MULTI-HOP SIGNED GRAPH CONVOLUTIONAL NETWORK UNDER LONG-TERM DEPENDENCIES SUPERVISION FOR NEURODEVELOPMENTAL DISORDER DIAGNOSIS

Anonymous authors
Paper under double-blind review

ABSTRACT

Brain network analysis leveraging resting-state functional magnetic resonance imaging offers promising prospects for the diagnosis of neurodevelopmental disorders (NDDs). Graph neural networks have demonstrated their powerful ability to extract informative latent features to facilitate diagnostic analysis. Nevertheless, existing signed multi-hop graph convolutional networks (GCNs) are plagued by insufficient prior supervision and the over-smoothing problem, which degrades diagnostic performance. In this study, we propose a long-term guided multi-hop signed graph convolutional network (LT-MSGCN) to identify autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD). We construct a dual-branch architecture consisting of an auto-encoding (AE) branch and a multi-hop GCN branch. Specifically, the transformer-augmented AE branch captures long-term dependencies among brain regions, and the resulting relationship matrices serve as regularizers for multi-hop adjacency matrices generated via approximate personalized propagation of neural predictions (APPNP) within the multi-hop GCN branch. After feature extraction by stacked signed graph convolution layers, hop-aware and region-adaptive attention fusion modules further boost discriminative graph representation learning. Under 10-fold cross-validation on the ABIDE-I and ADHD-200 datasets, LT-MSGCN achieves classification accuracies of 71.53% and 70.20%, respectively. The identified discriminative brain regions show good agreement with established neuroscientific findings. Experimental results demonstrate that LT-MSGCN yields competitive classification performance and favorable biological interpretability for the auxiliary diagnosis of NDDs.

1 INTRODUCTION

Neurodevelopmental disorders (NDDs) impose long-term adverse effects on individuals, families, and healthcare systems (Ahn & Hwang, 2017; Wang et al., 2022; Aljead et al., 2025). Among all NDD subtypes, attention deficit hyperactivity disorder (ADHD) and autism spectrum disorder (ASD) are the two most prevalent conditions. Current clinical diagnosis relies mainly on subjective behavioral rating scales and structured interviews, which lack quantitative neurobiological markers and frequently lead to underdiagnosis and misdiagnosis. This unmet clinical demand calls for objective diagnostic approaches assisted by neuroimaging techniques (Lohani et al., 2025; Rezaee, 2026).

Structural MRI has been widely validated to detect macroscopic morphological lesions in neurodegenerative diseases such as Alzheimer's disease (AD) and Parkinson's disease (PD) (Huang et al., 2026; Zhao et al., 2026), but it is inadequate for identifying subtle developmental brain abnormalities in ADHD and ASD, which rarely present obvious anatomical deficits. Resting-state functional MRI (rs-fMRI) has therefore emerged as a reliable alternative (Dong et al., 2025; Liu et al., 2024), quantifying spontaneous neuronal activity and interregional functional connectivity from blood oxygen level-dependent (BOLD) signals for comprehensive evaluation of coordinated developmental patterns across brain regions. Combined with machine learning and deep learning algorithms, biosignals extracted from BOLD time-series data can significantly improve NDD diagnostic accuracy (Shahini et al., 2025; Ul Ain et al., 2025; Huda et al., 2024).

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Local biosignals such as the amplitude of low-frequency fluctuation (ALFF), regional homogeneity (ReHo), and degree centrality (DC) quantitatively assess spontaneous neuronal activities within individual brain regions and have shown reliable performance in identifying disease-related abnormalities at the population level (Wang et al., 2025b; An et al., 2013; Liu et al., 2023). Nevertheless, solely relying on local abnormalities is insufficient for accurate individualized diagnosis. Functional connectivity (FC), conventionally computed as the Pearson correlation of BOLD time series between brain regions, quantifies pairwise interactions and is widely used for disease diagnosis (Huynh et al., 2024; Wen et al., 2026; Jung et al., 2024). Positive FC values reflect synchronized neural activities (Fox & Raichle, 2007), whereas negative FC values represent anti-correlated dynamics corresponding to functional antagonism (Chai et al., 2012). Beyond static FC, advanced variants such as effective connectivity (EC) (Dong et al., 2025), high-order FC (HOFC) (Kang et al., 2026), and dynamic FC (dFC) (Shehzad et al., 2025) serve as complementary information for NDD diagnosis.

By framing NDD diagnosis as a classification task, graph neural networks (GNNs) have emerged as powerful tools for mining diagnostic cues. The core advantage of GNNs lies in their ability to model message propagation across interconnected brain regions. In such graph-based diagnostic pipelines, FC features provide critical topological priors for constructing individual brain networks. Specifically, FC values between a target brain region and other brain areas can be concatenated to form the feature vector of each graph node. Meanwhile, these FC measurements also define graph edges, which can be further structured into adjacency matrices or simplified binarized matrices to establish brain topological models. To date, multiple GNN architectures have been successfully applied to NDD diagnosis, including graph isomorphism networks (GIN) Zhou et al. (2026), graph attention networks (GAT) Hu et al. (2021), graph sample and aggregate (GraphSAGE) Bousmaha et al. (2026), and graph convolutional networks (GCN) Guo et al. (2026); Wang et al. (2026b). To our knowledge, GIN, GAT, and GraphSAGE predominantly optimize edge weights through node-level data-driven learning. Most diagnostic classification frameworks adopt stacked cascaded GNN layers, and such automatic edge-learning approaches inevitably cause drastic fluctuations in edge representations across successive layers. This limitation severely degrades the stability of network modeling and impairs the biological interpretability of brain network analysis for NDD research. In comparison, GCN models that adopt fixed FC-based graph edges yield better robustness and are therefore more widely adopted in relevant studies. For example, a recently proposed multi-level correlation-aware and modal-aware GCN (MCM-GCN) achieves reliable NDD diagnosis by extracting discriminative diagnostic evidence at both individual and population levels Zuo et al. (2026). Another study focused on edge-node interactive learning and developed a dual-branch GCN framework. This model jointly learns high-level representative features from graph nodes and edges, thereby improving the classification accuracy for ADHD and ASD Chen et al. (2026).

Despite their promising diagnostic performance, existing GCN-based methods have two critical limitations. (1) Conventional GCNs rely exclusively on positive graph edges, which suits general scenarios such as social networks and protein interaction networks (Zhang et al., 2023; Wang et al., 2025a). Most GCN models for NDD diagnosis construct adjacency matrices using absolute FC values, retaining only positive edge information (Wang et al., 2026c; Zuo et al., 2026; Chen et al., 2026), which distorts the physiological significance of FC signals and limits biomarker identification. Signed GCNs (SGCNs) explicitly model both positive and negative FCs (Guo et al., 2026), yet effective message aggregation across the two edge types remains underexplored (Wang et al., 2026a). (2) All GCN variants suffer from the over-smoothing problem when stacked into deep architectures. One-hop neighbor-based aggregation restricts sufficient information mining, and deeper networks increase the risk of node feature homogenization. Multi-hop GCNs mitigate over-smoothing by reducing required depth (Abu-El-Haija et al., 2019b; Lee & Kim, 2022), but determining appropriate hop counts and weights remains challenging, and current studies assign fixed values through heuristics without effective prior guidance. A comprehensive review of signed GCNs and multi-hop GCNs is presented in Appendix A.

To tackle the above-mentioned drawbacks of conventional GCNs, this study develops a long-term guided multi-hop signed graph convolutional network (LT-MSGCN) for NDD diagnosis. The designed multi-hop aggregation strategy captures and integrates feature information from extensive neighboring nodes, while the introduced signed graph convolution operations preserve complete positive and negative FC information, improving both robustness and biological interpretability. The main contributions of this work are summarized as follows.

- We build an auto-encoding branch to capture long-term dependencies across brain regions and guide the subsequent multi-hop aggregation process. Specifically, a transformer block is embedded in the encoder to generate multi-head long-term (LT) relationship matrices. These matrices serve as reliable references to constrain and refine multi-hop aggregated adjacency matrices.
- A dedicated multi-hop feature learning branch is further devised. This branch adopts approximate personalized propagation of neural predictions (APPNP) to construct multi-hop signed adjacency matrices for SGCNs, with APPNP aggregation weights adaptively calibrated via learned long-term relationship matrices. Here, the multi-head relationship matrices in the auto-encoding branch enable diverse feature propagation paths, offering multi-faceted cues for identifying abnormal brain patterns in NDD diagnosis. Moreover, we embed elaborate feature fusion blocks to adaptively highlight the weights of different propagation hops and brain regions, further improving the classification performance of our model.
- Experimental evaluations on the ABIDE and ADHD-200 datasets verify the superiority of the proposed LT-MSGCN over baseline methods. The model achieves classification accuracies of 70.2% and 71.8%, respectively. Moreover, the FC biomarkers identified by our model are highly consistent with well-documented neuroscientific findings of ADHD and ASD, demonstrating good biological validity and generalization capability.

