LEARNING FROM BRAIN TOPOGRAPHY: A HIERARCHICAL LOCAL-GLOBAL GRAPH-TRANSFORMER NETWORK FOR EEG EMOTION RECOGNITION

Anonymous authors
Paper under double-blind review

ABSTRACT

Understanding how local neurophysiological patterns interact with global brain dynamics is essential for decoding human emotions from EEG signals. However, existing deep learning approaches often overlook the brain's intrinsic spatial organization, failing to simultaneously capture local topological relations and global dependencies. To address these challenges, we propose Neuro-HGLN, a Neurologically-informed Hierarchical Graph-Transformer Learning Network that integrates biologically grounded priors with hierarchical representation learning. Neuro-HGLN first constructs a spatial Euclidean prior graph based on physical electrode distances to serve as an anatomically grounded inductive bias. A learnable global graph is then introduced to model functional connectivity across the entire brain. In parallel, to capture fine-grained regional dependencies, Neuro-HGLN builds region-level local graphs using a multi-head attention mechanism. These graphs are processed synchronously through local-constrained parallel GCN layers to produce region-specific representations. Subsequently, an iTransformer encoder aggregates these features to capture cross-region dependencies under a dimension-as-token formulation. Extensive experiments demonstrate that Neuro-HGLN achieves state-of-the-art performance on multiple benchmarks, providing enhanced interpretability grounded in neurophysiological structure. These results highlight the efficacy of unifying local topological learning with cross-region dependency modeling for robust EEG emotion recognition.

1 INTRODUCTION

Electroencephalography (EEG)-based emotion recognition has attracted increasing attention due to its potential applications in affective computing and mental health assessment (Wang et al., 2024). As a direct measurement of neural activity, EEG provides rich temporal information reflecting dynamic brain processes underlying emotional states (Zheng & Lu, 2015). However, due to the low signal-to-noise ratio, high inter-subject variability, and the complex spatial–temporal dependencies inherent in neural dynamics, decoding emotions from EEG signals remains a challenging task.

A key characteristic of EEG signals is their strong dependency on the spatial organization of the brain, where the physical arrangement of electrodes reflects underlying neuroanatomical structure (Jang et al., 2021). Consequently, modeling spatial relationships is critical. Graph-based learning methods, particularly Graph Convolutional Networks (GCNs), have emerged as a promising paradigm by modeling electrodes as nodes (Song et al., 2018; 2021). Despite their effectiveness, most existing approaches rely on either static graph priors or globally learned connectivity patterns. This holistic processing often overlooks the intrinsic functional segregation of the brain, failing to capture fine-grained topological characteristics within specific anatomical regions.

In addition to spatial dependencies, EEG signals exhibit complex temporal dynamics that span multiple time scales. While Recurrent Neural Networks (RNNs) (Zhang et al., 2018) have been applied to model these dependencies, they often struggle with long-range interactions. Transformer-based models have recently demonstrated superior capability in modeling global contexts (Vaswani et al.,

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.

2017). Nevertheless, directly treating EEG channels as tokens in standard Transformers does not explicitly encode the underlying spatial topology or neurophysiological priors of EEG signals (Arjun et al., 2021). Such tokenization does not explicitly represent the structural correlations between spatially distant yet functionally coupled electrodes.

To address these limitations, we propose Neuro-HGLN, a Neurologically-informed Hierarchical Local–Global Graph–Transformer Learning Network. Neuro-HGLN emulates the neurological mechanism of functional segregation and integration by bridging biologically grounded priors with data-driven learning. Specifically, we construct a spatial Euclidean prior graph and a learnable global graph to capture holistic functional connectivity. Meanwhile, to respect anatomical distinctiveness, we design a hierarchical stream that builds region-level graphs processed by parallel GCNs. These local features are then synthesized by an inverted Transformer (iTransformer) encoder (Liu et al., 2023), which leverages a dimension-as-token strategy to model high-level cross-region dependencies. Extensive experiments demonstrate that Neuro-HGLN achieves state-of-the-art performance while offering enhanced interpretability aligned with neurophysiological structures.

The main contributions of this work are summarized as follows:

• We propose Neuro-HGLN, a neurologically informed hierarchical graph–transformer framework that jointly models local spatial topology and global cross-region feature dependencies for EEG emotion recognition.

