REVISITING GRAPH NEURAL NETWORKS FOR BRAIN DISORDERS DIAGNOSIS: A NEUROSCIENCE-INFORMED PERSPECTIVE

Anonymous authors
Paper under double-blind review

ABSTRACT

Brain disorders exhibit abnormal neural activity patterns and present significant challenges for accurate and objective diagnosis. Boasting a strong capacity for learning relationships, Graph Neural Networks (GNNs) are increasingly adopted to model multi-channel neural signals such as Electroencephalography (EEG) and functional Near-Infrared Spectroscopy (fNIRS) for the diagnosis of brain disorders. However, existing methods suffer from two fundamental limitations from a neuroscientific perspective on brain activity analysis. First, existing methods are path-agnostic and fail to capture the directed propagation flow of neural signals and overlook essential causal mechanisms of brain activity. Second, multi-scale modeling introduces message-passing conflict that prevents effective semantic unification across hierarchical representations of brain regions and the entire brain network, and leads to unstable feature fusion. To address these challenges, we propose CPHU, a neuroscience-informed and path-sensitive framework for recognizing brain disorders. CPHU introduces **C**ausal **P**ath Signal Encoding (CPSE) to extract path-aware node representations via predefined and learnable propagation operators, and **H**ierarchical Brain **S**emantic **U**nification (HBSU) to resolve message-passing conflicts by aligning semantic features across intra-regional, global, and fused graphs. Extensive experiments demonstrate the effectiveness and superiority of CPHU in brain disorder diagnosis. The code is available at `https://anonymous.4open.science/r/CPHU-3686/`.

1 INTRODUCTION

Brain disorders (Fornito et al., 2015; Insel & Cuthbert, 2015; Pape et al., 2019; Riou et al., 2025) encompass Alzheimer's disease and Parkinson's disease (Akram et al., 2025b; AlSharabi et al., 2022; Goni et al., 2022; Morris et al., 2024), with cognitive impairment as the primary characteristic. As of 2021, Alzheimer's disease became one of the top ten neurological disorders contributing to death and disability, which renders these mental disorders a significant challenge for public health. Consequently, the early and precise identification of and intervention for such disorders are essential for the recovery of patients. Current clinical diagnosis primarily relies on standardized assessment scales, such as the Clinical Dementia Rating scale (CDR) (Morris, 1997; Huang et al., 2021), and the experience of clinicians, which depend on the subjective cooperation of subjects and are susceptible to environmental interference, creating a risk of misdiagnosis. Safe, portable, and non-invasive neuroimaging technologies such as electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) (Teplan et al., 2002; Cohen, 2017; Abadal et al., 2025; Bunce et al., 2006; Abdollahpour et al., 2025)enable auxiliary diagnostic tools based on objective neurological signals.

Given the high complexity and environmental sensitivity of neurological signals, manual analysis is time-consuming and susceptible to subjective bias, which severely limits analytical effectiveness. Machine learning methods such as support vector machines (Singh et al., 2023; Jiao et al., 2023) and random forests (Bäck & Tannemaat, 2021; Zheng et al., 2023) have gradually emerged due to strong pattern recognition capabilities. However, signals such as EEG and fNIRS involve complex multi-channel dependencies that hinder the further application of traditional machine learning methods. To overcome these limitations, Graph Neural Networks (GNNs) (Lu et al., 2024; Demir et al., 2021;

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

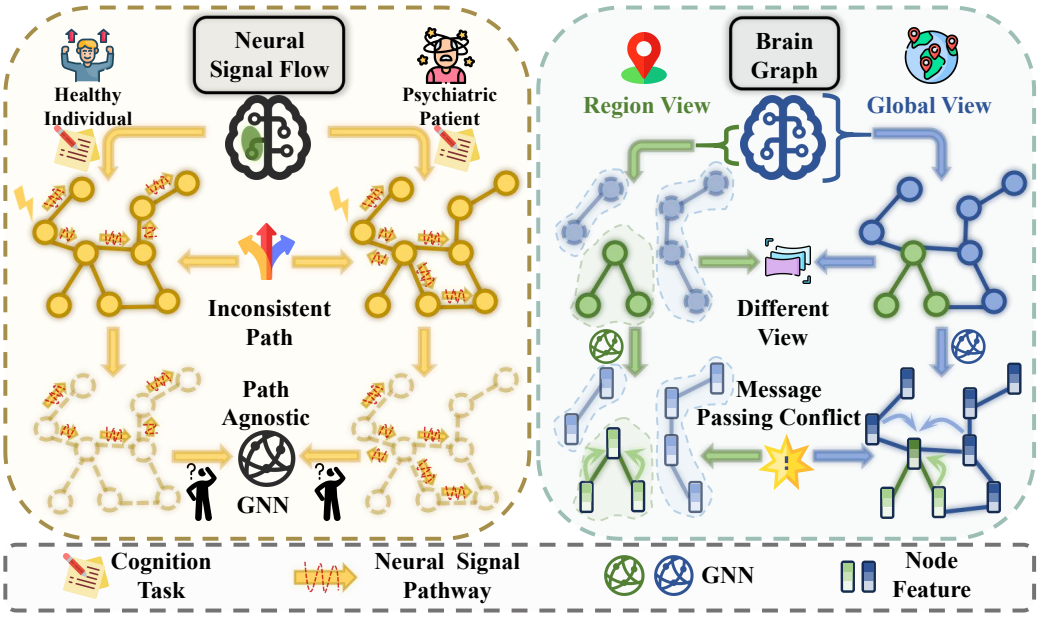


Figure 1: **Problem illustration.** (a) **Path Agnostic**: From a neurological perspective, brain activity involves the directed flow of neural signals across different regions and channels. Relying solely on static graph structures and statistical correlations is insufficient to capture this dynamic **neural signal flow**. (b) **Message Passing Conflict**: Neural activity can be analyzed from both regional and the entire brain scales, representing **different views**. Integrating these views in a GNN causes **message passing conflict**, hindering the effective unification of hierarchical representations across brain regions and the entire brain network.

Tang et al., 2021; Graña & Morais-Quilez, 2023) have gained prominence because of their capacity to capture neighborhood relationships, and current research (Guo et al., 2024; Li et al., 2024a; Liu et al., 2025) widely applies these networks to brain disorder diagnosis tasks. Nevertheless, existing methods still struggle to model the signal propagation mechanism during brain activity from a neurological perspective.

According to extensive neuroscientific research, the brain is a complex system of collaborative neuron clusters and regions rather than a collection of isolated processing units. During brain activity, neural signals transmit along specific pathways throughout the entire brain (Qaremohammadlou & Shamsollahi, 2025; Avvaru et al., 2021; Francisco-Vicencio et al., 2022). Capturing the **propagation paths** of neural signals among different channels is essential for characterizing the inherent patterns of neural activity in the brain. As shown in Fig. 2, patients exhibit abnormally enhanced propagation in frontal pathways (red, e.g., Fz/F3 $\rightarrow$ Fp1/Fp2) and a collapse of long-range causal driving from the temporal lobe (blue, e.g., T6 $\rightarrow$ Fp1/F3/Fz), confirming that healthy and disordered brains follow distinct propagation paths. Static statistical correlation cannot reveal such causal driving relationships between brain regions or channels, while tracking directional propagation paths can effectively distinguish healthy brains from impaired ones. In-depth analysis of propagation paths further distinguishes direct neural interaction from indirect information transmission, which restores the true flow of information within the brain network. This resolution of causal mechanisms captures the evolution process of propagation paths and provides biomarkers with more fundamental significance for the analysis of internal states of the brain than regional signal features. However, existing GNN-based methods do not consider the propagation paths of signals from a neurological perspective and are,