2 METHOD

2.1 DATASETS AND PREPROCESSING

We evaluated our LT-MSGCN on two publicly available multi-site neuroimaging datasets, ABIDE-I and ADHD-200, both downloaded from the Preprocessed Connectomes Project (PCP). Their demographic information is summarized in Table 3 in Appendix B.

- **ADHD-200**: It consists of resting-state fMRI data from children and adolescents diagnosed with ADHD and typically developing controls (TDs), collected from eight sites. We selected four representative sites that contain both ADHD and TD subjects: New York University Medical Center (NYU), Kennedy Krieger Institute (KKI), NeuroIMAGE (NI), and Peking University (PKU), comprising 541 subjects in total (243 ADHD and 298 TD). The preprocessed data were obtained from the Neuro Bureau ADHD200 Preprocessed Initiative via the Athena pipeline, which is based on AFNI and FSL software. More details are available at `http://preprocessed-connectomes-project.org/adhd200/`.
- **ABIDE-I**: This dataset includes resting-state fMRI data from individuals with ASD and TDs, collected from 17 international imaging sites. Following the quality-control criteria, 871 subjects (403 ASD, 468 TD) were included after excluding scans with incomplete brain coverage, severe artifacts and excessive head motion. We employed the preprocessed data released by the Configurable Pipeline for the Analysis of Connectomes (CPAC) from the PCP, which standardizes preprocessing across all sites. The preprocessed data are publicly available at `http://preprocessed-connectomes-project.org/abide/`.

After preprocessing, we extracted the mean time series from pre-defined ROIs using the automated anatomical labeling 116 (AAL116) atlas for both datasets. FC matrices were computed via Pearson's correlation analysis between time series from 90 cerebral regions, yielding a symmetric FC matrix $\boldsymbol{A}_0$ of size 90×90 for each subject. Each individual brain is modeled as an undirected graph $\boldsymbol{G} = (\boldsymbol{X}, \boldsymbol{A})$, where $\boldsymbol{X} = \boldsymbol{A}_0$ denotes the node feature matrix, with the i-th row representing the correlation pattern between the i-th ROI and all other ROIs. The sparse graph adjacency matrix $\boldsymbol{A}$ is derived from $\boldsymbol{A}_0$ by retaining, for each node, only the top 20% FCs with the highest absolute values and masking the remaining entries to zero.

2.2 OVERVIEW OF LT-MSGCN

The proposed LT-MSGCN is designed to capture long-range dependencies across brain regions and exploit their interactive effects on multi-hop feature aggregation. To achieve this goal, the model adopts a dual-branch architecture, as shown in Figure 1.

- An auto-encoding (AE) branch outputs multi-head LT relationship matrices $\{\boldsymbol{R}_i\}$, which serve as supervision signals for the multi-hop aggregated adjacency matrices $\{\boldsymbol{A}_i\}$. Both the encoder and decoder are implemented with transformer modules. The AE network

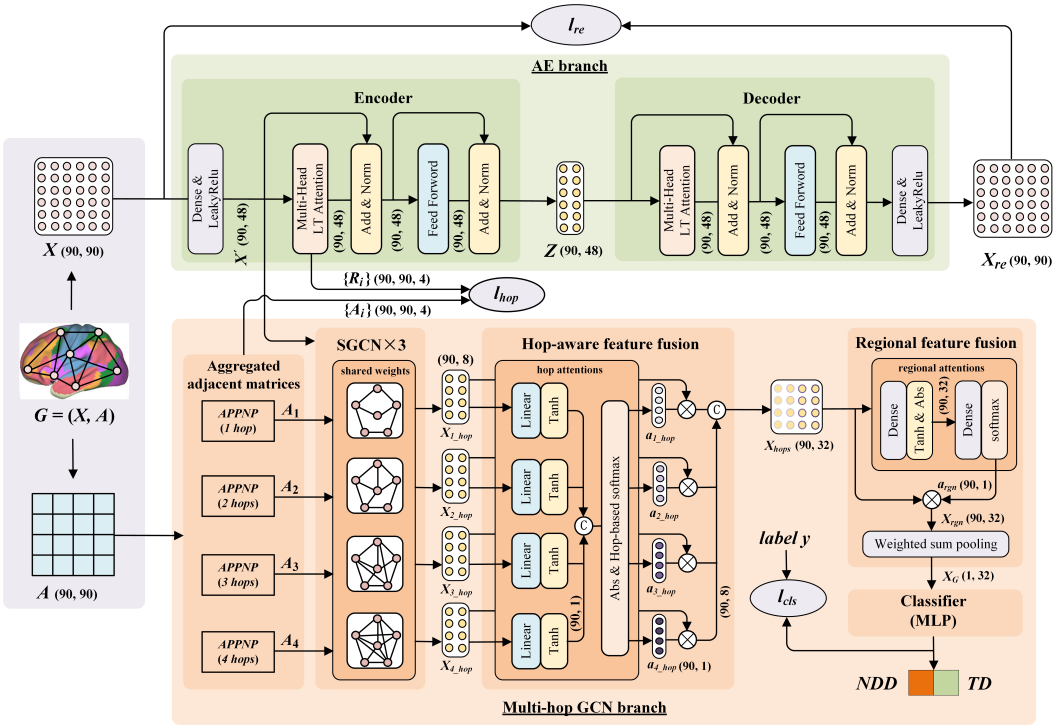

Figure 1: Architecture of the proposed LT-MSGCN.

regularizes intermediate node features $\boldsymbol{Z}$ to preserve input information, enabling $\{\boldsymbol{R}_i\}$ to encode rich long-range dependency patterns among brain nodes.

- A multi-hop GCN branch builds aggregated adjacency matrices $\{\boldsymbol{A}_i\}$ for different hop distances via the APPNP algorithm and establishes separate feature-learning pathways. Each pathway is equipped with three stacked SGCN layers (SGCN×3), and all pathways share SGCN weights layer-wise to guarantee uniform feature dimensions and cross-pathway feature alignment. Following hierarchical feature extraction, a hop-aware feature fusion module and a regional feature fusion module are applied sequentially. The former quantifies node-wise importance for each multi-hop pathway, and the latter assigns adaptive weights to individual brain regions and yields graph-level representations via graph pooling. Finally, a multilayer perceptron (MLP) classifier receives the graph-level representations to differentiate TD subjects from NDD patients.

2.3 Encoder and Decoder

Within the AE branch, the encoder produces multi-head LT relationship matrices $\{\boldsymbol{R}_i\}$, while the decoder supervises another encoder output $\boldsymbol{Z}$ to indirectly constrain $\{\boldsymbol{R}_i\}$. A dense layer equipped with the LeakyReLU activation first compresses the per-node features of the input node matrix $\boldsymbol{X}$, generating $\boldsymbol{X}'$ and reducing the feature dimension from 90 to 48. A modified transformer module is subsequently employed, whose multi-head LT attention block is elaborated below. In this block, the query (Q), key (K), and value (V) features are formulated as

$$\begin{aligned} \boldsymbol{Q} = \boldsymbol{K} &= \boldsymbol{X}'\boldsymbol{W}_{qk}, \\ \boldsymbol{V} &= \boldsymbol{X}'\boldsymbol{W}_{v}, \end{aligned} \tag{1}$$

where a learnable weight matrix $\boldsymbol{W}_{qk}$ is shared by $\boldsymbol{Q}$ and $\boldsymbol{K}$, while $\boldsymbol{W}_v$ denotes the exclusive learnable weight for $\boldsymbol{V}$. The dimensions of $\boldsymbol{W}_{qk}$ and $\boldsymbol{W}_v$ are 48×48 and 48×12, respectively. The shared $\boldsymbol{W}_{qk}$ facilitates the generation of symmetric LT relationship matrices, which matches the structural characteristics of multi-hop adjacency matrices. The single-head self-attention is then calculated as

$$\begin{aligned} head &= \boldsymbol{R}\boldsymbol{V}, \\ \boldsymbol{R} &= \tanh\left(\frac{\boldsymbol{Q}\boldsymbol{K}^T}{\sqrt{d_k}}\right), \end{aligned} \tag{2}$$

where the resulting LT relationship matrix $\boldsymbol{R}$ provides nonlinear dependencies among graph nodes, with all elements normalized within the range $(-1,\ 1)$. To fully exploit the multi-head mechanism,

we adopt four independent attention heads and concatenate their outputs to form the multi-head representation *Multihead*, which is written as

$$\textit{Multihead} = Concat(\{\textit{head}_i\}_{i=1}^{4})\boldsymbol{W}_o, \tag{3}$$

where $\boldsymbol{W}_o$ is a learnable weight of size 48×48. Each head has its own LT relationship matrix $\boldsymbol{R}_i$. We output $\{\boldsymbol{R}_i\}$ to supervise the multi-hop matrices $\{\boldsymbol{A}_i\}$. The feed-forward block consists of two dense layers paired with the GELU activation, transforming the features to 90×192 and then back to 90×48. Through the above encoder operations, the intermediate node feature $\boldsymbol{Z}$ of size 90×48 is finally obtained. The decoder adopts a symmetric structural design and ends with a dense layer equipped with the LeakyReLU activation, ultimately yielding the reconstructed node feature $\boldsymbol{X}_{re}$ that retains the same dimension as the original input $\boldsymbol{X}$.