• We introduce a spatial Euclidean prior graph together with a learnable global graph to capture anatomically grounded and functional connectivity patterns across EEG channels.

• We design parallel region-level local graphs with functional diversity regularization to capture complementary and spatially specialized connectivity patterns.

• Extensive experimental results demonstrate that the proposed Neuro-HGLN framework achieves superior performance while offering enhanced interpretability grounded in neurophysiological structure.

Additional discussion of related work is provided in Appendix A.

2 METHODOLOGY

To explicitly characterize the hierarchical topology of the brain for robust EEG-based emotion recognition, we propose the Neuro-HGLN framework, as illustrated in Fig. 1. The core motivation of Neuro-HGLN is to emulate the neurological mechanism of *functional segregation and integration*—where the brain processes information both within localized regions and through global communication across regions. Distinct from previous methods that rely on single-scale graph modeling or ignore spatial priors, Neuro-HGLN bridges biologically grounded spatial constraints with learnable topological and semantic representations. Through the synergistic fusion of local neurological priors and global semantic interactions, our method facilitates robust representation learning that strikes a balance between high accuracy and physiological interpretability.

As illustrated in Fig. 1, the overall architecture comprises two synergistic streams designed to capture complementary neurological patterns:

1. **Global Graph Learning Stream:** This pathway captures holistic brain dynamics by constructing a global graph that integrates a static spatial Euclidean prior with learnable adaptive weights, enabling the model to respect anatomical constraints while adapting to functional connectivity.
2. **Hierarchical Local-Region Stream:** This pathway models fine-grained regional characteristics. It utilizes parallel GCNs to extract intra-regional topological features and subsequently leverages an iTransformer encoder to model high-level cross-region dependencies under a dimension-as-token formulation.

2.1 PROBLEM FORMULATION

Formally, let $\mathbf{X}_{DE} \in \mathbb{R}^{N\times d}$ denote the input EEG Differential Entropy (DE) features, where N represents the number of electrodes and d is the feature dimension extracted across five standard

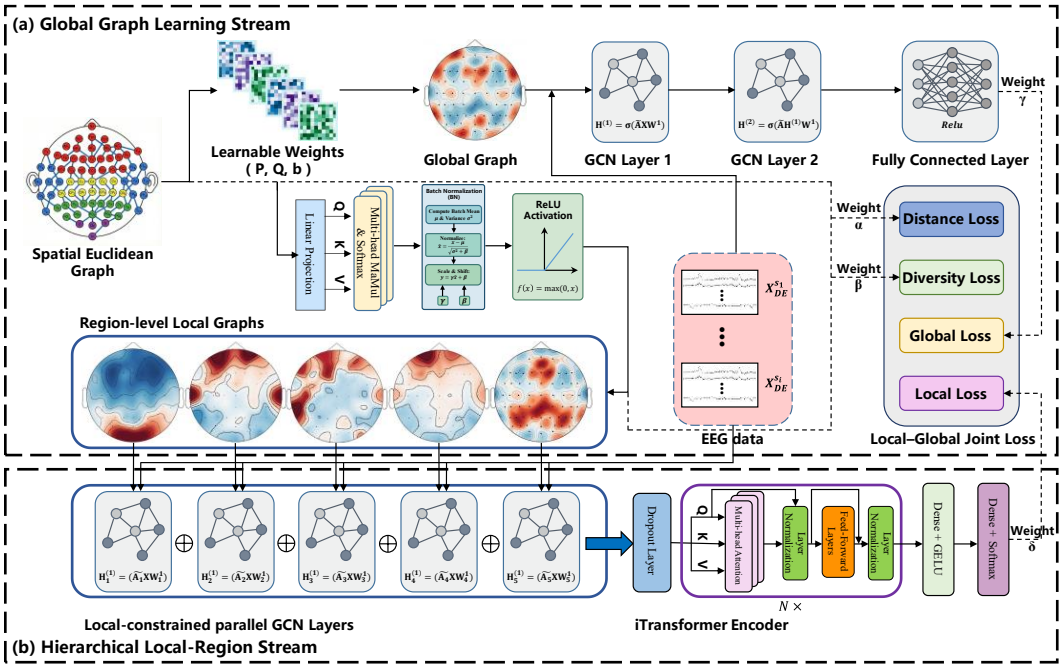


Figure 1: **The overall architecture of the proposed Neuro-HGLN framework.** The model orchestrates two synergistic streams to process EEG DE features: **(a) The Global Graph Learning Stream** captures holistic brain dynamics by integrating a static Spatial Euclidean Prior with learnable adaptive weights to construct a global graph, which is then processed by stacked GCN layers. **(b) The Hierarchical Local-Region Stream** employs parallel learnable region-level graphs to capture complementary localized topology, followed by an **iTransformer Encoder** for inter-regional dependency modeling. The final emotion prediction is achieved by fusing the logits from both streams.