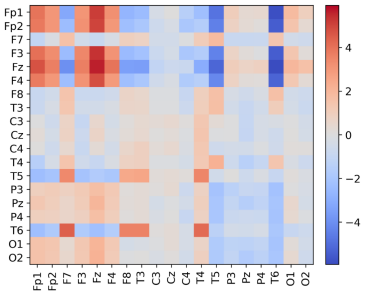


Figure 2: **EEG comparison between patients and healthy controls under cognitive activity.** Red and blue denote enhanced and weakened connectivity of patients relative to healthy controls. Electrodes Fp1–O2 follow the standard 10–20 system, where Fp, F, T, C, P, and O denote frontopolar, frontal, temporal, central, parietal, and occipital regions.

therefore, **path agnostic**. Based on the analysis above, we propose the following core question: **I)***How can we capture the **propagation paths** of neural signals?*

Secondly, neural activity in the brain can be regarded as **different views** across various scales within the spatial dimension (Kong et al., 2017; Wen et al., 2013; Li & Zhao, 2015). Regional-scale modeling captures fine electrophysiological patterns, reflecting synchronization features related to inhibitory processing and characterizing the activation of functional units through the power spectral density of spontaneous activity. In contrast, whole-brain modeling describes integration and coordination mechanisms between distal regions distributed across different spatial locations, revealing via cross-regional phase synchronization or causal information flow how functional networks reorganize during complex cognitive tasks. Jointly considering the energy distribution within regions and the topological evolution of the whole-brain network is critical for capturing multi-dimensional neural representations under different cognitive loads or functional states. However, existing GNN-based methods face the issue of **message passing conflict** when integrating multi-scale views. Based on this, this research proposes the following question: **II)***How can a **cross-view consistency** analysis be promoted to mitigate **message-passing conflicts**?*

To address the aforementioned issues, this paper proposes the **CPHU** framework. This framework utilizes structural encoding to extract path-aware node features that contain information regarding the propagation flow of electroencephalography signals. By aligning the most relevant semantic dimensions across different views, the framework achieves a unified representation of brain activity at different scales. Specifically, to address the first question **I)**, this research introduces **Causal Path Signal Encoding (CPSE)**. CPSE initializes a stochastic feature matrix to represent the spontaneous activity of neural signals and utilizes propagation matrices that explicitly define the connection pathways between graph nodes. Specifically, a predefined propagation operator models the static structural connections within local brain regions, and a learnable propagation operator adaptively captures dynamic functional interactions between distinct brain regions. By iteratively multiplying the stochastic matrix with these operators, the model mimics the step-by-step diffusion of signals along the network topology. This mechanism effectively generates path-aware representations. Subsequently, the framework concatenates original node features with path-aware node features and utilizes two graph neural networks to extract the embedding representation of this combined feature, thereby generating final node flow embeddings that contain rich propagation characteristics.

For Problem **II)**, this research proposes **Hierarchical Brain Semantic Unification (HBSU)**. First, this module integrates the knowledge representations of the model regarding the regional graph and the entire brain graph. During the fusion process, we select the top-k dimensions with the highest semantic correlation. Subsequently, the framework reduces the feature distance in the embedding space between the two original views, as well as between each original view and the fused view, across these critical dimensions. Finally, in the topology-aware pooling layer, this research utilizes semantic dimensions to calculate global scores and structural dimensions to calculate local scores. By combining global and local scores, the framework further achieves a deep unification of features from different views.

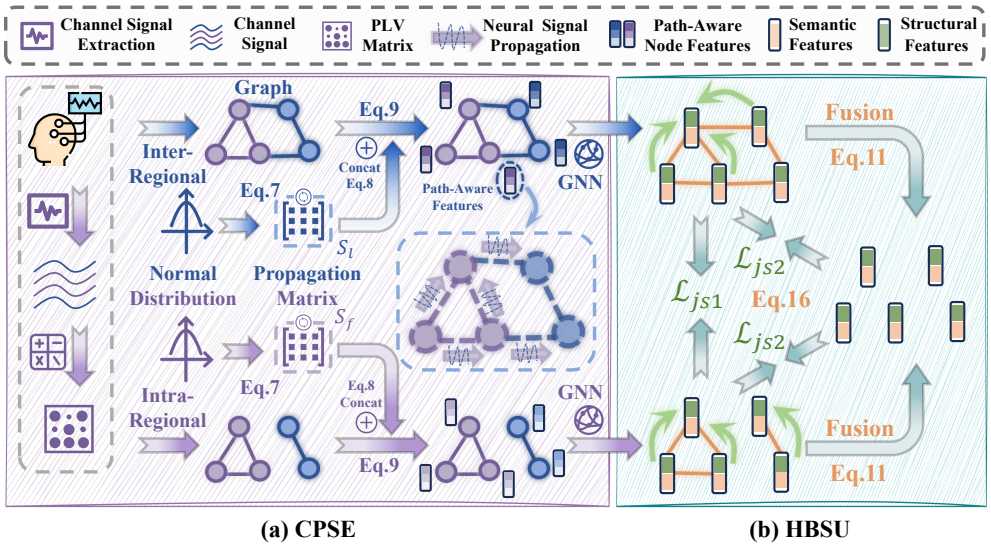
In summary, the contributions of this work are as follows:

- We identify two critical challenges in applying GNNs to brain disorder diagnosis from a neurological perspective.
- We propose CPHU, which is an effective automated diagnostic tool for brain disorders based on objective neurological signals.
- We conduct extensive experiments to verify the advanced performance of the proposed method.

2 Problem Statement

2.1 Preliminary

Notations. Define a graph as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where the vertex set $\mathcal{V}$ has a cardinality of $N = |\mathcal{V}|$ and the edge set is $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. The adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ represents the connectivity of the graph where $\mathbf{A}_{uv} = 1$ indicates the presence of an edge and $\mathbf{A}_{uv} = 0$ indicates the absence of an edge. The degree matrix $\mathbf{D}$ is a diagonal matrix with entries $D_{uu} = \sum_v \mathbf{A}_{uv}$. The normalized



(a) CPSE **(b) HBSU**

Figure 3: **Architecture illustration.** (a) We utilize predefined and learnable propagation operators to capture directed neural **signal flows** within and between brain regions and extract **path-aware** node representations. (b) This framework decomposes node features into semantic and structural components. By aligning **semantic representations** across **different views**, we resolve message passing conflicts between brain regions and the entire brain network to promote cross-view consistency in the analysis.

adjacency matrix is expressed as $\hat{\mathbf{A}} = \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$. Furthermore, the combinatorial Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$ and the symmetric normalized Laplacian matrix is $\tilde{\mathbf{L}} = \mathbf{I} - \hat{\mathbf{A}}$ where $\mathbf{I}$ denotes the identity matrix.

3 METHODOLOGY

Overview. In this section, we introduce the proposed CPHU framework, which comprises two core components: Causal Path Signal Encoding and Hierarchical Brain Semantic Unification. The overall workflow proceeds as follows. In Section section 3.1, the framework constructs intra-regional and global graphs based on phase synchronization to characterize the topological structure of the brain. Then we utilize fixed and learnable propagation matrices to model neurological signal flow, thereby generating path-aware node embeddings that capture directed neural transmission. In Section section 3.2, we address the message passing conflict issue by adaptively fusing multi-scale features and aligning semantic consistency across different views. Finally, the unified graph representation is fed into a classifier to identify brain disorders.