2.4 Aggregated Adjacency Matrix Generation

In the multi-hop GCN branch, the APPNP algorithm is adopted to generate aggregated adjacency matrices. Given the sparse graph adjacency matrix $\boldsymbol{A}$, the corresponding i-hop aggregated adjacency matrix $\boldsymbol{A}_i$ is defined as

$$\boldsymbol{A}_i = \alpha\boldsymbol{I} + \alpha\sum_{k=1}^{i-1}(1-\alpha)^k\boldsymbol{A}^k + (1-\alpha)^i\boldsymbol{A}^i, \tag{4}$$

where α denotes a learnable weight, and $\boldsymbol{I}$ represents a diagonal matrix with the same dimension as $\boldsymbol{A}$, which models self-loop connections of graph nodes. For illustration, the first three orders of aggregated adjacency matrices are provided as

$$\begin{aligned}
\boldsymbol{A}_1 &= \alpha\boldsymbol{I} + (1-\alpha)\boldsymbol{A}, \\
\boldsymbol{A}_2 &= \alpha\boldsymbol{I} + \alpha(1-\alpha)\boldsymbol{A} + (1-\alpha)^2\boldsymbol{A}^2, \\
\boldsymbol{A}_3 &= \alpha\boldsymbol{I} + \alpha(1-\alpha)\boldsymbol{A} + \alpha(1-\alpha)^2\boldsymbol{A}^2 + (1-\alpha)^3\boldsymbol{A}^3.
\end{aligned} \tag{5}$$

Notably, since all elements of $\boldsymbol{A}$ lie in $(-1, 1)$, the resulting $\boldsymbol{A}_i$ also retain values in the same interval, making them well-suited to serve as self-looped adjacency matrices for GCN modeling.

2.5 SGCNs

As illustrated in Figure 1, each multi-hop pathway processes the node feature matrix $\boldsymbol{X}'$ with its corresponding aggregated adjacency matrix $\boldsymbol{A}_i$, where three successive SGCN layers are deployed. Given that these matrices contain negative entries, a normalization operation is first performed as

$$\hat{\boldsymbol{A}}_i = \boldsymbol{D}_i^{-1/2}\boldsymbol{A}_i\boldsymbol{D}_i^{-1/2}, \tag{6}$$

where $\hat{\boldsymbol{A}}_i$ denotes the normalized adjacency matrix and $\boldsymbol{D}_i$ is the degree matrix of $\boldsymbol{A}_i$. Unlike the traditional degree matrix setting, the k-th diagonal entry of $\boldsymbol{D}_i$ is calculated based on the absolute row sum of $\boldsymbol{A}_i$, formulated as $\boldsymbol{D}_i(k,k) = \sum_j |\boldsymbol{A}_i(k,j)|$.

Subsequently, the graph convolution operation of SGCN implemented in the *i*-th multi-hop pathway is defined as

$$\boldsymbol{H}_i^{(l)} = \text{LeakyReLU}(\hat{\boldsymbol{A}}_i\boldsymbol{H}_i^{(l-1)}\boldsymbol{W}^{(l)}), \;\; l \in \{1,2,3\} \tag{7}$$

where $\boldsymbol{H}_i^{(l)}$ denotes the output of the *l*-th SGCN layer, and the initial input satisfies $\boldsymbol{H}_i^{(0)} = \boldsymbol{X}'$. $\boldsymbol{W}^{(l)}$ is the learnable weight of the *l*-th layer, shared across all multi-hop pathways for consistent feature alignment. The three stacked layers progressively transform $\boldsymbol{X}'$ to dimensions 90×24, 90×16, and 90×8. The final SGCN output of the *i*-th pathway is denoted $\boldsymbol{X}_{i_hop} = \boldsymbol{H}_i^{(3)}$ $(i = 1, \ldots, 4)$.

2.6 Feature Fusion Modules

Two feature fusion modules are designed for multi-scale feature integration. In the hop-aware module, each SGCN output $\boldsymbol{X}_{i_hop}$ is fed into a linear layer followed by a Tanh activation to generate a regional importance score vector $\boldsymbol{s}_{i_hop} \in \mathbb{R}^{90\times 1}$. Score vectors from all multi-hop pathways are concatenated column-wise into the integrated score matrix $\boldsymbol{S}$, where $\boldsymbol{S}(k,i)$ quantifies the attention score of the k-th brain region in the i-th pathway. Subsequently, a hop-wise softmax function is applied to the absolute values of all entries in $\boldsymbol{S}$,

$$\boldsymbol{a}_{i_hop}(k) = \frac{exp(|\boldsymbol{S}(k,i))|)}{\sum_{i'} exp(|\boldsymbol{S}(k,i'))|)}, \tag{8}$$

where $\boldsymbol{a}_{i_hop}$ is the attention weight vector for the i-th pathway. A comprehensive hop-aware feature $\boldsymbol{X}_{hops}$ of size 90×32 is then derived from the SGCN outputs $\{\mathbf{X}_{i_hop}\}$,

$$\boldsymbol{X}_{hops} = Concat(\{(\boldsymbol{a}_{i_hop}\boldsymbol{I}^T) \otimes \boldsymbol{X}_{i_hop}\}), \tag{9}$$

where $\otimes$ denotes the element-wise multiplication operation, and $\boldsymbol{I}$ is an all-one vector of size 1×8, employed to match the feature column dimension of $\boldsymbol{X}_{i_hop}$.

The regional feature fusion module is then implemented to complement feature learning. Two stacked dense units transform features for individual brain regions. The first is coupled with Tanh and absolute value operations to refine feature magnitudes, and the second computes per-region importance weights, aggregated via a softmax function into the regional weight vector $\boldsymbol{a}_{rgn} \in \mathbb{R}^{90\times 1}$. The final region-weighted feature is formulated as

$$\boldsymbol{X}_{rgn} = (\boldsymbol{a}_{rgn}\boldsymbol{I}^T) \otimes \boldsymbol{X}_{hops}, \tag{10}$$

where $\boldsymbol{I}$ now denotes an all-one vector of size 1×32. Finally, the graph-level feature vector $\boldsymbol{x}_G$ is obtained via a weighted sum pooling layer,

$$\boldsymbol{x}_G(k) = \sum_j \boldsymbol{X}_{rgn}(j,k), \tag{11}$$

where $\boldsymbol{x}_G(k)$ represents the k-th entry of $\boldsymbol{x}_G$, and $\boldsymbol{X}_{rgn}(j,k)$ refers to the entry located at the j-th row and k-th column of $\boldsymbol{X}_{rgn}$. For classification, a simple two-layer MLP is utilized to predict individual diagnostic labels (i.e., NDD or TD) from $\boldsymbol{x}_G$.

