
NeuroH-TGL: Neuro-Heterogeneity Guided Temporal Graph Learning Strategy for Brain Disease Diagnosis

Shengrong Li¹, Qi Zhu^{1†}, Chunwei Tian², Xinyang Zhang¹, Wei Shao¹,
Jie Wen^{2†}, Daoqiang Zhang¹

¹Nanjing University of Aeronautics and Astronautics

²Harbin Institute of Technology

lisrong@nuaa.edu.cn, zhuqi@nuaa.edu.cn, chunweitian@hit.edu.cn,
xinyang@nuaa.edu.cn, shaowei20022005@nuaa.edu.cn, wenjie@hit.edu.cn,
dqzhang@nuaa.edu.cn

Abstract

Dynamic functional brain networks (DFBNs) are powerful tools in neuroscience research. Recent studies reveal that DFBNs contain heterogeneous neural nodes with more extensive connections and more drastic temporal changes, which play pivotal roles in coordinating the reorganization of the brain. Moreover, the spatio-temporal patterns of these nodes are modulated by the brain’s historical states. However, existing methods not only ignore the spatio-temporal heterogeneity of neural nodes, but also fail to effectively encode the temporal propagation mechanism of heterogeneous activities. These limitations hinder the deep exploration of spatio-temporal relationships within DFBNs, preventing the capture of abnormal neural heterogeneity caused by brain diseases. To address these challenges, this paper proposes a **Neuro-H**eterogeneity guided **T**emporal **G**raph **L**earning strategy (NeuroH-TGL). Specifically, we first develop a spatio-temporal pattern decoupling module to disentangle DFBNs into topological consistency networks and temporal trend networks that align with the brain’s operational mechanisms. Then, we introduce a heterogeneity mining module to identify pivotal heterogeneity nodes that drive brain reorganization from the two decoupled networks. Finally, we design temporal propagation graph convolution to simulate the influence of the historical states of heterogeneity nodes on the current topology, thereby flexibly extracting heterogeneous spatio-temporal information from the brain. Experiments show that our method surpasses several state-of-the-art methods, and can identify abnormal heterogeneous nodes caused by brain diseases.

1 Introduction

Functional magnetic resonance imaging (fMRI) measures neural activity by detecting changes in blood oxygen level-dependent signals, and is commonly employed to construct functional brain networks [1, 2, 3]. In fact, the brain is constantly reorganizing even during the resting state [4, 5]. Obviously, compared with the static functional brain network, the dynamic functional brain network (DFBN) can more comprehensively describe the topological evolution of the brain. Studies have shown that brain diseases such as Alzheimer’s disease and Parkinson’s disease can change the spatio-temporal properties of DFBNs [6, 7]. Therefore, effectively analyzing the spatio-temporal structure of DFBNs is crucial for brain disease diagnosis and biomarker mining.

To capture the time-varying structure of DFBNs, they are usually modeled as a series of dynamic brain graphs [8, 9, 10, 11]. In dynamic brain graphs, nodes represent brain regions and edges

[†]Corresponding authors.

represent temporal connections between these regions. Existing dynamic brain graph analysis methods usually use graph convolution network (GCN) [12] to extract the topological feature, and then use temporal convolution to capture the temporal correlation between brain regions [10, 13, 14]. For example, STAGIN [4] first uses GCN to extract structural information in DFBNs, and then introduces transformer to capture the temporal dependence between brain graphs. Although significant progress has been made in the analysis of dynamic brain graphs, most methods overlook the crucial fact that the brain exhibits significant spatio-temporal heterogeneity: Certain neural nodes in DFBNs exhibit extensive connectivity or more active temporal evolution due to their functional properties. For instance, the posterior cingulate cortex forms stable and tight connections with the frontal and parietal lobes, while the connection strength between the primary motor cortex and supplementary motor area exhibits heightened temporal variability [15, 16]. These neural nodes with high spatio-temporal variability can flexibly adjust the reconstruction pattern of functional network, which is a key factor driving the reorganization of the brain. Therefore, identifying the spatio-temporal heterogeneity of neural nodes is significant for elucidating the evolution mechanism of the brain.

However, accurately capturing the spatio-temporal heterogeneity of neural nodes faces dual challenges.

(1) Spatio-temporal coordination of DFBNs. While maintaining stable connections of pivotal nodes, the brain network can dynamically adjust connections according to cognitive demands to achieve efficient information integration. This spatial consistency and temporal trend together constitute the neural basis supporting complex cognitive functions [17, 18, 19]. **(2) Sequential dependence of DFBNs.** Due to the continuity of brain activity and the lag in information interaction [20, 21, 5], current connectivity patterns are systematically influenced by prior network states. There is a rich sequential dependence exhibited among neural nodes.

To address these challenges, we propose a **Neuro-Heterogeneity** guided **Temporal Graph Learning** strategy (NeuroH-TGL) to comprehensively capture the intrinsic evolution mechanism of DFBNs. Specifically, to simulate the spatio-temporal coordination of DFBNs, we design a spatio-temporal pattern decoupling (STPD) module to disentangle the DFBNs into topological consistency networks and temporal trend networks. Then, we calculate the cross-window similarity of topologic consistency networks and temporal trend networks, and use them as the spatial and temporal heterogeneity weights, respectively. Subsequently, we apply spatio-temporal heterogeneity weighting to DFBNs, thereby highlighting the pivotal nodes driving network reorganization. Finally, we develop a temporal propagation graph convolution network (TPGCN) to further capture the propagation mechanisms of heterogeneous neural information, that is, to simulate the impact of historical states on the current topology, thereby flexibly capturing the spatio-temporal features within heterogeneous DFBNs. In summary, the main contributions of this paper are as follows:

- To accurately simulate the heterogeneous evolution mechanism of brain neural activities, we propose a NeuroH-TGL to identify neural nodes with high spatio-temporal variability that drive network reorganization, and construct brain networks that integrate heterogeneity.
- Since current functional network is persistently modulated by historical neural activity, we devise a TPGCN to model the propagation of neural information in the temporal dimension, thereby effectively extracting the spatio-temporal features from heterogeneous DFBNs.
- Experimental results show that the proposed method outperforms the current state-of-the-art methods, and can provide effective biomarkers for brain disease diagnosis.

2 Related Work

Brain Network Analysis. Brain network analysis aims to understand the organizational structure of the brain, thereby identifying its working mechanisms and abnormalities caused by neurological disorders [22, 1]. Current methods can be categorized into two types: static brain network analysis and dynamic brain network analysis. Static brain network analysis refers to extracting fixed connectivity between brain regions over a period of time. For example, BrainNetCNN [23] proposes edge-to-edge, edge-to-node, and node-to-graph convolutional filters to extract the local properties of structural brain networks. BNTransformer [24] employs a graph transformer to learn pairwise connection strengths between brain regions, and incorporates orthogonal clustering to identify discriminative node embeddings. Unlike these methods, dynamic brain network analysis focuses on capturing time-varying connectivity between brain regions. For instance, ACIFBN [25] leverages an attention mechanism to learn spatio-temporal interactions among fMRI sub-sequences. OT-MCSTGCN [6]

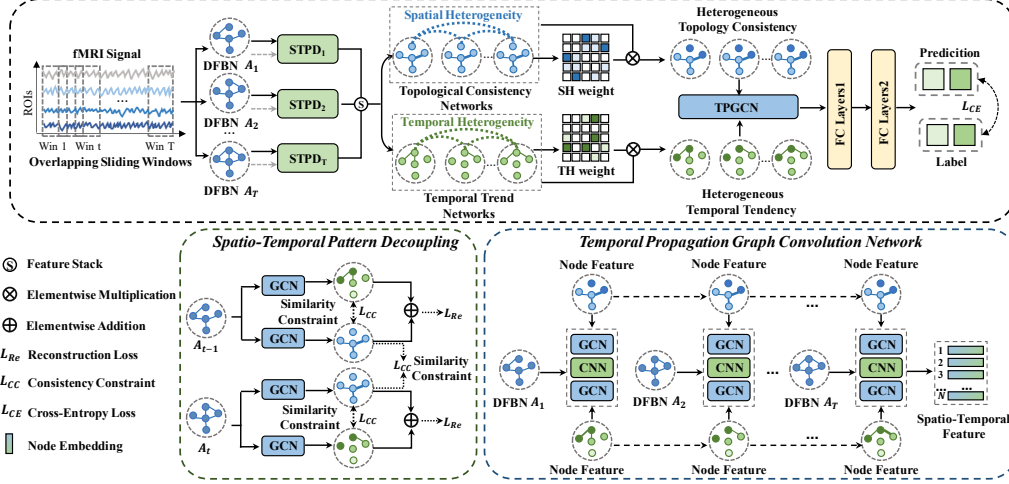


Figure 1: The overall framework of the proposed NeuroH-TGL model for brain disease diagnosis.

employs optimal transport to simulate the hubness propagation between adjacent brain graphs, thereby capturing high-order evolution in DFBNs. However, these methods overlook the inherent spatio-temporal heterogeneity of brain networks, failing to effectively model realistic dynamic dependencies in the brain. In this work, we introduce spatio-temporal decoupling and heterogeneity mining module to capture the connectivity density and temporal variability of brain structures, thereby accurately representing the heterogeneous brain activity.