3.1 CAUSAL PATH SIGNAL ENCODING

Phase-based Graph Construction The proposed framework constructs graph representations using phase coherence between EEG channels to quantify cortical functional connectivity and spatial topology. The phase locking value serves as the primary metric to evaluate synchronization through the analysis of instantaneous phase differences between neural signals, expressed as

$$P_{i,j} = \left| \frac{1}{T} \sum_{t=1}^{T} \exp[\sqrt{-1} \cdot \phi_i(t) - \phi_j(t)] \right|, \tag{1}$$

where $P_{i,j}$ is the weight between the i-th and j-th electrodes, T is the number of sampling points in the observation window, and $\phi_i(t)$, $\phi_j(t)$ are the instantaneous phases of the signals at time t extracted via the Hilbert transform. A global graph $\mathcal{G}_{gl}$ models the large-scale integration of the brain via the phase coherence matrix $\mathbf{P} \in \mathbb{R}^{N \times N}$. The adjacency matrix $\mathbf{A}_{gl}$ is defined as

$$\mathbf{A}_{gl} = [P_{i,j}]_{i,j=1}^{N}, \tag{2}$$

where N is the total count of channels and $P_{i,j}$ denotes the element representing the connection strength. Furthermore, an intra-regional graph $\mathcal{G}_{in}$ focuses on localized neural features by excluding connections between electrodes of different anatomical regions. The adjacency matrix $\mathbf{A}_{in}$ is computed as

$$\mathbf{A}_{in} = \mathbf{A}_{gl} \odot \mathbf{M}_{in}, \tag{3}$$

where $\odot$ is the Hadamard product and $\mathbf{M}_{in} \in \{0,1\}^{N\times N}$ is a binary mask matrix where $M_{i,j}$ equals 1 if the i-th and j-th channels belong to the same functional region. In this study, the adjacency matrix is utilized as the initial node feature matrix for the subsequent graph processing operations.

Graph Propagation Matrix Construction The integration of learnable and fixed propagation matrices serves the motivation to capture complex propagation paths across brain regions and internal channels. The learnable propagation matrix S_l is designed to adaptively identify relationships between nodes by utilizing a graph convolutional network followed by a multi-head attention mechanism. Initially, the features of the global graph undergo a transformation through a graph convolutional network layer to extract high-level representations. The operation is defined as follows:

$$H_{out} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H_{in}W\right), \tag{4}$$

where H_{in} denotes the input feature matrix, $\tilde{A} = A + I$ represents the adjacency matrix with added self-loops, $\tilde{D}$ signifies the diagonal degree matrix of $\tilde{A}$, W is a trainable weight matrix, and σ is the non-linear activation function. Subsequently, a multi-head attention mechanism processes the refined features to evaluate dependencies across multiple representation subspaces. The learnable propagation matrix is obtained by calculating the average of the attention scores from all heads:

$$S_l = \frac{1}{M}\sum_{m=1}^{M}\text{Softmax}\left(\frac{Q_m K_m^\top}{\sqrt{d_k}}\right), \tag{5}$$

where M indicates the total number of attention heads, $Q_m = H_{out}W_Q^m$ and $K_m = H_{out}W_K^m$ are the query and key projections for the m-th head, and d_k represents the scaling factor determined by the dimension of the keys. This adaptive mechanism enables the model to focus on task-relevant propagation paths within the global graph structure.

In parallel, the fixed propagation matrix S_f is employed to represent the inherent static structure of the intra-regional graph. This matrix is derived directly from the adjacency matrix via a normalized Laplacian transformation to ensure numerical stability and preserve the underlying topology. The calculation process for the normalized Laplacian is formulated as:

$$S_f = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \tag{6}$$

where I denotes the identity matrix, A represents the adjacency matrix of the intra-regional graph, and D is the diagonal degree matrix with entries defined as $D_{ii} = \sum_j A_{ij}$. By combining these learnable and fixed matrices, the proposed framework effectively captures both the functional interactions and the fundamental structural properties of the graph across different scales.

Physiological Signal Flow Embedding To capture this neural signal transmission, a stochastic feature matrix is initialized to represent the potential diversity of neural activity and is then modulated by the propagation matrix through an iterative diffusion process. This interaction simulates the evolution of signal flow through the brain network across multiple hops. The formulation for generating the propagated stochastic features is defined as follows:

$$\begin{aligned} &R^{(0)} \sim \mathcal{N}(0,1), \\ &R^{(p)} = \begin{cases} \text{QR}(S \cdot R^{(p-1)}) & \text{if } p \mod w = 0 \\ S \cdot R^{(p-1)} & \text{otherwise} \end{cases} \end{aligned} \tag{7}$$

where $R^{(0)} \in \mathbb{R}^{N\times d_r}$ is sampled from the standard normal distribution, S represents the propagation matrix, P is the total number of propagation iterations, w is the orthogonalization interval, and $\text{QR}(\cdot)$ denotes QR decomposition for numerical stability. The path-aware feature matrix is then constructed by concatenating features from all propagation steps:

$$V_{fl} = [R^{(0)} \parallel R^{(1)} \parallel \cdots \parallel R^{(P)}] \tag{8}$$

where $V_{fl} \in \mathbb{R}^{N\times(P+1)d_r}$ captures the multi-hop signal diffusion patterns. These path-aware features are subsequently integrated with the original node attributes to distinguish between direct neural interactions and indirect information transmission. This fusion ensures that the model is no longer path-agnostic by explicitly incorporating the directional mechanisms of the brain. The concatenated features are then transformed through a learnable embedding layer into a high-level latent space:

$$F_o = \sigma([X_r \parallel V_{fl}]W_e + b), \tag{9}$$

where $F_o \in \mathbb{R}^{N\times d_h}$ is the final node representation, $X_r \in \mathbb{R}^{N\times d_f}$ is the original node feature matrix, W_e and b are the trainable projection weight matrix and bias vector, and σ is the non-linear activation function. This strategy yields enriched embeddings that provide biomarkers with more fundamental significance for the analysis of internal brain states.

3.2 Hierarchical Brain Semantic Unification

Cross-Scale Adaptive Fusion Following the generation of path-aware node features, the framework employs graph attention networks to extract high-order semantic information from both the intra-regional and global graphs. This step updates the node representations by aggregating information from neighbors with dynamic attention weights that prioritize significant structural connections. The specific feature update rule for the node aggregation phase is formulated as:

$$H^{\ell+1} = \sigma\left(\sum_{j\in\mathcal{N}_i} \alpha_{ij} W H_j^{\ell}\right) \tag{10}$$

where H^{ℓ} represents the node features at layer ℓ, $\mathcal{N}_i$ denotes the neighborhood set of node i, W is the learnable weight matrix for the linear transformation, α_{ij} indicates the attention coefficient measuring the importance of neighbor j to node i, and σ signifies the non-linear activation function. Upon obtaining the refined representations H_{in} and H_{gl} from their respective graph encoders, an adaptive fusion module integrates these multi-view features, automatically weighing the contribution of regional channel interactions versus whole-brain communications for each node. The fusion weights are computed through a non-linear gating network followed by a weighted aggregation:

$$Z_{fu} = \Gamma_{in} \odot H_{in} + \Gamma_{gl} \odot H_{gl} \tag{11}$$

where the operator $\odot$ denotes the element-wise product, and $\Gamma = [\Gamma_{in}, \Gamma_{gl}]$ represents the fusion coefficient matrix derived via

$$\Gamma = \text{Softmax}(\tanh([H_{in} \parallel H_{gl}]W_{\Gamma}^{1})W_{\Gamma}^{2}), \tag{12}$$

in which $\parallel$ implies concatenation and $\{W_{\Gamma}^{1}, W_{\Gamma}^{2}\}$ are trainable projection matrices. The fused graph representation Z_{fu} subsequently undergoes a topology-aware pooling layer to reduce dimensionality while preserving essential structural and semantic information. This pooling strategy calculates a selection score for each node by combining a local voting mechanism based on neighborhood topology and a global voting mechanism based on feature semantics. The final graph-level embedding is obtained by aggregating the node features according to these calculated scores and is then fed into a multi-layer perceptron(MLP) for binary classification:

$$\hat{y} = \text{MLP}\left(\sum_{i=1}^{N} \text{Softmax}(s_i) \cdot Z_{fu,i}\right) \tag{13}$$

where $\hat{y}$ is the predicted probability, s_i is the combined importance score of node i derived from topological and semantic voting, and the summation aggregates the weighted node features for the final decision layer.