2.7 Loss Function Setting

A joint loss consisting of three components is adopted to train LT-MSGCN. For the AE branch, a reconstruction loss l_{re} calculates the mean squared error (MSE) between the reconstructed node feature $\boldsymbol{X}_{re}$ and the original input feature $\boldsymbol{X}$ as

$$l_{re} = \frac{1}{M}\sum_{j=1}^{M} \|\boldsymbol{X}_{re,j} - \boldsymbol{X}_j\|_F^2, \tag{12}$$

where M denotes the total number of subjects, and $\|\cdot\|_F$ denotes the Frobenius norm. For the multi-hop GCN branch, a label prediction loss l_{cls} is formulated as the binary cross-entropy between the predicted label $\hat{y}_j$ and its true label y_j for each training subject,

$$l_{cls} = -\frac{1}{M}\sum_{j=1}^{M}[y_j \log(\hat{y}_j) + (1-y_j)\log(1-\hat{y}_j)]. \tag{13}$$

A multi-hop guiding loss l_{hop} is introduced to bridge the two branches, modeling the interaction between the LT relational representations $\{\boldsymbol{R}_i\}$ and the aggregated adjacency matrices $\{\boldsymbol{A}_i\}$,

$$l_{hop} = \frac{1}{M}\sum_{j=1}^{M}\sum_{i} \|\boldsymbol{R}_{i,j} - \boldsymbol{A}_{i,j}\|_F^2, \tag{14}$$

where $\boldsymbol{R}_{i,j}$ and $\boldsymbol{A}_{i,j}$ denote the i-th relationship matrix and its corresponding adjacency matrix for the j-th subject, respectively. As a result, the overall training loss is defined as

$$l_{total} = l_{cls} + \alpha_1 l_{re} + \alpha_2 l_{hop}, \tag{15}$$

where the hyperparameters $\alpha_1 = 0.1$ and $\alpha_2 = 1$ are empirically set to balance the magnitudes of different loss terms. The entire LT-MSGCN framework is trained in an end-to-end manner by minimizing l_{total} using the Adam optimizer, with 200 training epochs, a learning rate of 1×10^{-3}, and a dropout ratio of 0.2 applied to both the SGCNs and the MLP-based classifier.

3 Experiments

All experiments are evaluated under 10-fold cross-validation, where subjects are randomly partitioned into ten equal subsets. For each fold, one subset is held out for testing and the remaining nine are used for training. Four metrics are reported for quantitative evaluation including accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under the receiver operating characteristic curve (AUC).

Table 1: Comparison with recent methods on ABIDE-I and ADHD-200 datasets (%).

GNN Method	Atlas	Description	ABIDE-I				ADHD-200			
			ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
ALTER	AAL116	multi-hop	70.69	72.46	68.15	72.56	63.51	55.79	69.81	64.71
MSSTAN	AAL116	multi-graph	68.70	69.80	65.80	73.40	69.00	69.50	68.70	66.20
BrainOSM	AAL116	multi-graph, SGCN	70.72	68.65	72.50	70.59	65.35	57.72	71.57	66.52
RostGNN	CC200	multi-graph	70.40	**75.70**	64.90	**73.70**	69.60	70.00	69.40	65.90
BrainIB	CC200	causal graph	69.21	65.32	**72.94**	69.02	63.58	52.75	69.52	60.39
BrainOOD	Schaefer100	causal graph	**72.06**	72.07	72.03	72.80	63.69	63.40	62.96	63.17
LT-MSGCN	AAL116	multi-hop, SGCN	71.53	70.54	72.75	71.96	**70.20**	**75.55**	62.65	**70.34**

3.1 Comparison with State-of-the-Art Methods

Table 1 summarizes the performance of LT-MSGCN against several state-of-the-art GNN methods. Specifically, BrainOSM Guo et al. (2026), adaptive long-range aware transformer (ALTER) Yu et al. (2024), and multi-scale spatial-temporal attention networks (MSSTAN) Kong et al. (2025) are evaluated using the AAL116 atlas. Brain network with the famed information bottleneck (BrainIB) Zheng et al. (2024) and robust spatio-temporal graph neural networks (RostGNN) Zhang et al. (2026b) are tested on the CC200 atlas. Meanwhile, brain network with out-of-distribution generation (BrainOOD) Xu et al. (2025) is implemented with the Schaefer100 atlas.

Most existing competing methods achieve an overall accuracy of about 70% or lower, with consistently poorer performance on ADHD than on ASD, primarily owing to the higher phenotypic heterogeneity of ADHD, which contains three distinct clinical subtypes. To boost diagnostic performance, existing studies adopt diverse feature enhancement strategies. RostGNN and MSSTAN construct multiple graph structures by segmenting regional BOLD signals in the spatio-temporal domain. BrainOSM generates diverse brain graphs by thresholding FC matrices and learns features from positive and negative subgraphs, but applies only a simple MLP classifier on top-k subgraph features without effective fusion, limiting its ADHD accuracy to about 65%. As a representative multi-hop method, ALTER adopts layer-stacked implicit multi-hop propagation with an adjacency matrix derived from conventional random walk modeling, which restricts its capability to capture complex connectivity patterns (ADHD accuracy 63%). The causal methods BrainIB and BrainOOD extract task-relevant causal subgraphs from whole-brain graphs. The in-domain BrainOOD achieves the highest ABIDE-I accuracy (72%) among baselines, attributable to its Schaefer100 functional parcellation, which yields more biologically plausible partitions than the anatomical AAL116 atlas.

LT-MSGCN achieves accuracies exceeding 70% on both the ABIDE-I and ADHD-200 datasets, demonstrating robust generalization across different NDD tasks. These results validate the effectiveness of our SGCN-based processing, which integrates explicit multi-hop graph construction and adaptive k-hop adjacency mixing under long-term supervision, thereby improving the reliability and interpretability of brain-network-based diagnosis.

Table 2: Ablation test for various modules in LT-MSGCN (%).

Variant	without	operation	ABIDE-I				ADHD-200			
			ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
1	multi-hop	single hop	65.61	68.77	62.53	66.62	64.45	67.13	60.32	66.21
2	AE branch	-	70.65	74.63	68.23	71.20	70.01	74.16	65.03	70.13
3	SGCN	GCN	70.27	76.51	63.03	71.03	67.84	76.15	62.30	69.12
4	hop-aware feature fusion	only concatenate	70.84	74.40	66.77	71.89	69.68	74.48	63.77	69.99
5	regional feature fusion	mean pool	70.56	75.21	69.43	71.18	69.46	74.54	67.65	70.49
LT-MSGCN	-	-	71.53	70.54	72.75	71.96	70.20	75.55	62.65	70.34

3.2 Ablation Study

To validate the contribution of each component in LT-MSGCN, we design five ablated variants and compare them against the full model on both datasets. Results are reported in Table 2. Variant 1 replaces the multi-hop structure with a conventional single-hop SGCN, removing the capability of aggregating information from multiple hops. Variant 2 removes the AE branch from the multi-hop network. Variant 3 employs conventional GCNs instead of the SGCNs. Variants 4 and 5 remove the

two feature fusion modules. Variant 4 merges features from different hops through simple concatenation instead of the hop-aware attention mechanism, whereas Variant 5 uses mean pooling instead of the regional attention strategy for graph pooling.

Variant 1 causes the largest degradation on both datasets. The accuracy drops from 71.53% to 65.61% on ABIDE-I and from 70.20% to 64.45% on ADHD-200 (5.92 and 5.75 points, respectively), demonstrating the importance of multi-hop aggregation. Variant 2 also decreases performance after removing the AE branch, indicating that it provides complementary regularization beyond the multi-hop SGCN. Variant 3 confirms the effectiveness of SGCNs, and Variants 4 and 5 show that both feature fusion modules provide complementary benefits for learning discriminative brain representations.

3.3 Optimization on Hops

We further investigate the effect of hop number on the performance of LT-MSGCN. As shown in Figure 2, the optimal classification accuracy is achieved when the hop number is set to 4. Fewer hops fail to aggregate sufficient multi-scale neighborhood information, limiting the exploitation of discriminative graph-structured cues, whereas more hops introduce excessive redundant and noisy neighborhood information, degrading the node-level signal-to-noise ratio (SNR) and deteriorating diagnostic performance.

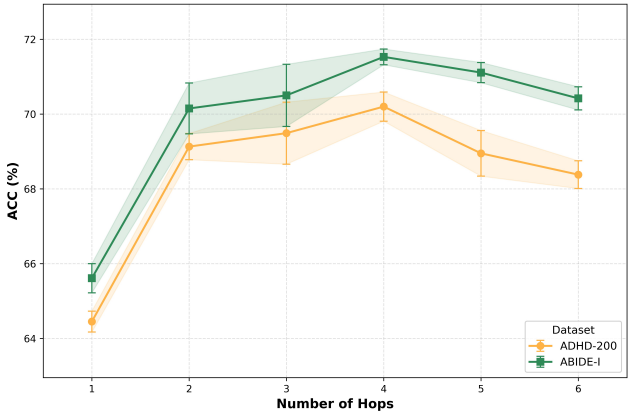

Figure 2: Classification performance under different hop numbers.