Spatio-Temporal Graph Convolution for DFBNs. Spatio-temporal graph convolution typically integrate GCN and temporal convolution within a unified architecture to extract time-varying structures from DFBNs. For instance, STAGIN [4] employs a GCN to extract structural features, followed by attention mechanisms to model temporal dynamics. ST-fMRI [26] integrates GCN with four parallel 1D convolutional filters to model long-range dynamic interactions between brain regions. ST-GCN [8] combines GCN with temporal convolution to capture the non-stationary properties of functional connectivity. OT-MCSTGCN [6] proposes a multi-channel spatio-temporal GCN to efficiently aggregate topological evolution information in DFBNs. Notably, DFBNs exhibit significant sequential dependence [25, 21]. Existing spatio-temporal models fail to incorporate historical brain states’ influence on current brain graphs, resulting in suboptimal diagnostic performance. In this paper, we design the TPGCN to model the temporal propagation mechanism of heterogeneous neural activity, comprehensively capturing spatio-temporal information within heterogeneous DFBNs.

3 Proposed Method

As shown in Figure 1, we develop a neuro-heterogeneity guided temporal graph learning strategy. This framework aims to identify neural nodes with high spatio-temporal heterogeneity that drive network reorganization, thereby enhancing the diagnostic performance of brain diseases.

3.1 Spatio-Temporal Patterns Decoupling

In this paper, we assume the rs-fMRI signal for each subject is represented as $X = (x_1, x_2, \dots, x_V) \in R^{V \times L}$, where x_i represents the time series signal of the i^{th} brain region, V denotes the number of neural nodes and L indicates the number of temporal signal points. To characterize the dynamic evolution patterns of brain activity, we employ T overlapping sliding windows of length S to partition the fMRI signal generating a series of sub-signals $F = (F_1, F_2, \dots, F_T) \in R^{T \times V \times S}$. For the sub-signals F_t under the t^{th} window, we use the Pearson correlation coefficient [4] to construct the brain network A_t :

$$A_{t(ij)} = \frac{Cov(F_{t(i)}, F_{t(j)})}{\sigma_{F_{t(i)}} \sigma_{F_{t(j)}}} \quad (1)$$

where i and j are indices of F_t , Cov indicates the cross covariance, and $\sigma_{F_{t(i)}}$ is the standard deviation of $F_{t(i)}$. Therefore, DFBNs can be represented as $A = (A_1, A_2, \dots, A_T) \in R^{T \times V \times V}$. To simulate the sparsity of brain networks, we further set elements with connection strengths lower than α to 0. Neuroscience research shows that the brain's cognitive function is supported by its intrinsic topological consistency and temporal trend [17, 18, 19]. Thus, decoupling spatio-temporal organizational patterns within DFBNs can help reveal network dysregulation caused by brain diseases. For each brain network A_t ($t = 1, 2, \dots, T$), we first employ two independent GCN [12] to extract topological consistency networks H_t^{top} and temporal trend networks H_t^{tem} :

$$H_t^{top} = GCN(A_t, F_t) = \hat{A}_t \left(ReLU \left(\hat{A}_t F_t W_t^{top(0)} \right) \right) W_t^{top(1)} \quad (2)$$

$$H_t^{tem} = GCN(A_t, F_t) = \hat{A}_t \left(ReLU \left(\hat{A}_t F_t W_t^{tem(0)} \right) \right) W_t^{tem(1)} \quad (3)$$

where $\hat{A}_t = \tilde{D}_t^{-\frac{1}{2}} (A_t + I) \tilde{D}_t^{-\frac{1}{2}}$, I is the identity matrix, \tilde{D}_t denotes the degree matrix after adding self-loops, $W_t^{top(0)}$, $W_t^{top(1)}$, $W_t^{tem(0)}$ and $W_t^{tem(1)}$ all represent the learnable parameters in the graph convolutional layers, and ReLU [27] denotes the nonlinear activation function.

To enhance the discriminability between topological consistency networks and temporal trend networks, we encourage the similarity constraint $L_{CC_{t1}}$ between them to gradually decrease throughout the training process:

$$L_{CC_{t1}} = \frac{1}{V} \sum_{i=1}^V \frac{H_{t(i)}^{top} \cdot H_{t(i)}^{tem}}{\|H_{t(i)}^{top}\|_2 \cdot \|H_{t(i)}^{tem}\|_2} \quad (4)$$

where the \cdot represents the dot product operation, and $\|\cdot\|_2$ denotes the L_2 norm. Additionally, to ensure the complementarity of topological consistency networks and temporal trend networks, we sum the decoupled features to obtain a reconstructed representation. Then, we introduce mean squared error (MSE) [28] as the reconstruction loss between the reconstructed representation and brain network A_t :

$$L_{Re_t} = \|H_t^{top} \oplus H_t^{tem} - A_t\|_2 \quad (5)$$

where \oplus represents element-wise addition. Notably, the topological consistency refers to the high stability of certain network structures across different windows. Therefore, we further impose a similarity constraint on topological consistency networks across adjacent windows, and encourage this similarity $L_{CC_{t2}}$ to increase as training progresses:

$$L_{CC_{t2}} = 1 - \frac{1}{V} \sum_{i=1}^V \frac{H_{t(i)}^{top} \cdot H_{t+1(i)}^{top}}{\|H_{t(i)}^{top}\|_2 \cdot \|H_{t+1(i)}^{top}\|_2} \quad (6)$$

The above operations ensure that the topological consistency networks and temporal trend networks within the same window are distinct and complementary, while also promoting the similarity of consistency networks across different windows, thereby better aligning with the intrinsic spatio-temporal coordination of the brain [17, 18, 19].

3.2 Spatio-Temporal Heterogeneity Mining

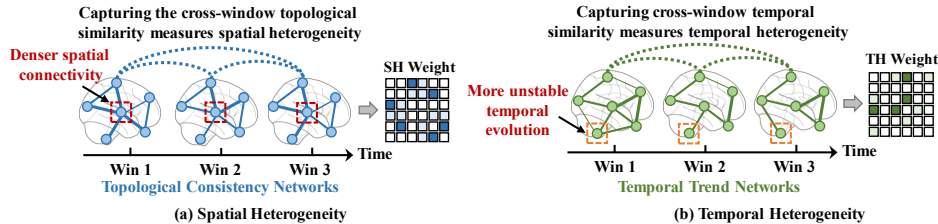


Figure 2: The schematic diagram of spatio-temporal heterogeneity mining. We calculate cross-window topological similarity and temporal similarity respectively to measure spatial heterogeneity and temporal heterogeneity.

There are some neural nodes with high spatio-temporal variability in DFBNs [6]. As shown in Figure 2, the neural node marked by red has denser spatial connections than other nodes, while the

node highlighted in orange shows more unstable temporal evolution. These nodes play an important role in coordinating the evolution of the brain, which will form a complex spatio-temporal network centered on these nodes. Moreover, the brain is a continuously evolving dynamic system with extensive asynchronous interactions [21, 6]. Therefore, we calculate the cross-window similarity for topological consistency networks and temporal trend networks separately to measure the connectivity density and temporal variability of the brain, thereby exploring the spatio-temporal heterogeneity of neural activity. The detailed process of mining spatio-temporal heterogeneity is shown in Figure 2. For spatial heterogeneity (SH), we first calculate the average correlation of topological consistency networks across all paired windows:

$$SH = \frac{2}{T(T-1)} \sum_{i=1}^{T-1} \sum_{j=i+1}^T \text{sim}(H_i^{top}, H_j^{top}) \quad (7)$$

where $\text{sim}()$ represents the cosine similarity [29]. Then, we apply spatial heterogeneity weighting to topological consistency networks: $Z_t^{top} = SH \otimes H_t^{top} (t = 1, 2, \dots, T)$, where \otimes denotes the element-wise multiplication. In contrast, for temporal heterogeneity (TH), lower cross-window similarity in temporal trend networks indicates more pronounced dynamic evolution. Therefore, we calculate TH using the following formula:

$$TH = \frac{2}{T(T-1)} \sum_{i=1}^{T-1} \sum_{j=i+1}^T (1 - \text{sim}(H_i^{tem}, H_j^{tem})) \quad (8)$$

Then, we apply temporal heterogeneity weighting to temporal trend networks: $Z_t^{tem} = TH \otimes H_t^{tem} (t = 1, 2, \dots, T)$. Spatio-temporal heterogeneity weighting not only preserves the dynamic topology of DFBNs, but also highlights important nodes and connections. Therefore, we can obtain topological consistency networks and temporal trend networks that fuse spatio-temporal heterogeneity.