Semantic Consistency Alignment To mitigate message passing conflicts and ensure a cohesive multi-dimensional neural representation, we implement a semantic feature selection mechanism. This process begins with the utilization of a focal loss function to supervise the classification task and identify the most discriminative feature dimensions through gradient sensitivity analysis:

$$\begin{aligned} \mathcal{L}_{fc} &= -\sum_{c=1}^{C} \alpha_c (1-\hat{p}_c)^{\gamma} \log(\hat{p}_c), \\ \mathcal{I}_s &= \text{Top}_k\left(\left|\frac{\partial \mathcal{L}_{fc}}{\partial \bar{H}_{fu}}\right|\right) \end{aligned} \tag{14}$$

In this formulation, $\hat{p}_c$ is the predicted probability for class c, while α_c and γ represent the class balancing and focusing hyperparameters, respectively. The term $\mathcal{I}_s$ contains the indices for the top-k feature dimensions with the highest gradient magnitude, where the gradient is evaluated on the pooled fused representation $\bar{H}_{fu}$ and averaged over the batch. To align these views while respecting the underlying topology, we first smooth the semantic features of both views along K-hop neighborhoods through random-walk propagation:

$$\tilde{H}_{in}^{\mathcal{I}_s} = \left(\hat{D}^{-1}\hat{A}\right)^K H_{in}^{\mathcal{I}_s}, \qquad \tilde{H}_{gl}^{\mathcal{I}_s} = \left(\hat{D}^{-1}\hat{A}\right)^K H_{gl}^{\mathcal{I}_s} \tag{15}$$

where $\hat{A} = A + I$ is the adjacency matrix with added self-loops, $\hat{D}$ is its diagonal degree matrix, and K denotes the number of propagation hops. To resolve conflicts between local and global information streams, the semantic alignment loss is partitioned into two components based on the Jensen-Shannon Divergence, denoted as D_{JS}. The first component, $\mathcal{L}_{js1}$, pulls the semantic features of the intra-regional and global graphs together. The second component, $\mathcal{L}_{js2}$, enforces consistency by aligning the features of both individual views with the fused representation:

$$\begin{aligned}
\mathcal{L}_{js1} &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} D_{JS}\left(\tilde{H}_{in,i}^{\mathcal{I}_s} \,\|\, \tilde{H}_{gl,j}^{\mathcal{I}_s}\right) \\
\mathcal{L}_{js2} &= \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Big(D_{JS}\left(\tilde{H}_{in,i}^{\mathcal{I}_s} \,\|\, H_{fu,j}^{\mathcal{I}_s}\right) \\
&\quad + D_{JS}\left(\tilde{H}_{gl,i}^{\mathcal{I}_s} \,\|\, H_{fu,j}^{\mathcal{I}_s}\right)\Big)
\end{aligned} \tag{16}$$

In these equations, $H_{\cdot,i}^{\mathcal{I}_s}$ signifies the feature vector of node i restricted to the semantic dimensions in $\mathcal{I}_s$, and the fused view H_{fu} remains unsmoothed. To optimize the network, the final objective function is constructed as a weighted combination of the classification loss and the dual semantic consistency terms:

$$\mathcal{L}_{tot} = \mathcal{L}_{fc} + \lambda_1\mathcal{L}_{js1} + \lambda_2\mathcal{L}_{js2} \tag{17}$$

The hyperparameters λ_1 and λ_2 balance the importance of the multi-view alignment against the primary binary classification task. This joint optimization strategy enables the precise capture of multi-dimensional neural representations while maintaining structural integrity across scales.

4 Experiments

4.1 Experimental Settings

Baselines. We compare CPHU with state-of-the-art approaches for brain time-series and graph learning: (1) K-Nearest Neighbors (KNN) and Random Forest (RF), two traditional machine learning models; (2) DMSGL (Pan et al., 2025), a spatiotemporal multi-view graph model for EEG emotion recognition; (3) Medformer (Wang et al., 2024b), a multi-granularity patching transformer; (4) LaBraM (Jiang et al., 2024), a large EEG representation model via vector quantization; (5) EvoBrain (Kotoge et al., 2025), a dynamic GNN for evolving brain networks; (6) EEGPT (Wang et al., 2024a), a self-supervised Transformer for universal EEG representations; (7) TarDiff (Deng et al., 2025), a target-oriented diffusion model for signal generation; (8) ALTER (Yu et al., 2024), a brain graph transformer with adaptive biased random walks; (9) STAGIN (Kim et al., 2021), a dynamic graph representation framework with timestamp-based attention; (10) MDGL (Ma et al., 2023), a multi-scale transformer GNN for dynamic brain connectomes.

Datasets. To evaluate CPHU, binary classification experiments are conducted on two public EEG datasets, ADFTD (Miltiadous et al., 2023) and APAVA (Escudero et al., 2006), plus a private fNIRS dataset, M-fNIRS, which differ in modality, channel count, and subject number, establishing a robust basis for evaluation across distinct brain signal characteristics. Detailed region definitions are provided in Appendix B

Evaluation Metrics. All tasks within this research are formulated as binary classification problems. Following established protocols for brain time-series learning (Deng et al., 2025), the performance evaluation utilizes two threshold-independent metrics: Area Under the Receiver Operating Characteristic curve (AUROC) and Area Under the Precision-Recall curve (AUPRC). The AUROC value

Table 1: Comparison with state-of-the-art methods on three datasets. AUROC: Area Under the Receiver Operating Characteristic curve (↑), AUPRC: Area Under the Precision-Recall Curve (↑). The best result is in **bold** and the second best is underlined.

Dataset	ADFTD		APAVA		M-fNIRS	
Method	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
KNN	50.75 ± 1.20	65.94 ± 2.87	41.77 ± 1.36	48.89 ± 4.57	52.83 ± 4.31	35.86 ± 2.79
Random Forest	59.18 ± 1.51	72.97 ± 1.36	44.76 ± 4.09	45.27 ± 5.14	75.74 ± 3.48	61.63 ± 3.91
STAGIN[NIPS21]	73.66 ± 2.98	83.99 ± 1.13	**92.97 ± 4.75**	95.40 ± 4.27	78.54 ± 2.98	66.56 ± 2.17
MDGL[IEEE23]	$\underline{78.75 \pm 2.87}$	$\underline{86.82 \pm 3.03}$	91.02 ± 4.10	$\underline{96.20 \pm 4.94}$	$\underline{84.38 \pm 3.69}$	$\underline{76.69 \pm 2.01}$
ALTER[NIPS24]	51.84 ± 0.45	67.09 ± 2.41	$\underline{92.70 \pm 3.10}$	95.31 ± 3.44	69.77 ± 1.48	59.42 ± 3.19
LaBraM[ICLR24]	51.37 ± 0.97	66.13 ± 2.30	50.87 ± 1.77	35.31 ± 4.10	50.08 ± 0.14	33.68 ± 3.54
EEGPT[NeurIPS24]	50.28 ± 0.73	63.68 ± 2.36	50.58 ± 1.03	65.03 ± 2.11	51.56 ± 5.86	42.01 ± 5.56
Medformer[NeurIPS24]	75.61 ± 3.73	70.77 ± 5.37	85.62 ± 4.95	83.30 ± 5.43	75.92 ± 2.89	73.50 ± 2.13
DMSGL[ACM MM25]	54.55 ± 5.16	69.61 ± 3.36	67.02 ± 2.99	69.78 ± 4.71	52.48 ± 0.76	32.88 ± 5.30
EvoBrain[NeurIPS25]	50.20 ± 0.96	67.60 ± 0.57	86.43 ± 3.39	92.97 ± 4.69	59.32 ± 3.26	36.72 ± 3.76
TarDiff[KDD25]	49.89 ± 0.23	65.73 ± 2.22	75.82 ± 4.14	78.56 ± 3.86	48.79 ± 1.37	29.97 ± 2.79
Ours	**79.59 ± 3.19**	**86.96 ± 2.47**	90.06 ± 2.65	**96.54 ± 4.14**	**87.57 ± 1.18**	**80.12 ± 1.58**

represents the integral of the True Positive Rate with respect to the False Positive Rate across all possible thresholds. These rates are defined as:

$$\begin{aligned} TPR &= \frac{TP}{TP+FN}, \quad FPR = \frac{FP}{FP+TN} \\ AUROC &= \int_0^1 TPR(FPR)dFPR \end{aligned} \tag{18}$$

The AUPRC value is determined by the integral of Precision relative to Recall:

$$\begin{aligned} Prec &= \frac{TP}{TP+FP}, \quad Rec = \frac{TP}{TP+FN} \\ AUPRC &= \int_0^1 Prec(Rec)dRec \end{aligned} \tag{19}$$

Implementation Details. All datasets are formulated as binary classification, with healthy controls assigned to class 0. ADFTD combines Alzheimer's disease and frontotemporal dementia patients as the positive class, APAVA uses Alzheimer's disease patients, and M-fNIRS aggregates all disease categories. To prevent data leakage, we adopt a subject-independent setting: following prior work (Deng et al., 2025), subject identifiers are randomly shuffled and split into training, validation, and test sets at an 8:1:1 ratio, keeping all samples of one subject in the same partition (with stratified splitting within each diagnostic category for ADFTD). As M-fNIRS samples are not organized by subject, we pool and shuffle them with a random seed before the same 8:1:1 sample-level split. All results are averaged over five random seeds.