frequency bands: δ (1–4 Hz), θ (4–8 Hz), α (8–14 Hz), β (14–30 Hz), and γ (30–50 Hz). The goal of EEG emotion recognition is to learn a mapping function $\mathcal{F} : \mathbf{X} \to \mathbf{y}$, parameterized by Θ, which projects the neural signals into the emotion space, where $\mathbf{y} \in \{1, \ldots, C\}$ represents the predicted emotion category among C classes.

2.2 Global Graph Learning Stream

To capture the holistic functional connectivity of the brain based on the anatomical structure, we construct a global graph $\mathcal{G}_g = (\mathcal{V}, \mathcal{E}_g, \mathbf{A}_g)$. While functional connectivity is dynamic, the physical proximity of electrodes imposes a fundamental constraint on signal propagation (Salvador et al., 2005). To incorporate this inductive bias, we explicitly construct a **Spatial Euclidean Prior Graph** based on standard 3D electrode positions.

2.2.1 Spatial Euclidean Prior Construction

Let $\mathcal{V}$ denote the set of N EEG electrodes. We map each electrode v_i to its standard 3D Cartesian coordinates $\mathbf{c}_i = (x_i, y_i, z_i) \in \mathbb{R}^3$ according to the International 10–10 System (Nuwer, 2018). The physical distance matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$ is computed using the pairwise Euclidean norm $\|\mathbf{c}_i - \mathbf{c}_j\|_2$.

To transform these physical distances into topological connection weights, we employ a Gaussian kernel function. This assumes that the strength of the neural interaction prior decays exponentially with the square of the physical distance. The static spatial adjacency matrix $\mathbf{A}_{prior} \in \mathbb{R}^{N \times N}$ is defined as:

$$\mathbf{A}_{prior,ij} = \exp\left(-\frac{\mathbf{D}_{ij}^2}{\tau}\right), \tag{1}$$

where τ is the temperature parameter that controls the sparsity and decay rate of the spatial connections.

2.2.2 ADAPTIVE GRAPH FUSION

However, physical distance alone is insufficient to capture the complex, nonlinear functional coupling associated with varying emotional states. Therefore, we refine the rigid spatial prior into an adaptive global graph $\mathbf{A}_g$ using learnable projections $\mathbf{P}_g, \mathbf{Q}_g \in \mathbb{R}^{N\times N}$ and bias $\mathbf{b}_g \in \mathbb{R}^{N\times N}$:

$$\mathbf{A}_g = \text{ReLU}\left((\mathbf{P}_g \mathbf{A}_{prior} + \mathbf{b}_g)\mathbf{Q}_g\right) \in \mathbb{R}^{N\times N}. \tag{2}$$

Here, $\mathbf{P}_g$ and $\mathbf{Q}_g$ facilitate the reshaping of the global connectivity patterns, the bias $\mathbf{b}_g$ allows for the generation of intrinsic connections independent of physical distance, and the ReLU activation function ensures non-negative edge weights.

The resulting matrix $\mathbf{A}_g$ serves as the adaptive adjacency matrix for the subsequent graph convolution operations. Given the input node representations $\mathbf{H}^{(0)} = \mathbf{X}_{DE}$, we progressively propagate information over the learned graph $\mathbf{A}_g$ through multiple GCN layers. At the l-th layer, the node representations $\mathbf{H}^{(l)}$ are updated according to

$$\mathbf{H}^{(l+1)} = \text{ReLU}\left(\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}_g\tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}\right), \tag{3}$$

where $\tilde{\mathbf{A}}_g = \mathbf{A}_g + \mathbf{I}$ incorporates self-loops, $\tilde{\mathbf{D}}$ is the corresponding degree matrix for normalization, and $\mathbf{W}^{(l)}$ is the learnable weight matrix of the l-th GCN layer.

