Biomedical Signal Processing and Control 123 (2026) 110628

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

MF-HGNN: A multi-view fusion heterogeneous graph neural network for psychiatric disorder diagnosis

Yuhang Liu [a], Rongbing Huang [a,*], Chang Liu [a], Jianwei Zhang [a], Fasheng Yi [a], Junchuan Chen [a], Hechun Li [b,c], Sisi Jiang [b,c], Cheng Luo [b,c]

[a] School of Computer Science, Chengdu University, Chengdu 610106, PR China
[b] Clinical Hospital of Chengdu Brain Science Institute, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 611731, PR China
[c] China-Cuba Belt and Road Joint Laboratory on Neurotechnology and Brain-Apparatus Communication, University of Electronic Science and Technology of China, Chengdu 611731, PR China

ARTICLE INFO

Keywords:
Mental illness
rs-fMRI
GNN
Heterogeneous GNN
Interpretability

ABSTRACT

Rapid and accurate diagnosis of psychiatric disorders remains a significant challenge in medicine. Known for their strength in modeling non-Euclidean data, Graph Neural Networks have been widely used in brain network analysis. Among existing methods, individual graph models can effectively identify disease-related abnormal brain regions, but they lack effective utilization of short-term blood oxygen level-dependent (BOLD) variation information. Population graph models incorporate non-imaging data but overlook the impact of site heterogeneity, resulting in lower accuracy and interpretability. In order to surmount the aforementioned issues, we propose an end-to-end Multi-View fusion heterogeneous graph neural network (MF-HGNN). The first stage extracts individual brain features from both topological and spatiotemporal views and employs a neighbor-distant pooling method to identify potential biomarkers. The second stage integrates imaging and non-imaging data to build a heterogeneous population graph that reflects site-level diversity. We further propose a heterogeneous graph hierarchical convolutional model that extracts complementary features from neighbors within and across sites for global classification of autism spectrum disorder. MF-HGNN attains an accuracy of 95.07% and an AUC of 97.36% on ABIDE I, and an accuracy of 99.16% and an AUC of 99.86% on Rest-MDD, outperforming existing methods. The identified biomarkers align with prior research, confirming the model's reliability. Code is available at: https://github.com/Lyhh01/MF-HGNN.

1. Introduction

In recent years, mental disorders have become a widespread global public health concern [1]. Studies show that such conditions often involve systemic impairment of health [2], making early diagnosis and intervention crucial. However, the lack of precise pathological mechanisms and reliable biomarkers makes accurate diagnosis difficult. Clinicians often rely on subjective judgment [3], leading to frequent misdiagnoses. Therefore, developing an accurate and reliable automated diagnostic method is crucial.

Functional magnetic resonance imaging (fMRI) has become widely applied to exploring brain functional networks [4,5]. Several studies have utilized functional connectomes derived from resting-state fMRI (rs-fMRI) data to investigate abnormal brain connectivity mechanisms. In early rs-fMRI studies, machine learning techniques were commonly employed for feature construction, feature selection, and model development [6,7]. Pang et al. [8] introduced an SVM-driven approach for automatic subtype identification, while Mousavian et al. [9] employed the Wilcoxon rank-sum test to extract informative patterns from rs-fMRI similarity matrices and applied classifiers for MDD prediction. As deep learning advances, its strength in modeling intricate structures from high-dimensional data has become more prominent [10]. Kawahara et al. [11] designed BrainNetCNN, a convolutional neural network tailored to topological brain features. The network is used to predict neurodevelopmental performance in premature infants. Although significant progress has been made, current methods still face challenges

* Corresponding author.
E-mail addresses: 212024085404011@cdu.edu.cn (Y. Liu), huangrb@cdu.edu.cn (R. Huang), liuchang@cdu.edu.cn (C. Liu), zhangjianwei@cdu.edu.cn (J. Zhang), yifs@163.com (F. Yi), 212025085404003@cdu.edu.cn (J. Chen), hechun_li@foxmail.com (H. Li), jss_uestc@163.com (S. Jiang), chengluo@uestc.edu.cn (C. Luo).

https://doi.org/10.1016/j.bspc.2026.110628
Received 6 November 2025; Received in revised form 27 March 2026; Accepted 10 May 2026
Available online 20 May 2026

in characterizing the intricate interactions within brain functional networks, hindering a comprehensive understanding of their dynamic connectivity and topological structures.

Graph Neural Networks (GNNs) have become mainstream tools for analyzing rs-fMRI brain networks, typically categorized into individual and population graph models. Individual graph models represent brain networks with nodes as regions of interest (ROIs) and edges as inter-regional connections, focusing on identifying disease-related abnormalities for subject classification [12–16]. However, they often overlook non-imaging data. Population graph models utilize genetic, phenotypic, and imaging data while integrating intrinsic relationships among these datasets to construct comprehensive graphs, thereby enabling disease diagnosis. [17–22]. However, nodes in the population graph represent only subject-level features without modeling brain networks, so they fail to locate key regions or biomarkers, limiting clinical trust. Recent studies have incorporated brain networks into population graphs [23,24], but they overlook the dynamic nature of brain connectivity over time. As a highly dynamic system, the brain shows time-varying functional patterns. Furthermore, the failure to account for variations among subjects across different testing sites, particularly the insufficient attention paid to heterogeneity issues arising from equipment and technical differences between sites, significantly impacted the accuracy of predictions.

To overcome these issues, we introduce a Multi-View Fusion Heterogeneous Graph Neural Network for accurate brain disorder classification and biomarker identification. The model operates in two stages. In the first stage, we introduce a Multi-View Node Feature Learning module (MVFL) that integrates dynamic spatiotemporal patterns and static clustering structures to capture individual-level features and model functional connectivity. MVFL consists of two branches: one extracts dynamic features via a sliding-window-based dynamic functional connectivity network (FCN) and multi-layer propagation to capture local node evolution; the other captures static topological features reflecting modular brain organization through functional clustering. To retain discriminative long-range features while filtering irrelevant regions, we propose a Neighbor-Distant dual-channel Pooling (NDP) method, which scores nodes and uses random walks to preserve distant but important nodes, aiding biomarker discovery. During the second phase, a heterogeneous Graph Neural Network is constructed to learn cross-attention features based on the imaging site. Edges are constructed using both imaging and phenotypic data. A Graph Transformer aggregates intra-site and inter-site neighbor features, while each layer integrates dual-branch features and applies global convolution to extract high-level representations. Finally, multilevel features are fused for classification. In summary, our key contributions are outlined below:

(1) We propose a MF-HGNN that unifies individual-level and group-level modeling. It extracts dynamic spatiotemporal and static topological features as node attributes, and fuses intra-site and cross-site information through a cross-graph attention convolution module to obtain the final feature embeddings.

(2) We introduce Neighbor-Distant Pooling to identify key biomarkers and enhance interpretability. NDP effectively highlights brain ROIs distinguishing patients from controls, showing high consistency with known disease-related brain regions.

(3) MF-HGNN demonstrates superior performance in cross-validation and transductive evaluation across autism spectrum disorder (ASD) and Resting-state fMRI Meta-MDD (REST-meta-MDD) datasets. Extensive ablation studies demonstrate the efficacy of each module as well as their optimal configuration.

The structure of this paper is as follows: Section 2 reviews related work, Section 3 details the proposed MF-HGNN model, Section 4 presents experimental results and biomarker analysis based on public datasets, Section 5 discusses the experimental setup, and finally, Section 6 summarizes the paper.

2. Related work

This section commences with an overview of conventional machine and deep learning techniques for rs-fMRI-based brain disorder analysis. It then narrows its focus to survey the development and use of GNNs within this domain.

2.1. Traditional machine learning and deep learning methods

Traditional machine learning studies on rs-fMRI often rely on handcrafted features, typically extracting ROIs to construct FCNs as input features [25]. Sen et al. [26] combined static and dynamic fMRI features and used an RBF-kernel SVM to distinguish MDD patients from controls. Khazaee et al. [27] selected optimal graph metrics for AD Classification using SVM. Plitt et al. [28] evaluated how different classifiers and ROI settings affect ASD classification, finding SVM performed best. However, these non-end-to-end methods depend heavily on initial feature extraction and struggle to capture topological and temporal patterns. Driven by computational progress, deep learning has become a prominent approach. Zhao et al. [29] employed deep and recurrent neural architectures to derive representations from functional connectivity (FC) and ROI temporal data, integrating classifiers to differentiate patients from controls. Zhang et al. [30] developed a diffusion kernel attention model to overcome limited interaction and flexibility among subtasks within conventional pipelines, facilitating robust brain network representation. Kawahara et al. [11] proposed BrainNetCNN, employing convolutions across edges, nodes, and graphs to classify brain disorders. Although these models bypass manual feature engineering, they remain constrained in capturing network topology and pinpointing biomarkers.

2.2. Methods based on graph neural network

Graph Neural Networks, leveraging their modeling advantages for non-Euclidean structures, have been increasingly applied in medical imaging. Lei et al. [31] proposed a GNN-capsule network integrating multi-head predictions to enhance lesion classification capabilities by modeling local retinal region relationships. In finger vein recognition, Chang et al. [32] developed FV-DGNN, which extracts vein features via convolutional autoencoders, constructs graphs using inter-sample distance vectors, and jointly models pairwise relationships to achieve robust matching. Functional connectivity derived from rs-fMRI exhibits complex, irregular graph structures, making GNNs particularly suitable for representing brain networks and learning their associations with individual characteristics. GNN-based methods for psychiatric disorder diagnosis generally fall into two categories: individual-level and population-level approaches.

Individual-level methods represent each brain as a graph, where nodes are brain regions. Li et al. [33] proposed an interpretable GNN for biomarker identification. Wang et al. [34] developed a residual graph-based network employing FCNs for capturing temporal relations and identifying critical regions. Zhao et al. [29] employed a dynamic GCN for disease classification. Xing et al. developed DS-GCN, combining spectral GCNs with LSTMs to model temporal patterns and predict subject-specific traits. Gadgil et al. [35] used a spatiotemporal GCN trained on short BOLD segments to capture non-stationary connectivity. Wen et al. [14] proposed a multi-view GNN with prior structural knowledge and graph structure learning for ASD detection. While these methods identify topological features and biomarkers, they often lack effective integration of temporal dynamics, limiting performance.

Population-level methods model all subjects as nodes in a single graph. Parisot et al. [36] built sparse graphs with phenotype-based edges and image-derived node features. Kunda et al. [37] introduced a second-order connectivity measure and used MIDA to reduce site bias. Huang et al. [18] developed a spectral GCN leveraging neuroimaging and clinical datasets. Kazi et al. [38] proposed InceptionGCN to capture

Table 1
Subject phenotypic information in both datasets.

Dataset	Site ID/Name	Subject (ND/HC)	Gender (M/F)	Age	Fiq	Edu (Years)
ABIDE	NYU	181/127	243/65	14.68 ± 7.32	110.78 ± 14.86	–
	UM	103/142	198/47	13.95 ± 3.28	109.76 ± 12.04	–
	LEUVEN	32/54	72/14	16.92 ± 5.15	112.43 ± 13.21	–
	UCLA	89/74	136/27	12.83 ± 2.71	103.48 ± 12.88	–
	MAX_MUN	45/49	88/6	29.15 ± 13.26	110.45 ± 11.84	–
	CALTECH	20/19	33/6	28.67 ± 10.58	107.71 ± 9.47	–
	PITT	34/41	64/11	18.79 ± 8.53	110.26 ± 11.99	–
	OLIN	23/20	38/5	17.21 ± 3.85	113.54 ± 17.01	–
	OHSU	13/16	27/2	10.86 ± 1.97	110.40 ± 17.08	–
	SDSU	18/22	34/6	14.32 ± 2.45	111.96 ± 12.90	–
	TRINITY	32/25	52/5	17.58 ± 4.12	109.70 ± 14.09	–
	USM	36/28	64/0	22.34 ± 8.91	105.31 ± 18.02	–
	YALE	35/41	62/14	12.15 ± 3.09	98.34 ± 20.21	–
	KKI	28/32	51/9	9.98 ± 1.37	107.00 ± 16.03	–
	STANFORD	15/18	28/5	9.63 ± 1.52	113.96 ± 15.10	–
	SBL	12/13	25/0	36.82 ± 10.75	106.50 ± 14.11	–
	CMU	10/12	19/3	26.58 ± 6.74	112.64 ± 11.28	–
MDD	20	282/251	148/385	38.61 ± 14.57	–	12.5 ± 3.2
	21	86/70	68/88	34.15 ± 12.82	–	13.2 ± 2.8
	25	89/63	45/107	66.42 ± 7.53	–	11.8 ± 2.5

A dash ("–") indicates that the corresponding non-imaging information is not available in the original dataset.

structural diversity. Rakhimberdina et al. [19] constructed multiple graphs using different feature combinations and fused them via a neural ensemble. Despite their strengths, these methods often struggle to jointly capture shared and modality-specific signals, limiting deep multimodal integration.

3. Methods

The proposed framework comprises two main components (see Fig. 1). First, a brain connectivity graph is constructed from imaging (e.g., rs-fMRI) and non-imaging data (e.g., gender, age, site) of m subjects to learn subject-level feature embeddings. This captures spatial and temporal brain correlations, simulates subnetwork organization, and identifies key biomarkers via the spatiotemporal topological feature extraction module and NDP. Second, a site-based heterogeneous population graph is built, where subjects are nodes and edges represent within-site and cross-site relationships weighted by phenotypic similarity. Cross-graph attention convolution is then applied to fuse features for node classification, distinguishing healthy controls (HC) from ASD patients.

3.1. Brain connectomic graph

The brain connectivity graph is a key representation in rs-fMRI analysis, modeling brain organization by dividing it into ROIs and capturing inter-regional interactions. We analyze and encode brain information from both topological and spatiotemporal perspectives, and employ a pooling method to further refine key features, enabling the precise extraction of meaningful information.

3.1.1. Graph construction

The rs-fMRI scan is preprocessed using a standard brain template, followed by extraction of the average BOLD signal from each region of interest (ROI).

Brain networks exhibit both static topological and dynamic temporal structural characteristics. Graph Neural Networks are well-suited for analyzing data represented in the form of nodes and edges. Accordingly, the FCN derived from rs-fMRI can be modeled as an individual brain connectivity graph $G_{\text{ind}} = (V, E, \mathcal{G})$. Here, $V = \{v_1, v_2, \dots, v_N\}$ denotes the set of nodes, each corresponding to a region of interest (ROI); E represents the set of edges, computed using the Pearson correlation (PC) coefficient of blood-oxygen-level-dependent (BOLD) signals; and $\mathcal{G} = \{G_{\text{topo}}, G_{\text{time}}\}$ represents the bi-view subgraph set of G_{ind}, where G_{topo} characterizes the static topological structure of the brain network, and G_{time} captures its dynamic temporal variations. Here, the BOLD signal is defined as $B = [b_1, \dots, b_N]^T \in \mathbb{R}^{N\times D}$, where $b_i \in \mathbb{R}^D$. N denotes the number of ROIs, and D represents the number of time points.