4.2 Performance Comparison

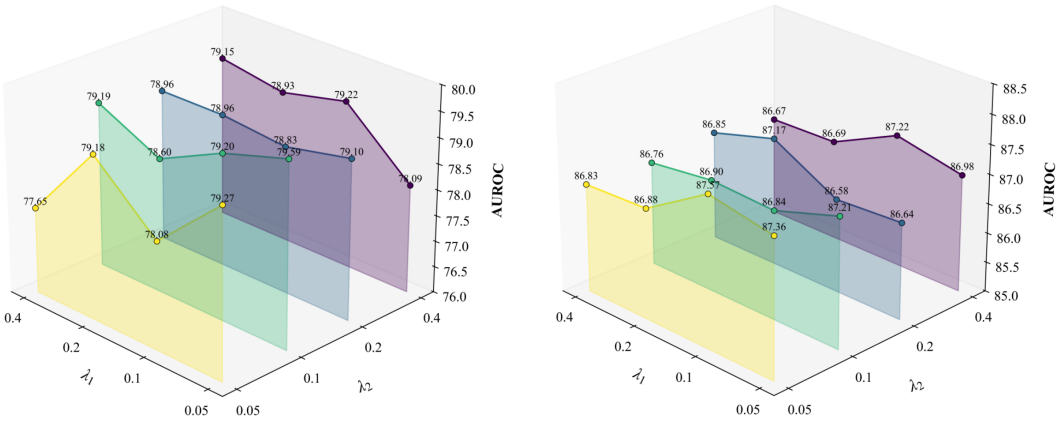
table 1 presents the performance comparison between CPHU and several state-of-the-art baseline models across three evaluation datasets. The results demonstrate that our approach achieves superior performance in terms of $AUROC$ and $AUPRC$ metrics under most experimental settings.

4.3 Ablation Study

Effects of Key Components in CPSE. To verify the efficacy of the path-aware features within CPSE, we conducted ablation studies on the ADFTD and M-fNIRS datasets by feeding raw node features directly into the graph neural network. Results in table 2 confirm that the application of CPSE consistently enhances both AUROC and AUPRC.

Table 2: Ablation study of CPSE and HBSU on ADFTD and M-fNIRS datasets.

CPSE	HBSU	ADFTD		M-fNIRS	
		AUROC	AUPRC	AUROC	AUPRC
✗	✓	78.33 ± 3.39	85.66 ± 4.09	84.16 ± 1.94	77.74 ± 0.17
✓	✗	76.20 ± 4.67	83.89 ± 4.22	87.43 ± 2.87	77.56 ± 5.09
✗	✗	75.84 ± 2.45	83.60 ± 4.04	83.76 ± 0.72	77.08 ± 1.05
✓	✓	$\mathbf{79.59 \pm 3.19}$	$\mathbf{86.96 \pm 2.47}$	$\mathbf{87.57 \pm 1.18}$	$\mathbf{80.12 \pm 1.58}$



(a) ADFTD (b) M-fNIRS

Figure 4: Hyper-parameter Sensitivity Analysis. Performance of CPHU under varying values of λ_1 and λ_2 on the ADFTD and M-fNIRS datasets.

Effects of Key Components in HBSU. To evaluate the ability of HBSU to mitigate message-passing conflicts, we removed λ_{js1} and λ_{js2} from the objective function in eq. (17). The performance gains documented in table 2 validate the contribution of HBSU toward achieving optimal classification accuracy.

4.4 Hyper-parameter Study

We compared the performance of the model across various values for the hyperparameters λ_1 and λ_2 in eq. (17). As illustrated in fig. 4, CPHU maintains competitive performance across a broad spectrum of λ_1 and λ_2.

5 Conclusion

In this paper, we investigated two critical challenges in graph neural network-based diagnosis of brain disorders from a neurological perspective: path agnostic and message passing conflict. Existing methodologies fail to effectively capture the directed propagation flow of neural signals and suffer from message-passing conflicts during multi-scale modeling, which hinders the enhancement of model performance. To address these issues, we proposed CPHU, a neuroscience-informed and path-sensitive framework for the recognition of brain disorders. The framework first introduces the CPSE module to extract path-aware features. Subsequently, it integrates the HBSU module to resolve message-passing conflicts by aligning semantic features among intra-regional, global, and fused graphs. Extensive experiments on multiple EEG and fNIRS datasets validate the superiority of CPHU over state-of-the-art methods.

AI USE STATEMENT

We used generative AI tools to polish the writing of this paper, without altering our intended meaning. We reviewed all AI-assisted text and take full responsibility for the final content of this work.

REFERENCES

Sergi Abadal, Pablo Galván, Alberto Mármol, Nadia Mammone, Cosimo Ieracitano, Michele Lo Giudice, Alessandro Salvini, and Francesco Carlo Morabito. Graph neural networks for electroencephalogram analysis: Alzheimer's disease and epilepsy use cases. *Neural Networks*, 181: 106792, 2025.

Neda Abdollahpour, Abhinav Kakaraparthi, Chakradhar Sai Valluru, and N Sertac Artan. Advancing functional near-infrared spectroscopy (fnirs) with transformers: a prisma-driven systematic review. *IEEE Access*, 2025.

Junaid Akram, Walayat Hussain, Rutvij H Jhaveri, Rajkumar Singh Rathore, and Ali Anaissi. Dynamic gnn-based multimodal anomaly detection for spatial crowdsourcing drone services. *Digital Communications and Networks*, 2025a.

Samina Akram, Muhammad Amjad Iqbal, Muhammad Rashid, Muhammad Shahid Bhatti, and Benish Fida. A hybrid deep learning framework for early-stage alzheimer's disease classification from neuro-imaging biomarkers. *IEEE Access*, 2025b.

Khalil AlSharabi, Yasser Bin Salamah, Akram M Abdurraqeeb, Majid Aljalal, and Fahd A Alturki. Eeg signal processing for alzheimer's disorders using discrete wavelet transform and machine learning approaches. *IEEE Access*, 10:89781–89797, 2022.

Sandeep Avvaru, Noam Peled, Nicole R Provenza, Alik S Widge, and Keshab K Parhi. Region-level functional and effective network analysis of human brain during cognitive task engagement. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:1651–1660, 2021.

Claudio Babiloni, Robert J Barry, Erol Başar, Katarzyna J Blinowska, Andrzej Cichocki, Wilhelmus HIM Drinkenburg, Wolfgang Klimesch, Robert T Knight, Fernando Lopes da Silva, Paul Nunez, et al. International federation of clinical neurophysiology (ifcn)–eeg research workgroup: Recommendations on frequency and topographic analysis of resting state eeg rhythms. part 1: Applications in clinical research studies. *Clinical Neurophysiology*, 131(1):285–307, 2020.

THW Bäck and MR Tannemaat. Machine learning for automated eeg-based biomarkers of cogniti e impairment during deep brain stimulation screening in patients ith parkinson's disease. *Clinical Neurophysiology*, 132:1041–1048, 2021.

