Figure 2: Two types of graphs in GNN-based ASD prediction.

As shown in Fig. 2, there are generally two types of graph-structured data in the GNN-based ASD prediction task: the brain graph and the population graph. The **Brain Connectivity Graph** is constructed from each individual's rs-fMRI data, where nodes represent regions of interest (ROIs) in the brain, and the edges indicate the connections linking these regions [18, 19]. This graph captures the unique connectivity patterns within an individual's brain, which can vary significantly between those with ASD and neurotypical individuals. The **Population Relation Graph** represents relationships across different individuals, where each node corresponds to an individual, and edges are based on similarities in their brain connectivity patterns or other non-imaging data [20]. This graph allows for the exploration of how different individuals relate to one another in terms of brain connectivity, providing insights into shared characteristics or differences across a population.

Although GNN-based methods have shown great potential in this field, there are still some significant challenges. **Firstly**, a large proportion of existing methods [18, 16, 21, 22] tend to focus on either the brain connectivity graph or the population relation graph, failing to fully leverage the complementary information that these two types of graphs can provide when used together. By only utilizing one type of graph, existing methods may miss out on valuable contextual information that could enhance the accuracy and robustness of ASD diagnosis. **Secondly**, due to the high heterogeneity in both data format and topological structure, it is still a challenging problem to effectively integrate the information from the above two graphs [23]. The brain connectivity graph is typically represented as a weighted complete graph derived from rs-fMRI data, while the population relation graph is usually a sparse graph with similarities in brain connectivity profiles and non-imaging attributes across the population. This disparity in graph representation poses significant difficulties in developing effective GNN-based rs-fMRI analysis methods. Therefore, it is a key challenge to make full use of the two types of data and effectively integrate heterogeneous graph-structured information.

To address the challenges of integrating heterogeneous graph data, we propose a novel approach called Brain-Population Graph-in-Graph Neural Network (BP-GiGNN) to effectively fuse the brain connectivity graph and the population relation graph for accurate and robust diagnosis of ASD. The proposed method consists of two components: the Internal Brain-GNN and the External Population-GNN. The **Internal Brain-GNN** processes the brain connectivity graph derived from rs-fMRI

data, capturing the intricate patterns of functional connectivity within an individual's brain. These connectivity patterns are crucial for identifying the subtle neural abnormalities associated with ASD. The **External Population-GNN** operates on the population relation graph, which encodes inter-individual relationships based on non-imaging attributes like age and gender. By embedding internal brain features into a population graph structure, BP-GiGNN provides a comprehensive analysis that integrates multi-level information. This multi-layered approach ensures that the model does not simply analyze brain connectivity in isolation but also contextualizes it within a broader population framework. The fusion of these two levels of information allows us to exploit the full spectrum of available data, offering a more holistic view of both individual and group-level patterns. The proposed method is evaluated on the ABIDE-UCLA and ABIDE-NYU datasets, which are widely used benchmarks in the study of ASD. Our BP-GiGNN demonstrates superior performance over state-of-the-art baseline methods in various evaluation settings. The results highlight the significant advantages of our BP-GiGNN method.

The key contributions of our work can be concluded as follows:

- We propose a novel Brain-Population Graph-in-Graph Neural Network (BP-GiGNN) that effectively integrates brain connectivity graphs and population relation graphs to enhance the ASD diagnosis. Our method leverages both internal brain-level GNNs and external population-level GNNs, enabling the simultaneous analysis of intricate brain connectivity patterns and inter-individual relationships.
- We propose a comprehensive strategy that embeds internal brain features into a population graph structure for better multi-level information fusion, which significantly improves predictive performance.
- We evaluate the effectiveness of BP-GiGNN on several public datasets and achieve state-of-the-art performance across various diagnostic settings and metrics, which offers robust and accurate diagnosis assistance.

2 Related Work

2.1 Autism Spectrum Disorder Prediction

Predicting neurodevelopmental disorders, such as Autism Spectrum Disorder (ASD), has been a critical area of research in medical image analysis [24]. Researchers have increasingly focused on using neuroimaging data, including structural and functional MRI, to identify biomarkers that can aid in the early detection and diagnosis of ASD [25, 26]. These imaging techniques provide insights into the anatomical and functional abnormalities associated with neurodevelopmental disorders, which are often subtle and difficult to detect through behavioral assessments alone.

In the field of ASD prediction, early methods [12, 13] focused on statistical analysis of rs-fMRI data to construct handcrafted features and traditional machine learning approaches. The authors of [12] employ functional connectivity analysis to classify individuals with ASD by pinpointing key brain embeddings. They use statistical techniques to analyze the rs-fMRI data and generate connectivity maps that highlight key differences in brain networks between ASD individuals and neurotypical controls, which are further exploited as the input of a support vector machine for diagnosis results. The researchers in [13] apply a functional random forest approach to explore the heterogeneity of executive function across individuals with ADHD and ASD. They use statistical analysis to extract functional connectivity features from rs-fMRI data, which are then input into a random forest classifier. While these handcrafted methods provide valuable insights into brain connectivity, they are limited by their reliance on manual feature selection and often struggle with capturing complex, non-linear patterns inherent in the data.

Recently, various deep learning-based methods have been proposed for ASD prediction [14, 27, 28, 15]. The study by [27] proposes a deep learning model to identify ASD using data from the ABIDE dataset, which trains a neural network on rs-fMRI data, permitting the model to derive pertinent embeddings from the inputs. The model demonstrates robustness across different sites and imaging protocols, highlighting the potential of deep learning to generalize well across diverse datasets and conditions. The ASD-SAENet method proposed by [15] combines a sparse autoencoder with a deep neural network for detecting ASD using fMRI data. The sparse autoencoder is designed to efficiently identify the most prominent features from the high-dimensional fMRI data by learning a compressed

representation, which is leveraged by the neural network for final decision. These deep learning-based methods show high effectiveness in automatically learning complex, non-linear features from rs-fMRI data, offering improved accuracy and robustness across diverse datasets.

2.2 GNN-based Methods for ASD prediction

Graph Neural Networks (GNNs) have attracted significant attention in the field of brain network analysis and Autism Spectrum Disorder (ASD) classification [29]. GNNs extend traditional neural networks to handle graph data, where nodes represent entities (brain regions), and edges represent relationships between these entities (connectivity in the brain). This ability to directly model the relational structure of data makes GNNs particularly well-suited for analyzing complex networks like those found in the human brain.