4 Biomarker Analysis on Brain Regions

In LT-MSGCN, the learned multi-hop aggregated adjacency matrices $\{\boldsymbol{A}_i\}$ serve as candidates for biomarker detection. Since they are inherently sparse with numerous zero-valued entries, making them incompatible with direct t-test comparison, we adopt their dense approximations, i.e., the long-term relationship matrices $\{\boldsymbol{R}_i\}$, for statistical analysis. Figure 3 visualizes the discriminative hop-aware brain relationships for ASD, and Figure 4 in Appendix C those for ADHD, with a strict p-value threshold of 0.001. The first subfigure of each figure illustrates the discriminative FCs, and the remaining subfigures present significant connections derived from $\{\boldsymbol{R}_i\}_{i=2}^{4}$, with the top five brain regions highlighted in red. All brain regions are categorized into seven functional subnetworks based on the Yeo7 atlas. As the hop number increases, the connections become sparser and concentrate within a restricted set of brain regions. These multi-hop refined connections correspond to stable, disease-specific topological patterns rather than noise fluctuations, verifying the reliability of the identified biomarkers.

For ASD analysis, the most discriminative connections at the fourth hop originate from the superior temporal gyrus (STG), middle temporal gyrus (MTG), middle temporal pole (TPOmid), and hippocampus (HIP), which belong to the sensorimotor network (SMN), limbic network (LN), and default mode network (DMN). The STG is responsible for auditory processing, speech perception, and social vocal cue integration. Atypical activation and disrupted connectivity in the STG are consistently observed in autism, and recent evidence further identifies increased gray matter volume (GMV) in the bilateral STG of ASD patients (Su et al., 2025). As a core component of the 'social brain', the MTG dominates semantic comprehension and social information processing, and its functional alterations in ASD have been elaborated in region-specific analyses (Xu et al., 2020). The HIP plays a vital role in episodic memory encoding and social associative learning. MRI radiomic

studies have identified HIP morphological and textural abnormalities as neuroanatomical signatures of ASD (Chaddad et al., 2017), and recent research reveals atypical sensory traits and abnormal white matter microstructures associated with hippocampal circuits (Kamiya et al., 2026). Collectively, these findings strongly validate the biological plausibility of our identified high-hop patterns for ASD diagnosis.

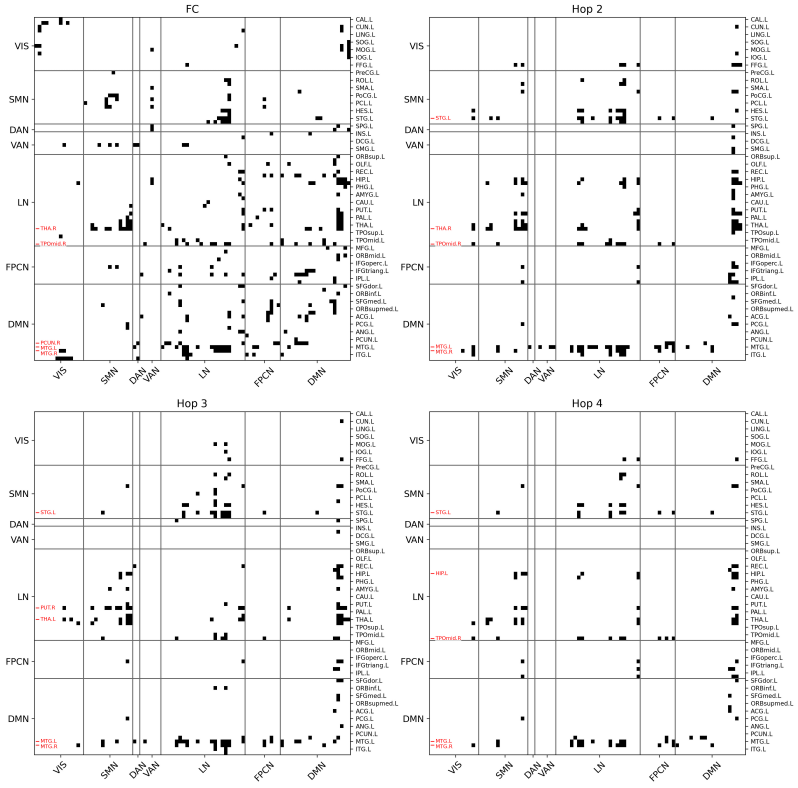

Figure 3: Discriminative hop-aware relationships among brain regions for ASD ($p < 0.001$). The 90 cerebral regions are grouped into seven functional subnetworks according to the Yeo7 brain network atlas. The top five brain regions with the most prominent discriminative relational values are highlighted in red.

5 Conclusion

In this study, we propose LT-MSGCN for the diagnosis of NDDs using resting-state fMRI data. To mitigate the inherent limitations of existing signed and multi-hop graph convolution models, namely inadequate prior guidance for multi-hop propagation and the over-smoothing problem, the framework incorporates an AE branch with multi-head transformer blocks to capture long-term dependencies across brain regions. The learned long-term relationship matrices act as soft priors to calibrate the APPNP-generated multi-hop adjacency matrices. Hop-aware and region-adaptive feature fusion modules dynamically weight different propagation hops and brain regions, strengthening discriminative graph representation learning. Experiments on ABIDE-I and ADHD-200 show that LT-MSGCN achieves competitive accuracy above 70% against state-of-the-art baselines, and the identified discriminative brain regions align well with established neuroscientific evidence, verifying its biological interpretability beyond classification performance.

AI use statement

During the preparation of this work, the authors used ChatGPT to assist with language polishing and improving readability. After using this tool, the authors reviewed and edited the manuscript as needed and take full responsibility for the content of the published article.

Reproducibility statement

Both datasets are publicly available from the Preprocessed Connectomes Project (PCP). All preprocessing steps, the architecture with all layer dimensions, and the hop setting (4) are described in Section 2 and Figure 1. Training hyperparameters and the evaluation protocol are given in Sections 2.7 and 3. Source code and preprocessed data will be released upon acceptance.

REFERENCES

Sami Abu-El-Haija, Bryan Perozzi, Amol Kapoor, Nazanin Alipourfard, Kristina Lerman, Hrayr Harutyunyan, Greg Ver Steeg, and Aram Galstyan. MixHop: Higher-order graph convolutional architectures via sparsified neighborhood mixing. In *International Conference on Machine Learning (ICML)*, volume 97, 2019a.

Sami Abu-El-Haija, Bryan Perozzi, Amol Kapoor, Nazanin Alipourfard, Kristina Lerman, Hrayr Harutyunyan, Greg Ver Steeg, and Aram Galstyan. MixHop: Higher-order graph convolutional architectures via sparsified neighborhood mixing. In *International Conference on Machine Learning (ICML)*, volume 97, 2019b.

Sinae Ahn and Sujin Hwang. Cognitive rehabilitation with neurodevelopmental disorder: A systematic review. *NeuroRehabilitation*, 41(4):707–719, 2017.

Mashal Aljead, Aya Qashta, Zahraa Jalal, and Alan M. Jones. Review of autism spectrum disorder (ASD): Epidemiology, aetiology, pathology, and pharmacological treatment. *Pharmaceuticals*, 18(11), 2025.

Li An, Qing-Jiu Cao, Man-Qiu Sui, Li Sun, Qi-Hong Zou, Yu-Feng Zang, and Yu-Feng Wang. Local synchronization and amplitude of the fluctuation of spontaneous brain activity in attention-deficit/hyperactivity disorder: A resting-state fMRI study. *Neuroscience Bulletin*, 29(5):603–613, 2013.

Rabab Bousmaha, Khouloud Meribai, Nardjes Bouchemal, Naila Bouchemal, and Galina Ivanova. Robust multi-site ADHD classification via GraphSAGE-based functional connectivity modeling from rs-fMRI. *Bioengineering-Basel*, 13(5), 2026.

S Carmona, O Vilarroya, A Bielsa, V Trèmols, JC Soliva, M Rovira, J Tomàs, C Raheb, JD Gispert, S Batlle, and A Bulbena. Global and regional gray matter reductions in ADHD: A voxel-based morphometric study. *Neuroscience Letters*, 389(2):88–93, 2005.

Ahmad Chaddad, Christian Desrosiers, Lama Hassan, and Camel Tanougast. Hippocampus and amygdala radiomic biomarkers for the study of autism spectrum disorder. *BMC Neuroscience*, 18, 2017.

Xiaoqian J. Chai, Alfonso Nieto Castanon, Dost Oenguer, and Susan Whitfield-Gabrieli. Anticorrelations in resting state networks without global signal regression. *NeuroImage*, 59(2):1420–1428, 2012.