Guy Bar-Shalom, Yam Eitan, Fabrizio Frasca, and Haggai Maron. A flexible, equivariant framework for subgraph gnns via graph products and graph coarsening. *Advances in Neural Information Processing Systems*, 37:101168–101222, 2024.

Deyu Bo, Chuan Shi, Lele Wang, and Renjie Liao. Specformer: Spectral graph neural networks meet transformers. *arXiv preprint arXiv:2303.01028*, 2023.

Khac-Hoai Nam Bui, Jiho Cho, and Hongsuk Yi. Spatial-temporal graph neural network for traffic forecasting: An overview and open research issues. *Applied Intelligence*, 52(3):2763–2774, 2022.

Scott C Bunce, Meltem Izzetoglu, Kurtulus Izzetoglu, Banu Onaral, and Kambiz Pourrezaei. Functional near-infrared spectroscopy. *IEEE engineering in medicine and biology magazine*, 25(4): 54–62, 2006.

Haibo Chen, Xin Wang, Zeyang Zhang, Haoyang Li, Ling Feng, and Wenwu Zhu. Autogfm: Automated graph foundation model with adaptive architecture customization. In *Forty-second International Conference on Machine Learning*, 2025.

Michael X Cohen. Where does eeg come from and what does it mean? *Trends in neurosciences*, 40 (4):208–218, 2017.

Andac Demir, Toshiaki Koike-Akino, Ye Wang, Masaki Haruna, and Deniz Erdogmus. Eeg-gnn: Graph neural networks for classification of electroencephalogram (eeg) signals. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 1061–1067. IEEE, 2021.

Andac Demir, Toshiaki Koike-Akino, Ye Wang, and Deniz Erdoğmuş. Eeg-gat: graph attention networks for classification of electroencephalogram (eeg) signals. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 30–35. IEEE, 2022.

Bowen Deng, Chang Xu, Hao Li, Yu-hao Huang, Min Hou, and Jiang Bian. Tardiff: Target-oriented diffusion guidance for synthetic electronic health record time series generation. In *SIGKDD*, pp. 474–485, 2025.

Federico Errica, Henrik Christiansen, Viktor Zaverkin, Takashi Maruyama, Mathias Niepert, and Francesco Alesiani. Adaptive message passing: A general framework to mitigate oversmoothing, oversquashing, and underreaching. *arXiv preprint arXiv:2312.16560*, 2023.

J Escudero, Daniel Abásolo, Roberto Hornero, Pedro Espino, and Miguel López. Analysis of electroencephalograms in alzheimer's disease patients with multiscale entropy. *Physiological measurement*, pp. 1091, 2006.

Alex Fornito, Andrew Zalesky, and Michael Breakspear. The connectomics of brain disorders. *Nature reviews neuroscience*, 16(3):159–172, 2015.

Manuel A Francisco-Vicencio, Fernando Góngora-Rivera, Xochitl Ortiz-Jimenez, and Dulce Martinez-Peon. Sustained attention variation monitoring through eeg effective connectivity. *Biomedical Signal Processing and Control*, 76:103650, 2022.

Johannes Gasteiger, Aleksandar Bojchevski, and Stephan Günnemann. Predict then propagate: Graph neural networks meet personalized pagerank. *arXiv preprint arXiv:1810.05997*, 2018.

Maria Goni, Simon B Eickhoff, Mehran Sahandi Far, Kaustubh R Patil, and Juergen Dukart. Smartphone-based digital biomarkers for parkinson's disease in a remotely-administered setting. *IEEE access*, 10:28361–28384, 2022.

Manuel Graña and Igone Morais-Quilez. A review of graph neural networks for electroencephalography data analysis. *Neurocomputing*, 562:126901, 2023.

Yi Guo, Chao Tang, Hao Wu, and Badong Chen. Gnn-based multi-source domain prototype representation for cross-subject eeg emotion recognition. *Neurocomputing*, 609:128445, 2024.

Hui-Chuan Huang, Yu-Min Tseng, Yi-Chun Chen, Pin-Yuan Chen, and Hsiao-Yean Chiu. Diagnostic accuracy of the clinical dementia rating scale for detecting mild cognitive impairment and dementia: A bivariate meta-analysis. *International journal of geriatric psychiatry*, 36(2):239–251, 2021.

Thomas R Insel and Bruce N Cuthbert. Brain disorders? precisely. *Science*, 348(6234):499–500, 2015.

Ju Jia, Renjie Li, Cong Wu, Siqi Ma, Lina Wang, and Robert H Deng. Sigfinger: A subtle and interactive gnn fingerprinting scheme via spatial structure inference perturbation. *IEEE Transactions on Dependable and Secure Computing*, 22(4):3629–3646, 2025.

Wei-Bang Jiang, Liming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous eeg data in bci. In *International Conference on Learning Representations*, volume 2024, pp. 16405–16426, 2024.

Bin Jiao, Rihui Li, Hui Zhou, Kunqiang Qing, Hui Liu, Hefu Pan, Yanqin Lei, Wenjin Fu, Xiaoan Wang, Xuewen Xiao, et al. Neural biomarker diagnosis and prediction to mild cognitive impairment and alzheimer's disease using eeg technology. *Alzheimer's research & therapy*, 15(1):32, 2023.

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.

Wanzeng Kong, Zhanpeng Zhou, Bei Jiang, Fabio Babiloni, and Gianluca Borghini. Assessment of driving fatigue based on intra/inter-region phase synchronization. *Neurocomputing*, 219:474–482, 2017.

Rikuto Kotoge, Zheng Chen, Tasuku Kimura, Yasuko Matsubara, Takufumi Yanagisawa, Haruhiko Kishima, and Yasushi Sakurai. Evobrain: Dynamic multi-channel eeg graph modeling for time-evolving brain networks. In *NeurIPS*, 2025.

Gaotang Li, Marlena Duda, Xiang Zhang, Danai Koutra, and Yujun Yan. Interpretable sparsification of brain graphs: Better practices and effective designs for graph neural networks. In *SIGKDD*, pp. 1223–1234, 2023.

Guohao Li, Matthias Müller, Bernard Ghanem, and Vladlen Koltun. Training graph neural networks with 1000 layers. In *International conference on machine learning*, pp. 6437–6449. PMLR, 2021.

Ling Li and Dandan Zhao. Age-related inter-region eeg coupling changes during the control of bottom–up and top–down attention. *Frontiers in aging neuroscience*, 7:223, 2015.

Ting Li, Chuantao Yin, Yanmei Chai, Hui Chen, Wenge Rong, and Yuanxin Ouyang. An rsb-gnn-based eeg approach for exploring students' affective states in e-learning. In *2024 IEEE Frontiers in Education Conference (FIE)*, pp. 1–9. IEEE, 2024a.

Yang Li, Yang Yang, Qinghe Zheng, Yunxia Liu, Hongjun Wang, Shangling Song, and Penghui Zhao. Dynamical graph neural network with attention mechanism for epilepsy detection using single channel eeg. *Medical & Biological Engineering & Computing*, 62(1):307–326, 2024b.

Zhengdao Li, Kai Hwang, Keqin Li, Jie Wu, and Tongkai Ji. Graph-generative neural network for eeg-based epileptic seizure detection via discovery of dynamic brain functional connectivity. *Scientific reports*, 12(1):18998, 2022.

Renjie Liao, Zhizhen Zhao, Raquel Urtasun, and Richard S Zemel. Lanczosnet: Multi-scale deep graph convo-lutional networks. In *7th International Conference on Learning Representations, ICLR 2019*, 2019.

Guanting Liu, Ying Yan, Jun Cai, Edmond Qi Wu, Shencun Fang, Adrian David Cheok, and Aiguo Song. Gcd: Graph contrastive denoising module for gnns in eeg classification. *Expert Systems with Applications*, 265:126013, 2025.