Generally, there are two categories of GNN-based ASD prediction approaches: brain-based methods [16, 30] and population-based methods [21, 22]. The brain-based methods for ASD prediction treat the problem as a graph classification task, where the brain is represented as a connectivity graph, and each node is a region of interest (ROI). This modeling allows for the analysis of complex neural connections and interactions that are central to understanding ASD. For instance, BrainGNN [16] utilizes GNNs to learn complex representations of these brain connectivity graphs and incorporates an interpretable mechanism that identifies the most critical ROIs and connections. The population-based methods approach ASD prediction as a node classification task, where each node in the graph represents an individual, and the entire population is modeled as a population relation graph. These methods enable the analysis of relationships and similarities between individuals within a population, incorporating both neuroimaging data and non-imaging attributes such as age and gender. For example, the work by [21] develops a model that utilizes spectral graph convolutions to predict diseases like ASD in a population-based context and captures the relationships between individuals, where individuals with greater similarity have stronger connections.

More recently, researchers have attempted to adopt GNN for the joint learning of both brain connectivity and population relation graphs, integrating the strengths of both approaches to enhance the ASD prediction [8, 23]. Hi-GCN [23] designs a hierarchical graph convolutional network to learn graph feature embeddings while considering both the topology of individual brain networks and the relationships between subjects in the global population. LG-GNN [8] proposes a local-to-global graph neural network to bridge the gap between local brain region analysis and global population-level classification through a dual-stage process. Although demonstrating promising performance in classifying brain disorders, these joint learning methods still face the inherent challenge of effectively integrating the diverse types of information [8]. Thus, it remains an open challenge for ASD prediction to effectively integrate ROI-level and individual-level information from the heterogeneous graph structures.

3 Brain-Population Graph-in-Graph Neural Network

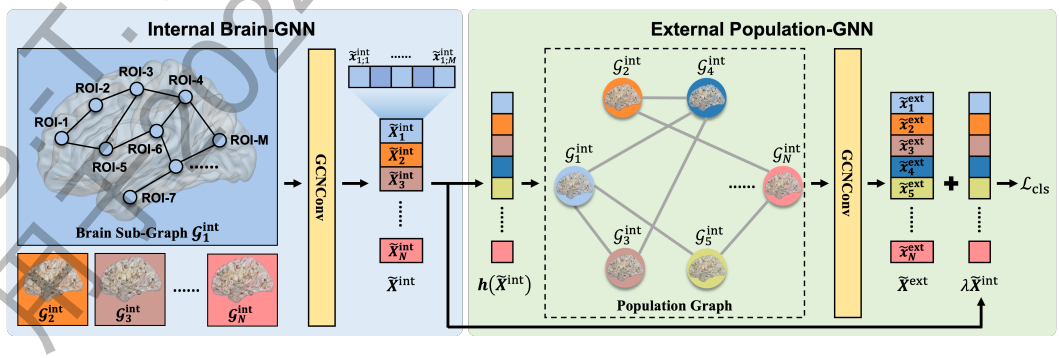


Figure 3: Illustration of the proposed BP-GiGNN method. The input rs-fMRI data is first processed by the **Internal Brain-GNN** to extract brain-level features. These brain-level subgraphs are then incorporated and processed by the **External Population-GNN** to obtain the overall prediction.

3.1 Overall Framework

In this paper, we propose a novel framework named Brain-Population Graph-in-Graph Neural Network (BP-GiGNN) to enhance the performance of neurodevelopmental disorder diagnosis. The BP-GiGNN framework is designed to integrate both individual brain connectivity information and broader population-level relationships. As shown in Fig. 3, the proposed BP-GiGNN method consists of two interconnected components: an internal brain-GNN and an external population-GNN.

The **Internal Brain-GNN** focuses on modeling each individual's rs-fMRI data as a brain connectivity graph. In this component, each region of interest (ROI) within the brain is represented as a node, and the functional connections between these regions are represented as edges. The Internal Brain-GNN is responsible for capturing intricate patterns of brain activity and interactions among different brain regions, which are crucial for identifying disease-related biomarkers. By analyzing the connectivity structure within an individual's brain, the Internal Brain-GNN extracts features that are indicative of neurodevelopmental disorders, allowing the model to pinpoint specific neural abnormalities that may be associated with ASD.

Built upon the brain-level embeddings of the internal graphs, the **External Population-GNN** extends the analysis to a broader context by constructing a population relation graph across different individuals. This component leverages the shared feature patterns across individuals by connecting nodes (individuals) based on similarities in their brain connectivity profiles and non-imaging data. The External Population-GNN captures inter-individual relationships and patterns that are not apparent when considering brain data in isolation. By integrating these relationships, the model can identify commonalities and differences across the population, enhancing the robustness and accuracy of the diagnostic predictions.

By embedding individual brain graphs into a larger population graph, BP-GiGNN effectively captures both internal patterns (within the brain) and external patterns (across the population). This dual-level integration provides a comprehensive understanding of neurodevelopmental disorders, enabling the model to make more informed and accurate predictions. The ability to simultaneously analyze local brain connectivity and global population relationships ensures that BP-GiGNN can account for the complex and heterogeneous nature of brain disorders, ultimately resulting in enhanced diagnostic accuracy and a deeper understanding of the fundamental neural processes.

3.2 Internal Brain-GNN

3.2.1 Brain Graph Construction

Constructing the brain connectivity graph is a critical step in capturing the intricate functional interactions within the brain, which are essential for ASD diagnosis [5]. The brain's functional connectivity, which reflects the coordinated activity between different regions, provides valuable insights into how various parts of the brain communicate and how these communications may be disrupted in the presence of neurodevelopmental disorders.

To construct this internal brain graph, we start with the rs-fMRI data of each individual. The entire brain can be systematically divided into M regions of interest (ROIs) according to a certain brain atlas. Each ROI represents a distinct area of the brain, and its corresponding node, denoted as v_i^{int}, is associated with the rs-fMRI time series data $\boldsymbol{x}_i^{\text{int}}$.

In this way, the internal brain connectivity graph is constructed as $\mathcal{G}^{\text{int}} = \{\boldsymbol{V}^{\text{int}}, \boldsymbol{A}^{\text{int}}\}$, where $\boldsymbol{V}^{\text{int}} = \{v_1^{\text{int}}, \cdots, v_M^{\text{int}}\}$ represents the set of nodes corresponding to the M ROIs. The node features $\boldsymbol{X}^{\text{int}} = (\boldsymbol{x}_1^{\text{int}}, \cdots, \boldsymbol{x}_M^{\text{int}})^\top$ encapsulate the time series data for each ROI, providing a detailed representation of neural activity across the brain. The edges in this graph, represented by the adjacency matrix $\boldsymbol{A}^{\text{int}} \in \mathbb{R}^{M \times M}$, are defined based on the functional relationships between pairs of ROIs as the following:

$$\boldsymbol{A}_{i,j}^{\text{int}} = \begin{cases} |\text{corr}(\boldsymbol{x}_i^{\text{int}}, \boldsymbol{x}_j^{\text{int}})| & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \tag{1}$$

where $\text{corr}(\cdot, \cdot)$ is the Pearson correlation coefficient. Specifically, for each pair of ROIs i and j (where $i \neq j$), the weight of the edge $\boldsymbol{A}_{i,j}^{\text{int}}$ is calculated as the absolute value of correlation coefficient between their respective time series data $\boldsymbol{x}_i^{\text{int}}$ and $\boldsymbol{x}_j^{\text{int}}$. This correlation quantifies the degree of synchronized activity between the two regions, reflecting how strongly they are functionally

connected. Notably, self-loops are removed from the graph, setting $\boldsymbol{A}^{\text{int}}_{i,j} = 0$ when $i = j$, to focus solely on inter-regional interactions.