3.3 Temporal Propagation Graph Convolution Network

The heterogeneous information in the brain propagate continuously along the temporal dimension, which means that each brain network gradually influences the state of the adjacent brain network [6, 4]. To flexibly capture cross-temporal interactions in heterogeneous brain networks, we further design a temporal propagation graph convolution network. This framework utilizes the heterogeneous spatio-temporal features of the brain networks from the previous moment to guide the information aggregation of the brain networks in the next moment. In this framework, each spatio-temporal convolutional block consists of two GCNs [12] and one 2D convolutional network (CNN) [30], so as to efficiently aggregate dynamic structure information. To reduce the number of parameters, we use two parameter-sharing spatio-temporal convolutional blocks to extract features from topological consistency networks and temporal trend networks, respectively. The topological consistency feature E_t^{top} and temporal trend feature E_t^{tem} can be learned as follows:

$$E_t^i = GCN(A_t, CNN(GCN(A_t, Z_t^i \oplus Z_{t-1}^i))) \quad (9)$$

where $i \in \{top, tem\}$. Then, we add the two types of features to obtain the fused spatio-temporal representation: $E_t = E_t^{top} \oplus E_t^{tem}$. After performing the same operation on the two networks for each window, we concatenate the features of all windows to obtain the global spatio-temporal representation E :

$$E = \text{Concatenate}(\{E_t | t \in \{1, 2, \dots, T\}\}) \quad (10)$$

Finally, we feed the features E into a multi-layer perceptron to predict the diagnostic results, and use cross-entropy loss L_{CE} to supervise the update of model parameters. The overall training objective can be formulated as: $L = L_{CE} + \lambda_1 L_{CC} + \lambda_2 L_{Re}$, where λ_1 and λ_2 are hyperparameters that control the relative importance of different loss terms, L_{CC} and L_{Re} represent the similarity loss and reconstruction loss for the decoupling of the STPD module across all windows, respectively.

4 Experiments

4.1 Experimental Settings

Datasets. We conduct experiments on both the public ADNI dataset (<https://adni.loni.usc.edu/>) and the Parkinson’s disease (PD) dataset collected by the Affiliated Hospital of Nanjing Medical

University. The ADNI dataset comprises 140 normal controls (NC), 268 patients with mild cognitive impairment (MCI), and 102 patients with Alzheimer’s disease (AD). The PD dataset includes 54 NC, 44 tremor dominant Parkinson’s disease (TDPD) patients, and 64 postural instability and gait disorder Parkinson’s disease (PGPD) patients.

Preprocessing. All fMRI data are preprocessed using SPM8 implemented in the DPARSF toolbox [31]. Specifically, we first correct and align the original images based on the EPI template. Then, we utilize detrending techniques to alleviate the effects of head motion as well as the interference from cerebrospinal fluid and white matter. Finally, we use the automated anatomical labeling atlas [32] to divide the fMRI of ADNI dataset into 90 brain regions with 140 time points, and the fMRI of PD dataset into 90 brain regions with 220 time points.

Metrics. For the ADNI dataset, we conduct the following classification tasks: NC vs. MCI vs. AD, NC vs. MCI, NC vs. AD, and MCI vs. AD. For the PD dataset, we conduct the following classification tasks: NC vs. TDPD vs. PGPD, NC vs. TDPD, NC vs. PGPD, and TDPD vs. PGPD. We employ 10-fold cross-validation to evaluate the diagnostic performance of all methods on different tasks. For the three-class task, we adopt macro-averaged metrics [33] to ensure equitable evaluation across all categories. We report the mean values of 10 runs.

Implementation Details. All experiments are implemented in PyTorch and trained on an NVIDIA GeForce RTX 3080 GPU with 12GB. We employ Adam optimizer [34] with a learning rate of $7e-4$ to optimize the proposed method. The size of the convolution kernel is 3×3 . Additionally, we adopt an early stopping mechanism that 80 epochs patience in total 300 epochs. The batch size is set to 8. The threshold α is set to 0.6. For the ADNI dataset, $T = 6$ and $S = 90$. For the PD dataset, $T = 8$ and $S = 80$. The hyperparameters λ_1 and λ_2 are varied within the range $\{1e-6, 1e-5, 1e-3, 1e-2, 1e-1, 1, 10\}$, with the optimal combination determined through grid search. The source code has been uploaded to the supplementary material.

4.2 Performance Analysis

Comparison Methods. To validate the effectiveness of the proposed method, we compare it with 11 state-of-the-art brain network analysis approaches. These methods can be categorized into two types: static brain network analysis and dynamic brain network analysis. Static brain network analysis methods include BrainNetCNN [23], FBNetGen [35], BNTransformer [24], BrainGNN [36], LSGNN [37], and ALTER [38]. Dynamic brain network analysis methods include ACIFBN [25], DRAT [39], ST-GCN [8], ST-fMRI [26], STAGIN [4], OT-MCSTGCN [6], and MGNN [40].

Classification Result. Table 1 and Table 2 show the diagnostic results of different methods on the ADNI and PD datasets, respectively. The standard deviations can be referred to in Appendix A. Obviously, the proposed NeuroH-TGL significantly outperforms the comparison methods. Specifically, on the ADNI dataset, NeuroH-TGL achieves accuracy improvements of 4.69%, 1.04%, 0.15%,

Table 1: Classification results of different methods on the ADNI dataset (%).

Type	Method	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC
		NC vs. MCI vs. AD			NC vs. MCI			NC vs. AD			MCI vs. AD		
Static	BrainNetCNN	57.06	38.06	54.81	71.34	81.16	55.54	71.47	50.00	62.69	68.92	75.32	53.29
	FBNetGen	57.50	56.97	67.75	61.25	66.46	58.47	63.33	58.50	57.53	65.83	73.71	64.07
	BNTransformer	60.21	40.94	61.96	69.75	79.65	57.58	74.00	71.72	75.89	74.50	81.56	57.64
	BrainGNN	58.82	45.33	61.66	74.25	83.20	62.88	74.05	62.90	68.66	79.19	87.36	64.12
	LSGNN	58.75	36.96	58.20	60.27	68.62	54.04	70.42	53.07	66.26	64.05	73.72	56.11
	ALTER	63.73	51.68	68.98	76.42	82.20	71.11	75.23	59.18	74.25	82.43	88.45	73.50
Dynamic	ACIFBN	62.35	57.88	70.42	71.25	78.96	64.49	79.17	71.53	78.44	83.75	90.91	62.88
	DRAT	60.83	51.09	70.11	71.50	79.78	57.93	72.92	55.79	68.62	80.31	88.02	70.01
	ST-GCN	67.50	41.68	59.07	73.44	83.88	54.22	67.92	51.46	60.51	82.08	89.81	61.79
	ST-fMRI	57.45	38.79	64.84	71.81	81.71	61.41	76.45	68.40	78.96	84.16	84.61	62.16
	STAGIN	56.46	27.75	54.76	69.25	80.97	56.44	67.08	47.99	55.49	74.44	85.35	53.71
	OT-MCSTGCN	59.02	41.31	58.52	76.08	80.14	60.89	73.75	57.61	66.41	76.26	85.40	68.15
	MGNN	61.76	51.38	66.74	77.46	83.16	70.51	81.35	77.46	80.84	80.54	86.80	75.43
	NeuroH-TGL	72.19	57.81	70.49	78.50	84.35	72.61	81.50	72.12	83.01	86.67	92.46	76.01

Table 2: Classification results of different methods on the PD dataset (%).