Ying Chen, Yilan Li, Zhixin Li, Xiaojin Meng, and Yibin Tang. Gated edge-node interaction graph convolutional network for adhd classification. *Journal of Neural Engineering*, 23(3), 2026.

Jun-En Ding, Shihao Yang, Anna Zilverstand, Kaustubh R. Kulkarni, Xiaosi Gu, and Feng Liu. Spatial craving patterns in marijuana users: Insights from fMRI brain connectivity analysis with high-order graph attention neural networks. *IEEE Journal of Biomedical and Health Informatics*, 29(1):358–370, 2025.

Wenhao Dong, Yueyang Li, Weiming Zeng, Lei Chen, Hongjie Yan, Wai Ting Siok, and Nizhuan Wang. STARFormer: A novel spatio-temporal aggregation reorganization transformer of FMRI for brain disorder diagnosis. *Neural Networks*, 192, 2025.

Xiaolong Fan, Maoguo Gong, and Yue Wu. Markov clustering regularized multi-hop graph neural network. *Pattern Recognition*, 139, JUL 2023.

Michael D. Fox and Marcus E. Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9):700–711, 2007.

Johannes Gasteiger, Aleksandar Bojchevski, and Stephan Günnemann. Predict then propagate: Graph neural networks meet personalized pagerank. In *International Conference on Learning Representations*, 2019.

Guiliang Guo, Guangqi Wen, Lingwen Liu, Ruoxian Song, Peng Cao, Jinzhu Yang, and Osmar R. Zaiane. BrainOSM: Outlier screening for multi-view functional brain network analysis. *Computer Methods and Programs in Biomedicine*, 273, 2026.

Chaobo He, Hao Cheng, Jiaqi Yang, Yong Tang, and Quanlong Guan. Signed graph embedding via multi-order neighborhood feature fusion and contrastive learning. *Neural Networks*, 182, FEB 2025.

Jinlong Hu, Lijie Cao, Tenghui Li, Shoubin Dong, and Ping Li. GAT-LI: a graph attention network based learning and interpreting method for functional brain network classification. *BMC Bioinformatics*, 22(1), 2021.

Fei Huang, Nanguang Chen, and Anqi Qiu. 3D-CNN enhanced multiscale progressive vision transformer for AD diagnosis. *IEEE Journal of Biomedical and Health Informatics*, 30(1):402–412, 2026.

Shiza Huda, Danish Mahmood Khan, Komal Masroor, Ayesha Warda, Ayesha Rashid, and Mariam Shabbir. Advancements in automated diagnosis of autism spectrum disorder through deep learning and resting-state functional mri biomarkers: A systematic review. *Cognitive Neurodynamics*, 18(6, SI):3585–3601, 2024.

Nguyen Huynh, Da Yan, Yueen Ma, Shengbin Wu, Cheng Long, Mirza Tanzim Sami, Abdullateef Almudaifer, Zhe Jiang, Haiquan Chen, Michael N. Dretsch, Thomas S. Denney, Rangaprakash Deshpande, and Gopikrishna Deshpande. The use of generative adversarial network and graph convolution network for neuroimaging-based diagnostic classification. *Brain Sciences*, 14(5), 2024.

Wonsik Jung, Eunjin Jeon, Eunsong Kang, and Heung-Il Suk. EAG-RS: A novel explainability-guided ROI-selection framework for ASD diagnosis via inter-regional relation learning. *IEEE Transactions on Medical Imaging*, 43(4):1400–1411, 2024.

Taku Kamiya, Kaie Habata, Yongjeon Cheong, Daichi Shiotsu, Takuya Makino, Kotaro Kowada, Riku Sanada, Ichiro M. Omori, Sohei Ide, Hidehiko Okazawa, Minyoung Jung, and Hirotaka Kosaka. Atypical sensory traits and changes in white matter microstructures connected to the amygdala and hippocampus of the autistic brain. *Journal of Neural Transmission*, 2026.

Jiannan Kang, Yuqi Li, Juanmei Wu, Wenqin Mao, Xiaoli Li, Xin Li, and Rui Su. Multiscale static and dynamic brain functional network analysis reveals aberrant connectivity patterns in preschool children with autism spectrum disorder. *Behavioural Brain Research*, 498, 2026.

Youyong Kong, Xiaotong Zhang, Wenhan Wang, Yue Zhou, Yueying Li, and Yonggui Yuan. Multi-scale spatial-temporal attention networks for functional connectome classification. *IEEE Transactions on Medical Imaging*, 44(1):475–488, 2025.

Zhihui Lan, Yunkai Sun, Lei Zhao, Yang Xiao, Changxiao Kuai, and Shao-Wei Xue. Aberrant effective connectivity of the ventral putamen in boys with attention-deficit/hyperactivity disorder. *Psychiatry Investigation*, 18(8):763–769, 2021.

Minjae Lee and Seoung Bum Kim. HAPGNN: Hop-wise attentive pagerank-based graph neural network. *Information Sciences*, 613:435–452, 2022.

Jingming Li, Zhengyuan Lyu, Hu Yu, Si Fu, Ke Li, Li Yao, and Xiaojuan Guo. Signed curvature graph representation learning of brain networks for brain age estimation. *IEEE Journal of Biomedical and Health Informatics*, 28(12):7491–7502, DEC 2024.

Guanlu Liu, Weizhao Lu, Jianfeng Qiu, and Liting Shi. Identifying individuals with attention-deficit/hyperactivity disorder based on multisite resting-state functional magnetic resonance imaging: A radiomics analysis. *Human Brain Mapping*, 44(8):3433–3445, 2023.

Haoxiang Liu, Junquan Zhang, Zhi Fang, Xiaonan Zhang, Xiyue Sun, Li Feng, Dingyang Liu, Zhiquan Yang, JingLiang Cheng, Huafu Chen, Wei Huang, Xi Jiang, and Rong Li. Multi-hop spatio-temporal graph convolutional networks for brain disorder diagnosis and prognosis. *Pattern Recognition*, 179(D), NOV 2026.

Rui Liu, Zhi-An Huang, Yao Hu, Lei Huang, Ka-Chun Wong, and Kay Chen Tan. Spatio-temporal hybrid attentive graph network for diagnosis of mental disorders on fMRI time-series data. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 8(6):4046–4058, 2024.

Dhruv Chandra Lohani, Vaishali Chawla, and Bharti Rana. A systematic literature review of machine learning techniques for the detection of attention-deficit/hyperactivity disorder using MRI and/or EEG data. *Neuroscience*, 570:110–131, 2025.

Mansooreh Pakravan. Explainable graph attention network on resting state fMRI dynamic functional connectivity reveals cerebellar hub breakdown in chronic knee osteoarthritis pain. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, 50(1):289–300, 2026.

Pei Quan, Lei Zheng, Wen Zhang, Yang Xiao, Lingfeng Niu, and Yong Shi. ExGAT: Context extended graph attention neural network. *Neural Networks*, 181, JAN 2025.

Khosro Rezaee. Machine learning in automated diagnosis of autism spectrum disorder: A comprehensive review. *Computer Science Review*, 61, 2026.

Ariadna Albajara Saenz, Thomas Villemonteix, Peter Van Schuerbeek, Simon Baijot, Mathilde Septier, Pierre Defresne, Veronique Delvenne, Gianfranco Passeri, Hubert Raeymaekers, Laurent Victoor, Eric Willaye, Philippe Peigneux, Nicolas Deconinck, and Isabelle Massat. Motor abnormalities in attention-deficit/hyperactivity disorder and autism spectrum disorder are associated with regional grey matter volumes. *Frontiers in Neurology*, 12, 2021.

KP Schulz, CY Tang, J Fan, DJ Marks, AM Cheung, JH Newcorn, and JM Halperin. Differential prefrontal cortex activation during inhibitory control in adolescents with and without childhood attention-deficit/hyperactivity disorder. *Neuropsychology*, 19(3):390–402, 2005.

Alen Shahini, Aditya Prabhakara Kamath, Ekta Sharma, Massimo Salvi, Ru-San Tan, Siuly Siuly, Silvia Seoni, Rahul Ganguly, Aruna Devi, Ravinesh Deo, Prabal Datta Barua, and U. Rajendra Acharya. A systematic review for artificial intelligence-driven assistive technologies to support children with neurodevelopmental disorders. *Information Fusion*, 124, 2025.

Ahsan Shehzad, Dongyu Zhang, Shuo Yu, Shagufta Abid, and Feng Xia. Dynamic graph transformer for brain disorder diagnosis. *IEEE Journal of Biomedical and Health Informatics*, 29(6):4388–4400, 2025.