By representing the ROIs as nodes and their functional relationships as edges, the brain connectivity graph encapsulates the complex network of neural interactions within the brain. This graph-based representation enables further analysis of brain activity patterns, facilitating the identification of potential biomarkers and the characterization of functional disruptions that are associated with brain disorders. The constructed brain connectivity graph provides a rich source of information that can be leveraged by GNNs to enhance the accuracy and interpretability of brain disorder diagnoses.

3.2.2 Internal GNN Model

Once the brain connectivity graph is constructed, the next step is to process these graphs using the internal GNN to extract meaningful features that capture the complex neural interactions within the brain. Operating on the initial node features derived from the brain connectivity graph, the internal GNN can transform these features through multiple layers to produce rich embeddings that are crucial for downstream tasks.

The internal GNN model in our framework utilizes two Graph Convolutional (GCNConv) layers, which are central to the feature extraction process. GCNConv layers extend the principles of convolutional neural networks to graph-structured data, enabling the model to aggregate information from a node's neighbors in the graph, allowing the GNN to capture both local and global patterns within the brain's connectivity structure.

Specifically, the operation of the i-th $(i = 1, \cdots, K)$ GCNConv layer is calculated as:

$$\begin{aligned} \boldsymbol{X}^{(i)} &= \text{GCNConv}(\boldsymbol{X}^{(i-1)}, \boldsymbol{A}^{\text{int}}; \boldsymbol{W}^{(i-1)}) \\ &= \sigma(\boldsymbol{D}^{-1/2}\boldsymbol{A}^{\text{int}}\boldsymbol{D}^{-1/2}\boldsymbol{X}^{(i-1)}\boldsymbol{W}^{(i-1)}) \end{aligned} \tag{2}$$

where $\boldsymbol{X}^{(i-1)}$ is the input node features to the i-th layer, $\boldsymbol{X}^{(i)}$ is the output node embeddings after applying the graph convolution, $\boldsymbol{W}^{(i-1)}$ is the learnable parameter, $\sigma(\cdot)$ is the activation function like ReLU. Here, $\boldsymbol{D}_{i,i} = \sum_j \boldsymbol{A}^{\text{int}}_{i,j}$ is the diagonal degree matrix of the brain connectivity graph, which normalizes the adjacency matrix $\boldsymbol{A}^{\text{int}}$ to account for variations in node connectivity.

Initially, the input to the first GCNConv layer is the matrix $\boldsymbol{X}^{(0)}$, which corresponds to the raw input data $\boldsymbol{X}^{\text{int}}$. As these features pass through the successive GCNConv layers, the model progressively refines them, resulting in higher-level embeddings that encapsulate the most relevant aspects of the brain's functional connectivity for disorder diagnosis. The output of the internal GNN model is denoted as $\tilde{\boldsymbol{X}}^{\text{int}} = \boldsymbol{X}^{(K)}$, which represents the extracted internal embeddings of the brain graph $\mathcal{G}^{\text{int}}$. By effectively transforming the raw input data into meaningful embeddings, the internal GNN model lays the foundation for accurate and interpretable diagnosis within the broader BP-GiGNN framework.

3.3 External Population-GNN

3.3.1 Population Graph Construction

In addition to the internal brain connectivity within an individual, it is also essential to consider the broader context of how different individuals within a population are connected. These inter-individual associations, coupled with individual-level non-imaging data, also contribute significantly to the overall accuracy and robustness of disease prediction models [21]. By integrating both neurological and non-neurological factors, we can capture a more comprehensive view of the disorder, accounting for both individual variability and shared characteristics across the population.

To systematically explore these inter-individual correlations, we construct the external population graph, which is designed to model the relationships between different individuals based on their internal brain subgraphs and extra non-imaging phenotypic data. The external population graph is denoted as $\mathcal{G}^{\text{ext}} = \{\boldsymbol{V}^{\text{ext}}, \boldsymbol{A}^{\text{ext}}\}$, where the node set $\boldsymbol{V}^{\text{ext}} = \{\mathcal{G}^{\text{int}}_1, \cdots, \mathcal{G}^{\text{int}}_N\}$ represents the internal brain subgraphs of all N individuals in the study. For each node in the external population graph, it corresponds to an internal brain subgraph $\mathcal{G}^{\text{int}}_j$ of a specific individual j, as defined and constructed in Section 3.2. The ROI-level features extracted from each individual's internal brain graph, represented

by the embeddings $\tilde{\boldsymbol{X}}^{\text{int}}$, serve as the foundational data for constructing the nodes in the external population graph.

To facilitate the transition from individual brain connectivity to population-level analysis, we derive external node embeddings $\boldsymbol{X}^{\text{ext}}$ from the brain-level embeddings $\tilde{\boldsymbol{X}}^{\text{int}}$ obtained from the internal GNN as

$$\boldsymbol{X}^{\text{ext}} = (\boldsymbol{h}(\tilde{\boldsymbol{X}}_1^{\text{int}}), \cdots, \boldsymbol{h}(\tilde{\boldsymbol{X}}_N^{\text{int}})) \tag{3}$$

$$\boldsymbol{h}(\tilde{\boldsymbol{X}}_j^{\text{int}}) = \left(\bigcup_{k=1}^{M-1} \bigcup_{l=k+1}^{M} \text{corr}(\tilde{\boldsymbol{x}}_{j;k}^{\text{int}}, \tilde{\boldsymbol{x}}_{j;l}^{\text{int}}) \right) \tag{4}$$

where $\text{corr}(\cdot, \cdot)$ is the correlation function, M is the number of ROIs. Specifically, the internal correlation function $\boldsymbol{h}(\cdot)$ is designed to capture the pairwise correlations between all ROI-level features within each individual's brain, effectively summarizing the internal connectivity patterns in a single vector. For a specific individual j, the correlation between every pair of ROI embeddings, $\tilde{\boldsymbol{x}}_{j;k}^{\text{int}}$ and $\tilde{\boldsymbol{x}}_{j;l}^{\text{int}}$, is calculated. The resulting coefficients are concatenated into a $M(M-1)/2$-dim vector, which serves as the input embedding for the j-th brain subgraph in the population graph.