Topology graph construction. To capture the static topological characteristics of brain networks, PC coefficients are computed between the time series of each pair of ROIs using the full-length BOLD signals, yielding the full-period PC matrix $\mathbf{P}_{\text{whole}}$. Taking $\mathbf{P}_{\text{whole}}$ as the node feature representation, the topological subgraph feature matrix is constructed as $\mathbf{X}_{\text{topo}} \in \mathbb{R}^{N\times N}$. Only the top 65% of the most strongly correlated edges are retained (with their weights set to 1), while the remaining edges are set to 0, forming the adjacency matrix $\mathbf{A}_{\text{topo}} \in \{0,1\}^{N\times N}$. Consequently, the topological subgraph constructed from the full time period is represented as $G_{\text{topo}} = (\mathbf{X}_{\text{topo}}, \mathbf{A}_{\text{topo}})$.

Time-varying graph construction. A sliding window technique is employed to extract dynamic brain network patterns from fMRI time series. The BOLD time series is partitioned into $T = \left\lfloor \frac{D-r}{s} \right\rfloor$ intervals, with r and s representing the window length and step size, respectively. The signal from the ith region of interest (ROI) during the tth interval is represented by $b_i(t) \in \mathbb{R}^r$ $(t = 1, 2, \dots, T)$. Subsequently, the PC coefficient between nodes is calculated over each interval to form the FCN, defined as:

$$u_{ij}(t) = \frac{\operatorname{cov}(b_i(t), b_j(t))}{\sigma(b_i(t))\,\sigma(b_j(t))} \tag{1}$$

Here, $\operatorname{cov}(\cdot,\cdot)$ represents the covariance of BOLD signals from a pair of ROIs, while $\sigma(\cdot)$ indicates the standard deviation. This yields a sequence of functional connectivity matrices: $U(t) = (u_{ij}(t)) \in \mathbb{R}^{N\times N}$ $(t = 1, 2, \dots, T)$. Each ROI's correlation with others is used as its node feature, i.e., $\mathbf{X}_{time}(t) = P(t)$, where $P(t)$ is the dense FCN. We retain only the top 30% of edges (set to 1), and set the rest to zero, forming a binary adjacency matrix $\mathbf{A}_{time}(t) = \{u_{ij} \geq \delta(t)\} \in \{0,1\}$, where $\delta(t)$ corresponds to the 30% threshold. Thus, the graph at time t is $G_{time}(t) = (\mathbf{X}_{time}(t), \mathbf{A}_{time}(t))$, and the dynamic sequence is $\mathcal{G}_{\text{time}} = G_{time}(t)_{t=1}^{T}$.

3.1.2. Multi-view node feature learning

The MVFL module comprises two components: Multi-View Feature Encoding (MVFE) and Cross-View Attention Fusion. MVFE incorporates both temporal and topological feature extraction, which jointly capture complementary information embedded in brain network representations. Traditional approaches typically focus solely on static topological features, thereby overlooking the complex dynamic changes and interactions among distinct subnetworks within the brain. By integrating dynamic spatiotemporal features with static topological

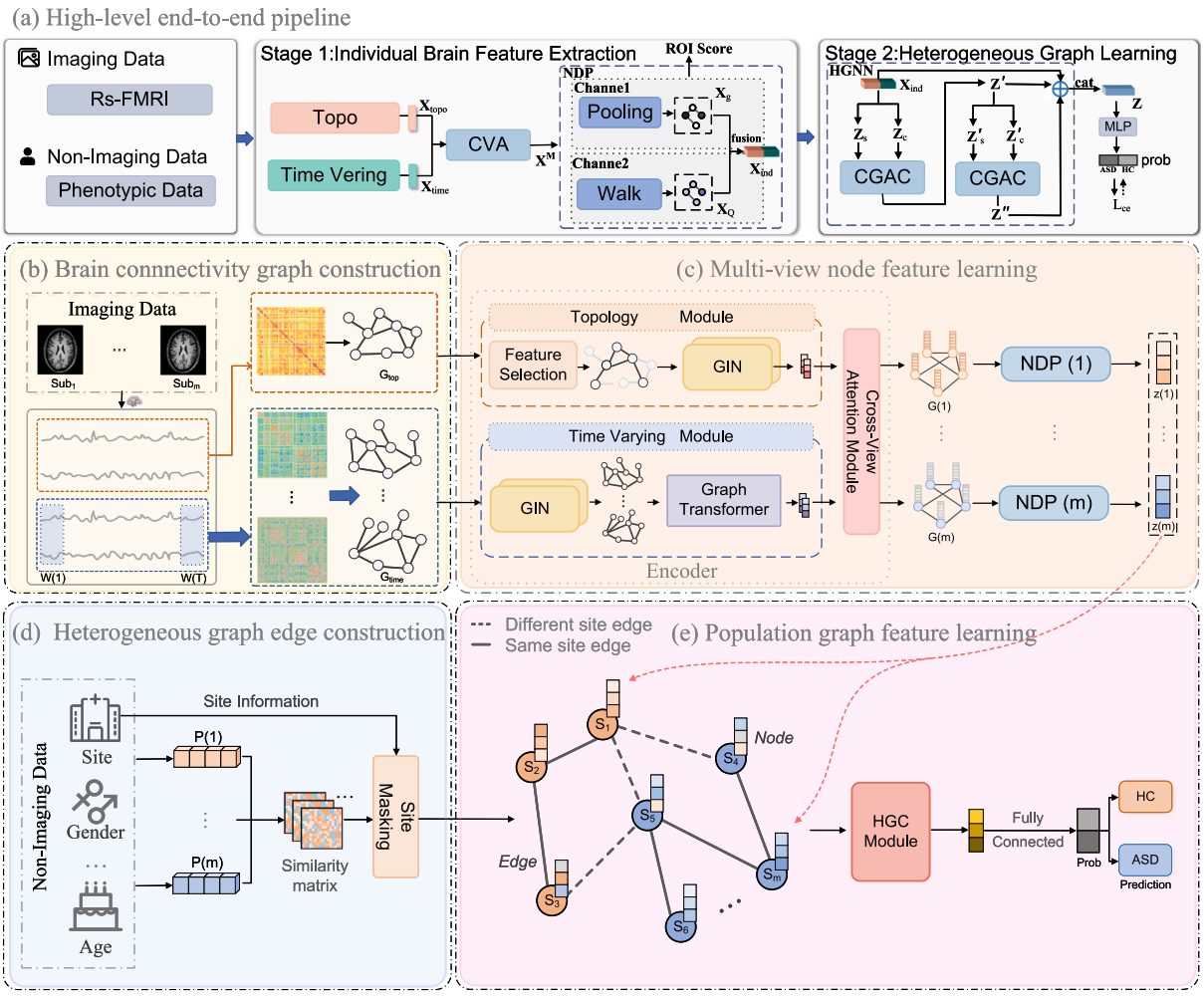


Fig. 1. Overview of the MF-HGNN framework. (a) illustrates the overall end-to-end flowchart of our proposed model. (b) illustrates the construction of brain connectivity graphs from topological and time-varying views. (c) depicts the dual-branch multi-view feature learning module, where the cross-view attention module and the neighbor–distant pooling (NDP) module are used to fuse and enhance features, respectively. (d) shows the construction of edges in the site-aware heterogeneous population graph from phenotypic information, using a site mask to distinguish within-site and cross-site connections. (e) shows the population-based graph neural network for node classification, where each subject-level embedding is treated as a node.

features, the proposed framework enhances the understanding of brain network characteristics, enabling the model to simultaneously capture both microscopic structural details and macroscopic network evolution patterns. A detailed description follows.

Topology feature fusion module. The Brain ROI-aware Graph Isomorphism Network (BrainRGIN) [39] is designed to jointly encode node and edge attributes by integrating the Relational Graph Convolutional Network (RGCN) and the Graph Isomorphism Network (GIN). It is defined as:

$$\mathbf{h}_i^{(l)} = \mathrm{MLP}^{(l)}\left(\left(1+\epsilon^{(l)}\right)\cdot \mathbf{W}_i^{(l)} \cdot \mathbf{h}_i^{(l-1)} + \sum_{j\in\mathcal{N}^{(l)}(i)} \mathbf{W}_j^{(l)} \cdot \mathbf{e}_{i,j}^{(l-1)} \cdot \mathbf{h}_j^{(l-1)} \right) \tag{2}$$

where $\mathbf{h}_i^{(l)}$ denotes the feature of node i at layer l, $e_{i,j}$ represents the edge feature from node i to node j, and $\mathcal{N}^{(l)}(i)$ is the index set of neighboring nodes of node i. Specifically, $\mathbf{W}_i^{(l)}$ and $\mathbf{W}_j^{(l)}$ are functions defined based on node clustering assignments:

$$\mathbf{W}_i^{(l)} = \theta_2^{(l)} \cdot \sigma\left(\theta_1^{(l)} r_i\right) + b^{(l)} \tag{3}$$

where r_i denotes the positional encoding of node i, θ_1 and θ_2 are parameters for learning node clusters, b is a bias term, and l is the number of layers.

The topological feature extraction module aims to explore the topological characteristics of different functional clusters in brain networks. During this process, local features of the brain undergo significant changes, leading to variations in the global feature distribution. Based on this, BrainRGIN convolutional layers are employed to extract important topological features. First, a two-layer dropout is applied to the original input features, performing feature selection by randomly discarding subsets of node features multiple times. This process both disrupts redundant dependencies among features and mitigates overfitting, thereby enhancing the model's generalization ability:

$$\mathbf{X}_{drop}^{(k)} = \mathrm{Dropout}^{(k)}(\mathbf{X}_{topo}) \quad (k=1,2) \tag{4}$$

where $\mathbf{X}_{drop}^{(k)}$ denotes the features after the kth dropout, and $\mathbf{X}_{\mathrm{topo}}$ represents the initial node features of the topological graph. Then, BrainRGIN convolutional layers are applied separately in the two layers to capture the diverse structural properties of brain networks, with the representation of one layer formulated as follows:

$$\mathbf{X}_{\mathrm{rgin}} = \mathrm{BrainRGIN}\left(\mathbf{X}_{drop}, \mathbf{A}_{topo}\right) \tag{5}$$

where $\mathbf{X}_{\mathrm{rgin}}$ represents the node features obtained after processing through the BrainRGIN convolutional layers. Finally, the resulting multi-layer topological features are concatenated to generate the intermediate node embeddings of the topological graph, denoted as $\mathbf{X}_{\mathrm{topo}}^{M}$,

which is formulated as follows:

$$\mathbf{X}_{topo}^{M} = \text{cat}\left(\mathbf{X}_{rgin}^{(l)} \mid k = 1, 2\right) \tag{6}$$

Time-varying feature extraction module. For feature extraction, the constructed time-varying graph sequence serves as input. We use the Spatio-Temporal Attention Graph Isomorphism Network (STAGIN) [40] as the core dynamic representation learner, comprising multiple GIN and Transformer layers (see Fig. 1). In this study, a single layer suffices for dynamic graph feature extraction. Graph Neural Networks remain widely adopted models for processing graph-structured data, generally involving two fundamental operations: aggregation and propagation. The Graph Isomorphism Network (GIN) has demonstrated strong capabilities in graph representation learning, particularly in graph classification tasks. The primary difference between GIN and traditional GNNs pertains to the aggregation process. The aggregation operation in GIN is expressed as:

$$\mathbf{h}_{v}^{(l)} = \text{MLP}^{(l)}\left(\left(1 + \epsilon^{(l)}\right)\mathbf{h}_{v}^{(l-1)} + \sum_{u \in \mathcal{N}(v)} \mathbf{h}_{u}^{(l-1)}\right) \tag{7}$$

where $h_v^{(l)}$ indicates the feature embedding of node v at layer k, $\mathcal{N}(v)$ denotes the neighbors of v, MLP refers to a multilayer perceptron, and ϵ is a trainable scalar. Eq. (7) illustrates GIN's ability to assign adaptive weights per node, differing from the mean aggregation in earlier approaches.

Over a specific time window, the revised node feature matrix derived from the GIN layer is formulated as:

$$\mathbf{X}' = \sigma(\epsilon I + \mathbf{A}_{\mathbf{time}})\mathbf{X}_{time}\mathbf{W} \tag{8}$$

where $\mathbf{W}$ denotes a trainable parameter matrix. To obtain graph-level representations, a readout layer is applied after the GIN layer to aggregate node-level features. Specifically, we employ a readout module built upon a squeeze-and-excitation framework using a multi-layer perceptron (MLP). It is formulated as:

$$\mathbf{sa} = \text{Sigmoid}\left(\mathbf{P}^1\sigma\left(\mathbf{P}^0 \cdot \xi\left(\mathbf{X}'\right)\right)\right) \tag{9}$$

where $\mathbf{sa} \in (0, 1)^N$ represents the spatial attention vector, while $\mathbf{P}^0$ and $\mathbf{P}^1$ denote trainable parameter matrices, and $\xi(\cdot)$ denotes the pooling operation along the channel dimension. The readout yields a sequence of spatial graph embeddings per sample $\mathbf{h}(t) = \{\mathbf{X}'(t)\mathbf{sa}(t)\}_{t=1}^{T}$.

Given the Transformer's strength in modeling semantics over extended sequences, we utilize a single-head variant equipped with self-attention to capture temporal dependencies at various ranges. The spatial graph features $\mathbf{h}(t)$ are passed into the Transformer, yielding spatiotemporal embeddings $\{\mathbf{z}(t)\}_{t=1}^{T}$.

Finally, temporal embeddings from each subject are concatenated to derive a comprehensive graph representation of the individual:

$$\mathbf{X}_{time}^{M} = \|_{t=1}^{T}[\mathbf{z}(t)] \tag{10}$$

where $\mathbf{X}_{\text{time}}^{M}$ denotes the intermediate node embeddings of the time-varying graph, with $\|$ representing the concatenation operation across the sequence of time-varying graphs.

Cross-view attention module. The Cross-View attention (CVA) mechanism fuses information from two complementary perspectives to enhance subject-level representation. Unlike self-attention, which receives the query, key, and value all from a single modality, cross-view attention involves three input vectors: the query from one view and the key and value from another view.