Yanan Su, Guangfei Li, Shanmei Wang, Dongmei Hao, Clara S. Li, Yiyao Ye-Lin, Xiaolin Wang, Ruolin Zhang, Lin Yang, and Chiang-Shan R. Li. Gray matter volumes of the superior temporal gyrus link preterm birth and developmentally disordered eye gazing patterns in toddlers at eighteen months. *Progress in Neuro-psychopharmacology & Biological Psychiatry*, 143, 2025.

Haoteng Tang, Lei Guo, Xiyao Fu, Yalin Wang, Scott Mackin, Olusola Ajilore, Alex D. Leow, Paul M. Thompson, Heng Huang, and Liang Zhan. Signed graph representation learning for functional-to-structural brain network mapping. *Medical Image Analysis*, 83, 2023.

Qurat Ul Ain, Soyiba Jawed, Ahmad Rauf Subhani, Wasi Haider Butt, and Muhammad Usman Akram. Examining AI-powered ADHD diagnosis: Current trends, key challenges, and future directions in the field. *IEEE Access*, 13:93148–93177, 2025.

Chenfei Wang, Yunping Wang, Qinghan Xue, Zhiheng Zhou, Guiying Yan, and Xingqin Qi. Classification of Alzheimer's disease by modeling brain networks as signed networks under deep learning frameworks. *IEEE Transactions on Computational Biology and Bioinformatics*, 23(2): 717–726, 2026a.

Chuang Wang, Dongyan Li, Sixiang Sun, Junli Liu, and Yuanyuan Li. KSNet: Advancing autism prediction via KAN-based graph convolution and multi-source data fusion. *Neural Networks*, 201, 2026b.

Likai Wang, Tao Zhu, and Yipu Zhang. An interpretable deep graph neural network based on attentional multi-scale feature fusion for fMRI analysis. In *2024 IEEE International Conference on Image Processing, ICIP*, pp. 3003–3009, 2024.

Rong Wang, Yongchen Fan, Ying Wu, Yu-Feng Zang, and Changsong Zhou. Lifespan associations of resting-state brain functional networks with ADHD symptoms. *iScience*, 25(7), 2022.

Wei Wang, Linchong Ma, Dong Liu, Hongjun Zhang, Guangsheng Wu, and Yun Zhou. AdptDilatedGCN: Protein-ligand binding affinity prediction based on multi-scale interaction fusion mechanism and dilated GCN. *International Journal of Biological Macromolecules*, 311(1), 2025a.

Xiaotong Wang, Yibin Tang, Yuan Gao, Xiaojing Meng, Ying Chen, and Aimin Jiang. ADHD classification with GCN via joint feature learning among nodes and edges. *IEEE Transactions on Medical Imaging*, 45(5):2481–2492, 2026c.

Zedong Wang, Liqin Zheng, Lijuan Yang, Shunjie Yin, Shiqi Yu, Kai Chen, Tao Zhang, Hesong Wang, and Yong Zhang. Structural and functional whole brain changes in autism spectrum disorder at different age stages. *European Child & Adolescent Psychiatry*, 34(5):1589–1602, 2025b.

Xin Wen, Shijie Guo, Li Dong, Xiaobo Liu, Wenbo Ning, Jie Shi, Songhua Liu, Cheng Luo, Rui Cao, and Dezhong Yao. A consistency-driven pseudo-labeling framework for robust functional connectivity modeling in neuropsychiatric disorder diagnosis. *Engineering Applications of Artificial Intelligence*, 166(A), 2026.

Xuyun Wen, Qumei Cao, Yunxi Zhao, Xia Wu, and Daoqiang Zhang. D-mhgcn: An end-to-end individual behavioral prediction model using dual multi-hop graph convolutional network. *IEEE Journal of Biomedical and Health Informatics*, 28(10):6130–6140, 2024.

Xubin Wu, Yan Niu, Xia Li, Jie Xiang, and Yidi Li. A prior causality-guided multi-view diffusion network for brain disorder classification. *CAAI Transactions on Intelligence Technology*, 10(6): 1731–1744, DEC 2025.

Jiaxing Xu, Yongqiang Chen, Xia Dong, Mengcheng Lan, Tiancheng Huang, Qingtian Bian, James Cheng, and Yiping Ke. Brainood: Out-of-distribution generalizable brain network analysis. In *International Conference on Learning Representations*, 2025.

Jinping Xu, Chao Wang, Ziyun Xu, Tian Li, Fangfang Chen, Kai Chen, Jingjing Gao, Jiaojian Wang, and Qingmao Hu. Specific functional connectivity patterns of middle temporal gyrus subregions in children and adults with autism spectrum disorder. *Autism Research*, 13(3):410–422, 2020.

Shuo Yu, Shan Jin, Ming Li, Tabinda Sarwar, and Feng Xia. Long-range brain graph transformer. In A Globerson, L Mackey, D Belgrave, A Fan, U Paquet, J Tomczak, and C Zhang (eds.), *Advances in Neural Information Processing Systems (NEURIPS 2024)*, 2024.

Cai Zhang, Weimin Li, Dingmei Wei, Yanxia Liu, and Zheng Li. Network dynamic GCN influence maximization algorithm with leader fake labeling mechanism. *IEEE Transactions on Computational Social Systems*, 10(6):3361–3369, 2023.

Han Zhang, Huan Wang, and Mingjing Han. HopGAT: A multi-hop graph attention network with heterophily and degree awareness. *Pattern Recognition*, 172(A), 2026a.

Yupei Zhang, Yuxin Li, Shuhui Liu, and Xuequn Shang. Robust spatio-temporal graph neural networks with sparse structure learning. *Pattern Recognition*, 172, 2026b.

Jingyu Zhao, Rizhi Ding, and Manhua Liu. KGT: Knowledge-guided graph transformer for neurodegenerative disease diagnosis and brain age prediction with MRI. *Medical Image Analysis*, 113, 2026.

Kaizhong Zheng, Shujian Yu, Baojuan Li, Robert Jenssen, and Badong Chen. Brainib: Interpretable brain network-based psychiatric diagnosis with graph information bottleneck. *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

Shufeng Zhou, Lina Zhou, Yueying Zhou, Hongyan Han, Hongxia Zheng, and Lishan Qiao. GIN-transformer based pairwise graph contrastive learning framework. *Neural Networks*, 198, 2026.

Shijia Zuo, Yu Li, Yinbao Qi, and Aiping Liu. Multilevel correlation-aware and modal-aware graph convolutional network for diagnosing neurodevelopmental disorders. *IEEE Transactions on Biomedical Engineering*, 73(5):1863–1876, 2026.

A Related Work

A.1 Multi-Hop GNNs for Brain Connectivity Analysis

Multi-hop GNNs expand the receptive field of graph nodes to capture indirect interregional brain interactions, enabling the simultaneous extraction of local unimodal and long-range transmodal connectivity signatures. Such multi-scale representations greatly enhance individual brain fingerprinting and improve classification performance for NDDs. According to their feature encoding pipelines, existing brain-oriented multi-hop GNNs can be divided into two mainstream categories: layer-stacked implicit multi-hop propagation, and explicit k-hop adjacency mixing with attention modulation.

Layer-stacked implicit multi-hop propagation. This method derives multi-order topological features through cascaded graph propagation layers (Abu-El-Haija et al., 2019a; Ding et al., 2025; Zhang et al., 2026a; Quan et al., 2025). Each layer only aggregates features from one-hop neighbors, while multi-hop information is implicitly accumulated with the increase of network depth. Multi-scale outputs from different layers are concatenated or fused via jump connections to integrate latent multi-hop signals, without explicitly computing high-order adjacency matrices. Several brain diagnostic models have adopted this paradigm. For behavioral prediction in healthy populations, the dual multi-hop GCN (D-MHGCN) (Wen et al., 2024) employs a dual-branch structure to extract hemisphere-specific cortical features through stacked propagation layers. The bilateral cortical embeddings are further fed into a whole-brain GCN module to regress individual cognitive metrics. For NDD diagnosis, the interpretable deep graph neural network (IDGNN) (Wang et al., 2024) adopts stacked multi-hop convolution and layer-wise global pooling to preserve multi-scale region-of-interest (ROI) features. Combined with multi-head attention for cross-layer fusion, this model obtains competitive accuracy in ASD classification. In addition, the causality-guided multi-view diffusion network (CG-MVD) (Wu et al., 2025) stacks diffusion attention blocks constrained by pre-estimated causal edges to capture implicit multi-hop brain interactions, achieving promising diagnostic performance for both ADHD and ASD. Despite these advances, such layer-based multi-scale embedding still suffers from inherent over-smoothing, especially for features learned in deep layers.