Once the external node embeddings $\boldsymbol{X}^{\text{ext}}$ are constructed for all N individuals, the adjacency matrix $\boldsymbol{A}^{\text{ext}}$ of the external population graph is computed similarly to Eq. (1). Furthermore, to fully capture the nuances of inter-individual relationships, we follow the previous work [8] to refine this adjacency matrix by incorporating non-imaging information metrics, including gender, age, and site. By constructing this external population graph, we provide a comprehensive framework that not only accounts for the detailed brain connectivity within individuals but also leverages the shared characteristics and relationships across the entire population.

3.3.2 External GNN Model

After constructing the population graph, the next step is the implementation of the external GNN model. This model should effectively aggregate and leverage information from correlated individuals, integrating both brain-level and population-level data to enhance the predictive power of the system.

Specifically, we employ two Graph Convolutional Network (GCNConv) layers that incorporate the jumping knowledge mechanism [31]. Jumping knowledge allows the model to dynamically select and aggregate information from different layers of the network, effectively capturing both local and global patterns within the graph. This mechanism is particularly useful in complex graph structures, such as population-level analyses with multi-scale relationships. The result of this process is a set of external embeddings $\tilde{\boldsymbol{X}}^{\text{ext}} = (\tilde{\boldsymbol{x}}_1^{\text{ext}}, \cdots, \tilde{\boldsymbol{x}}_N^{\text{ext}})^\top$, where each vector $\tilde{\boldsymbol{x}}_j^{\text{ext}}$ represents the aggregated feature embedding for the j-th subgraph $\mathcal{G}_j^{\text{int}}$ in the population graph.

Recognizing the multi-level nature of neurodevelopmental disorder indicators, we further enhance the discriminative performance of our model by fusing the internal and external embeddings. The internal embeddings $\tilde{\boldsymbol{X}}^{\text{int}}$ derived from the brain graph, are combined with the external embeddings $\tilde{\boldsymbol{X}}^{\text{ext}}$ derived from the population graph, to obtain the final embeddings as

$$\tilde{\boldsymbol{X}}^{\text{final}} = \tilde{\boldsymbol{X}}^{\text{ext}} + \lambda \tilde{\boldsymbol{X}}^{\text{int}} \tag{5}$$

where λ is a learnable weight parameter that determines the relative contribution of the internal and external embeddings. By allowing the model to adjust this parameter during training, we ensure that the final embeddings optimally reflect the most relevant information from both levels of analysis.

The BP-GiGNN is supervised by the cross-entropy objective

$$\mathcal{L} = -\sum_{i=1}^{N} [(1 - y_i) \log(1 - \hat{y}_i) + y_i \log(\hat{y}_i)] \tag{6}$$

where $\hat{y}_i$ is the prediction output, y_i denotes the ground truth label of individual i. Through this comprehensive approach, the external GNN model learns to effectively leverage both the internal brain-level and external population-level embeddings.

4 Experimental Setup

4.1 Datasets

In the experiments, we evaluate the proposed method on the widely recognized Autism Brain Imaging Data Exchange (ABIDE) dataset [32], which offers a valuable dataset of neuroimaging information specifically assembled for researching Autism Spectrum Disorder (ASD). The ABIDE initiative compiles neuroimaging datasets from a wide range of research institutions worldwide. To thoroughly assess the performance of our method, we focus on two distinct subsets of the ABIDE dataset: ABIDE-UCLA and ABIDE-NYU. These subsets were selected based on their consistency in data collection methods and their representativeness of the ABIDE dataset. The raw sequences are then preprocessed by the widely adopted DPARSF toolbox [33] to enhance the quality of the brain connectivity graphs used in our experiments. The details of the datasets are as follows:

- **ABIDE-NYU** includes 262 subjects, with 127 having an ASD diagnosis and 135 as normal ones. Similar to ABIDE-UCLA, ABIDE-NYU is also a merged dataset, bringing together the NYU Sample from ABIDE-I and the NYU Sample 1 from ABIDE-II. These samples were collected at NYU Langone Medical Center and exhibit overlaps in phenotypic characterization and scan parameters.
- **ABIDE-UCLA** includes data from 141 participants, with 78 individuals diagnosed with ASD and 63 healthy controls. The ABIDE-UCLA subset is a composite dataset, formed by merging several related samples: UCLA Sample 1 and UCLA Sample 2 from ABIDE-I, along with the UCLA Sample from ABIDE-II. These samples were all collected at the University of California, Los Angeles, which share similar inclusion criteria.

4.2 Baselines & Metrics

The baseline methods selected for comparison in this paper include general graph neural network (GNN) models and GNN-based brain disorder prediction models. **General GNN models** include GCN [34], GraphSAGE [35], and GAT [36]. GCN is a seminal model in the field of graph-based learning that applies convolutional operations directly to graph-structured data. GraphSAGE extends the GCN framework by introducing an inductive learning approach, which generates node embeddings through a process of neighborhood sampling and aggregation. GAT enhances graph neural networks by incorporating an attention mechanism that allows the model to assign varying levels of importance to different neighbors during feature aggregation. **GNN-based brain disorder prediction models** consist of LG-GNN [8]. LG-GNN is a specialized GNN model for brain disorder prediction, which combines local brain region analysis with global population-level relationships.

We assess the efficacy of our model and other methods by employing a wide array of metrics commonly used in neuroimaging-based classification tasks, such as Accuracy, Sensitivity, and Specificity. Additionally, we also calculate F1-Score and Area Under the ROC Curve to provide a comprehensive evaluation.

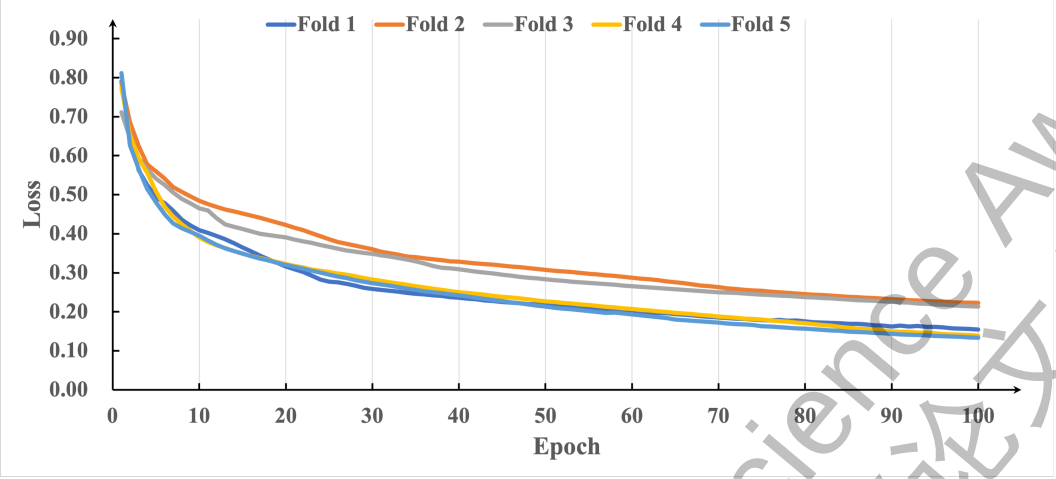
4.3 Implementation Details

Setting the learning rate to 1×10^{-4} and implementing a weight decay of 5×10^{-5}, the model undergoes training for 200 epochs. The model is trained for 200 epochs. In order to ensure thorough evaluation, we employ 5-fold cross-validation on both datasets. For each fold, the data is split into training and validation subsets, and the model is fit and evaluated based on this partitioning. We report the results across the five folds, providing a comprehensive assessment of the model's performance and its generalizability across different subsets of the data. The proposed BP-GiGNN is implemented by PyTorch.