The CVA module is composed of two parallel branches. Specifically, the first branch extracts the query vector from the topological embedding features and performs matching with the key and value vectors derived from the temporal features. In contrast, the second branch obtains the query vector from the temporal embedding features and matches it with the key and value vectors from the topological embeddings. Based on this dual-branch design, the detailed computation process of the CVA module is defined as follows:

$$\mathbf{Att}_{T\to I} = \text{normalize}(\mathbf{Q}_T\mathbf{K}_I^{\top})\mathbf{V}_I\mathbf{W}_{T\to I} \tag{11}$$

$$\mathbf{Att}_{I\to T} = \text{normalize}(\mathbf{Q}_I\mathbf{K}_T^{\top})\mathbf{V}_T\mathbf{W}_{I\to T} \tag{12}$$

where $\mathbf{Q}_T, \mathbf{Q}_I, \mathbf{K}_T, \mathbf{K}_I, \mathbf{V}_T$, and $\mathbf{V}_I$ denote the queries, keys, and values derived from the topological intermediate embeddings (mathbfT) and temporal intermediate embeddings (mathbfI), respectively. $\mathbf{Att}_{T\to I}$ and $\mathbf{Att}_{I\to T}$ represent the attention outputs of each branch within the CVA module, while $\mathbf{W}_{T\to I}$ and $\mathbf{W}_{I\to T}$ are learnable weight matrices.

After the attention computation, each branch of the CVA module adds the original features from its respective modality through a residual connection. Finally, the outputs from both branches are concatenated to preserve modality-specific information. The CVA module function is formulated as follows:

$$\mathbf{X}^{M} = \text{Concat}(\mathbf{Att}_{T\to I} + \mathbf{I},\ \mathbf{Att}_{I\to T} + \mathbf{T}) \tag{13}$$

where $\mathbf{X}^M$ represents the fused intermediate node embeddings, and $\mathbf{Att}_{T\to I}$ and $\mathbf{Att}_{I\to T}$ correspond to the attention computation results of the two branches within the CVA module, respectively.

Algorithm 1: Neighbor_Distant_Pooling($\mathbf{X}_M, \mathbf{A}_{topo}, len$)

Input:
Intermediate node embeddings $\mathbf{X}^M$
Topological adjacency matrix $\mathbf{A}_{\text{topo}}$
Biased random walk length *len*

Output:
Individual node embeddings $\mathbf{X}_{ind}$

Initialize $P = \{\}, M = \{\}, S_{loc} = \{\}, S_{glob} = \{\},\ flag[v] = 0$
Compute $S_{loc} = GCN(\mathbf{X}^M, \mathbf{A}_{topo})$
Select $V' = \text{TOP}_k(S_{loc})$
Construct a pooling subgraph $g(X_g, A_g)$ from V'
Compute $S_{glob} = MLP(\mathbf{X}^M)$
for *each node i in* V' **do**
  Compute $P_{ik} = \frac{I(v_k)}{\sum_{v_j \in \Gamma(v)} I(v_j)}$ for each $k \in \Gamma(i)$
  Update $M[i, k] = P_{ik}$
end
Sort nodes in V' by descending order of S_{glob}
for *each node i in sorted* V' **do**
  Initialize $path = [i]$,Set $start\_node = v_i, k = 1$
  while $k < len$ **do**
    Sample node j from $\Gamma(v)$ based on M
    if $flag[i] = 0$ **then**
      Append v_j to path
      $k = k + 1$
    end
  end
  Collect $Q = Q \cup \{path\}$
end
for *each path i in* Q **do**
  Extract $f(v_{i,j})$ for each $j \in len(path)$
  $\mathbf{A}_{cross}^{(i)} = \frac{I(v_j)}{\sum_{k=1}^{len} I(v_k)}$ for each $k \in len(p_i)$
  $\mathbf{X}_{Q_i} = \sum_{j=1}^{len} w_{ij} \cdot f(v_{i,j})$
  $\mathbf{X}_{ind}^{(i)} = \mathbf{X}_g^{(i)} + \mathbf{X}_{Q_i} \cdot \mathbf{A}_{cross}^{(i)}$
end
return $\mathbf{X}_{ind}$

Neighbor-distant dual-path pooling module. Different neurological disorders affect distinct brain regions; therefore, not all regions contribute equally to diagnosis. In ASD, long-range functional connectivity is significantly disrupted, reflecting altered coherence between distant brain regions. Traditional pooling methods often remove redundant nodes but fail to effectively preserve long-range dependencies. In response to this issue, we propose the NDP method, which retains key nodes and edges while maintaining global long-range information (see Fig. 2). The first path prunes nodes according to their scores, a process that inevitably leads to a compromise in long-range contextual

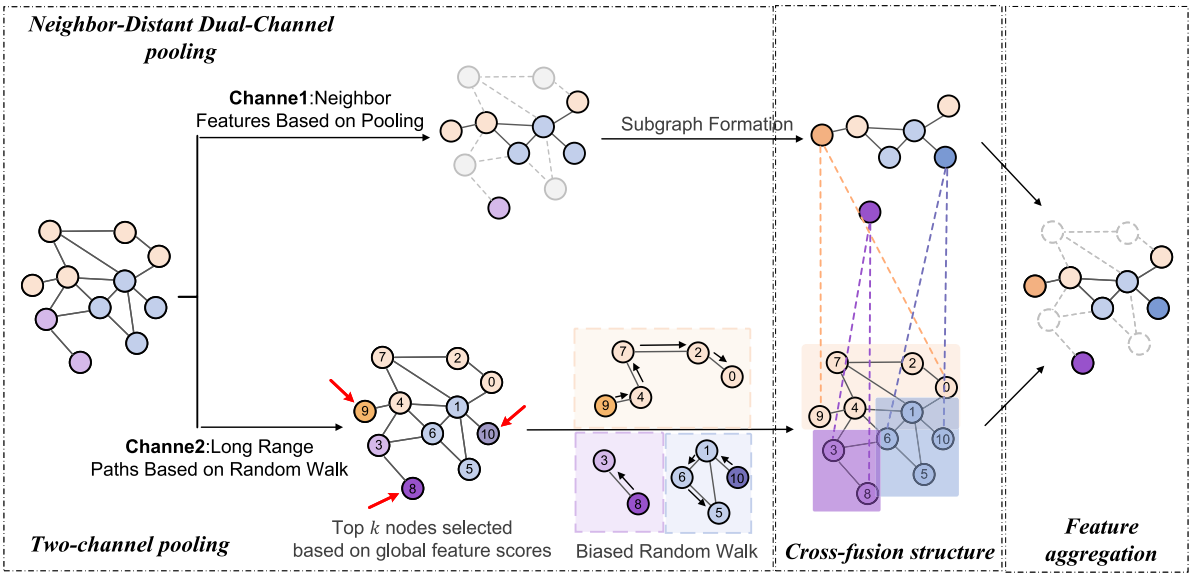


Fig. 2. The NDP Module consists of two channels: one employs TopKPooling to extract locally relevant features, while the other utilizes a biased random walk to capture long-range dependencies. The features obtained from both channels are then fused and used as the node representations in the population graph.

information. The second path employs biased random walks to aggregate and transfer global information, including that from discarded nodes, to the important ones, thereby recovering the lost dependencies. Consequently, NDP preserves both local and long-range structural characteristics, while emphasizing that ROIs play a critical role in distinguishing neurodevelopmental disorders (ND) from HC. To provide a clearer explanation of the proposed method, Algorithm 1 presents the detailed key steps.

Channel 1: The neighbor-feature pooling phase utilizes a locally feature-driven node-dropping strategy for the construction of the pooled graph. This method ranks and filters nodes based on their topology scores, enabling the generation of a pooled graph that accurately captures local neighborhood feature interactions. First, the topology scores of n nodes are computed using the node feature matrix as follows:

$$S_{loc}^{(l)} = \mathrm{GCN}\left(\mathbf{X}_M^{(l)}, \mathbf{A}_{topo}\right) \in \mathbb{R}^{n\times 1} \tag{14}$$

where S_{loc} denotes the local topology scores corresponding to each node feature. Then, based on these scores, the top-k nodes are obtained as follows:

$$idx^{(l)} = \mathrm{TOP}_k\left(S_{loc}^{(l)}\right) \tag{15}$$

$$\mathbf{X}_g^{(l+1)} = \mathbf{X}_{M,idx(l)}^{(l)} \odot \mathrm{sigmoid}\left(S_{loc,idx(l)}^{(l)}\right) \tag{16}$$

$$\mathbf{A}_g^{(l+1)} = \mathbf{A}_{topo\left(idx^{(l)},\ idx^{(l)}\right)}^{(l)} \tag{17}$$

Then, we apply a top-k algorithm to retain the 100 highest-scoring nodes. The updated $\mathbf{X}_g$ and $\mathbf{A}_g$ are obtained via indexing. Here, k denotes the pooling ratio, which is set to 90% of all nodes based on experimental results. A detailed discussion of the pooling ratio's specific impact on model performance is reserved for Section 4.9.

Channel 2: Random walks are essential tools in network analysis and are typically categorized as either biased or simple. Simple random walks move randomly between nodes, whereas biased walks adjust transition probabilities to better capture complex structures. In this channel, we employ biased random walks to extract long-range paths of important nodes. First, the global scores of n nodes are computed as follows:

$$S_{glob}^{(l)} = \mathbf{X}_M^{(l)} w^{(l)} + b^{(l)} \tag{18}$$

where w denotes a trainable weight matrix, b represents a bias vector, and the significance of node v, denoted as $I(v)$, is defined as follows:

$$I(v) = S_{glob} \tag{19}$$

Subsequently, the vertices are ranked from highest to lowest according to the scores $S_{glob}^{(l)}$, and the top-k highest-scoring nodes are sequentially selected as the starting points for the random walks. The index set of these top-k nodes is denoted as $idx_{\mathrm{glob}} = \{\mathrm{rank}(S_{glob}) \le k\}$, where $\mathrm{rank}(s_i)$ indicates the ranking of node i's score among all nodes in the graph.

The walks proceed along the edges connecting the nodes. For the current node v, its set of neighbors is denoted as $\Gamma(v)$. For each neighboring node pair (v,u) with $u \in \Gamma(v)$, the weight is defined as $w_{vu} = I(u)$. The transition probability matrix $\mathcal{P}$ is then constructed by normalizing the neighbor weights of node v. The matrix element $\mathcal{P}_{vu}$ represents the probability of transitioning from node v to its neighbor u, computed as follows:

$$\mathcal{P}_{vu} = \frac{w_{vu}}{\sum_{k\in\Gamma(v)} w_{vk}} = \frac{w_{vu}}{\sum_{k\in\Gamma(v)} I(k)} \tag{20}$$

During the random walk along the same path, to prevent information coverage from being restricted due to path redundancy, each node is allowed to be visited at most once, thereby encouraging the walk to extend into unexplored neighborhoods. Consequently, a set of visited nodes V_{visited} must be recorded. The walk terminates once the predetermined length is reached. The collection of all paths can be expressed as $Q = \{Q_1, Q_2, \ldots, Q_l\}$, where l denotes the total number of paths, and P_i represents the sequence of node features for the ith path, $\mathbf{X}_{Q_i} = \{f(v_{i,1}), f(v_{i,2}), \ldots, f(v_{i,j})\}$, with $f(v_{i,j})$ indicating the feature vector of the jth node in the ith path.

At this stage, both the initial graph and the long-range paths have been constructed. To effectively integrate information from the two channels, we introduce a cross-channel convolution, defined as follows:

$$\mathbf{X}_{ind} = \mathbf{X}_g + \sum_{i=1}^{l} \mathbf{A}_{\mathrm{cross}}^{(i)} \cdot \mathbf{X}_{Q_i} \tag{21}$$

where $\mathbf{X}_{\mathrm{ind}}$ represents the individual node embeddings. In addition, the cross-channel connection matrix $\mathbf{A}_{\mathrm{cross}}$ is obtained as follows:

$$\mathbf{A}_{\mathrm{cross}}^{(i)} = \mathrm{softmax}(S_{\mathrm{path}}^{(i)}) \tag{22}$$

where S_{path} represents the scores corresponding to the nodes in the path sequence Q_i.

3.2. Heterogeneous population graph

The heterogeneous population graph is a powerful tool for integrating multi-site subject data and capturing group-level associations. It models complex inter-subject relationships and mitigates site-related biases, supporting robust analysis in cross-site settings.

3.2.1. Building heterogeneous population graph

Research on diverse population networks that account for site variability remains limited. Recent studies have revealed systematic fMRI discrepancies across imaging sites [41]. To address this issue, we propose a Site-based Heterogeneous Population Graph (SHG), which represents imaging centers as node attributes and classifies edges into *intra-site* and *inter-site* categories.

The SHG is denoted as $G_{\text{pop}}(V', E')$, where nodes V' represent subjects, and edges E' consist of intra-site edges E'_s and inter-site edges E'_c. Each subject's feature vector, derived from the brain connectivity graph, forms the heterogeneous feature matrix $\mathbf{X}_p \in \mathbb{R}^{n\times h}$, where $h = N \times D$. The adjacency matrix is defined as $\mathbf{A}_p = \text{Thr}(S^p)$, where $S^p \in \mathbb{R}^{n\times n}$ encodes phenotypic similarity. The threshold function $\text{Thr}(\cdot)$ ensures sparsity by preserving edges whose weights exceed γ. By doing so, it seamlessly unifies imaging and non-imaging data within a coherent network architecture.

3.2.2. Heterogeneous graph convolution module

A Transformer-inspired graph convolution module [42] serves as the backbone for feature extraction. For given node features $X_p^l = (x_1^l, \ldots, x_n^l)^T$, the multi-head attention score $a_{c,i,j}^{(l)}$ directed from node j to node i is first computed.

Once the attention between vertices is computed, message aggregation is performed on node i 's neighbors:

$$e_{c,i,j} = H_{c,e} e_{ij} + b_{c,e} \tag{23}$$

$$v_{c,j} = H_{c,v} x_j + b_{c,v} \tag{24}$$

$$x_i^{(l+1)} = \|_{c=1}^{\text{Num}} \left[\sum_{j\in N(i)} a_{ij}^{(l)} \left(v_{c,j}^{(l)} + e_{c,ij} \right) \right] \tag{25}$$

where $\|$ denotes concatenation based on the first attention head's scores. $H_{c,e}$, $H_{c,v}$, $b_{c,e}$, and $b_{c,v}$ are trainable parameters. Within the Graph Transformer, the scaled adjacency array from a standard GCN functions as the propagation array. For weighted aggregation, the target feature $x_j^{(l)}$ is transformed to $v_j^{(l)}$ [43].

To reduce site heterogeneity, we design a Cross Graph Attention Convolution (CGAC) layer that captures shared features within sites and distinctive features across sites for richer representation. Global modeling is then applied to all samples for cross-site feature sharing. This module has two branches (see Fig. 3), each using CGAC layers. Specifically, the $(l+1)$th layer of the within-site branch incorporates features from the lth layer of the cross-site branch, and vice versa. This cross-branch interaction helps mitigate heterogeneity.