Ane López-González, Rajanikant Panda, Adrián Ponce-Alvarez, Gorka Zamora-López, Anira Escrichs, Charlotte Martial, Aurore Thibaut, Olivia Gosseries, Morten L Kringelbach, Jitka Annen, et al. Loss of consciousness reduces the stability of brain hubs and the heterogeneity of brain dynamics. *Communications biology*, 4(1):1037, 2021.

Jielong Lu, Zhihao Wu, Luying Zhong, Zhaoliang Chen, Hong Zhao, and Shiping Wang. Generative essential graph convolutional network for multi-view semi-supervised classification. *IEEE transactions on multimedia*, 26:7987–7999, 2024.

Yunling Ma, Qianqian Wang, Liang Cao, Long Li, Chaojun Zhang, Lishan Qiao, and Mingxia Liu. Multi-scale dynamic graph learning for brain disorder detection with functional mri. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:3501–3512, 2023.

Andreas Miltiadous, Katerina D Tzimourta, Theodora Afrantou, Panagiotis Ioannidis, Nikolaos Grigoriadis, Dimitrios G Tsalikakis, Pantelis Angelidis, Markos G Tsipouras, Euripidis Glavas, Nikolaos Giannakeas, et al. A dataset of scalp eeg recordings of alzheimer's disease, frontotemporal dementia and healthy subjects from routine eeg. *Data*, pp. 95, 2023.

Shibing Mo, Kai Wu, Qixuan Gao, Xiangyi Teng, and Jing Liu. Autosgnn: automatic propagation mechanism discovery for spectral graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 19493–19502, 2025.

Huw R Morris, Maria Grazia Spillantini, Carolyn M Sue, and Caroline H Williams-Gray. The pathogenesis of parkinson's disease. *The Lancet*, 403(10423):293–304, 2024.

John C Morris. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the alzheimer type. *International psychogeriatrics*, 9:173–176, 1997.

Lyle Muller, Frédéric Chavane, John Reynolds, and Terrence J Sejnowski. Cortical travelling waves: mechanisms and computational principles. *Nature Reviews Neuroscience*, 19(5):255–268, 2018.

Yue Pan, Cunbo Li, Peiyang Li, Fali Li, Feng Wan, Dezhong Yao, Zehong Cao, and Peng Xu. Real-time eeg emotion recognition from dynamic mixed spatiotemporal graph learning. In *ACMMM*, pp. 5697–5706, 2025.

Katrin Pape, Ryad Tamouza, Marion Leboyer, and Frauke Zipp. Immunoneuropsychiatry—novel perspectives on brain disorders. *Nature Reviews Neurology*, 15(6):317–328, 2019.

Mohaddese Qaremohammadlou and Mohammad Bagher Shamsollahi. Investigating brain connectivity and information flow in mental workload using eeg and fnirs integration. *arXiv preprint arXiv:2511.02457*, 2025.

Taoyang Qin, CHEN Ke-Jia, and Zheng Liu. Enhancing spectral gnns: From topology and perturbation perspectives. In *Forty-second International Conference on Machine Learning*, 2025.

Khadijeh Raeisi, Mohammad Khazaei, Pierpaolo Croce, Gabriella Tamburro, Silvia Comani, and Filippo Zappasodi. A graph convolutional neural network for the automated detection of seizures in the neonatal eeg. *Computer methods and programs in biomedicine*, 222:106950, 2022.

Aurélien Riou, Aline Broeglin, and Amandine Grimm. Mitochondrial transplantation in brain disorders: Achievements, methods, and challenges. *Neuroscience & Biobehavioral Reviews*, 169: 105971, 2025.

Shan H Siddiqi, Konrad P Kording, Josef Parvizi, and Michael D Fox. Causal mapping of human brain function. *Nature reviews neuroscience*, 23(6):361–375, 2022.

Arun Singh, Rachel C Cole, Arturo I Espinoza, Jan R Wessel, James F Cavanagh, and Nandakumar S Narayanan. Evoked mid-frontal activity predicts cognitive dysfunction in parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 94(11):945–953, 2023.

Siyi Tang, Jared A Dunnmon, Khaled Saab, Xuan Zhang, Qianying Huang, Florian Dubost, Daniel L Rubin, and Christopher Lee-Messer. Self-supervised graph neural networks for improved electroencephalographic seizure analysis. *arXiv preprint arXiv:2104.08336*, 2021.

Michal Teplan et al. Fundamentals of eeg measurement. *Measurement science review*, 2(2):1–11, 2002.

Neeraj Wagh and Yogatheesan Varatharajah. Eeg-gcnn: Augmenting electroencephalogram-based neurological disease diagnosis using a domain-guided graph convolutional neural network. In *Machine Learning for Health*, pp. 367–378. PMLR, 2020.

Guangyu Wang, Wenchao Liu, Yuhong He, Cong Xu, Lin Ma, and Haifeng Li. Eegpt: Pretrained transformer for universal and reliable representation of eeg signals. *NeurIPS*, pp. 39249–39280, 2024a.

Shiping Wang, Zhaoliang Chen, Shide Du, and Zhouchen Lin. Learning deep sparse regularizers with applications to multi-view clustering and semi-supervised classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9):5042–5055, 2021.

Yihe Wang, Nan Huang, Taida Li, Yujun Yan, and Xiang Zhang. Medformer: A multi-granularity patching transformer for medical time-series classification. *NeurIPS*, pp. 36314–36341, 2024b.

Xuyun Wen, Xia Wu, Rui Li, Adam S Fleisher, Eric M Reiman, Xiaotong Wen, Kewei Chen, and Li Yao. Alzheimer's disease-related changes in regional spontaneous brain activity levels and inter-region interactions in the default mode network. *brain research*, 1509:58–65, 2013.

Miles A Whittington, Roger D Traub, Nancy Kopell, Bard Ermentrout, and Eberhard H Buhl. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *International journal of psychophysiology*, 38(3):315–336, 2000.

Yujun Yan, Jiong Zhu, Marlena Duda, Eric Solarz, Chandra Sripada, and Danai Koutra. Groupinn: Grouping-based interpretable neural network for classification of limited, noisy brain data. In *SIGKDD*, pp. 772–782, 2019.

Yujun Yan, Milad Hashemi, Kevin Swersky, Yaoqing Yang, and Danai Koutra. Two sides of the same coin: Heterophily and oversmoothing in graph convolutional neural networks. In *2022 IEEE International Conference on Data Mining (ICDM)*, pp. 1287–1292. IEEE, 2022.

Mingqi Yang, Yanming Shen, Rui Li, Heng Qi, Qiang Zhang, and Baocai Yin. A new perspective on the effects of spectrum in graph neural networks. In *International Conference on Machine Learning*, pp. 25261–25279. PMLR, 2022.

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.

Youbing Zeng, Jiaying Lin, Zhuoshuo Li, Zehui Xiao, Chen Wang, Xinting Ge, Cheng Wang, Gui Huang, and Mengting Liu. Adaptive node feature extraction in graph-based neural networks for brain diseases diagnosis using self-supervised learning. *NeuroImage*, 297:120750, 2024.

Tong Zhang, Tingting Hu, Mengqi Wu, Zihua Xu, and CL Philip Chen. Acm-gnn: Adaptive cluster-oriented modularity graph neural network for eeg depression detection. *IEEE Transactions on Computational Social Systems*, 2025.

Xiaowei Zheng, Bozhi Wang, Hao Liu, Wencan Wu, Jiamin Sun, Wei Fang, Rundong Jiang, Yajie Hu, Cheng Jin, Xin Wei, et al. Diagnosis of alzheimer's disease via resting-state eeg: integration of spectrum, complexity, and synchronization signal features. *Frontiers in aging neuroscience*, 15:1288295, 2023.

Meiqi Zhu, Xiao Wang, Chuan Shi, Houye Ji, and Peng Cui. Interpreting and unifying graph neural networks with an optimization framework. In *Proceedings of the web conference 2021*, pp. 1215–1226, 2021.