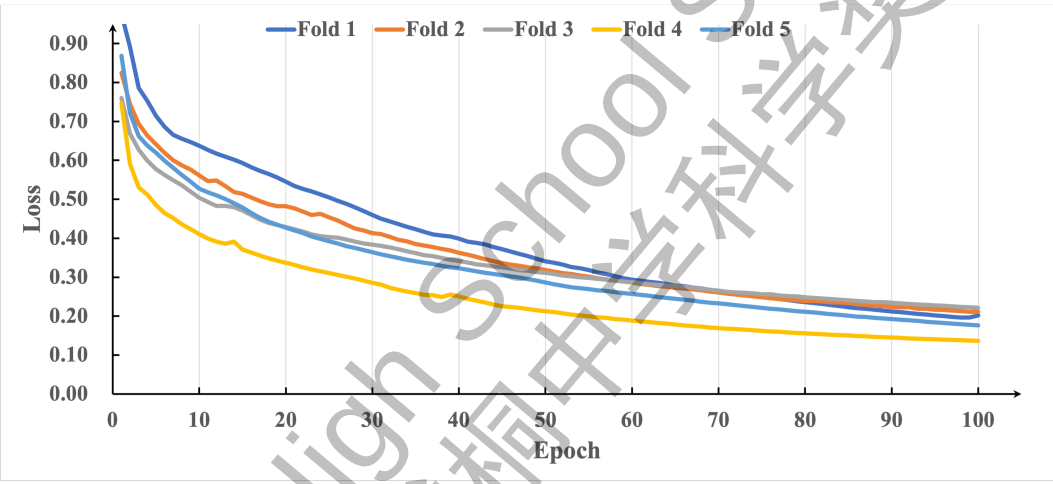
5 Experimental Results

5.1 Comparison Study

To assess the quantitative performance of our BP-GiGNN model, we conduct an extensive evaluation with several leading baselines on two benchmark datasets: ABIDE-NYU and ABIDE-UCLA. We



(a) on the ABIDE-NYU dataset.



(b) on the ABIDE-UCLA dataset.

Figure 4: Training loss of the proposed BP-GiGNN across 5 folds.

depict the loss function values in Fig. 4. We detail the results in Table 1 and Table 2. Based on the results, we see three conclusions:

(1) The BP-GiGNN approach consistently surpasses all leading baseline methods across various evaluation settings and metrics. Specifically, on the ABIDE-NYU dataset, BP-GiGNN exceeds the second-best method, LG-GNN, by a significant margin with improvements of 11.38% in accuracy and comparable gains in other metrics. Similarly, on the ABIDE-UCLA dataset, BP-GiGNN demonstrates an increase of 15.80% in accuracy compared to LG-GNN. These substantial improvements can be attributed to the advantages of our model, which integrates both internal brain-level and external population-level information through a graph-in-graph architecture. This dual-level integration allows BP-GiGNN to capture complex intra-individual and inter-individual patterns more effectively, leading to more accurate and robust predictions.

(2) General GNN models, such as GCN, GraphSAGE, and GAT, perform relatively poorly when applied directly to ASD diagnosis. This underperformance is evident in their lower accuracy and AUC scores compared to specialized brain disorder prediction models. The primary reason for this is that these general GNN models are not designed to fully exploit the rich, multi-level relationships inherent in neuroimaging data. Specifically, they struggle to effectively model the intricate associations between internal brain connectivity (brain-level information) and external phenotypic relationships

Table 1: The comparison results of the compared approaches on ABIDE-NYU.

Method	Acc (%)	Sen (%)	Spe (%)	AUC (%)	F1 (%)
GCN [34]	$54.89_{\pm 4.40}$	$52.50_{\pm 7.64}$	$57.05_{\pm 5.37}$	$61.07_{\pm 6.23}$	$52.55_{\pm 6.06}$
SAGE [35]	$60.96_{\pm 3.69}$	$55.87_{\pm 5.47}$	$65.69_{\pm 4.68}$	$64.18_{\pm 3.83}$	$57.78_{\pm 4.48}$
GAT [36]	$59.35_{\pm 3.43}$	$56.81_{\pm 5.35}$	$61.85_{\pm 8.16}$	$61.11_{\pm 4.16}$	$57.22_{\pm 3.24}$
LG-GNN [8]	$74.38_{\pm 3.10}$	$70.29_{\pm 13.74}$	$78.06_{\pm 9.61}$	$85.83_{\pm 1.34}$	$71.79_{\pm 6.28}$
Internal-GNN	$58.97_{\pm 5.72}$	$57.68_{\pm 6.27}$	$60.15_{\pm 6.63}$	$63.87_{\pm 5.85}$	$57.42_{\pm 5.67}$
Internal-GNN+GCN	$79.23_{\pm 12.57}$	$92.36_{\pm 4.95}$	$66.86_{\pm 27.75}$	$93.57_{\pm 3.04}$	$82.07_{\pm 7.55}$
Ours	$\mathbf{85.76}_{\pm 5.50}$	$\mathbf{92.50}_{\pm 8.08}$	$\mathbf{79.94}_{\pm 15.78}$	$\mathbf{98.41}_{\pm 1.41}$	$\mathbf{86.45}_{\pm 3.97}$

Table 2: The comparison results of the compared approaches on ABIDE-UCLA.