Type	Method	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC
		NC vs. TDPD	vs. PGPD		NC vs. TDPD			NC vs. PGPD			TDPD vs. PGPD		
Static	BrainNetCNN	55.70	46.09	67.16	85.67	81.98	83.52	79.55	75.32	73.10	74.00	79.43	69.78
	FBnetGen	62.50	57.65	69.23	74.56	52.80	65.65	74.55	78.39	51.82	72.00	73.05	66.14
	BNTransformer	63.75	56.47	71.75	82.50	70.88	74.83	79.00	77.18	78.62	76.73	79.50	70.20
	BrainGNN	63.49	57.58	72.51	83.67	78.47	84.85	83.93	75.70	73.67	75.00	74.37	63.06
	LSGNN	56.88	46.44	64.46	73.75	74.84	77.08	74.17	76.93	68.47	72.00	76.61	58.62
Dynamic	ALTER	62.28	48.47	63.23	86.56	80.83	78.38	80.53	78.55	74.91	79.55	83.11	68.64
	ACIFBN	61.25	48.72	66.33	82.89	77.91	79.80	76.97	73.90	68.59	74.18	73.36	63.79
	DRAT	61.88	54.72	64.66	74.75	67.53	79.20	72.83	66.37	61.49	71.27	79.67	59.69
	ST-GCN	58.13	54.93	65.56	82.67	81.21	80.18	79.70	72.80	77.15	74.45	72.67	67.44
	ST-fMRI	58.75	52.10	66.10	75.33	74.04	73.64	78.33	81.79	74.27	71.67	72.96	44.33
	STAGIN	55.63	43.52	63.71	81.89	74.20	76.55	82.20	80.89	79.30	78.00	63.00	59.85
	OT-MCSTGCN	59.38	38.24	62.37	81.56	79.31	81.59	82.20	83.59	75.50	75.82	73.39	63.05
	MGNN	59.89	51.56	64.25	85.78	83.17	81.10	78.79	80.14	73.22	78.73	74.98	69.43
	NeuroH-TGL	66.25	61.71	73.85	91.25	91.00	94.21	87.17	89.38	88.42	83.75	86.80	82.91

and 2.51% over the suboptimal results across four classification tasks, respectively. On the PD dataset, NeuroH-TGL achieves accuracy improvements of 2.50%, 4.69%, 3.24%, and 4.20% over the suboptimal results across four classification tasks, respectively. The reason for the performance improvement is that the proposed method exploits the spatio-temporal heterogeneous activity patterns of the brain, thereby highlighting the pivotal neural nodes involved in the evolution of DFBN. On this basis, we further design the TPGCN to model the sequential dependencies between heterogeneous neural nodes, thereby collaboratively extracting the time-varying topological features within them. Additionally, we conduct an analysis of computational efficiency. The proposed method requires only 0.0924M parameters and 0.2479M FLOPs, which indicates that the method can achieve excellent diagnostic performance with a minimal amount of computational resources.

T-SNE Visualization. To visually demonstrate the performance of different methods, we use t-SNE [41] to visualize their learned features. In Figure 3(a), the original feature distribution is chaotic. BrainNetCNN forms two clusters with confusion. The feature distributions learned by ACIFBN, DART and OT-MCSTGCN are loose and fail to establish clear inter-class boundaries. In contrast, our method effectively aggregates intra-class features while maintaining distinct inter-class separation. This is because BrainNetCNN ignores the dynamic features of the brain. Although DART, ACIFBN and OT-MCSTGCN capture spatio-temporal information to a certain extent, they all neglect the heterogeneity of neural activity. Unlike them, the proposed NeuroH-TGL not only effectively captures

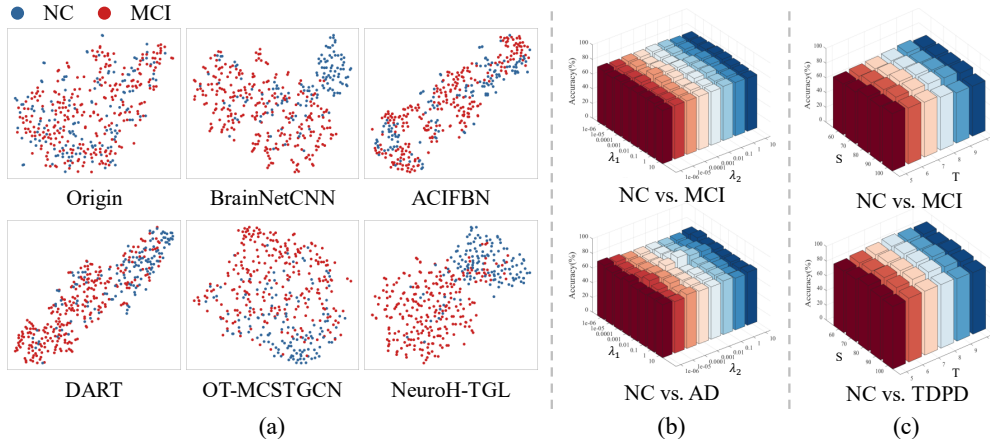


Figure 3: (a) T-SNE visualization for different methods on the NC vs. MCI task. (b) The impact of λ_1 and λ_2 on different diagnostic tasks. (c) The impact of T and S on different diagnostic tasks.

the pivotal heterogeneous nodes driving the network reorganization, but also comprehensively encodes the temporal dependence between historical neural activity and the current brain topology.

Parameter Sensitivity Analysis. The hyperparameters λ_1 and λ_2 in the objective function are vary within the range $\{1e-6, 1e-5, 1e-3, 1e-2, 1e-1, 1, 10\}$. In Figure 3(b), we systematically investigate the impact of different parameter combinations on disease diagnosis. The experimental results demonstrate that the accuracy remains stable across varying values of λ_1 and λ_2 . Therefore, the proposed method exhibits robustness and is not sensitive to hyperparameters. Additionally, we also explore the impact of T and S on diagnostic performance, and the experimental results are shown in Figure 3(c). Specifically, T changes within the set $\{5, 6, 7, 8, 9, 10\}$, and S varies within the set $\{60, 70, 80, 90, 100\}$. We conduct a grid search on both S and T to ensure the optimal parameter combination. For NC vs. MCI, the best performance is achieved when $T=6$ and $S=90$. For NC vs. TDPD, the best performance is achieved when $T=8$ and $S=80$. Larger values of T and S lead to longer overlapping sequences across windows, which might smooth out valuable dynamic information in the brain. Conversely, smaller values of T and S result in shorter time sequences per window, potentially making the statistical correlation between brain regions unreliable. Therefore, a moderate window size can balance reliable statistical correlation with temporal evolution.

4.3 Ablation Study

To validate the effectiveness of each module, we conduct ablation experiments on ADNI and PD datasets, with the results presented in Table 3. The standard deviations can be referred to in Appendix B. The simplified models included: (1) w/o STPD: The STPD module is removed (i.e., $\lambda_1 = \lambda_2 = 0$). (2) w/o STHW: Spatial and temporal heterogeneity weights (STHW) are replaced with all-ones matrices. (3) w/o TPGCN: The TPGCN is substituted with GCN. The experimental results indicate that removing any module will lead to a decrease in performance. For instance, in the NC vs. MCI task, removing STPD, STHW, and TPGCN led to accuracy decreases of 3.75%, 3.25%, and 3.50%, respectively. The reasons for this phenomenon include: The STPD module effectively decouples the topological consistency features and temporal trend features aligned with brain dynamics from DFBN, thus laying a foundation for spatio-temporal heterogeneity mining. The STHW module effectively captures the heterogeneous activity patterns of each brain region, making it possible to identify abnormal brain states associated with neurological disorders. Moreover, TPGCN outperforms GCN, proving its effectiveness in simulating the temporal propagation mechanisms of heterogeneous neural information, thereby comprehensively capturing the spatio-temporal dependencies in heterogeneous DFBNs. Therefore, all the proposed modules are effective and promote each other.

Table 3: Ablation results of the proposed method on the ADNI and PD datasets (%).

Datasets	Method	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC
		NC vs. MCI vs. AD			NC vs. MCI			NC vs. AD			MCI vs. AD		
ADNI	w/o STPD	66.88	35.00	57.22	73.75	83.20	62.16	76.67	71.42	76.87	81.67	88.77	72.95
	w/o STHW	65.21	54.88	67.21	74.25	82.94	63.90	77.08	70.36	74.69	78.06	86.63	67.01
	w/o TPGCN	65.80	60.13	70.98	74.00	82.58	64.76	77.91	72.77	77.35	82.08	89.00	68.66
	NeuroH-TGL	72.19	57.81	70.49	78.50	84.35	72.61	81.50	72.12	83.01	86.67	92.46	76.01
PD		NC vs. TDPD vs. PGPD			NC vs. TDPD			NC vs. PGPD			TDPD vs. PGPD		
	w/o STPD	65.00	55.15	64.57	86.25	87.06	83.54	83.75	85.76	86.12	81.25	86.21	70.50
	w/o STHW	63.60	52.48	65.53	83.50	80.96	79.78	79.33	82.47	79.93	78.75	81.96	69.91
	w/o STPGC	64.38	55.18	64.40	84.75	83.73	88.53	80.50	82.64	81.16	76.00	78.23	70.65
	NeuroH-TGL	66.25	61.71	73.85	91.25	91.00	94.21	87.17	89.38	88.42	83.75	86.80	82.91

w/o means without.