APPENDIX

A Related Work

A.1 Graph Neural Network

Graph Neural Networks (GNNs) enable the learning of powerful node representations via recursive message passing among neighboring nodes (Wang et al., 2021; Yan et al., 2022; Raeisi et al., 2022; Li et al., 2024b). Existing research typically classifies GNN models into spatial-domain (Bui et al., 2022; Jia et al., 2025; Akram et al., 2025a) and spectral-domain (Gasteiger et al., 2018; Zhu et al., 2021; Liao et al., 2019; Yang et al., 2022) approaches. Spatial-domain methods define graph convolutions by aggregating feature information within local neighborhoods. Within this category, AMP(Errica et al., 2023) utilizes variational inference to adaptively regulate message passing depth for the mitigation of over-smoothing and over-squashing. RevGNN(Li et al., 2021) implements reversible residual connections to enable the efficient training of architectures with one thousand layers. AutoGFM(Chen et al., 2025) disentangles graph patterns to customize optimal architectures across diverse domains. Additionally, CS-GNN(Bar-Shalom et al., 2024) associates subgraphs with node clusters to enhance computational efficiency while maintaining equivariance. In contrast, spectral-domain methods utilize the graph Laplacian matrix to define filters in the frequency domain. Within this category, PSL-GNN(Qin et al., 2025) employs a perturbed sheaf Laplacian to increase eigenvalue diversity and enhance model expressiveness. Specformer(Bo et al., 2023) incorporates Transformer self-attention into the spectral domain to enable global set-to-set filtering. Furthermore, AutoSGNN(Mo et al., 2025) leverages large language models and evolutionary strategies to automatically discover optimal propagation mechanisms for diverse graph types. Despite these advancements, the direct application of these methodologies to the detection of brain diseases often results in performance bottlenecks because the standard architectures ignore the underlying biological signal representations.

A.2 Electroencephalogram Signals

Electroencephalogram signals provide a non-invasive modality for monitoring neural activity with high temporal resolution. These signals are categorized into intra-region and inter-region components based on the scope of neural communication. Intra-region signals(Whittington et al., 2000; López-González et al., 2021; Muller et al., 2018) represent regional electrical activities within a specific brain area that do not communicate with channels in other regions. These periodic rhythms emerge from the synchronous discharge of large neuronal populations regulated by inhibitory networks, providing a time-gating framework for cognitive processing and serving as vital indicators of functional states. Conversely, inter-region signals(Babiloni et al., 2020; Siddiqi et al., 2022) reflect the functional integration and information coupling across the entire brain. These signals follow structured and heritable cyclical patterns that coordinate the activation sequence of functional networks for efficient cognitive performance. Such connectivity reveals the information exchange and causal interactions within brain networks, providing essential neurophysiological markers for the diagnosis of neuropsychiatric disorders. Although these signals contain rich information, manual analysis is inefficient and susceptible to subjective bias. Because the spatial arrangement of electrodes naturally defines a graph, Graph Neural Networks offer a robust framework for modeling topological dependencies.

A.3 GNN for Brain Disorders Recognition

GNN for Brain Disorders(Yan et al., 2019; Li et al., 2023) combines the non-invasive nature of electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) with the ability of GNN to model complex graph-structured data. Existing methods can be categorized into Prior-based or Adaptive Structure learning according to graph construction dependency. Prior-based models determine topology before training.EEG-GCNN (Wagh & Varatharajah, 2020) integrates physical coordinates with coherence connectivity to detect abnormal signals. Additionally, the framework(Zeng et al., 2024) introduces a task-induced self-supervised learning strategy to adaptively extract discriminative node features from raw EEG. Conversely, Adaptive Structure models optimize connections dynamically during training. ACM-GNN (Zhang et al., 2025) introduces channel-pair contrastive

learning to cluster channels for cross-subject detection. EEG-GAT (Demir et al., 2022) applies self-attention to compute interaction weights and remove manual topology dependence. Furthermore, GGN(Li et al., 2022) dynamically generates brain functional connectivity graphs from EEG signals to capture evolving spatio-temporal patterns. Despite these achievements, existing methods overlook neurological information. This oversight results in Path Agnostic and Message Passing Conflict problems. This study addresses these issues through a Physiology-Informed Perspective.

B Datasets Details

ADFTD. This publicly available EEG dataset contains 19-channel recordings from 88 participants categorized into three groups: 36 individuals with Alzheimer's disease, 23 with frontotemporal dementia, and 29 healthy controls. The raw data collection utilized a sampling rate of 500 Hz. Each subject contributed a single trial with mean durations of 13.5 minutes for the Alzheimer's disease group, 12 minutes for the frontotemporal dementia group, and 13.8 minutes for the healthy control group. We applied a bandpass filter between 0.5 and 45 Hz to each trial before downsampling the signals to 256 Hz. The processed trials were segmented into 1-second non-overlapping windows of 256 timestamps, resulting in 69,752 samples after excluding segments shorter than 1 second. The 19 electrodes are organized into 5 functional brain regions representing the frontopolar-frontal, lateral frontal-temporal, central, posterior temporal, and parieto-occipital areas.

APAVA. The APAVA dataset is a public EEG resource comprising 16-channel time series from 23 subjects, including 12 patients diagnosed with Alzheimer's disease and 11 healthy control participants. The data structure includes 30.0 ± 12.5 trials per subject on average. Every trial consists of a 5-second sequence with 1,280 timestamps. Following a standard scaling procedure for each trial, we implemented a sliding window to generate 9 half-overlapping samples per trial, where each sample contains 256 timestamps over a 1-second duration. This preprocessing pipeline produced a total of 5,967 samples, each associated with a unique subject ID. The 16 channels are distributed across 5 brain regions covering the prefrontal, frontal-temporal, temporal, central, and parieto-occipital zones.

M-fNIRS. M-fNIRS is a proprietary functional near-infrared spectroscopy dataset designed for the diagnosis of mental disorders, with public release planned upon the acceptance of this work. The collection includes 1,090 subjects across five categories: healthy controls, schizophrenia, bipolar I disorder, bipolar II disorder, and major depressive disorder. Detailed class distributions are available in Appendix D. By measuring oxygenated hemoglobin concentration changes, fNIRS provides enhanced spatial resolution compared to EEG, particularly within the prefrontal cortex. Each recording contains 53 measurement channels and 3,400 temporal samples. The 53 channels are organized into 10 anatomical regions, with 5 regions in each hemisphere representing the premotor and supplementary motor area, Broca area, frontal eye field, dorsolateral prefrontal cortex, and frontopolar area.

C Brain Region Definitions

Table 3: Overview of brain regions and corresponding channels for each dataset.

Dataset	Brain Region	Channels / Electrodes
ADFTD	Frontopolar-Frontal	Fp1, Fp2, F3, F4, Fz
	Lateral Frontal-Temporal	F7, F8, T3, T4
	Central	C3, Cz, C4
	Posterior Temporal	T5, T6
	Parieto-Occipital	P3, Pz, P4, O1, O2
APAVA	Prefrontal	Fp1, Fp2, F3, F4
	Frontal-Temporal	F7, F8, T3, T4
	Temporal	T5, T6
	Central	C3, C4
	Parieto-Occipital	P3, P4, O1, O2
M-fNIRS	L-PreM&SMA	1, 4, 10
	L-BROCA	2, 3, 5, 7, 8, 13
	L-FEF	12, 24
	L-DLPFC	6, 11, 14, 17, 18, 20
	L-FPA	9, 15, 16, 19, 21, 22, 23, 27
	R-PreM&SMA	30, 33, 35, 36, 37, 41, 43, 48
	R-BROCA	31, 32, 34, 39, 42, 45
	R-FEF	26, 38
	R-DLPFC	44, 46, 49, 50, 51, 53
	R-FPA	40, 47, 52

D Diagnostic Category Distribution of M-fNIRS

Table 4: Diagnostic category distribution within the M-fNIRS dataset.

Label	Diagnostic Category	Number of Subjects
0	Healthy Controls	739
1	Schizophrenia	36
2	Bipolar I Disorder	99
3	Bipolar II Disorder	100
4	Major Depressive Disorder	116
Total		**1,090**