First, a Graph Transformer is employed to integrate data from intra-site and inter-site neighbors to generate the final node embeddings matrix $\mathbf{Z}$. For ease of notation, we denote $\mathbf{Z}_s^{(0)}$ and $\mathbf{Z}_c^{(0)}$ as the initial node features for intra-site and inter-site neighbors, respectively:

$$\mathbf{Z}_s^{(0)} = \text{GraphTransformer}(\mathbf{X}_{ind}, E'_s) \tag{26}$$

$$\mathbf{Z}_c^{(0)} = \text{GraphTransformer}(\mathbf{X}_{ind}, E'_c) \tag{27}$$

Subsequently, a global convolution is applied to explore and integrate the distinctive information between intra-site and inter-site features, yielding the final node features $Z^{(0)}$ of the first layer, which are formally expressed as follows:

$$\mathbf{Z}^{(0)} = \text{GraphTransformer}\left([\mathbf{Z}_s^{(0)}; \mathbf{Z}_c^{(0)}], E'\right) \tag{28}$$

where $[o; o]$ denotes vertical stacking of two matrices. In addition, the CGAC layers at the l^{th} layer are represented as:

$$\mathbf{Z}_s^{(l)} = \text{GraphTransformer}\left([\mathbf{Z}_s^{(l-1)}; \mathbf{Z}_c^{(l-1)}], E'_s\right) \tag{29}$$

$$\mathbf{Z}_c^{(l)} = \text{GraphTransformer}\left([\mathbf{Z}_c^{(l-1)}; \mathbf{Z}_s^{(l-1)}], E'_c\right) \tag{30}$$

$$\mathbf{Z}^{(l)} = \text{GraphTransformer}\left([\mathbf{Z}_s^{(l)}; \mathbf{Z}_c^{(l)}], E'\right) \tag{31}$$

After each convolution layer, a normalization operation is performed to ensure that their sum equals one. The same operation is applied during initialization to prevent significant gaps and negative values. Finally, an MLP predicts the output labels:

$$\hat{y}_i = \text{MLP}(\|_{l=1}^{L} \mathbf{Z}_i^{(l)}) \tag{32}$$

where L denotes the number of heterogeneous graph convolution layers.

The process of message passing and feature extraction at each layer of the heterogeneous graph convolution module is illustrated in Fig. 3. The model generates outputs using the combined features from all layers, and parameter optimization is guided by the cross-entropy loss:

$$L_{ce} = -\frac{1}{n} \sum_{i=1}^{n} \left[y_i \cdot \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right] \tag{33}$$

4. Experiments and analyses

To comprehensively assess the effectiveness and reliability of the MF-HGNN framework, we first conducted experiments under two commonly used learning paradigms, namely inductive learning and transductive learning. At the same time, in order to improve the interpretability and soundness of the model, we designed experiments focusing on module necessity verification, hyperparameter optimization, and biomarker visualization, thereby providing comprehensive support for the model's performance.

4.1. Dataset

Validating MF-HGNN's generalization, experiments were performed using two public datasets: ABIDE [44] and REST-meta-MDD [45]. Both datasets contain rich imaging and non-imaging data, and we performed various classification tasks on them.

4.1.1. ABIDE dataset

The ABIDE I dataset comprises 1112 participants from 17 distinct international research sites worldwide, providing imaging data (rs-fMRI) alongside non-imaging (phenotypic) information such as age, gender, and research site. This study utilized 871 subjects who passed the image quality check, including 403 individuals with ASD and 468 healthy controls, with detailed demographics presented in Table 1. The initial preprocessing of the neuroimaging data was carried out with the Configurable Pipeline for the Analysis of Connectomes (C-PAC) [46]. This involved normalizing the functional images to the MNI152 standard space. Subsequently, the mean time series for all HO brain regions were derived based on the Harvard–Oxford atlas [47].

4.1.2. MDD dataset

This study utilized 841 subjects from the publicly available REST-meta-MDD cohort, including 457 individuals with MDD and 384 controls from the three primary sites, as detailed in Table 1. The data were preprocessed using DPARSF [48], following a procedure similar to C-PAC, involving temporal alignment, motion correction, and spatial filtering. Finally, mean temporal signals across cerebral regions were derived according to the Harvard-Oxford (HO) atlas.

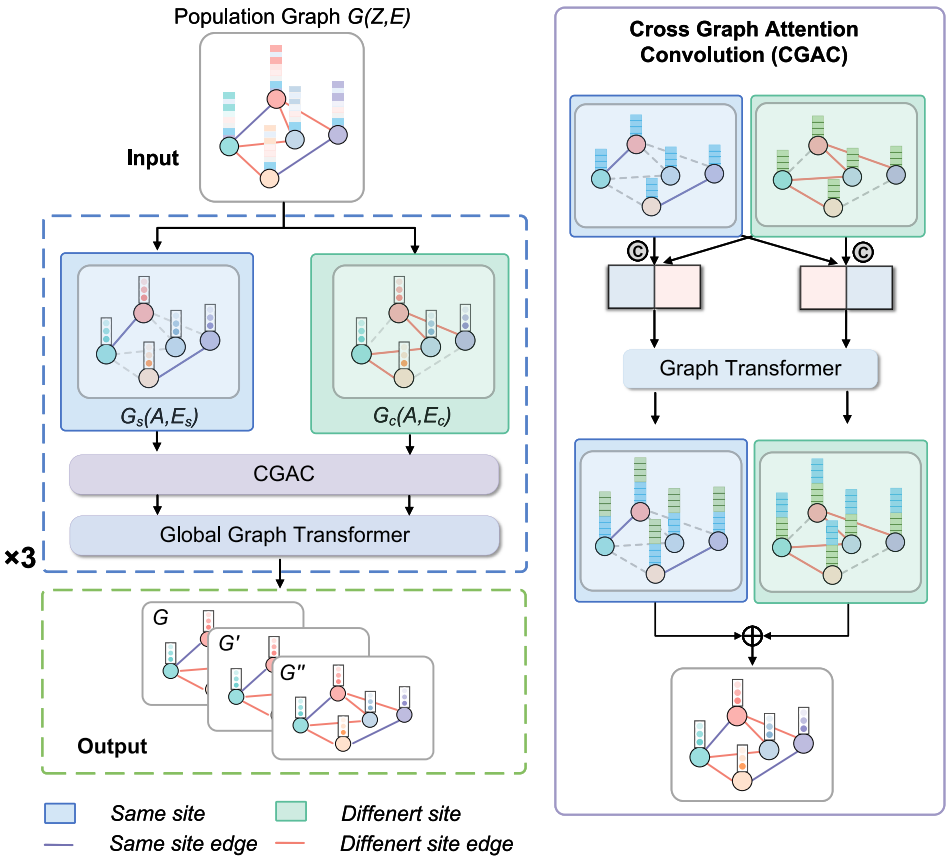


Fig. 3. The HPG module. It employs a Graph Transformer to capture site-specific homogeneous and heterogeneous features, subsequently integrating them for classification. Before graph convolution, feature complementarity is achieved through a cross-concatenation mechanism.

Table 2
Classification results on ABIDE and MDD datasets.

Type	Method	ABIDE					MDD				
		ACC	AUC	SEN	SPE	F1	ACC	AUC	SEN	SPE	F1
M.L.	SVM	65.48 ± 2.34	64.81 ± 2.46	70.66 ± 2.12	64.79 ± 3.25	68.65 ± 2.15	57.31 ± 3.16	54.86 ± 4.17	58.99 ± 3.45	43.39 ± 2.62	47.23 ± 2.30
	RF	60.16 ± 3.89	66.24 ± 3.10	69.06 ± 2.55	66.47 ± 3.25	67.78 ± 2.21	60.73 ± 3.98	60.10 ± 4.21	59.48 ± 5.65	51.27 ± 4.01	54.32 ± 4.68
D.L.	BrainNetCNN	69.75 ± 2.79	71.90 ± 3.21	66.72 ± 9.96	72.58 ± 8.16	70.65 ± 5.64	66.59 ± 2.89	68.74 ± 3.37	63.49 ± 8.21	69.32 ± 8.84	67.83 ± 5.47
Ind.GCN	BrainGNN(2021)	71.05 ± 4.23	65.47 ± 6.54	62.04 ± 5.52	71.23 ± 4.83	67.70 ± 7.44	68.19 ± 5.46	61.55 ± 7.82	58.16 ± 6.31	67.47 ± 6.35	63.79 ± 8.61
	RGTNet(2024)	70.57 ± 5.76	67.64 ± 5.02	58.95 ± 5.65	65.79 ± 5.72	62.42 ± 4.91	70.25 ± 5.34	67.84 ± 5.03	59.17 ± 5.51	65.88 ± 5.69	62.31 ± 5.08
Pop.GCN	PopulationGCN(2017)	71.64 ± 4.91	72.72 ± 6.95	75.01 ± 10.25	67.83 ± 11.09	71.34 ± 4.96	65.51 ± 2.94	66.64 ± 6.83	69.11 ± 8.35	61.96 ± 10.14	66.43 ± 4.89
	MMGL(2022)	89.78 ± 7.70	89.62 ± 7.92	91.89 ± 6.23	87.35 ± 11.18	89.78 ± 7.70	84.26 ± 2.06	86.23 ± 3.45	87.09 ± 2.24	81.27 ± 2.38	85.29 ± 2.72
	MIDA(2022)	71.83 ± 3.32	77.31 ± 4.87	72.64 ± 2.85	69.15 ± 2.36	72.04 ± 3.46	67.92 ± 8.62	73.26 ± 5.43	58.71 ± 3.63	65.39 ± 3.41	68.92 ± 4.75
	MAMF-GCN(2023)	89.89 ± 2.31	89.28 ± 2.30	91.44 ± 3.79	88.69 ± 3.35	89.40 ± 2.25	95.28 ± 0.07	95.71 ± 0.05	95.21 ± 0.05	95.20 ± 0.09	95.19 ± 0.07
	LG-GNN(2023)	81.53 ± 2.24	87.65 ± 3.26	83.69 ± 4.69	78.01 ± 5.11	83.21 ± 3.69	91.86 ± 1.48	92.24 ± 1.78	91.00 ± 3.39	92.49 ± 2.88	91.51 ± 1.70
	FC-HGNN(2025)	92.08 ± 0.16	96.01 ± 0.11	90.99 ± 0.24	94.87 ± 0.09	92.83 ± 0.12	85.02 ± 0.07	89.48 ± 0.12	81.75 ± 0.09	86.71 ± 0.25	84.07 ± 0.08
Ours	MF-HGNN(ours)	**95.07 ± 0.03**	**97.36 ± 0.02**	**94.46 ± 0.08**	**96.58 ± 0.06**	**95.47 ± 0.03**	**99.16 ± 0.01**	**99.86 ± 0.00**	**99.73 ± 0.01**	**98.44 ± 0.04**	**99.07 ± 0.02**

4.2. Implementation details and methods of comparison

The model was built, trained, and tested using PyTorch on a platform with NVIDIA A16 GPUs. Adam optimizer [49] was applied, starting at a learning rate of 0.01, with dropout of 0.3 and up to 500 epochs. The Transformer encoder with GIN had a hidden dimension of 64. Five metrics evaluated classification performance: accuracy, sensitivity, specificity, area under the curve, and F1 score. Experiments were grouped into four categories: machine learning, deep learning, individual graph models, and population graph models. Data were mapped using the Harvard-Oxford atlas. Phenotypic features and hyperparameters matched those in the original papers. The experimental setup was consistent across methods, though preprocessing and inputs varied. Machine learning (SVM) and traditional deep learning models (BrainNetCNN) used the input of average time series from brain regions as input. Graph neural networks (BrainGNN, PopulationGCN, etc.) used PC matrices as adjacency matrices to build graphs for input.

SVM [50] and RF [51] are classical supervised learning classifiers. SVM finds an optimal hyperplane to maximize class margins, while RF combines multiple decision trees via majority voting to improve accuracy and reduce overfitting.

BrainNetCNN [52]: Utilizes a novel convolutional network framework that leverages the topological locality of brain networks to predict clinical neurodevelopmental outcomes.

BrainGNN [33]: Implements this via a dedicated node pooling layer.

RGTNet [34]: This method integrates information from the entire graph using a virtual global node, helping to identify brain regions with global diagnostic significance.

Table 3
Leave-group-out classification results on ABIDE and MDD datasets.

Type	Method	ABIDE					MDD				
		ACC	AUC	SEN	SPE	F1	ACC	AUC	SEN	SPE	F1
M.L.	SVM	64.28 ± 2.51	64.27 ± 2.63	69.32 ± 2.37	63.13 ± 3.42	65.91 ± 2.45	57.65 ± 3.16	55.10 ± 4.17	59.25 ± 3.45	43.62 ± 2.62	47.45 ± 2.30
	RF	61.08 ± 1.34	64.43 ± 1.45	79.69 ± 1.52	39.45 ± 2.45	68.74 ± 1.05	57.43 ± 1.52	58.59 ± 1.64	40.14 ± 2.05	71.97 ± 1.82	46.17 ± 2.00
D.L.	BrainNetCNN	68.05 ± 2.97	71.18 ± 3.38	65.86 ± 10.09	70.29 ± 8.43	67.21 ± 5.87	66.90 ± 2.89	68.95 ± 3.37	63.76 ± 8.21	69.56 ± 8.84	68.04 ± 5.47
Ind.GCN	BrainGNN(2021)	68.95 ± 4.41	64.60 ± 6.78	60.79 ± 5.89	69.04 ± 5.15	63.45 ± 7.75	68.50 ± 5.46	61.77 ± 7.82	58.42 ± 6.31	67.71 ± 6.35	64.01 ± 8.61
	RGTNet(2024)	70.36 ± 5.89	67.26 ± 5.29	57.83 ± 5.94	65.43 ± 6.07	59.67 ± 5.31	70.57 ± 5.34	68.04 ± 5.03	59.41 ± 5.51	66.10 ± 5.69	62.52 ± 5.08
Pop.GCN	PopulationGCN(2017)	71.18 ± 5.03	72.36 ± 7.21	74.15 ± 10.51	66.24 ± 11.48	69.58 ± 5.17	65.21 ± 2.94	66.41 ± 6.83	68.86 ± 8.35	61.72 ± 10.14	66.19 ± 4.89
	MMGL(2022)	89.32 ± 7.87	89.03 ± 8.22	92.49 ± 5.66	85.57 ± 13.34	89.32 ± 7.87	84.68 ± 2.06	86.46 ± 3.45	87.37 ± 2.24	81.50 ± 2.38	85.51 ± 2.72
	MIDA(2022)	71.05 ± 3.54	76.83 ± 5.15	71.72 ± 3.19	67.41 ± 2.76	69.88 ± 3.68	67.62 ± 8.62	73.02 ± 5.43	58.47 ± 3.63	65.15 ± 3.41	68.69 ± 4.75
	MAMF-GCN(2023)	89.33 ± 2.54	88.46 ± 2.57	90.88 ± 4.03	87.71 ± 3.75	89.26 ± 2.51	95.65 ± 0.07	95.94 ± 0.05	95.48 ± 0.05	95.44 ± 0.09	95.43 ± 0.07
	LG-GNN(2023)	80.20 ± 2.49	87.06 ± 3.52	82.55 ± 5.03	75.89 ± 5.47	78.91 ± 3.97	92.22 ± 1.48	92.46 ± 1.78	91.25 ± 3.39	92.71 ± 2.88	91.73 ± 1.70
	FC-HGNN(2025)	90.59 ± 0.09	95.24 ± 0.07	90.83 ± 0.14	90.30 ± 0.17	90.56 ± 0.07	96.79 ± 0.02	98.72 ± 0.01	95.84 ± 0.04	97.59 ± 0.02	96.71 ± 0.02
Ours	MF-HGNN(ours)	**94.72 ± 0.04**	**97.06 ± 0.03**	**96.19 ± 0.08**	**93.05 ± 0.05**	**94.59 ± 0.03**	**99.53 ± 0.00**	**99.89 ± 0.00**	**98.96 ± 0.02**	**100.00 ± 0.00**	**99.48 ± 0.00**

Table 4
Results of Shapiro–Wilk normality test.