5 Discussions

Heterogeneity Weights Visualization. To explore the impact of brain diseases on neural activity, Figure 4 displays the spatial and temporal heterogeneity weights across all brain regions of different groups. Based on the experimental results, we can draw the following conclusions. First, brain diseases alter spatio-temporal properties of the brain. For instance, MCI and AD groups exhibit lower spatial heterogeneity but higher temporal heterogeneity compared to the NC group. Second,

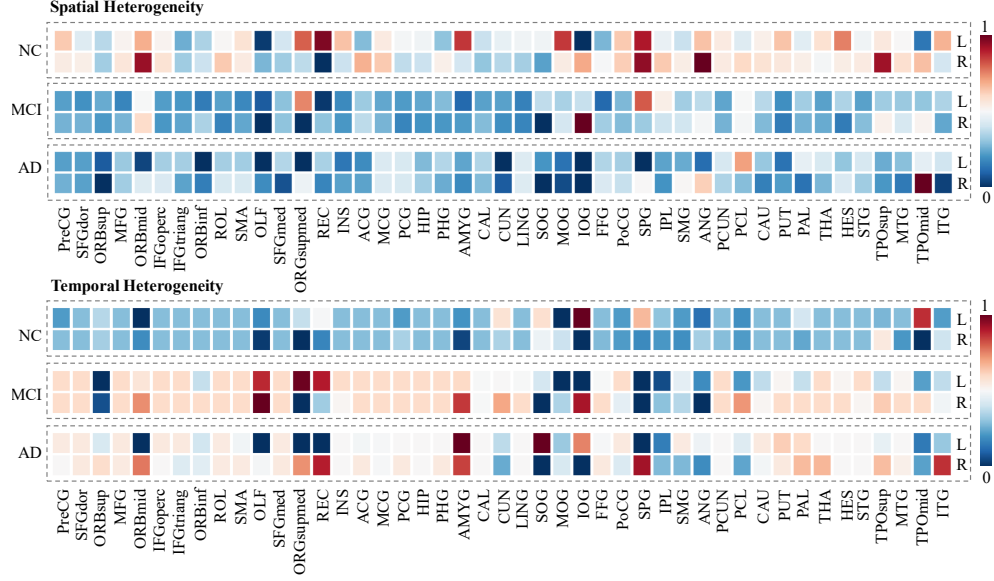


Figure 4: Visualization of spatio-temporal heterogeneity weights across different groups. Each square represents a brain region, and abbreviations are provided for each brain region. 'L' indicates the left brain region, and 'R' indicates the right brain region.

the spatio-temporal heterogeneity of each brain region in the same group is different. The NC group exhibits higher spatial heterogeneity but lower temporal heterogeneity of brain regions, while the MCI and AD groups show the opposite pattern. This may be because neurodegenerative changes reduce the complexity of the brain, thus decreasing the spatial heterogeneity [25, 42]. Moreover, brain diseases can trigger compensatory mechanisms that increase variability in temporal activities, enhancing the temporal heterogeneity [6, 43]. Notably, we also find that the spatio-temporal heterogeneity of supplementary motor area, hippocampus and amygdala in the patient group is significantly different from that in the NC group. Therefore, these brain regions may serve as potential biomarkers for the diagnosis of MCI and AD.

Discriminative Brain Regions. To further evaluate the efficacy of the proposed method in identifying biomarkers, we employ t-test on spatio-temporal feature of each brain region, thereby identifying the 10 most discriminative regions ($p < 0.05$). The visualization results are shown in Figure 5. For NC vs. MCI, the significant brain regions are concentrated in the middle temporal gyrus and parahippocampal gyrus, among others. This may be because MCI leads to visual impairments and memory deficits, thereby causing abnormalities in the related brain regions [6, 44]. For NC vs. AD, key brain regions include the amygdala and superior frontal gyrus, which are responsible for emotion regulation and

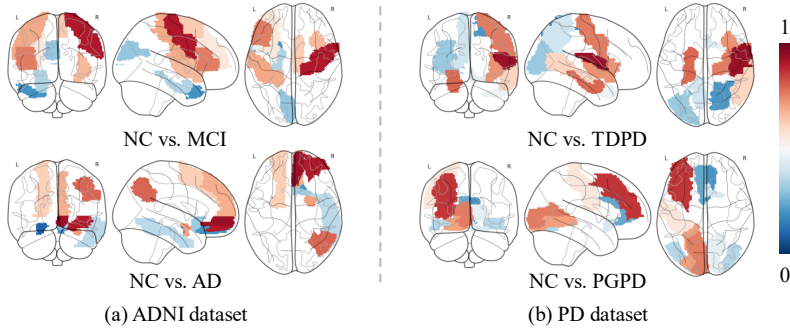


Figure 5: Distribution of the 10 most discriminative brain regions on different diagnostic tasks. Different colors indicate the relative importance of these brain regions.

behavioral control and are closely related to the occurrence of AD [45]. For NC vs. TDPD, important brain regions include the precentral gyrus and rolandic operculum, possibly because motor disorders in TDPD patients disrupt the normal functioning of these motor control areas [46]. For NC vs. PGPD, significant brain regions include the inferior occipital gyrus and lingual gyrus. This is because PGPD patients exhibit abnormalities in processing complex visual scenes [7]. Therefore, the proposed method can provide reasonable biomarkers for brain disease diagnosis.

Conclusion. In this paper, we propose the NeuroH-TGL to collaboratively capture neural nodes in the brain that exhibit spatial density and significant temporal variability, addressing the shortcomings of existing methods that overlook the spatio-temporal heterogeneity of nodes. Specifically, we first decouple the DFBNs into topological consistency networks and temporal trend networks based on their spatio-temporal coordination. Then, we measure the spatial density of topological consistency networks and the temporal variability of temporal trend networks across global time domains, respectively, to emphasize the significant spatio-temporal associations driven by these heterogeneous nodes. Finally, we develop the TPGCN to model the influence of the historical state of heterogeneous nodes on the current network configuration, enabling a comprehensive capture of dynamic topological features. Extensive experiments show that NeuroH-TGL not only significantly enhances diagnostic performance but also identifies abnormal spatio-temporal features caused by brain diseases.

Limitations and Future Work. The proposed method is based solely on a single fMRI modality, overlooking the complementary information from other modalities. In future research, we will develop a heterogeneity reorganization mechanism for DFBNs under structural connectivity constraints. This framework will be capable of integrating complementary heterogeneity features between function and structure to improve diagnostic accuracy and provide interpretability.

6 Acknowledgements

This work was supported in part by Key Research and Development Plan of Jiangsu Province (No. BE2022842), National Natural Science Foundation of China (Nos. 62371234, 62076129, 62136004, 62272226 and 62276130), Natural Science Foundation of Jiangsu Province (No. BK20231438), and also National Key R&D Program of China (No. 2023YFF1204803).

References

- [1] Hongting Ye, Yalu Zheng, Yueying Li, Ke Zhang, Youyong Kong, and Yonggui Yuan. Rh-brainfs: regional heterogeneous multimodal brain networks fusion strategy. *Advances in Neural Information Processing Systems*, 36:59286–59303, 2023.
- [2] Zuozhen Zhang, Junzhong Ji, and Jinduo Liu. Metarlec: Meta-reinforcement learning for discovery of brain effective connectivity. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 10261–10269, 2024.
- [3] Huzheng Yang, James Gee, and Jianbo Shi. Brain decodes deep nets. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 23030–23040, 2024.
- [4] Byung-Hoon Kim, Jong Chul Ye, and Jae-Jin Kim. Learning dynamic graph representation of brain connectome with spatio-temporal attention. *Advances in Neural Information Processing Systems*, 34:4314–4327, 2021.
- [5] Davide Momi, Zheng Wang, Sara Parmigiani, Ezequiel Mikulan, Sorenza P Bastiaens, Mohammad P Oveisi, Kevin Kadak, Gianluca Gaglioti, Allison C Waters, Sean Hill, et al. Stimulation mapping and whole-brain modeling reveal gradients of excitability and recurrence in cortical networks. *Nature Communications*, 16(1):3222, 2025.
- [6] Qi Zhu, Shengrong Li, Xiangshui Meng, Qiang Xu, Zhiqiang Zhang, Wei Shao, and Daoqiang Zhang. Spatio-temporal graph hubness propagation model for dynamic brain network classification. *IEEE Transactions on Medical Imaging*, 43(6):2381–2394, 2024.
- [7] Javier Pagonabarraga, Helena Bejr-Kasem, Saul Martinez-Horta, and Jaime Kulisevsky. Parkinson disease psychosis: From phenomenology to neurobiological mechanisms. *Nature Reviews Neurology*, 20(3):135–150, 2024.