Explicit k-hop adjacency mixing with attention modulation. Unlike implicit layer-based accumulation, this category explicitly constructs power-normalized aggregated adjacency matrices within a single convolution layer. It integrates feature information from self-loops, direct one-hop neighbors, and distant multi-hop nodes. Trainable attention modules are commonly embedded to assign task-adaptive weights to different hop-level propagation paths within a unified adjacency tensor. Here, the approximate personalized propagation of neural predictions (APPNP) (Gasteiger et al., 2019) serves as a typical backbone for implementing weighted multi-hop aggregation in such frameworks. Representative studies have validated this explicit paradigm. For example, the Markov clustering regularized multi-hop GNN (MCMGN) (Fan et al., 2023) introduces Markov regularization to calibrate learnable APPNP weights and achieves robust performance on general graph classification tasks. Targeting NDD diagnosis, the multi-hop spatio-temporal GCN (MSTGCN) (Liu et al., 2026) proposes a task-specific explicit multi-hop attention architecture. It integrates an attention-driven diffusion kernel into the APPNP propagation process to learn discriminative multi-hop embeddings, which serve as reliable neuroimaging biomarkers for ADHD and ASD identification. Nevertheless, this explicit paradigm suffers from obvious hyperparameter sensitivity. There is no unified criterion for determining optimal hop numbers and fusion coefficients, which requires exhaustive tuning and substantially increases computational costs in practical neuroimaging experiments.

A.2 Signed GNNs for Mental Disorder Diagnosis

Conventional unsigned GNNs typically ignore negative topological correlations in brain networks and thus fail to capture disease-specific inhibitory structural patterns. In contrast, SGNNs disentangle positive and negative message propagation streams to model complementary brain interaction relationships, yielding more biologically meaningful graph representations. To the best of our knowledge, SGNNs have evolved through single-hop and multi-hop learning paradigms and have achieved promising performance in fMRI analysis.

Single-hop SGNNs: They aggregate feature information from first-order positive and negative neighbors in a decoupled manner, without expanding receptive fields via multi-layer stacking. A dual-branch architecture is commonly adopted to perform feature learning on predefined positive graphs (PGs) and negative graphs (NGs), followed by a cascaded feature fusion module that integrates branch-specific features to form a task-oriented comprehensive representation. For example, dual-edge GATs equipped with self-attention pooling are implemented on PGs and NGs to explore the latent associations between FCs and chronic knee osteoarthritis pain (Pakravan, 2026). For AD diagnosis, the signed curvature GNN (SCGNN) characterizes positive and negative edges via curvature representation and conducts branch-wise feature extraction using signed curvature convolutions (Li et al., 2024). Various message passing strategies have also been benchmarked for PG and NG modeling, among which signed GATs exhibit superior performance for AD classification (Wang et al., 2026a). Most recently, a brain-oriented ordered signed graph mining framework (BrainOSM) has been proposed for dual-task analysis of ASD and AD, which leverages shared GCNs coupled with multi-graph pooling to hierarchically process PGs and NGs (Guo et al., 2026).

Multi-hop SGNNs: To remedy single-hop constraints on the receptive field, recent studies have integrated multi-hop GNNs with signed message passing mechanisms to model hierarchical propagation pathways for both PGs and NGs. The multi-order signed graph convolutional network (MOSGCN) has been proposed to aggregate and fuse multi-hop feature outputs from dual positive and negative branches, which was validated for analyzing trust and distrust relationships among users on the Bitcoin transaction platform (He et al., 2025). This method highlights balanced and unbalanced graph embeddings to fully capture intrinsic multi-hop structural characteristics. Inspired by this, such a technical paradigm has also been applied to the deep signed brain graph mining (DSBGM) network. In detail, additional embeddings, namely positive and negative structural embeddings, are elaborately designed to construct multi-view representations of brain topological networks. Benefiting from this design, the DSBGM enables effective mapping between functional and structural brain networks, thereby achieving superior prediction performance for brain age estimation and AD diagnosis (Tang et al., 2023). Despite these advances, existing multi-hop SGNNs still suffer from a common limitation, that is, multi-hop message propagation and feature fusion lack refined prior knowledge guidance, leading to suboptimal multi-hop feature learning.

B Dataset Demographics

Table 3: Demographic information of the datasets used in this study.

Dataset	Subgroup	Number	Gender (F/M)	Age (mean $\pm$ std.)
ADHD-200	TD	298	163 / 135	12.00 ± 2.84
	ADHD	243	198 / 45	12.08 ± 2.93
ABIDE-I	TD	468	90 / 378	16.84 ± 7.23
	ASD	403	54 / 349	17.07 ± 7.95

C ADDITIONAL EXPERIMENTAL RESULTS

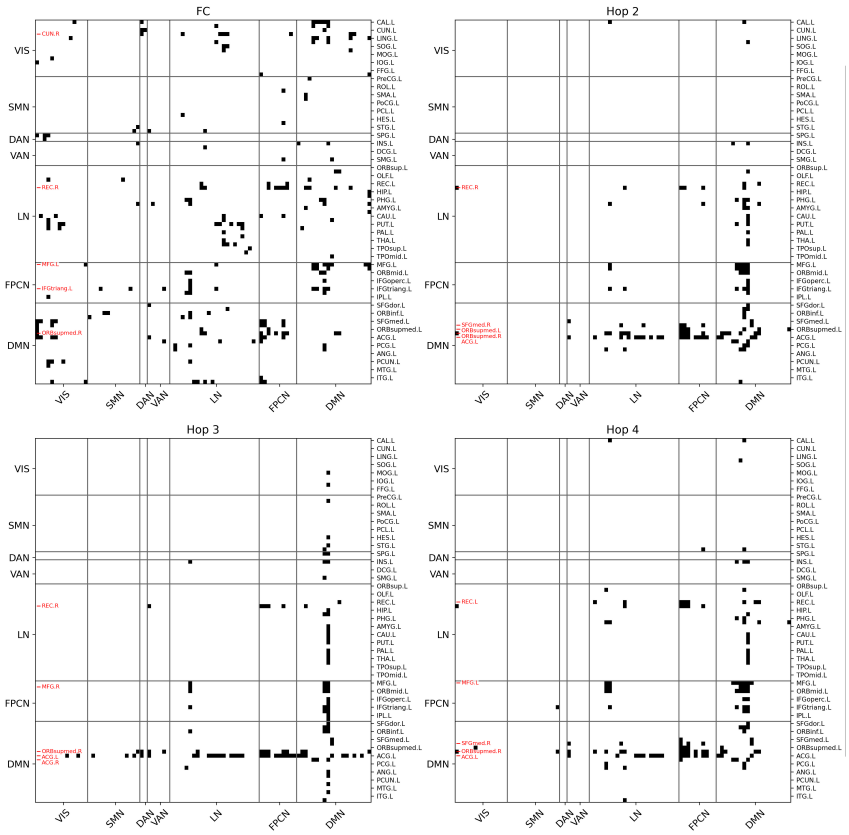

Figure 4: Discriminative hop-aware relationships among brain regions for ADHD ($p < 0.001$). The 90 cerebral regions are grouped into seven functional subnetworks according to the Yeo7 brain network atlas. The top five brain regions with the most prominent discriminative relational values are highlighted in red.

For ADHD analysis, the discriminative connections are found in the medial superior frontal gyrus (SFGmed), medial orbital part of the superior frontal gyrus (ORBsupmed), and anterior cingulate gyrus (ACG), all of which belong to the DMN. The SFGmed supports self-monitoring, goal maintenance, and attentional regulation, and motor deficits in ADHD have been associated with GMV abnormalities within it (Saenz et al., 2021). The ORBsupmed participates in reward valuation, affective regulation, and value-based decision-making, functions closely tied to core ADHD symptoms, and aberrant interregional connections in both SFGmed and ORBsupmed have been detected in boys diagnosed with ADHD (Lan et al., 2021). The ACG acts as a critical hub connecting the frontal lobe and limbic system, mediating conflict monitoring and cognitive control. FMRI studies have reported significantly elevated activation of the left ACG among ADHD participants during conflict tasks (Schulz et al., 2005), and voxel-based morphometry findings demonstrate reduced regional GMV in the left ACG for ADHD cohorts (Carmona et al., 2005). These prior neuroimaging observations corroborate the biological interpretability of our identified brain regional biomarkers.