Dataset	Total	Norm	Non-norm	Conformity (%)
ABIDE	120	115	5	95.8
MDD	120	102	18	85.0
Overall	240	217	23	90.4

Conformity rate (%) = (Normally Distributed Cases/Total Tests) ×100%.

Table 5
Wilcoxon signed-rank test on cross-validation results.

Type	Comparative model	Subject-wise cross-validation				Leave-group-out cross-validation			
		ABIDE		MDD		ABIDE		MDD	
		ACC Gain (p-val)	AUC Gain (p-val)	ACC Gain (p-val)	AUC Gain (p-val)	ACC Gain (p-val)	AUC Gain (p-val)	ACC Gain (p-val)	AUC Gain (p-val)
Ind.GCN	BrainGNN(2021)	32.45 (0.002, **)	41.99 (0.002, **)	45.19 (0.002, **)	57.89 (0.002, **)	36.37 (0.002, **)	39.76 (0.002, **)	37.87 (0.002, **)	63.99 (0.002, **)
	RGTNet(2024)	37.06 (0.002, **)	41.58 (0.002, **)	45.89 (0.002, **)	45.68 (0.002, **)	32.40 (0.002, **)	45.62 (0.002, **)	41.61 (0.002, **)	50.55 (0.002, **)
Pop.GCN	PopulationGCN(2017)	33.47 (0.002, **)	33.95 (0.002, **)	51.92 (0.002, **)	53.07 (0.002, **)	32.94 (0.002, **)	31.22 (0.002, **)	50.98 (0.002, **)	51.00 (0.002, **)
	MMGL(2022)	5.50 (0.0462, *)	8.30 (0.0059, **)	17.27 (0.002, **)	16.20 (0.002, **)	6.04 (0.0309, *)	9.02 (0.0039, **)	17.21 (0.002, **)	14.42 (0.002, **)
	LG-GNN(2023)	15.00 (0.002, **)	9.73 (0.002, **)	7.80 (0.002, **)	7.49 (0.002, **)	19.14 (0.002, **)	14.41 (0.002, **)	8.32 (0.002, **)	7.94 (0.002, **)
	FC-HGNN(2025)	2.78 (0.0098, **)	1.05 (0.002, **)	17.06 (0.002, **)	11.70 (0.002, **)	4.56 (0.0039, **)	1.92 (0.002, **)	2.82 (0.0039, **)	1.19 (0.002, **)

Performance gain (%) = (MF-HGNN's mean - Comparative model's mean)/Comparative model's mean ×100%.
Significance levels: * $p < 0.05$ (significant), ** $p < 0.01$ (highly significant), ns (non-significant, p≥0.05).

PopulationGCN [36]: An early and influential study in this domain.

MIDA [37]: Employs a domain adaptation method to minimize site-related feature variability by extracting second-order functional connectivity features for autism classification.

MMGL [21]: A comprehensive framework designed to learn multimodal features associated with specific diseases through modality-aware learning.

MAMF-GCN [22]: An attention-based multi-scale multi-channel fusion model integrating multimodal brain imaging with diverse preprocessing.

LG-GNN [24]: Includes a specialized functional connectivity extraction component, enabling effective learning from subjects' brain imaging data. It also considers subjects' phenotypic features during final classification.

FC-HGNN [53]: A heterogeneous graph neural network model based on brain functional connectivity that integrates imaging and non-imaging data in a two-stage process, significantly improving psychiatric disorder classification accuracy by leveraging gender-specific feature aggregation.

4.3. Visualization and analysis

To highlight the strengths of MF-HGNN, t-distributed stochastic neighbor embedding (t-SNE) [54] was utilized for dimensionality reduction and visualization of feature distributions before and after model processing on the ABIDE cohort. As illustrated in Fig. 4(a), the feature distributions of the two classes are intermixed without distinct separation. Fig. 4(b) to (f) display feature visualizations learned by various representative methods. From these, BrainGNN and PopulationGNN exhibit some clustering tendencies, but the features of different classes remain highly similar, resulting in many misclassifications and limited classification performance. Meanwhile, FC-HGNN and MMGL show more uniform distributions, but the distances between different classes are small, making effective separation difficult. Compared to these representative methods, our model MF-HGNN achieves excellent classification results: features of different classes are well separated, while features within the same class are closely clustered. This indicates the model effectively distinguishes between categories, substantially improving classification discriminability.

4.4. Classification performance comparison

To ensure the model's generalization capability, this study employs a subject-independent 10-fold cross-validation strategy. To prevent data leakage, all samples are partitioned at the individual subject level, ensuring that all data from a single subject are assigned exclusively to one of the 10 mutually exclusive and collectively exhaustive subsets. During training, nine subsets are used for model training, while the remaining one serves as the test set. This process is repeated 10 times, ensuring each fold is used exactly once as the independent test set, as

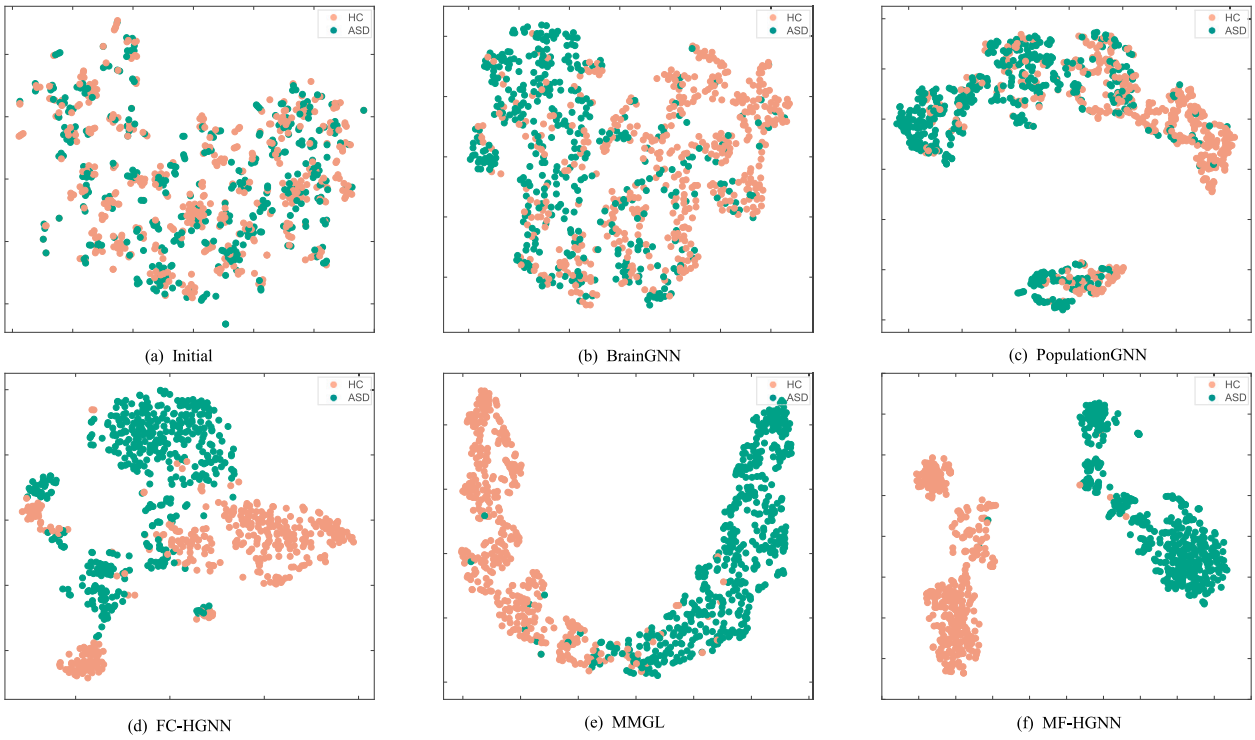


Fig. 4. Visualization results of different methods on the ABIDE dataset.

Table 6
Transductive learning results on ABIDE and MDD datasets.

Type	Method	ABIDE		MDD	
		ACC	AUC	ACC	AUC
Ind.GCN	BrainGNN(2021)	66.42 ± 4.77	68.32 ± 5.28	63.39 ± 5.61	59.75 ± 7.93
	RGTNet(2024)	65.04 ± 5.03	66.92 ± 5.44	65.85 ± 5.38	63.17 ± 5.13
Pop.GCN	PopulationGCN(2017)	67.81 ± 4.13	66.32 ± 3.64	61.27 ± 2.85	62.68 ± 6.42
	MMGL(2022)	85.68 ± 3.16	89.86 ± 2.91	80.63 ± 2.26	82.26 ± 3.77
	LG-GNN(2023)	76.95 ± 4.23	82.63 ± 4.34	86.68 ± 1.74	87.91 ± 1.86
	FC-HGNN(2025)	85.87 ± 0.28	86.54 ± 0.69	81.14 ± 0.08	85.56 ± 0.10
	MF-HGNN(ours)	**88.41 ± 0.14**	**95.87 ± 0.04**	**95.46 ± 0.02**	**96.89 ± 0.01**

shown in Table 2. Results are reported as mean ± standard deviation, with the optimal performance metric highlighted in bold. Based on the experimental findings, we draw the following key conclusions:

1. Traditional machine learning models, which depend on handcrafted features, are often surpassed by deep learning approaches. This is because conventional models grapple with the high-dimensionality and complexity of rs-fMRI data, leading to limited efficacy in binary classification.

2. Many individual graph models match BrainNetCNN. Graph-based analysis aids biomarker exploration from rs-fMRI. BrainGNN, using ROI-aware graph convolutions, achieves top performance in this group. Population graph models generally surpass individual ones by modeling subject relationships via imaging and non-imaging data (e.g., gender, site). Multimodal models like MAMF-GCN and MMGL perform better than unimodal ones.

3. This section evaluates major population-level graph methods, including PopulationGCN, MMGL, MIDA, MAMF-GCN, LG-GNN, and FC-HGNN. On the ABIDE dataset, the proposed method achieves 23.43% and 23.24% higher accuracy than PopulationGCN and MIDA, respectively, with significant improvements across all evaluation metrics, including specificity and F1 score. This improvement is primarily attributed to MF-HGNN's precise modeling of brain temporal and topological features, along with the integration of multiple phenotypic data sources. Specifically, the implementation of site-heterogeneous neural networks enables individual-level classification, significantly enhancing the model's ability to capture comprehensive representations.

4. MF-HGNN achieved the highest accuracy and AUC on both the ABIDE and REST-meta-MDD datasets. On ABIDE, it reached 95.07% accuracy and 97.36% AUC, while on REST-meta-MDD, it achieved 99.16% accuracy and 99.86% AUC. Its performance even surpassed that of MMGL, which leverages multiple modalities, and MAMF-GCN, which integrates various neuroimaging data types. This demonstrates the effectiveness of our method in the first stage, where the extracted dynamic and static brain information can generate robust features for downstream classification tasks. Furthermore, in the second stage, the CGAC layer fully captures both shared and complementary information from intra-site and cross-site neighbors, significantly enhancing the model's ability to learn holistic representations.

5. Some methods offer important biomarkers that enhance clinical interpretability. While approaches including FC-HGNN and MAMF-GCN demonstrate strong predictive performance, these models do not simultaneously optimize prediction accuracy and biomarker identification. In contrast, only the MAMF-GCN model achieves a breakthrough in ASD classification performance while providing new directions for biomarker discovery and clinical diagnosis in psychiatric disorders.

In addition to individual-level 10-fold cross-validation, to further mitigate potential data leakage risks, we also employed a leave-group-out cross-validation strategy stratified by diagnostic labels and data

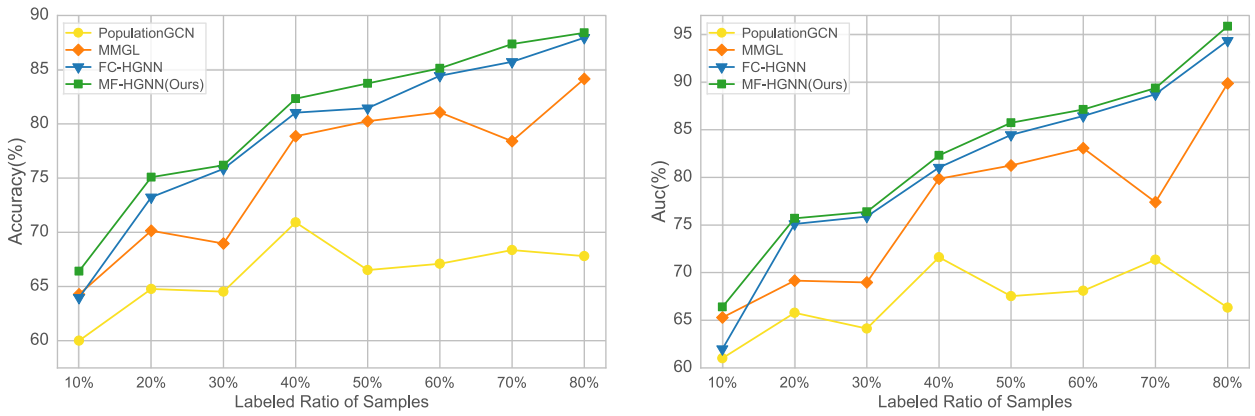


Fig. 5. Classification results on ABIDE dataset with varying training sample proportions.

acquisition sites. Specifically, stratified sampling based on the joint distribution of diagnoses and sites was performed to construct 10 disjoint sample groups, thereby evaluating the model's generalization capability under more stringent independence conditions.