- [8] Soham Gadgil, Qingyu Zhao, Adolf Pfefferbaum, Edith V Sullivan, Ehsan Adeli, and Kilian M Pohl. Spatio-temporal graph convolution for resting-state fmri analysis. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII* 23, pages 528–538. Springer, 2020.
- [9] Rui Liu, Yao Hu, Jibin Wu, Ka-Chun Wong, Zhi-An Huang, Yu-An Huang, and Kay Chen Tan. Dynamic graph representation learning for spatio-temporal neuroimaging analysis. *IEEE Transactions on Cybernetics*, 55(3):1121–1134, 2025.
- [10] Youyong Kong, Xiaotong Zhang, Wenhan Wang, Yue Zhou, Yueying Li, and Yonggui Yuan. Multi-scale spatial-temporal attention networks for functional connectome classification. *IEEE Transactions on Medical Imaging*, 44(1):475–488, 2025.
- [11] Sin-Yee Yap, Junn Yong Loo, Chee-Ming Ting, Fuad Noman, Raphaël C.-W. Phan, Adeel Razi, and David L. Dowe. A deep probabilistic spatiotemporal framework for dynamic graph representation learning with application to brain disorder identification. In *IJCAI*, pages 5353–5361, 2024.
- [12] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
- [13] Dongdong Chen and Lichi Zhang. Fe-stgcn: Spatio-temporal graph neural network with functional and effective connectivity fusion for mci diagnosis. In Hayit Greenspan, Anant Madabhushi, Parvin Mousavi, Septimiu Salcudean, James Duncan, Tanveer Syeda-Mahmood, and Russell Taylor, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, pages 67–76, Cham, 2023. Springer Nature Switzerland.
- [14] Tiago Azevedo, Alexander Campbell, Rafael Romero-Garcia, Luca Passamonti, Richard AI Bethlehem, Pietro Lio, and Nicola Toschi. A deep graph neural network architecture for modelling spatio-temporal dynamics in resting-state functional mri data. *Medical Image Analysis*, 79:102471, 2022.
- [15] Matthew Ainsworth, Zhemeng Wu, Helen Browncross, Anna S Mitchell, Andrew H Bell, and Mark J Buckley. Frontopolar cortex shapes brain network structure across prefrontal and posterior cingulate cortex. *Progress in Neurobiology*, 217:102314, 2022.
- [16] Sonia Turrini, Francesca Fiori, Naomi Bevacqua, Chiara Saracini, Boris Lucero, Matteo Candidi, and Alessio Avenanti. Spike-timing-dependent plasticity induction reveals dissociable supplementary–and premotor–motor pathways to automatic imitation. *Proceedings of the National Academy of Sciences*, 121(27):e2404925121, 2024.
- [17] Georg Northoff and Dusan Hirjak. Is depression a global brain disorder with topographic dynamic reorganization? *Translational Psychiatry*, 14(1):278, 2024.
- [18] Andrea Santoro, Federico Battiston, Maxime Lucas, Giovanni Petri, and Enrico Amico. Higher-order connectomics of human brain function reveals local topological signatures of task decoding, individual identification, and behavior. *Nature Communications*, 15(1):10244, 2024.
- [19] Tal Seidel Malkinson, Dimitri J Bayle, Brigitte C Kaufmann, Jianghao Liu, Alexia Bourgeois, Katia Lehongre, Sara Fernandez-Vidal, Vincent Navarro, Virginie Lambrecq, Claude Adam, et al. Intracortical recordings reveal vision-to-action cortical gradients driving human exogenous attention. *Nature Communications*, 15(1):2586, 2024.
- [20] Xiaolong Peng, Qi Liu, Catherine S Hubbard, Danhong Wang, Wenzhen Zhu, Michael D Fox, and Hesheng Liu. Robust dynamic brain coactivation states estimated in individuals. *Science Advances*, 9(3):eabq8566, 2023.
- [21] Diego Vidaurre, Stephen M Smith, and Mark W Woolrich. Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48):12827–12832, 2017.

- [22] Song Wang, Zhenyu Lei, Zhen Tan, Jiaqi Ding, Xinyu Zhao, Yushun Dong, Guorong Wu, Tianlong Chen, Chen Chen, Aiying Zhang, et al. Brainmap: Learning multiple activation pathways in brain networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 14432–14440, 2025.
- [23] Jeremy Kawahara, Colin J Brown, Steven P Miller, Brian G Booth, Vann Chau, Ruth E Grunau, Jill G Zwicker, and Ghassan Hamarneh. Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment. *NeuroImage*, 146:1038–1049, 2017.
- [24] Xuan Kan, Wei Dai, Hejie Cui, Zilong Zhang, Ying Guo, and Carl Yang. Brain network transformer. *Advances in Neural Information Processing Systems*, 35:25586–25599, 2022.
- [25] Jianjia Zhang, Xiaotong Wu, Xiang Tang, Luping Zhou, Lei Wang, Weiwen Wu, and Dinggang Shen. Asynchronous functional brain network construction with spatiotemporal transformer for mci classification. *IEEE Transactions on Medical Imaging*, 44(3):1168–1180, 2025.
- [26] Simon Dahan, Logan ZJ Williams, Daniel Rueckert, and Emma C Robinson. Improving phenotype prediction using long-range spatio-temporal dynamics of functional connectivity. In *Machine Learning in Clinical Neuroimaging: 4th International Workshop, MLCN 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 4*, pages 145–154. Springer, 2021.
- [27] Yuhan Bai. Relu-function and derived function review. In *SHS web of conferences*, volume 144, page 02006. EDP Sciences, 2022.
- [28] Jiawei Ren, Mingyuan Zhang, Cunjun Yu, and Ziwei Liu. Balanced mse for imbalanced visual regression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7926–7935, 2022.
- [29] Harald Steck, Chaitanya Ekanadham, and Nathan Kallus. Is cosine-similarity of embeddings really about similarity? In *Companion Proceedings of the ACM Web Conference 2024*, pages 887–890, 2024.
- [30] An Yan, Shuo Cheng, Wang-Cheng Kang, Mengting Wan, and Julian McAuley. Cosrec: 2d convolutional neural networks for sequential recommendation. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pages 2173–2176, 2019.
- [31] Shuo Yu, Shan Jin, Ming Li, Tabinda Sarwar, and Feng Xia. Long-range brain graph transformer. *Advances in Neural Information Processing Systems*, 37:24472–24495, 2024.
- [32] Edmund T Rolls, Chu-Chung Huang, Ching-Po Lin, Jianfeng Feng, and Marc Joliot. Automated anatomical labelling atlas 3. *NeuroImage*, 206:116189, 2020.
- [33] Yu Wang, Junxian Mu, Pengfei Zhu, and Qinghua Hu. Exploring diverse representations for open set recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 5731–5739, 2024.
- [34] Zijun Zhang. Improved adam optimizer for deep neural networks. In *2018 IEEE/ACM 26th international symposium on quality of service (IWQoS)*, pages 1–2. IEEE, 2018.
- [35] Xuan Kan, Hejie Cui, Joshua Lukemire, Ying Guo, and Carl Yang. Fbnetgen: Task-aware gnn-based fmri analysis via functional brain network generation. In *International Conference on Medical Imaging with Deep Learning*, pages 618–637. PMLR, 2022.
- [36] Xiaoxiao Li, Yuan Zhou, Nicha Dvornek, Muhan Zhang, Siyuan Gao, Juntang Zhuang, Dustin Scheinost, Lawrence H Staib, Pamela Ventola, and James S Duncan. Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, 74:102233, 2021.
- [37] Dongdong Chen, Mengjun Liu, Zhenrong Shen, Xiangyu Zhao, Qian Wang, and Lichi Zhang. Learnable subdivision graph neural network for functional brain network analysis and interpretable cognitive disorder diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 56–66. Springer, 2023.