Method	Acc (%)	Sen (%)	Spe (%)	AUC (%)	F1 (%)
GCN [34]	$61.77_{\pm 7.60}$	$72.56_{\pm 11.25}$	$46.67_{\pm 10.89}$	$71.52_{\pm 9.70}$	$68.45_{\pm 7.04}$
SAGE [35]	$59.70_{\pm 7.00}$	$64.36_{\pm 8.98}$	$53.33_{\pm 12.96}$	$66.62_{\pm 6.62}$	$64.74_{\pm 6.87}$
GAT [36]	$57.97_{\pm 6.50}$	$64.36_{\pm 22.09}$	$48.89_{\pm 20.61}$	$58.28_{\pm 10.96}$	$62.02_{\pm 12.12}$
LG-GNN [8]	$72.90_{\pm 8.80}$	$78.97_{\pm 12.62}$	$64.44_{\pm 4.44}$	$80.91_{\pm 10.08}$	$76.73_{\pm 8.32}$
Internal-GNN	$62.60_{\pm 5.28}$	$67.69_{\pm 11.68}$	$55.56_{\pm 9.94}$	$65.90_{\pm 9.79}$	$67.23_{\pm 6.69}$
Internal-GNN+GCN	$75.67_{\pm 9.12}$	$69.10_{\pm 24.42}$	$84.44_{\pm 15.07}$	$87.56_{\pm 8.24}$	$74.19_{\pm 14.31}$
Ours	$\mathbf{88.70}_{\pm 9.77}$	$\mathbf{87.05}_{\pm 13.23}$	$\mathbf{91.11}_{\pm 10.89}$	$\mathbf{96.01}_{\pm 5.30}$	$\mathbf{89.53}_{\pm 9.42}$

(population-level information). As a result, they fail to capture the full spectrum of relevant features needed for accurate ASD classification.

(3) GNN-based brain disorder prediction models like LG-GNN exhibit advancements compared to traditional GNN models, but there remains room for further enhancement. These models do achieve better performance by leveraging multi-level relational information, which helps in capturing the complex structures within brain networks and across individuals. However, they still face significant challenges in the effective fusion of heterogeneous graph information across different levels. The integration of diverse data types (e.g., neuroimaging features and demographic information) within a unified model is complex, and current methods struggle with this aspect, leading to suboptimal performance. BP-GiGNN addresses these challenges more comprehensively, but the results suggest that ongoing research is needed to further refine the methods for integrating multi-layered and heterogeneous data sources.

5.2 Ablation Study

Furthermore, an ablation study is conducted in our BP-GiGNN by systematically removing or modifying parts of the model and observing the resulting changes in performance. Specifically, three variations of the BP-GiGNN model are designed as:

- **Internal-GNN**: The model of internal brain-GNN only.
- **Internal-GNN+GCN**: The model of internal brain-GNN and plain GCN, i.e., replacing the external population-GNN of BP-GiGNN with plain GCN.
- **Ours**: The model of both internal brain-GNN and external population-GNN, i.e., the proposed BP-GiGNN.

The results are presented in Table 1 and Table 2. Several observations can be made as follows:

(1) Only using internal-GNN leads to suboptimal performance, similar to the results observed with general GNN models. This underperformance can be attributed to its inability to leverage the multi-level relational information that is crucial for accurately diagnosing ASD. By focusing solely on brain-level connectivity within individuals, the internal-GNN fails to capture the broader population-level relationships that provide additional context and insights necessary for more accurate predictions. This limitation highlights the importance of integrating multi-level data for comprehensive modeling in neurodevelopmental disorder diagnosis.

(2) The performance is improved when adding a plain GCN to the internal-GNN. This enhancement indicates that incorporating population-level relational information, even in a basic form, is beneficial for ASD diagnosis. The plain GCN introduces an additional layer of analysis that considers the similarities and differences between individuals at the population level, which helps to refine the model's predictions. This result underscores the value of including population-level data in the diagnostic process, as it complements the brain-level features captured by the internal-GNN and contributes to a more accurate overall model.

(3) The BP-GiGNN model, which combines both the internal-GNN and the external population-GNN, achieves the best performance across all metrics. This superior performance demonstrates the effectiveness of integrating both internal brain-level and external population-level information within a unified framework. By leveraging the strengths of both components, BP-GiGNN can capture the intricate, multi-layered relationships that are essential for accurately diagnosing neurodevelopmental disorders. The model's ability to synthesize detailed brain connectivity patterns with broader population-level interactions allows it to provide more robust and reliable predictions. This finding reinforces the critical importance of a holistic, multi-level approach in advancing the accuracy and effectiveness of neurodevelopmental disorder models.

5.3 Visualization

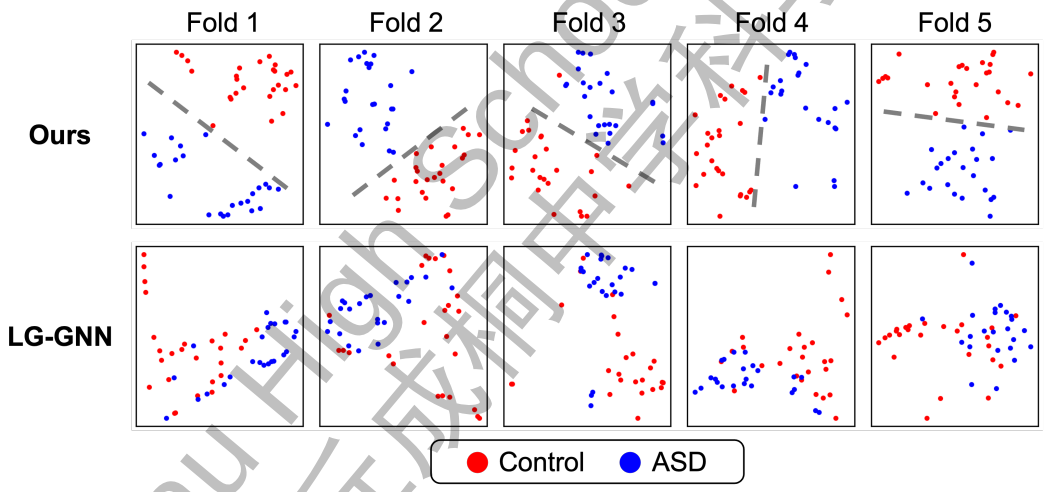


Figure 5: Scatter plot of the embeddings extracted by our BP-GiGNN and the compared LG-GNN.

To qualitatively assess the efficacy of our BP-GiGNN model, we visualize the embeddings extracted by our BP-GiGNN and the compared LG-GNN. To qualitatively assess the efficacy of the proposed approach, we extract and visualize the features Specifically, we extract the features from the penultimate layer of both our method and LG-GNN and reduce their dimensionality to 2D space using t-SNE [37]. The scatter plot of the extracted features for all 5 folds is presented in Fig. 5.

It is demonstrated that the features learned by our method exhibit significantly better separability between healthy controls and ASD patients, compared with the features extracted by LG-GNN. In the visualization of our method, the clusters corresponding to different classes are more distinct with tighter intra-class grouping and greater inter-class separation. This indicates that the embeddings extracted by BP-GiGNN have stronger discriminative power, leading to more accurate classification. In contrast, the feature representation from LG-GNN shows greater overlap between the two classes, suggesting that it struggles to fully capture the underlying differences between ASD and healthy individuals. This further underscores the effectiveness of BP-GiGNN's graph-in-graph approach, which integrates both internal brain connectivity and external population relations to provide a more comprehensive and robust feature extraction process.