The experimental results on the ABIDE and REST-meta-MDD datasets under this validation strategy are summarized in Table 3. The results demonstrate that, within the leave-group-out cross-validation framework, MF-HGNN consistently achieves the best performance among all compared methods. On the ABIDE dataset, MF-HGNN attained an accuracy of 94.72% and an AUC of 97.06%, comprehensively outperforming PopulationGCN, MMGL, LG-GNN, and FC-HGNN. On the REST-meta-MDD dataset, the model similarly maintained an accuracy and AUC above 95%, while the performance of other comparative methods exhibited a more pronounced decline. Notably, FC-HGNN demonstrates significantly improved performance on this dataset compared to individual-level cross-validation, which can be attributed to the stratified leave-group-out partitioning strategy that preserves site distribution structure and results in a more balanced distribution of phenotypic features across groups. Meanwhile, the cross-graph attention convolution module effectively mitigates site heterogeneity by modeling intra-site and inter-site feature interactions, while the relatively small technical differences across sites further enhance cross-group prediction performance. Taken together, these findings indicate that the superiority of MF-HGNN is not derived from a specific data partitioning scheme, but rather reflects its strong generalization capability across different acquisition sites and subject groups, while also validating the soundness of our experimental design.

To conduct rigorous statistical comparisons of model performance, we first employed the Shapiro–Wilk test to assess the normality of performance metrics under the two data partitioning strategies. As shown in Table 4, the normality compliance rate was 95.8% for the ABIDE dataset but only 85.0% for the MDD dataset, indicating that the normality assumption required for parametric tests was not fully satisfied. Therefore, we employed the non-parametric Wilcoxon signed-rank test to perform statistical significance tests on core metrics comparing MF-HGNN with mainstream models under both validation settings. As presented in Table 5, the Wilcoxon signed-rank test p-values for MF-HGNN compared to all baseline models were below 0.05 under both data partitioning strategies, indicating that the performance differences between models are statistically significant. The majority of these comparisons yielded p-values of 0.002 or 0.0039, substantially below the 0.01 significance threshold. Only in the two comparisons with MMGL on the ACC metric for the ABIDE dataset did the p-values reach 0.0462 and 0.0309, respectively. Although slightly above 0.01, these values remain within the 0.05 significance level.

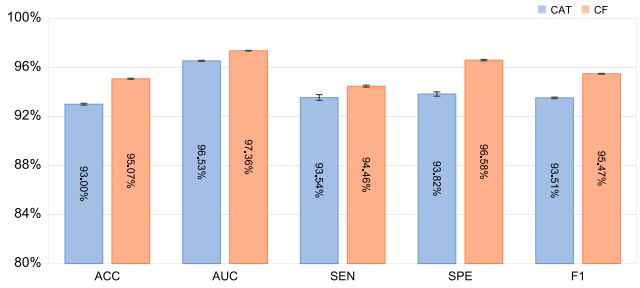


Fig. 6. Comparison between Transformer-based cross-fusion strategy (orange) and direct feature concatenation strategy (blue).

4.5. Assessment of transductive learning

The study enhanced the usual 10-fold cross-validation by splitting the dataset into three separate parts for evaluating the model's transductive learning and generalization performance. The data were allocated as 80% training, 10% validation, and 10% testing subsets, with all subsets mutually exclusive. This procedure was repeated 10 times.

The proposed approach exceeds competing methods across the two cohorts, as demonstrated in Table 6. In the transductive learning setting, for classification between ND and HC, population graph models generally outperform individual graph models. The MF-HGNN model attains an accuracy of 88.41% in differentiating ASD patients from healthy controls, surpassing models like PopulationGCN, MMGL, and FC-HGNN. Only MF-HGNN maintains accuracy above 95% within the MDD cohort, while alternative techniques exhibit a notable decline in effectiveness, highlighting the robust generalizability of the proposed framework.

To further evaluate the model's generalization ability, we increased the proportion of labeled training data from 10% to 80% of the total patient count. We conducted evaluations under the previously described transductive learning framework. PopulationGCN, MMGL, FC-HGNN, and MF-HGNN models' experimental results using varying proportions of training data on the ABIDE dataset are presented in Fig. 5. As shown, the accuracy (ACC) and AUC of all models improve as training data increases. Notably, when the labeled training data increases from 10% to 20%, all methods achieve significant improvements. Even with only 10% training data, MF-HGNN still attains an ACC of 67.34%, outperforming all methods. This clearly demonstrates that our method has stronger generalization ability and is better suited for semi-supervised node classification tasks on the ABIDE dataset.

Table 7
Evaluation of ablation studies using the ABIDE dataset.

	ACC	AUC	SEN	SPE	F1
w/o MVFL	93.46 ± 0.07	97.15 ± 0.03	93.46 ± 0.19	94.67 ± 0.07	93.98 ± 0.06
w/o NDP	93.69 ± 0.09	96.71 ± 0.07	93.88 ± 0.22	94.67 ± 0.07	94.19 ± 0.07
w/o Heterogeneous GNN	91.50 ± 0.07	95.77 ± 0.04	91.10 ± 0.19	93.60 ± 0.10	92.24 ± 0.05
Only Topology View	88.63 ± 0.08	94.62 ± 0.03	87.91 ± 0.10	91.68 ± 0.36	89.59 ± 0.08
Only Time-Varying View	92.77 ± 0.13	96.41 ± 0.05	92.69 ± 0.24	94.23 ± 0.11	93.36 ± 0.10
Only SAGPooling	93.33 ± 0.06	96.29 ± 0.03	93.21 ± 0.14	95.15 ± 0.06	94.16 ± 0.05
Only CGAC	93.46 ± 0.05	96.76 ± 0.05	93.56 ± 0.18	94.67 ± 0.20	93.95 ± 0.04
MF-HGNN	**95.07 ± 0.03**	**97.36 ± 0.02**	**94.46 ± 0.08**	**96.58 ± 0.06**	**95.47 ± 0.03**

Table 8
The results of different node scoring methods and pooling strategies.

Methods	ACC	AUC	SEN	SPE	F1
Score in GCN	93.39	96.45	95.36	91.38	94.27
Score in GAT	91.62	97.15	92.18	92.32	92.21
TopkPooling	92.65	94.02	92.51	94.02	93.24
Diffpooling	92.33	94.32	93.05	93.36	93.02
MF-HGNN	**95.07**	**97.36**	**94.46**	**96.58**	**95.47**

4.6. Ablation experiment

Ablation studies were conducted on the ABIDE dataset to evaluate the effectiveness of the proposed Multi- View Feature Encoding module, NDP, and individual-feature-based heterogeneous graph strategies. As shown in Table 7, model performance significantly deteriorates when site heterogeneity is ignored, with ACC and the AUC decreasing by 3.57% and 1.59%, respectively. This indicates that feature variations caused by different sites substantially impact disease diagnosis, while the proposed cross-graph attention convolution method effectively extracts information from both intra-site and cross-site data. Furthermore, replacing both the topological view and the time-varying view in the MVFL module with standard convolutional layers resulted in decreases of 1.61% and 0.21% in ACC and AUC, respectively. This demonstrates that the MVFL module possesses distinct advantages in capturing complex information from diverse perspectives. Through its bidirectional cross-attention mechanism, it effectively extracts shared information, enabling the learning of more discriminative feature representations for each individual subject.

We also designed a hierarchical pooling method, and its removal led to a 1.58% decrease in ACC and a 0.65% decrease in AUC. This indicates that the proposed pooling strategy effectively suppresses noise that may interfere with downstream tasks while preserving both short-range and long-range critical nodes and edges. Overall, each module makes a substantial contribution to the final diagnostic performance, and integrating all modules yields the optimal results.

Additionally, we conducted independent analyses of each view within the MVFL module. As shown in Table 7, using only the topological view or the time-varying view resulted in ACC decreases of 6.44% and 2.3%, respectively, compared to the full model, indicating that both views make significant contributions to performance. We also investigated the effectiveness of directly concatenating the two views for feature fusion. As shown in Fig. 6, this led to reductions in ACC and AUC by 2.07% and 0.83%, respectively. This demonstrates that our proposed fusion method effectively captures common features across perspectives while reducing inter-view feature discrepancies. The performance gain stems from the integration of dynamic temporal information with static topological structures, thereby enabling the generation of high-quality individualized features for population-based graph modeling.

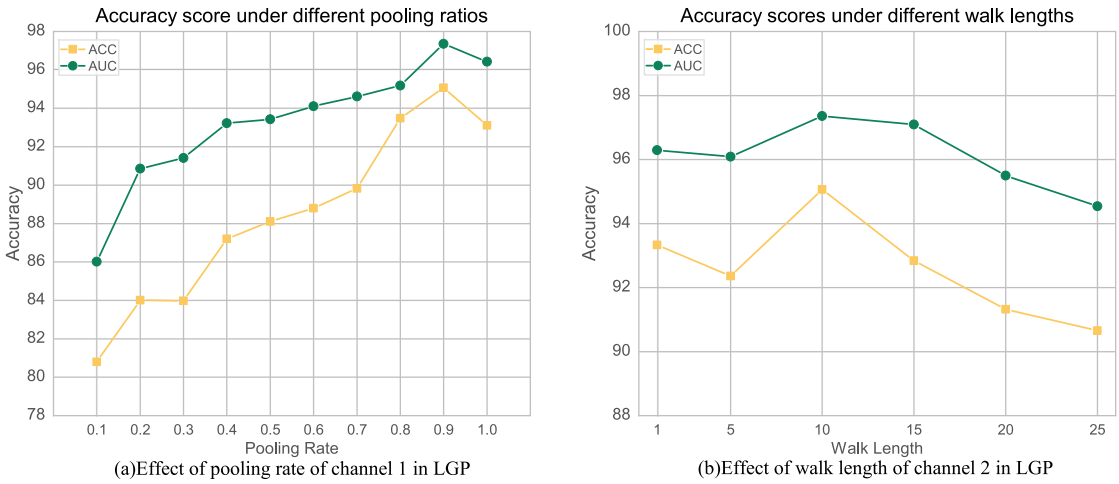
As shown in Table 7, when only the first channel of the NDP module is retained, the ACC drops to 93.33% and the AUC decreases to 96.29%. This indicates that relying solely on pooling without extracting long-range dependencies results in suboptimal model performance. Additionally, we evaluated a heterogeneous group graph convolution module containing only CGAC. Compared to the full model, ACC and AUC decrease by 1.61% and 0.6%, respectively, indicating a significant degradation in classification performance in the absence of global information aggregation.

4.7. Node score and pooling technology in NDP

We conducted experiments on different node scoring methods for random walks in the NDP module. As shown in Table 8, our method achieved the best results among the three tested scoring methods. When using GCN to compute node scores, the model performance fell short of expectations, likely because GCN relies too heavily on neighborhood features and has limited capability in modeling long-range dependencies. Surprisingly, the attention-based GAT [55] performed the worst, achieving only 91.62% accuracy. This may be attributed to GAT's high sensitivity to sample features and hyperparameters, leading to significant fluctuations during training and difficulty in convergence. Furthermore, we investigated the impact of different pooling strategies on classification performance. Compared to SAGPooling [56], DiffPooling [57] performed worse, likely due to noise amplification during hierarchical pooling. Meanwhile, TopKPooling [58], which relies on a fixed node selection strategy, may fail to adequately preserve crucial local information. Overall, employing SAGPooling effectively captures local topological correlations to identify important neighboring nodes, thereby enhancing the model's disease classification performance.

4.8. Fusion technology in heterogeneous GNN

To evaluate the impact of different graph convolution techniques on model performance, we replaced the TransformerConv layer in our model with several graph convolutional operators, including GCNConv [59], ChebConv [60], GATConv [55], and SAGEConv [61], while keeping all other experimental settings unchanged. The results presented in Table 9 show that GATConv yielded the poorest performance, achieving only 84.94% ACC and 90.27% AUC. This may be attributed to its high sensitivity to phenotypic noise, which leads to unstable training. Both ChebConv and GCNConv performed worse than SAGEConv, possibly because they fail to effectively capture global dependencies. The proposed Graph Transformer method achieved the best overall performance among all approaches. This advantage likely arises from its ability to dynamically model long-range dependencies and accurately distinguish the semantic differences between intra-site and inter-site edges through attention weights. Among all compared methods, SAGEConv ranked second only to the Graph Transformer, achieving an ACC of 94.47% and an AUC of 96.77%. This experiment confirms the necessity of introducing a dynamic global weighting mechanism into heterogeneous graph convolutions and provides an optimal solution for effective brain network feature extraction.

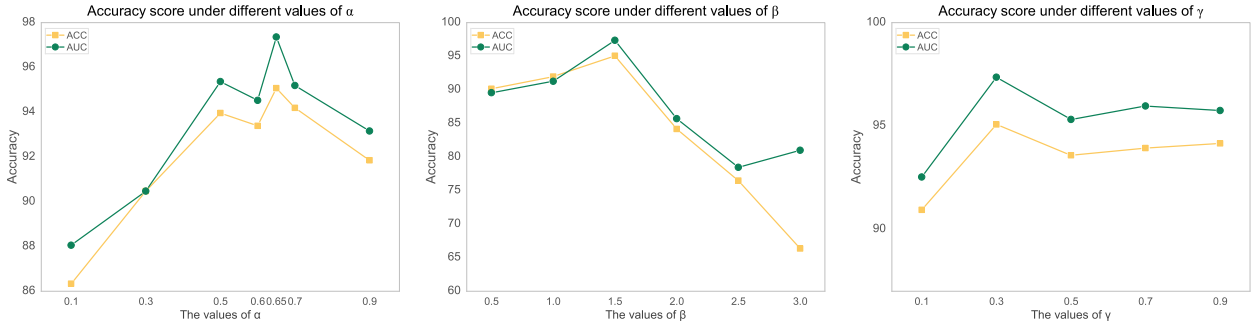


(a)Effect of pooling rate of channel 1 in LGP (b)Effect of walk length of channel 2 in LGP

Fig. 7. Experimental results for hyperparameter sensitivity in the model on the ABIDE dataset.

Table 9
Classification results of different convolution methods in the SHG.