- [38] Shuo Yu, Shan Jin, Ming Li, Tabinda Sarwar, and Feng Xia. Long-range brain graph transformer. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 24472–24495. Curran Associates, Inc., 2024.
- [39] Xuan Kan, Antonio Aodong Chen Gu, Hejie Cui, Ying Guo, and Carl Yang. Dynamic brain transformer with multi-level attention for functional brain network analysis. In *2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 1–4. IEEE, 2023.
- [40] Qianqian Wang, Wei Wang, Yuqi Fang, Pew-Thian Yap, Hongtu Zhu, Hong-Jun Li, Lishan Qiao, and Mingxia Liu. Leveraging brain modularity prior for interpretable representation learning of fmri. *IEEE Transactions on Biomedical Engineering*, 71(8):2391–2401, 2024.
- [41] Dmitry Kobak and Philipp Berens. The art of using t-sne for single-cell transcriptomics. *Nature Communications*, 10(1):5416, 2019.
- [42] Sebastian Moguilner, Sandra Baez, Hernan Hernandez, Joaquín Migeot, Agustina Legaz, Raul Gonzalez-Gomez, Francesca R Farina, Pavel Prado, Jhosmary Cuadros, Enzo Tagliazucchi, et al. Brain clocks capture diversity and disparities in aging and dementia across geographically diverse populations. *Nature Medicine*, pages 3646–3657, 2024.
- [43] Zhengwang Xia, Tao Zhou, Saqib Mamoon, and Jianfeng Lu. Inferring brain causal and temporal-lag networks for recognizing abnormal patterns of dementia. *Medical Image Analysis*, 94:103133, 2024.
- [44] Zhijian Yang, Ilya M Nasrallah, Haochang Shou, Junhao Wen, Jimit Doshi, Mohamad Habes, Guray Erus, Ahmed Abdulkadir, Susan M Resnick, Marilyn S Albert, et al. A deep learning framework identifies dimensional representations of alzheimer’s disease from brain structure. *Nature Communications*, 12(1):7065, 2021.
- [45] Kaitlin M Stouffer, Xenia Grande, Emrah Düzel, Maurits Johansson, Byron Creese, Menno P Witter, Michael I Miller, Laura EM Wisse, and David Berron. Amidst an amygdala renaissance in alzheimer’s disease. *Brain*, 147(3):816–829, 2024.
- [46] Han Liu, Zeqi Hao, Shasha Qiu, Qianqian Wang, Linlin Zhan, Lina Huang, Youbin Shao, Qing Wang, Chang Su, Yikang Cao, et al. Grey matter structural alterations in anxiety disorders: A voxel-based meta-analysis. *Brain Imaging and Behavior*, 18(2):456–474, 2024.

NeurIPS Paper Checklist

1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [\[Yes\]](#)

Justification: The paper introduces a novel Neuro-Heterogeneity guided Temporal Graph Learning strategy (NeuroH-TGL) to identify the pivotal neural nodes that drive network reorganization in the brain. This approach effectively captures the heterogeneous spatio-temporal features within the brain, thereby enhancing the diagnostic performance of brain diseases. Extensive experimental results validate the effectiveness of the proposed method and identify effective biomarkers for brain disease diagnosis. Therefore, the main claims made in the abstract and introduction accurately reflect the contributions and scope of the paper.

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [\[Yes\]](#)

Justification: The paper includes a separate section titled 'Limitations and Future Work' to discuss the limitations of the proposed method and future research plans. In this section, the authors explicitly point out that due to the challenges in collecting multimodal neuroimaging data, the study focuses solely on brain network analysis based on the fMRI modality. Secondly, the paper proposes future directions for extending the framework to construct multimodal dynamic brain networks, with particular emphasis on investigating reorganization mechanisms of DFBNs under structural connectivity constraints to achieve improved results. These efforts indicate that the authors have recognized the limitations of the current research and have proposed specific measures for improvement.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.

- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [NA]

Justification: Our paper does not include theoretical results. But we formulate our question and method with detailed formulas.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: The article provides a detailed description of the experimental setup and validates it on public datasets. Most importantly, in the supplemental materials, we provide the source code. Therefore, the experimental results are reproducible.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example

- (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
- (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
- (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
- (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [\[Yes\]](#)

Justification: In this paper, we provide the link to the publicly available ADNI dataset, and we upload the source code to the supplementary material. Therefore, the data and code are accessible.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [\[Yes\]](#)

Justification: The paper provides a detailed description of the dataset splits and hyperparameter settings in the Implementation Details section.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: The paper provides detailed information about the standard deviations of comparison experiments and ablation studies in Appendix A and Appendix B, respectively, offering readers sufficient information to assess the reliability and validity of the experimental results.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.).
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: The paper explicitly states that the proposed method is trained on an NVIDIA GeForce RTX 3080 GPU with 12GB of video memory.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. Code of ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: The ADNI dataset is a publicly available dataset that has been used in numerous previous studies, and it is clearly free of ethical concerns. The Parkinson's disease dataset is collected in collaboration with our partner hospitals and has not been made public. However,

the data collection is carried out with the consent of the participants, who are explicitly informed about the purpose of the sample collection, and all personal information related to the samples is anonymized. Therefore, it does not adversely affect any individuals. Consequently, there are no ethical or moral issues present.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [\[Yes\]](#)

Justification: The experimental results in the paper clearly demonstrate that the proposed method can effectively enhance the diagnostic performance of brain diseases and provide reasonable biomarkers.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [\[NA\]](#)

Justification: The paper poses no such risks.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.

- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [\[Yes\]](#)

Justification: We cited the source code paper we used properly.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. New assets

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [\[Yes\]](#)

Justification: The code for the model and our analysis will be well documented, and will be public on github.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. Crowdsourcing and research with human subjects

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [\[NA\]](#)

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.

- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. Institutional review board (IRB) approvals or equivalent for research with human subjects

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. Declaration of LLM usage

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [NA]

Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.

A Appendix A: Comparison Results with Standard Deviation

Due to text layout and page limitations, only the mean values from 10 tests are presented in the main text. To ensure a comprehensive presentation of the experimental results, we list the mean values and standard deviations (std) for each evaluation metric in Table 4 and Table 5.

Table 4: Classification results (mean/std) of different methods on the ADNI dataset (%).

Type	Method	ACC	F1	AUC	ACC	F1	AUC
		NC vs. MCI vs. AD			NC vs. MCI		
Static	BrainNetCNN	57.06/03.87	38.06/05.63	54.81/05.50	71.34/04.51	81.16/04.21	55.54/07.73
	FBNetGen	57.50/06.88	56.97/04.69	67.75/06.11	61.25/09.29	66.46/11.22	58.47/12.03
	BNTransformer	60.21/03.78	40.94/07.51	61.96/06.90	69.75/06.17	79.65/04.68	57.58/09.83
	BrainGNN	58.82/05.48	45.33/09.87	61.66/04.67	74.25/03.72	83.20/02.31	62.88/11.82
	LSGNN	58.75/02.91	36.96/09.56	58.20/04.99	60.27/08.83	68.62/13.00	54.04/08.45
	ALTER	63.73/05.70	51.68/07.54	68.98/05.91	76.42/08.96	82.20/07.31	71.11/12.26
Dynamic	ACIFBN	62.35/05.02	57.88/06.12	70.42/03.40	71.25/04.37	78.96/05.37	64.49/08.50
	DRAT	60.83/06.44	51.09/09.14	70.11/08.83	71.50/05.15	79.78/09.09	57.93/19.89
	ST-GCN	67.50/05.27	41.68/09.89	59.07/09.93	73.44/02.52	83.88/17.67	54.22/11.43
	ST-fMRI	57.45/04.30	38.79/12.07	64.84/07.52	71.81/02.95	81.71/03.72	61.41/06.51
	STAGIN	56.46/05.62	27.75/05.83	54.76/05.34	69.25/03.54	80.97/01.53	56.44/07.40
	OT-MCSTGCN	59.02/06.82	41.31/10.78	58.52/09.15	76.08/07.28	80.14/03.83	60.89/08.05
	MGNN	61.76/05.63	51.38/09.65	66.74/06.20	77.46/06.32	83.16/06.23	70.51/12.48
	NeuroH-TGL	72.19/06.31	57.81/13.21	70.49/12.30	78.50/05.27	84.35/03.88	72.61/07.13
Type	Method	ACC	F1	AUC	ACC	F1	AUC
		NC vs. AD			MCI vs. AD		
Static	BrainNetCNN	71.47/08.48	50.00/32.92	62.69/16.07	68.92/13.63	75.32/23.49	53.29/11.14
	FBNetGen	63.33/07.86	58.50/10.03	57.53/11.74	65.83/10.40	73.71/08.32	64.07/12.72
	BNTransformer	74.00/10.68	71.72/07.97	75.89/09.99	74.50/09.07	81.56/08.43	57.64/18.27
	BrainGNN	74.05/08.36	62.90/10.90	68.66/08.10	79.19/04.37	87.36/02.79	64.12/12.39
	LSGNN	70.42/06.83	53.07/18.88	66.26/13.00	64.05/09.21	73.72/10.06	56.11/07.38
	ALTER	75.23/07.32	59.18/24.16	74.25/10.48	82.43/05.70	88.45/04.11	73.50/10.80
Dynamic	ACIFBN	79.17/05.27	71.53/10.66	78.44/08.56	83.75/05.09	90.91/02.81	62.88/15.06
	DRAT	72.92/08.79	55.79/24.21	68.62/10.45	80.31/07.79	88.02/04.79	70.01/17.71
	ST-GCN	67.92/04.19	51.46/20.23	60.51/10.43	82.08/02.67	89.81/02.28	61.79/05.12
	ST-fMRI	76.45/11.22	68.40/27.38	78.96/10.92	84.16/05.16	84.61/04.91	62.16/09.45
	STAGIN	67.08/04.73	47.99/12.31	55.49/15.81	74.44/01.11	85.35/00.73	53.71/11.38
	OT-MCSTGCN	73.75/08.34	57.61/23.67	66.41/12.45	76.26/06.02	85.40/03.93	68.15/06.89
	MGNN	81.35/06.62	77.46/07.68	80.84/08.44	80.54/04.65	86.80/03.69	75.43/08.22
	NeuroH-TGL	81.50/06.73	72.12/11.49	83.01/10.12	86.67/04.08	92.46/02.25	76.01/16.12