6 Conclusion

In this paper, we propose a novel Brain-Population Graph-in-Graph Neural Network (BP-GiGNN) to address the challenges associated with predicting Autism Spectrum Disorder (ASD), using resting-state functional MRI (rs-fMRI) data. Our method offers a unified approach that effectively integrates both brain-level internal connectivity information and population-level external attributes. In BP-GiGNN, the internal brain-GNN captures the intricate neural interactions within the brain, while the external population-GNN models relationships between individuals, leveraging both brain connectivity embeddings and non-imaging attributes. The comprehensive experimental results on two public datasets, ABIDE-NYU and ABIDE-UCLA, demonstrate the superior performance of BP-GiGNN over several state-of-the-art models, including traditional graph neural networks and specialized brain disorder diagnosis approaches.

In the future, there are several directions for further enhancing the BP-GiGNN model and its applications: 1) Exploring interpretability of GNNs in ASD diagnosis. While BP-GiGNN has demonstrated strong performance, a key area for future work is improving the interpretability of the model. By explaining the model in identifying specific brain regions or population-level relationships that contribute to ASD diagnosis, we can gain deeper insights into the neural processes and offer clinicians more practical guidance. 2) Incorporating multi-modal data for more accurate diagnosis. Combining rs-fMRI data with other sources of information (genetic data, behavioral assessments) could enhance the model's ability to capture the full complexity of ASD, which is expected to result in more precise and resilient diagnosis results. 3) Extending to other neurodevelopmental disorders. Although BP-GiGNN is focused on ASD, future work could involve applying the model to other neurodevelopmental disorders, such as ADHD. By validating the model on different datasets and extending its application to other conditions, we can assess its generalizability and further expand its utility into a wider range.

References

[1] Anita Thapar, Miriam Cooper, and Michael Rutter. Neurodevelopmental disorders. *The Lancet Psychiatry*, 4(4):339–346, 2017.

[2] Bibiana Restrepo, Janice Enriquez, and Robin L Hansen. Signs and symptoms of autism spectrum disorder. *Developmental-Behavioral Pediatrics E-Book*, page 431, 2022.

[3] Qi Zhou, Yuling Lei, Hang Du, and Yuexian Tao. Public concerns and attitudes towards autism on chinese social media based on k-means algorithm. *Scientific Reports*, 13(1):15173, 2023.

[4] Matthew J Maenner. Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2016. *MMWR. Surveillance Summaries*, 69, 2020.

[5] Noriaki Yahata, Jun Morimoto, Ryuichiro Hashimoto, Giuseppe Lisi, Kazuhisa Shibata, Yuki Kawakubo, Hitoshi Kuwabara, Miho Kuroda, Takashi Yamada, Fukuda Megumi, et al. A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nature communications*, 7(1):11254, 2016.

[6] Geraldine Dawson, Sally Rogers, Jeffrey Munson, Milani Smith, Jamie Winter, Jessica Greenson, Amy Donaldson, and Jennifer Varley. Randomized, controlled trial of an intervention for toddlers with autism: the early start denver model. *Pediatrics*, 125(1):e17–e23, 2010.

[7] Annette Estes, Jeffrey Munson, Sally J Rogers, Jessica Greenson, Jamie Winter, and Geraldine Dawson. Long-term outcomes of early intervention in 6-year-old children with autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(7):580–587, 2015.

[8] Hao Zhang, Ran Song, Liping Wang, Lin Zhang, Dawei Wang, Cong Wang, and Wei Zhang. Classification of brain disorders in rs-fmri via local-to-global graph neural networks. *IEEE transactions on medical imaging*, 42(2):444–455, 2022.

[9] Torbjörn Falkmer, Katie Anderson, Marita Falkmer, and Chiara Horlin. Diagnostic procedures in autism spectrum disorders: a systematic literature review. *European child & adolescent psychiatry*, 22:329–340, 2013.

[10] Meenakshi Khosla, Keith Jamison, Gia H Ngo, Amy Kuceyeski, and Mert R Sabuncu. Machine learning in resting-state fmri analysis. *Magnetic resonance imaging*, 64:101–121, 2019.

[11] Dongren Yao, Jing Sui, Mingliang Wang, Erkun Yang, Yeerfan Jiaerken, Na Luo, Pew-Thian Yap, Mingxia Liu, and Dinggang Shen. A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity. *IEEE transactions on medical imaging*, 40(4):1279–1289, 2021.

[12] Mark Plitt, Kelly Anne Barnes, and Alex Martin. Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *NeuroImage: Clinical*, 7:359–366, 2015.

[13] Michaela Cordova, Kiryl Shada, Damion V Demeter, Olivia Doyle, Oscar Miranda-Dominguez, Anders Perrone, Emma Schifsky, Alice Graham, Eric Fombonne, Beth Langhorst, et al. Heterogeneity of executive function revealed by a functional random forest approach across adhd and asd. *NeuroImage: Clinical*, 26:102245, 2020.

[14] Md Delowar Hossain, Muhammad Ashad Kabir, Adnan Anwar, and Md Zahidul Islam. Detecting autism spectrum disorder using machine learning. *arXiv preprint arXiv:2009.14499*, 2020.

[15] Fahad Almuqhim and Fahad Saeed. Asd-saenet: a sparse autoencoder, and deep-neural network model for detecting autism spectrum disorder (asd) using fmri data. *Frontiers in Computational Neuroscience*, 15:654315, 2021.

[16] Xiaoxiao Li, Yuan Zhou, Nicha Dvornek, Muhan Zhang, Siyuan Gao, Juntang Zhuang, Dustin Scheinost, Lawrence H Staib, Pamela Ventola, and James S Duncan. Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, 74:102233, 2021.

[17] Xuegang Song, Alejandro Frangi, Xiaohua Xiao, Jiuwen Cao, Tianfu Wang, and Baiying Lei. Integrating similarity awareness and adaptive calibration in graph convolution network to predict disease. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23*, pages 124–133. Springer, 2020.

[18] Sofia Ira Ktena, Sarah Parisot, Enzo Ferrante, Martin Rajchl, Matthew Lee, Ben Glocker, and Daniel Rueckert. Metric learning with spectral graph convolutions on brain connectivity networks. *NeuroImage*, 169:431–442, 2018.

[19] Kamalaker Dadi, Mehdi Rahim, Alexandre Abraham, Darya Chyzhyk, Michael Milham, Bertrand Thirion, Gaël Varoquaux, Alzheimer's Disease Neuroimaging Initiative, et al. Benchmarking functional connectome-based predictive models for resting-state fmri. *NeuroImage*, 192:115–134, 2019.

[20] Sarah Parisot, Sofia Ira Ktena, Enzo Ferrante, Matthew Lee, Ricardo Guerrero, Ben Glocker, and Daniel Rueckert. Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer's disease. *Medical image analysis*, 48:117–130, 2018.