Methods	ACC	AUC	SEN	SPE	F1
MF-HGNN(Ours)	**95.07 ± 0.03**	**97.36 ± 0.02**	**94.46 ± 0.08**	**96.58 ± 0.06**	**95.47 ± 0.03**
GCNConv	91.83 ± 0.03	95.14 ± 0.08	91.91 ± 0.23	94.16 ± 0.32	92.81 ± 0.08
ChebConv	90.59 ± 0.04	94.09 ± 0.03	92.44 ± 0.09	89.98 ± 0.15	91.11 ± 0.04
GATConv	84.94 ± 0.21	90.27 ± 0.61	87.24 ± 0.30	85.68 ± 1.03	85.94 ± 0.26
SAGEConv	94.47 ± 0.12	96.77 ± 0.05	94.09 ± 0.21	96.74 ± 0.16	95.29 ± 0.10



(a)Effect of threshold α on the experiment (b)Effect of threshold β on the experiment (c)Effect of threshold γ on the experiment

Fig. 8. Experimental results at different α, β, and γ values on ABIDE and MDD data.

Table 10
Results of computational efficiency comparison.

Methods	Params (M)	Time/epoch (s)	Inference (ms)
BrainGNN	0.066	0.966	0.315
RGTNet	0.149	1.515	0.622
LG-GNN	0.152	8.003	6.776
MMGL	0.546	0.229	0.011
FC-HGNN	0.351	24.191	12.485
PopGCN	24.016	2.561	1.658
MF-HGNN	20.739	34.900	23.793

4.9. Parameter sensitivity analysis

To investigate the impact of key hyperparameters, we conducted a sensitivity analysis on the pooling ratio of the NDP module to evaluate its effect on model performance. As shown in Fig. 7(a), the NDP module significantly improved model performance. We further examined the influence of the random walk distance in the second channel of the pooling module. Experimental results (see Fig. 7(b)) indicate that the random walk distance affects model performance: excessively long walks capture numerous disease-irrelevant nodes, leading to decreased accuracy. Moreover, increasing the number of discarded nodes has a more pronounced negative impact on the model. Additionally, we conducted an experimental analysis on the edge retention thresholds for both the brain graph convolution and the two views in the group graph, as illustrated in Fig. 8. Results reveal that setting the threshold too high or too low degrades performance: excessively high thresholds introduce substantial noise, while excessively low thresholds omit critical data, ultimately impairing the model's ability to effectively learn brain graph features.

4.10. Impact of learning rate

Although a learning rate scheduler was used during training, the initial learning rate still significantly impacted model performance. We explored how different initial learning rates affect the model's performance, as shown in Fig. 9.

When the initial learning rate was set to 0.01, the model achieved the best performance, especially with notable improvements in specificity and F1 score. This may be because, at this learning rate, the model

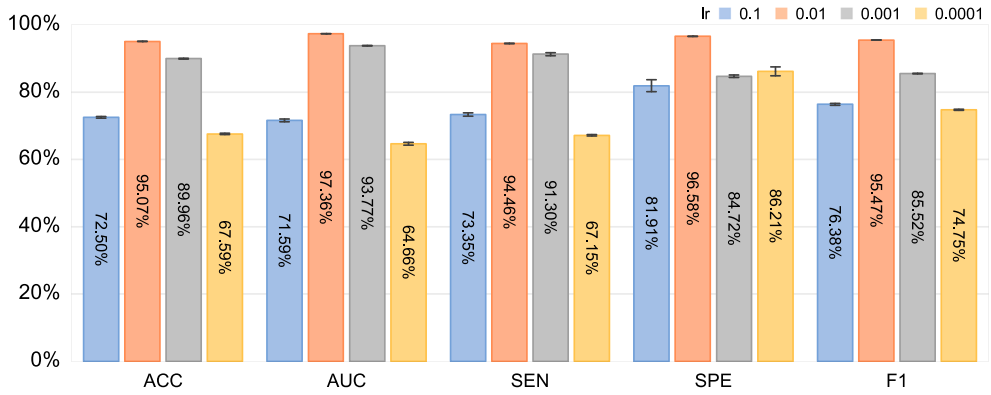


Fig. 9. The impact of different initial learning rates on model performance on the ABIDE dataset.

better fits the spatiotemporal features of the brain network without skipping over optimal parameter regions due to too large steps or getting stuck in local minima due to too small steps. However, learning rates that are too high or too low caused a sharp decline in overall performance, indicating that an appropriate learning rate is essential for ensuring model stability.

4.11. Computational efficiency analysis

To evaluate the clinical utility of MF-HGNN, we conducted a comparative analysis with mainstream methods across three dimensions: model parameter size, training time per epoch, and inference speed per sample. The results are presented in Table 10.

BrainGNN and RGTNet exhibit the smallest parameter count and optimal speed, but fail to integrate multi-site heterogeneous information. MMGL achieves the fastest inference through optimized multimodal fusion. FC-HGNN has 0.351 M parameters, with its sex-specific aggregation module introducing additional overhead, resulting in a training time of 24.191 s per epoch. PopGCN has 24.016 M parameters, based on spectral graph convolution and sparse graph construction, with a training time of 2.561 s per epoch. Our MF-HGNN has 20.739 M parameters, lower than PopGCN, with a training time of 34.900 s per epoch and an inference time of 23.793 ms per sample—slightly higher than FC-HGNN. This is primarily due to the MVFL module, which models dynamic and static brain network features from two distinct views. Notably, MF-HGNN's parameter count and inference speed remain suitable for clinical deployment, adapting to real-time diagnostic scenarios. With approximately 20 M parameters, it can be directly deployed on various edge computing devices, demonstrating its feasibility and practical value for clinical implementation.

5. Discussion

5.1. Explainability investigation

To identify the most discriminative biomarkers for disease classification, we specifically analyzed the random walk node scores from the NDP module. Table 11 details the top 10 brain regions selected based on their highest scores. Additionally, we visualized these important regions using BrainNet Viewer [62], as illustrated in Fig. 10 and Fig. 11.

In the ASD classification task, experimental results indicate that the high-scoring brain regions are primarily located in the temporal lobe, amygdala, and parahippocampal areas [63,64]. These regions are mainly associated with social, emotional processing, and memory functions. Notably, abnormalities in frontal lobe circuits have been explicitly linked to social deficits in ASD [65]. The orbitofrontal cortex significantly correlates with default mode network connectivity abnormalities, which may affect motor-related functions [66]. These

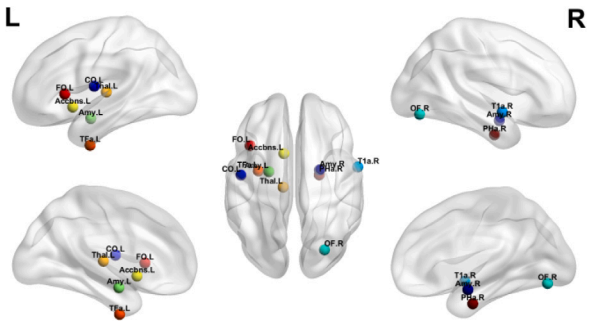


Fig. 10. The top-10 important brain regions for ASD categorization.

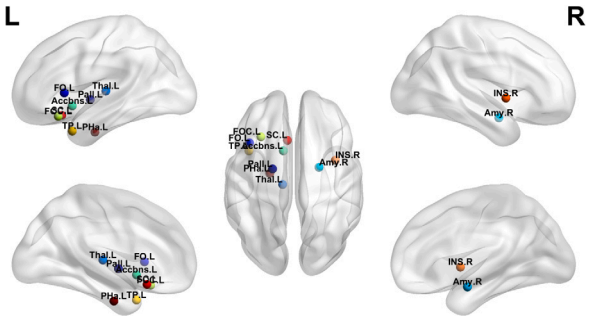


Fig. 11. The top-10 important brain regions for MDD categorization.

regions are highly relevant to language processing, social cognition, motor execution and regulation, and behavioral expression [67,68]. Therefore, these findings demonstrate the reliability of our method in detecting ASD.

In the MDD classification task, the high-scoring regions are mainly located in the frontal cortex, amygdala, and thalamus, all of which are key areas involved in emotion regulation. Studies show that MDD patients exhibit impairments in reward circuitry, leading to difficulties in emotional regulation [69]. The amygdala, related to anxiety and emotional dysregulation, is typically hyperactive in MDD [70]. The cingulate gyrus, a default mode network core node associated with self-referential thinking and emotion regulation, exhibits abnormal connectivity with the prefrontal cortex and hippocampal regions in MDD patients [71]. The parahippocampal gyrus, involved in memory and emotional processing, may indirectly participate in pathological processes through structural and functional abnormalities [72]. The brain regions identified by our method highly overlap with those implicated in depression, indicating that our approach is also practical for MDD diagnosis.

Table 11
Specific ROIs of particular interest in both datasets (HO atlas).

ABIDE		MDD	
ROIs_index	ROIs_abbr	ROIs_index	ROIs_abbr
13	Amy.R	105	Pall.L
6	Amy.L	111	Accbns.L
48	PHa.R	67	PHa.L
22	T1a.R	99	Thal.L
95	FO.L	53	SC.L
104	CO.L	81	FO.L
99	TFa.L	4	INS.R
1	Thal.L	15	TP.L
54	OF.R	110	Amy.R
7	Amybns.L	65	FOC.L

Table 12
Evaluation results of AAL-based parcellation and node feature inputs on the ABIDE dataset.

Methods	ACC	AUC	SEN	SPE	F1
MF-HGNN	**95.07**	**97.36**	**94.46**	**96.58**	**95.47**
With AAL Atlas	93.57	97.05	93.72	94.45	94.05
With Time Series	89.46	92.83	88.81	91.23	90.56

5.2. Selection of brain atlases

We also explored the impact of different brain atlases and graph node features on model classification performance. We used data processed with the Automated Anatomical Labeling (AAL) atlas [73] as node and edge features for the graph. Additionally, we employed the original time series as node features under the HO atlas. Since sample acquisition times vary across different sites, we standardized using the first 78 ms of data from each ROI as the node feature. The experimental results are shown in Table 12.

The performance with the AAL atlas slightly decreased, possibly because although it provides a more precise brain region segmentation, it does not match the critical subregions and the HO atlas, resulting in only a small performance drop. When using the raw time series as node features, performance dropped significantly. This is likely due to variations in equipment and standards across sites, causing substantial differences in time series lengths, which prevented effective utilization of the full time series information and significantly reduced feature effectiveness.

5.3. Selection of phenotype data

Different datasets contain various types of non-imaging data. To verify which phenotypic information positively influences model performance, we conducted experiments combining different data types. The results are shown in Table 13.

The differences in disease mechanisms and population distributions may affect model classification. For example, gender information significantly improves classification performance on the ABIDE dataset, whereas in the MDD dataset, including gender information leads to a decline in performance. This may be due to differing disease mechanisms between genders. In the ABIDE dataset, adding age caused a noticeable performance drop; age features failed to provide helpful information and interfered with core patterns related to gender and site, indicating that more features do not necessarily improve the model.

Even when using only site information in both datasets, there was still a clear discriminative power, demonstrating the significant impact of site heterogeneity on performance in multi-center studies.

5.4. Limitation and future work

Although multiple datasets were employed in this study for psychiatric disorder diagnosis, there remains considerable room to expand the comprehensiveness of data sources and the diversity of data types. Therefore, enlarging the scale of high-quality datasets will be a key focus of our future research.

Meanwhile, the current model contains a large number of parameters, leading to extended training time. While its inference speed meets clinical offline analysis requirements, real-time diagnostic scenarios demand higher efficiency. Therefore, optimizing the model for lightweight deployment — through techniques such as knowledge distillation, sparse graph convolutions, and parameter pruning — will be a key direction in future work to reduce computational costs while maintaining high classification accuracy, thereby enhancing the model's deployment efficiency in resource-constrained clinical settings.

In addition, many hyperparameters in the current model still rely on manual tuning, which to some extent limits the model's applicability. In the future, we plan to explore adaptive network architectures driven by intelligent algorithms to achieve automatic hyperparameter optimization, thereby reducing manual intervention and enhancing the model's capacity for automated deployment.

Finally, given that the current model incorporates relatively limited input modalities, we plan to further integrate heterogeneous data types — including imaging, physiological signals, and clinical text — and employ cross-modal feature fusion techniques to expand the effective feature space, thereby improving the reliability and robustness of disease prediction.

6. Conclusion

This paper proposes a multi-view fusion rs-fMRI heterogeneous graph model, MF-HGNN, which serves as an end-to-end disease prediction framework based on multi-view heterogeneous graph learning. The model organically integrates dynamic spatiotemporal features with static topological features through a multi-view fusion strategy, capturing both the instantaneous interaction patterns of brain networks and their stable modular structures. In addition, the proposed NDP method effectively identifies relevant biomarkers and significantly enhances overall model performance. Finally, the population graph is constructed by incorporating both within-site and cross-site relationships, which effectively reduces the impact of site heterogeneity on classification results. Compared with existing methods, MF-HGNN achieves superior predictive performance on multiple public datasets. Experimental results indicate that the model has great potential to improve clinical diagnostic accuracy and support clinical decision-making. Moreover,

Table 13
Experimental results based on combinations of phenotypic data from the ABIDE and MDD datasets.

ABIDE			MDD		
Phenotypic information	ACC	AUC	Phenotypic information	ACC	AUC
Site	94.20	96.95	Site	75.74	81.52
Site + Sex	**95.07**	**97.36**	Site + Age	**99.17**	**99.87**
Site + Age	92.13	96.18	Site + Age + Sex	97.38	98.85
Site + Sex + Age	93.78	95.23	Site + Age + Edu	98.10	99.27
Site + Sex + Fiq	91.14	94.95			

the key biomarkers identified by MF-HGNN, including the amygdala, parahippocampal gyrus, and frontal cortex, are consistent with findings from existing psychiatric disorder studies, further validating the reliability of the model.

In the future, we will focus on enhancing the model's adaptability to real-world clinical scenarios by investigating few-shot and zero-shot learning techniques to address the scarcity of specific psychiatric disorder samples in clinical settings. Additionally, we plan to design a lightweight, highly adaptive application to provide clinicians with more convenient diagnostic support tools.

CRediT authorship contribution statement

Yuhang Liu: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Rongbing Huang:** Writing – review & editing, Validation. **Chang Liu:** Writing – review & editing, Validation. **Jianwei Zhang:** Supervision, Funding acquisition. **Fasheng Yi:** Resources, Investigation. **Junchuan Chen:** Resources. **Hechun Li:** Validation. **Sisi Jiang:** Validation. **Cheng Luo:** Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the [National Key R&D Program of China] (Grant number 2024YFE0215100), the [Natural Science Foundation of Sichuan Province, China] (Grant number 24NSFSC2722), the [National Natural Science Foundation of China] (Grant number 62406043), the [Open Project Fund of the Key Laboratory of Digital Innovation of Tianfu Culture, Sichuan Provincial Department of Culture and Tourism], and the [Transformation Foundation of Tianfu Jincheng Laboratory] (Grant number 2025ZH013).