Table 5: Classification results (mean/std) of different methods on the PD dataset (%).

Type	Method	ACC	F1	AUC	ACC	F1	AUC
		NC vs. TDPD vs. PGPD			NC vs. TDPD		
Static	BrainNetCNN	55.70/11.68	46.09/12.90	67.16/10.13	85.67/06.80	81.98/08.07	83.52/11.72
	FBnetGen	62.50/13.11	57.65/16.24	69.23/13.49	74.56/08.13	52.80/31.54	65.65/23.42
	BNTransformer	63.75/05.48	56.47/08.10	71.75/03.40	82.50/06.12	70.88/24.91	74.83/12.12
	BrainGNN	63.49/07.36	57.58/10.14	72.51/06.90	83.67/08.06	78.47/11.21	84.85/10.96
	LSGNN	56.88/08.59	46.44/11.69	64.46/10.71	73.75/27.07	74.84/17.88	77.08/16.43
	ALTER	62.28/07.85	48.47/13.40	63.23/13.95	86.56/09.63	80.83/16.55	78.38/18.57
Dynamic	ACIFBN	61.25/10.75	48.72/17.28	66.33/16.23	82.89/10.93	77.91/16.31	79.80/18.10
	DRAT	61.88/08.13	54.72/07.69	64.66/07.81	74.75/12.47	67.53/24.90	79.20/17.45
	ST-GCN	58.13/08.41	54.93/10.88	65.56/08.51	82.67/09.02	81.21/09.85	80.18/13/04
	ST-fMRI	58.75/08.48	52.10/10.15	66.10/08.19	75.33/13.92	74.04/14.31	73.64/17.50
	STAGIN	55.63/07.63	43.52/10.77	63.71/04.28	81.89/09.71	74.20/26.10	76.55/21.18
	OT-MCSTGCN	59.38/08.95	38.24/08.67	62.37/14.32	81.56/10.89	79.31/13.10	81.59/13.40
	MGNN	59.89/06.23	51.56/11.43	64.25/12.15	85.78/11.27	83.17/15.54	81.10/19.85
	NeuroH-TGL	66.25/14.58	61.71/16.88	73.85/15.14	91.25/08.00	91.00/08.20	94.21/08.27
Type	Method	ACC	F1	AUC	ACC	F1	AUC
		NC vs. PGPD			TDPD vs. PGPD		
Static	BrainNetCNN	79.55/08.81	75.32/16.97	73.10/21.24	74.00/09.95	79.43/09.10	69.78/19.29
	FBnetGen	74.55/07.51	78.39/04.73	51.82/25.05	72.00/12.49	73.05/14.63	66.14/19.14
	BNTransformer	79.00/09.60	77.18/15.50	78.62/11.08	76.73/09.82	79.50/09.28	70.20/14.43
	BrainGNN	83.93/11.05	75.70/27.87	73.67/18.62	75.00/11.83	74.37/25.70	63.06/19.10
	LSGNN	74.17/08.98	76.93/09.53	68.47/16.28	72.00/07.48	76.61/08.45	58.62/13.83
	ALTER	80.53/07.60	78.55/15.06	74.91/13.03	79.55/08.18	83.11/07.73	68.64/20.59
Dynamic	ACIFBN	76.97/13.27	73.90/26.74	68.59/24.62	74.18/14.29	73.36/26.45	63.79/10.20
	DRAT	72.83/12.89	66.37/27.64	61.49/25.87	71.27/07.79	79.67/05.16	59.69/21.38
	ST-GCN	79.70/10.05	72.80/20.98	77.15/13.92	74.45/14.45	72.67/27.77	67.44/18.90
	ST-fMRI	78.33/09.28	81.79/08.50	74.27/11.15	71.67/07.93	72.96/13.02	44.33/23.31
	STAGIN	82.20/06.04	80.89/10.86	79.30/09.40	78.00/13.27	63.00/36.08	59.85/21.41
	OT-MCSTGCN	82.20/08.69	83.59/09.09	75.50/15.85	75.82/11.14	73.39/25.59	63.05/18.53
	MGNN	78.79/09.38	80.14/10.92	73.22/16.48	78.73/07.00	74.98/25.74	69.43/16.27
	NeuroH-TGL	87.17/08.40	89.38/05.29	88.42/12.64	83.75/09.76	86.80/08.57	82.91/15.43

B Appendix B: Ablation Results with Standard Deviation

Similarly, to ensure a comprehensive presentation of the ablation results, we list the mean values and standard deviations for each evaluation metric in Table 6 and Table 7.

Table 6: Ablation results (mean/std) of the proposed method on the ADNI dataset (%).

Method	ACC	F1	AUC	ACC	F1	AUC
	NC vs. MCI vs. AD			NC vs. MCI		
w/o STPD	66.88/02.50	35.00/05.06	57.22/06.68	73.75/04.64	83.20/02.15	62.16/08.43
w/o STHW	65.21/04.85	54.88/11.18	67.21/04.89	74.25/05.25	82.94/02.96	63.90/07.58
w/o STPGC	65.80/04.33	60.13/06.91	70.98/05.18	74.00/04.06	82.58/02.27	64.76/09.32
NeuroH-TGL	72.19/06.31	57.81/13.21	70.49/12.30	78.50/05.27	84.35/03.88	72.61/07.13
Method	ACC	F1	AUC	ACC	F1	AUC
	NC vs. AD			MCI vs. AD		
w/o STPD	76.67/06.24	71.42/10.93	76.87/09.74	81.67/03.33	88.77/02.59	72.95/13.90
w/o STHW	77.08/05.97	70.36/10.05	74.69/08.72	78.06/04.72	86.63/02.69	67.01/13.40
w/o STPGC	77.91/06.47	72.77/08.43	77.35/09.11	82.08/06.47	89.00/03.65	68.66/20.30
NeuroH-TGL	81.50/06.73	72.12/11.49	83.01/10.12	86.67/04.08	92.46/02.25	76.01/16.12

Table 7: Ablation results (mean/std) of the proposed method on the PD dataset (%).

Method	ACC	F1	AUC	ACC	F1	AUC
	NC vs. TDPD vs. PGPD			NC vs. TDPD		
w/o STPD	65.00/03.00	55.15/04.44	64.57/06.56	86.25/10.38	87.06/09.10	83.54/15.76
w/o STHW	63.60/03.56	52.48/09.00	65.53/06.12	83.50/07.68	80.96/08.49	79.78/12.56
w/o STPGC	64.38/07.42	55.18/08.81	64.40/09.95	84.75/07.94	83.73/07.99	88.53/09.50
NeuroH-TGL	66.25/14.58	61.71/16.88	73.85/15.14	91.25/08.00	91.00/08.20	94.21/08.27
Method	ACC	F1	AUC	ACC	F1	AUC
	NC vs. PGPD			TDPD vs. PGPD		
w/o STPD	83.75/09.21	85.76/09.05	86.12/09.29	81.25/10.08	86.21/08.81	70.50/14.04
w/o STHW	79.33/09.13	82.47/77.58	79.93/11.89	78.75/08.80	81.96/15.01	69.91/29.99
w/o STPGC	80.50/09.13	82.64/09.05	81.16/09.47	76.00/08.00	78.23/09.10	70.65/13.63
NeuroH-TGL	87.17/08.40	89.38/05.29	88.42/12.64	83.75/09.76	86.80/08.57	82.91/15.43