[21] Sarah Parisot, Sofia Ira Ktena, Enzo Ferrante, Matthew Lee, Ricardo Guerrero Moreno, Ben Glocker, and Daniel Rueckert. Spectral graph convolutions for population-based disease prediction. In *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III 20*, pages 177–185. Springer, 2017.

[22] Yongxiang Huang and Albert CS Chung. Edge-variational graph convolutional networks for uncertainty-aware disease prediction. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23*, pages 562–572. Springer, 2020.

[23] Hao Jiang, Peng Cao, MingYi Xu, Jinzhu Yang, and Osmar Zaiane. Hi-gcn: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction. *Computers in Biology and Medicine*, 127:104096, 2020.

[24] Monica D Rosenberg, BJ Casey, and Avram J Holmes. Prediction complements explanation in understanding the developing brain. *Nature communications*, 9(1):589, 2018.

[25] Jared A Nielsen, Brandon A Zielinski, P Thomas Fletcher, Andrew L Alexander, Nicholas Lange, Erin D Bigler, Janet E Lainhart, and Jeffrey S Anderson. Multisite functional connectivity mri classification of autism: Abide results. *Frontiers in human neuroscience*, 7:599, 2013.

[26] Heather Cody Hazlett, Hongbin Gu, Brent C Munsell, Sun Hyung Kim, Martin Styner, Jason J Wolff, Jed T Elison, Meghan R Swanson, Hongtu Zhu, Kelly N Botteron, et al. Early brain development in infants at high risk for autism spectrum disorder. *Nature*, 542(7641):348–351, 2017.

[27] Anibal Sólon Heinsfeld, Alexandre Rosa Franco, R Cameron Craddock, Augusto Buchweitz, and Felipe Meneguzzi. Identification of autism spectrum disorder using deep learning and the abide dataset. *NeuroImage: Clinical*, 17:16–23, 2018.

[28] Taban Eslami, Vahid Mirjalili, Alvis Fong, Angela R Laird, and Fahad Saeed. Asd-diagnet: a hybrid learning approach for detection of autism spectrum disorder using fmri data. *Frontiers in neuroinformatics*, 13:70, 2019.

[29] 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.

[30] Xiaodan Xing, Qingfeng Li, Hao Wei, Minqing Zhang, Yiqiang Zhan, Xiang Sean Zhou, Zhong Xue, and Feng Shi. Dynamic spectral graph convolution networks with assistant task training for early mci diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 639–646. Springer, 2019.

[31] Keyulu Xu, Chengtao Li, Yonglong Tian, Tomohiro Sonobe, Ken-ichi Kawarabayashi, and Stefanie Jegelka. Representation learning on graphs with jumping knowledge networks. In *International conference on machine learning*, pages 5453–5462. PMLR, 2018.

[32] Adriana Di Martino, Chao-Gan Yan, Qingyang Li, Erin Denio, Francisco X Castellanos, Kaat Alaerts, Jeffrey S Anderson, Michal Assaf, Susan Y Bookheimer, Mirella Dapretto, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, 19(6):659–667, 2014.

[33] Chaogan Yan and Yufeng Zang. Dparsf: a matlab toolbox for" pipeline" data analysis of resting-state fmri. *Frontiers in systems neuroscience*, 4:1377, 2010.

[34] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

[35] Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017.

[36] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018.

[37] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

致谢

1. **研究背景，指导老师与参赛学生的关系，在论文写作过程中所起的作用，指导是否有偿；以及他人协助完成的研究成果**

本研究的选题是我关注自闭症群体而产生的。通过电视、网络等渠道对世界自闭症日的报道，我了解到自闭症群体——“来自星星的孩子”们在日常生活中所面临的困扰和挑战。我还了解到，早期的发现和干预对自闭症的治疗具有十分关键的意义和巨大的作用。与此同时，在参与实验室的学习和研究过程中，我接触到了图神经网络及其在脑科学领域的应用。了解到图神经网络能够有效分析复杂的神经网络和大脑功能连接，这让我萌生了利用图神经网络及其相关技术来帮助自闭症的早期诊断的想法。我希望设计出更早、更准确的自闭症诊断方法，通过自己的研究，更及时地对自闭症儿童进行干预，帮助他们提高治疗效果和生活质量。

本研究的第一位指导老师是清华大学的高跃老师，他是我在中学生“英才计划”中的指导老师。在整个调研过程中，高跃老师在调研选题、搭建神经网络框架、实验设计、撰写论文等环节，帮助我理清调研思路，确保方法的科学性、可行性，并给予全方位的指导。第二位指导老师是我高中的科技老师，北师大实验附中的马静老师，主要在写论文阶段对我提出宝贵的建议和指导，帮助我在论文结构和表达逻辑上进行优化和提升。感谢两位指导老师的无偿指导，他们的支持不仅帮助我顺利完成了研究，更让我在学术能力和思维方式上得到了极大的提升。

本研究的所有成果均为我在导师指导下独立完成，感谢 PyTorch Geometric 库提供的图神经网络代码实现，感谢 LG-GNN 方法的作者在 GitHub 上开源的代码库。

2. **详细的分工说明;指导老师和参赛学生的具体分工及每一个队员在研究报告撰写中承担的工作以及贡献；不同环节遇到的困难及解决问题的经过**

在研究过程中，本人在导师的指导下确定选题方向、调研相关工作、制定研究方案，完成了数据获取、模型搭建与训练、结果整理与分析，以及最终的研究报告的撰写工作。高跃老师在初期与我讨论确定了研究方向的可行性，在图神经网络、核磁共振数据分析、算法优化等方面提供了宝贵的指导意见。马静老师在研究报告撰写上提供了专业的指导，尤其是帮助我完善了结构组织和表达逻辑。

在初期遇到的主要困难是要确定选题可行性。我自己对于图神经网络的理解仅限于学习过程中的接触，要把它应用到自闭症诊断中还涉及到复杂的神经影像数据，这对于当时的我来说都是不熟悉的领域。为了解决这一问题，我与高跃老师进行了多次的讨论，老师建议我深入阅读与医学影像处理相关的文献，了解该领域现有的研究成果与技术手段。在老师的指导下，我查阅了大量 fMRI 影像分析和神经发育障碍诊断方面的论文，认识到图神经网络在处理脑网络连接中的广阔应用前景。这不仅帮助我更好地理解了研究课题的前沿进展，还让我有了更多的信心去探索这一选题，并且从技术上评估其可行性。在实验阶段遇到的主要困难是模型框架代码的搭建。模型框架涉及许多复杂的步骤，涉及到 rs-fMRI 的脑影像数据，包括数据读取、图结构构建、神经网络的设计与搭建、模型优化与训练、以及测试评估等一系列过程。为了解决这个问题，我通过调研发现了 LG-GNN 等前沿方法的开源代码仓库，对于我的工作具有非常高的参考价值。通过分析这些代码的思路，我成功搭建了 BP-GiGNN 的基础框架，为后续的模型调整和优化打下了重要基础。