Data availability

Data will be made available on request.

References

[1] D. Vigo, G. Thornicroft, R. Atun, Estimating the true global burden of mental illness, Lancet Psychiatry 3 (2) (2016) 171–178.
[2] Y.E. Tian, M.A. Di Biase, P.E. Mosley, et al., Evaluation of brain-body health in individuals with common neuropsychiatric disorders, JAMA Psychiatry 80 (6) (2023) 567–576.
[3] G.I. Papakostas, Managing partial response or nonresponse: switching, augmentation, and combination strategies for major depressive disorder, J. Clin. Psychiatry 70 (suppl 6) (2009) 16–25.
[4] F.V. Farahani, W. Karwowski, N.R. Lighthall, Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review, Front. Neurosci. 13 (2019) 585.
[5] P.M. Matthews, A. Hampshire, Clinical concepts emerging from fMRI functional connectomics, Neuron 91 (3) (2016) 511–528.
[6] M. Liu, B. Li, D. Hu, Autism spectrum disorder studies using fMRI data and machine learning: a review, Front. Neurosci. 15 (2021) 697870.
[7] Y. Du, Z. Fu, V.D. Calhoun, Classification and prediction of brain disorders using functional connectivity: promising but challenging, Front. Neurosci. 12 (2018) 525.
[8] H. Pang, Z. Yu, H. Yu, J. Cao, Y. Li, M. Guo, C. Cao, G. Fan, Use of machine learning method on automatic classification of motor subtype of Parkinson's disease based on multilevel indices of rs-fMRI, Park. Relat. Disord. 90 (2021) 65–72.
[9] M. Mousavian, J. Chen, Z. Traylor, S. Greening, Depression detection from sMRI and rs-fMRI images using machine learning, Journal of Intelligent Information Systems 57 (2021) 395–418.
[10] W. Yin, L. Li, F.-X. Wu, Deep learning for brain disorder diagnosis based on fMRI images, Neurocomputing 469 (2022) 332–345.
[11] J. Kawahara, C.J. Brown, S.P. Miller, B.G. Booth, V. Chau, R.E. Grunau, J.G. Zwicker, G. Hamarneh, BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment, NeuroImage 146 (2017) 1038–1049.
[12] J.H. Oh, D.J. Lee, C.H. Ji, D.H. Shin, J.W. Han, Y.H. Son, T.E. Kam, Graph-based conditional generative adversarial networks for major depressive disorder diagnosis with synthetic functional brain network generation, IEEE J. Biomed. Health Inf. 28 (3) (2024) 1504–1515.
[13] D.J. Lee, D.H. Shin, Y.H. Son, J.W. Han, J.H. Oh, D.H. Kim, J.H. Jeong, T.E. Kam, Spectral graph neural network-based multi-atlas brain network fusion for major depressive disorder diagnosis, IEEE J. Biomed. Health Inf. 28 (5) (2024) 2967–2978, http://dx.doi.org/10.1109/JBHI.2024.3366662.
[14] G. Wen, P. Cao, H. Bao, W. Yang, T. Zheng, O. Zaiane, MVS-GCN: A prior brain structure learning-guided multi-view graph convolution network for autism spectrum disorder diagnosis, Comput. Biol. Med. 142 (2022) 105239.
[15] K. Zhao, B. Duka, H. Xie, et al., A dynamic graph convolutional neural network framework reveals new insights into connectome dysfunctions in ADHD, Neuroimage 246 (2022) 118774.
[16] F. Noman, C.M. Ting, H. Kang, et al., Graph autoencoders for embedding learning in brain networks and major depressive disorder identification, IEEE J. Biomed. Health Inf. 28 (3) (2024) 1644–1655.
[17] S. Liu, S. Wang, C. Sun, B. Li, S. Wang, F. Li, DeepGCN based on variable multi-graph and multimodal data for ASD diagnosis, CAAI Trans. Intell. Technol. 9 (4) (2024) 879–893, http://dx.doi.org/10.1049/cit2.12340.
[18] Y. Huang, A.C. 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, vol. 23, Springer, 2020, pp. 562–572.
[19] Z. Rakhimberdina, T. Murata, Linear graph convolutional model for diagnosing brain disorders, in: Complex Networks and their Applications VIII: Volume 2 Proceedings of the Eighth International Conference on Complex Networks and their Applications COMPLEX NETWORKS 2019, vol. 8, Springer, 2020, pp. 815–826.
[20] M. Cao, M. Yang, C. Qin, et al., Using DeepGCN to identify the autism spectrum disorder from multi-site resting-state data, Biomed. Signal Process. Control. 70 (2021) 103015.
[21] S. Zheng, Z. Zhu, Z. Liu, et al., Multi-modal graph learning for disease prediction, IEEE Trans. Med. Imaging 41 (9) (2022) 2207–2216.
[22] J. Pan, H. Lin, Y. Dong, Y. Wang, Y. Ji, MAMF-GCN: Multi-scale adaptive multi-channel fusion deep graph convolutional network for predicting mental disorder, Comput. Biol. Med. 148 (2022) 105823.
[23] H. Jiang, P. Cao, M. Xu, J. Yang, O. Zaiane, Hi-GCN: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction, Comput. Biol. Med. 127 (2020) 104096.
[24] H. Zhang, R. Song, L. Wang, L. Zhang, D. Wang, C. Wang, W. Zhang, Classification of brain disorders in rs-fMRI via local-to-global graph neural networks, IEEE Trans. Med. Imaging 42 (2) (2023) 444–455.
[25] S. Wang, M.E. Celebi, Y.D. Zhang, X. Yu, S. Lu, X. Yao, Q. Zhou, M. Martinez-Garcia, Y. Tian, J.M. Gorriz, et al., Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects, Inf. Fusion 76 (2021) 376–421.
[26] B. Sen, K.R. Cullen, K.K. Parhi, Classification of adolescent major depressive disorder via static and dynamic connectivity, IEEE J. Biomed. Health Inf. 25 (7) (2021) 2604–2614.
[27] A. Khazaee, A. Ebrahimzadeh, A. Babajani-Feremi, Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease, Brain Imaging and Behavior 10 (2016) 799–817.
[28] M. Plitt, K.A. Barnes, A. Martin, Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards, Neuroimage Clin. 7 (2015) 359–366.
[29] M. Zhao, W. Yan, R. Xu, J. Sui, et al., An attention-based hybrid deep learning framework integrating temporal coherence and dynamics for discriminating schizophrenia, in: 2021 IEEE 18th International Symposium on Biomedical Imaging, ISBI, IEEE, 2021, pp. 118–121.
[30] J. Zhang, L. Zhou, L. Wang, M. Liu, D. Shen, Diffusion kernel attention network for brain disorder classification, IEEE Trans. Med. Imaging 41 (10) (2022) 2814–2827.
[31] Y. Lei, S. Lin, Z. Li, et al., GNN-fused CapsNet with multi-head prediction for diabetic retinopathy grading, Eng. Appl. Artif. Intell. 133 (2024) 107994.
[32] J. Chang, T. Lai, L. Yang, et al., FV-DGNN: A distance-based graph neural network for finger vein recognition, IEEE Trans. Instrum. Meas. 72 (2023) 1–11.
[33] X. Li, Y. Zhou, N. Dvornek, M. Zhang, S. Gao, et al., Braingnn: Interpretable brain graph neural network for fMRI analysis, Med. Image Anal. 74 (2021) 102233.
[34] Y. Wang, H. Long, T. Bo, et al., Residual graph transformer for autism spectrum disorder prediction, Comput. Methods Programs Biomed. 247 (2024) 108065.

[35] S. Gadgil, Q. Zhao, A. Pfefferbaum, E.V. Sullivan, E. Adeli, K.M. Pohl, Spatio-temporal graph convolution for resting-state fMRI analysis, in: Proc. Int. Conf. Med. Image. Comput. Springer, 2020, pp. 528–538.
[36] S. Parisot, S.I. Ktena, E. Ferrante, et al., 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, vol. 20, Springer, 2017, pp. 177–185.
[37] M. Kunda, S. Zhou, G. Gong, et al., Improving multi-site autism classification via site-dependence minimization and second-order functional connectivity, IEEE Trans. Med. Imaging 42 (1) (2023) 55–65.
[38] A. Kazi, S. Shekarforoush, S.A. Krishna, H. Burwinkel, G. Vivar, K. Kortüm, S.A. Ahmadi, S. Albarqouni, N. Navab, Inceptiongcn: receptive field aware graph convolutional network for disease prediction, in: Inf. Process. Med. Imaging. Springer, 2019, pp. 73–85.
[39] B. Thapaliya, E. Akbas, J. Chen, et al., Brain networks and intelligence: A graph neural network based approach to resting state fMRI data, Med. Image Anal. 101 (2025) 103433.
[40] B.H. Kim, J.C. Ye, J.J. Kim, Learning dynamic graph representation of brain connectome with spatio-temporal attention, Adv. Neural Inf. Process. Syst. 34 (2021) 4314–4327.
[41] L. Kang, J. Chen, J. Huang, et al., Autism spectrum disorder recognition based on multi-view ensemble learning with multi-site fMRI, Cogn. Neurodynamics 17 (2) (2023) 345–355.
[42] A. Vaswani, N. Shazeer, N. Parmar, et al., Attention is all you need, Adv. Neural Inf. Process. Syst. 30 (2017).
[43] Y. Shi, Z. Huang, S. Feng, et al., Masked label prediction: Unified message passing model for semi-supervised classification, 2020, arXiv preprint arXiv:2009.03509.
[44] A. Di Martino, C.G. Yan, Q. Li, et al., The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism, Mol. Psychiatry 19 (6) (2014) 659–667.
[45] C.G. Yan, X. Chen, L. Li, et al., Reduced default mode network functional connectivity in patients with recurrent major depressive disorder, Proc. Natl. Acad. Sci. 116 (18) (2019) 9078–9083.
[46] C. Craddock, S. Sikka, et al., Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (C-PAC), Front. Neuroinform. Conference Abstract: Neuroinformatics 2013 (2013) http://dx.doi.org/10.3389/conf.fninf.2013.09.00042.
[47] R.S. Desikan, F. Ségonne, et al., An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, Neuroimage 31 (3) (2006) 968–980.
[48] C.G. Yan, X.D. Wang, X.N. Zuo, Y.F. Zang, DPABI: data processing & analysis for (resting-state) brain imaging, Neuroinformatics 14 (2016) 339–351.
[49] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, in: International Conference on Learning Representations, 2015, ICLR, San Diego, California.
[50] V. Vapnik, The support vector method of function estimation, in: Nonlinear Modeling: Advanced Black-Box Techniques, Springer, 1998, pp. 55–85.
[51] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[52] Y. Li, R. Chen, Y. Zhang, et al., A CNN-GCN framework for multi-label aerial image scene classification, in: IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2020, pp. 1353–1356.
[53] Y. Gu, S. Peng, Y. Li, et al., FC-HGNN: A heterogeneous graph neural network based on brain functional connectivity for mental disorder identification, Inf. Fusion 113 (2025) 102619.
[54] L. Van der Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.
[55] P. Veličković, G. Cucurull, A. Casanova, et al., Graph attention networks, 2017, arXiv preprint arXiv:1710.10903.
[56] J. Lee, I. Lee, J. Kang, Self-attention graph pooling, in: International Conference on Machine Learning, PMLR, 2019, pp. 3734–3743.
[57] Z. Ying, J. You, C. Morris, et al., Hierarchical graph representation learning with differentiable pooling, Adv. Neural Inf. Process. Syst. 31 (2018).
[58] C. Cangea, P. Veličković, N. Jovanović, et al., Towards sparse hierarchical graph classifiers, 2018, arXiv preprint arXiv:1811.01287.
[59] T.N. Kipf, Semi-supervised classification with graph convolutional networks, 2016, arXiv preprint arXiv:1609.02907.
[60] M. Defferrard, X. Bresson, P. Vandergheynst, Convolutional neural networks on graphs with fast localized spectral filtering, Adv. Neural Inf. Process. Syst. 29 (2016).
[61] W. Hamilton, Z. Ying, J. Leskovec, Inductive representation learning on large graphs, Adv. Neural Inf. Process. Syst. 30 (2017).
[62] M. Xia, J. Wang, Y. He, Brainnet viewer: a network visualization tool for human brain connectomics, PLoS One 8 (7) (2013) e68910.
[63] C.S. Monk, S.J. Peltier, J.L. Wiggins, et al., Abnormalities of intrinsic functional connectivity in autism spectrum disorders, Neuroimage 47 (2) (2009) 764–772.
[64] D. Li, H.O. Karnath, X. Xu, Candidate biomarkers in children with autism spectrum disorder: a review of MRI studies, Neurosci. Bull. 33 (2017) 219–237.
[65] M. Schuetze, M.T.M. Park, I.Y. Cho, F.P. MacMaster, M.M. Chakravarty, S.L. Bray, Morphological alterations in the thalamus, striatum, and pallidum in autism spectrum disorder, Neuropsychopharmacology 41 (11) (2016) 2627–2637.
[66] N. Yahata, J. Morimoto, R. Hashimoto, G. Lisi, K. Shibata, Y. Kawakubo, H. Kuwabara, M. Kuroda, T. Yamada, F. Megumi, et al., A small number of abnormal brain connections predicts adult autism spectrum disorder, Nat. Commun. 7 (1) (2016) 11254.
[67] D. Kim, J.Y. Lee, et al., Overconnectivity of the right Heschl's and inferior temporal gyrus correlates with symptom severity in preschoolers with autism spectrum disorder, Autism Res. 14 (11) (2021) 2314–2329.
[68] C. Press, N. Weiskopf, J.M. Kilner, Dissociable roles of human inferior frontal gyrus during action execution and observation, NeuroImage 60 (3) (2012) 1671–1677.
[69] C. Nauczyciel, S. Robic, T. Dondaine, et al., The nucleus accumbens: a target for deep brain stimulation in resistant major depressive disorder, J. Mol. Psychiatry 1 (2013) 17, http://dx.doi.org/10.1186/2049-9256-1-17.
[70] M. Bellani, M. Baiano, P. Brambilla, Brain anatomy of major depression II. Focus on amygdala, Epidemiol. Psychiatr. Sci. 20 (1) (2011) 33–36.
[71] M.S. Korgaonkar, A. Fornito, L.M. Williams, et al., Abnormal structural networks characterize major depressive disorder: a connectome analysis, Biol. Psychiatry 76 (7) (2014) 567–574.
[72] L. Cui, S. Li, S. Wang, et al., Major depressive disorder: hypothesis, mechanism, prevention and treatment, Signal Transduct. Target. Ther. 9 (1) (2024) 30.
[73] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Joliot, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, Neuroimage 15 (1) (2002) 273–289.