Following the stacked GCN layers, the final node representations, denoted as $\mathbf{H}^{(L)} \in \mathbb{R}^{N\times F_{out}}$, contain high-level semantic features distributed across the topological structure, where F_{out} denotes the output feature dimension of the final GCN layer. To map these features into the emotion space, we introduce a Fully Connected (FC) layer. The output logits of the global stream are computed as:

$$\mathbf{y}_{global} = \text{Flatten}(\mathbf{H}^{(L)})\mathbf{W}_{fc} + \mathbf{b}_{fc} \in \mathbb{R}^{C}, \tag{4}$$

where $\mathbf{W}_{fc}$ and $\mathbf{b}_{fc}$ denote the weight matrix and bias of the dense layer, respectively, and C is the number of emotion classes. This output vector $\mathbf{y}_{global}$ effectively encapsulates the holistic spatial distribution of brain activity grounded in physical topology, serving as the prediction output of the global stream for joint optimization.

2.3 HIERARCHICAL LOCAL-REGION STREAM

While the Global Graph Stream captures holistic patterns, brain activity is fundamentally characterized by functional segregation—where distinct anatomical regions process information locally before integrating it globally (Sporns, 2013). To emulate this mechanism, we propose the **Hierarchical Local-Region Stream** (as shown in Fig. 1 (b)), which orchestrates a two-stage learning process: intra-regional topological extraction followed by inter-regional semantic integration.

2.3.1 INTRA-REGIONAL TOPOLOGY LEARNING VIA PARALLEL GCNS

To capture fine-grained and potentially localized connectivity patterns without imposing hard anatomical partitions, we learn K complementary region-level graphs from the shared spatial prior. Each graph maintains a full $N \times N$ adjacency matrix over all electrodes, allowing regional specialization to emerge adaptively rather than relying on predefined anatomical partitions. The functional diversity regularization introduced in Section 2.4.3 further discourages redundant connections across these graphs and promotes complementary specialization.

Specifically, we first project the spatial prior $\mathbf{A}_{prior}$ into a learnable latent topology space:

$$\mathbf{A}_{latent} = \text{ReLU}\left((\mathbf{P}_{latent}\mathbf{A}_{prior} + \mathbf{b}_{latent})\mathbf{Q}_{latent}\right) \in \mathbb{R}^{N\times N}, \tag{5}$$

where $\mathbf{P}_{latent}, \mathbf{Q}_{latent} \in \mathbb{R}^{N\times N}$ and $\mathbf{b}_{latent} \in \mathbb{R}^{N\times N}$ are learnable parameters. The resulting $\mathbf{A}_{latent}$ provides a shared latent representation of the spatial topology from which multiple connectivity patterns are derived.

To capture complementary connectivity patterns, we then employ a multi-head attention mechanism to generate K region-level graphs from the shared latent topology. Specifically, $\mathbf{A}_{latent}$ is first projected into query and key representations:

$$\mathbf{Q} = \mathbf{A}_{latent}\mathbf{W}_Q, \qquad \mathbf{K} = \mathbf{A}_{latent}\mathbf{W}_K, \tag{6}$$

where $\mathbf{W}_Q, \mathbf{W}_K \in \mathbb{R}^{N\times d_m}$ are learnable projection matrices. The projected representations are then split along the feature dimension into K attention heads:

$$\mathbf{Q} = [\mathbf{Q}^1, \ldots, \mathbf{Q}^K], \qquad \mathbf{K} = [\mathbf{K}^1, \ldots, \mathbf{K}^K], \tag{7}$$

where $\mathbf{Q}^k, \mathbf{K}^k \in \mathbb{R}^{N\times d_h}$, $d_h = d_m/K$, and $k = 1, \ldots, K$.

For each head, pairwise attention scores are computed as

$$\mathbf{S}^k = \frac{\mathbf{Q}^k(\mathbf{K}^k)^T}{\sqrt{d_h}} \in \mathbb{R}^{N\times N}. \tag{8}$$

Rather than aggregating the attention heads into a single representation, we retain the attention map produced by each head as an individual region-level graph:

$$\mathbf{A}^k_{local} = \mathrm{ReLU}\left(\mathrm{BN}\left(\mathrm{Softmax}(\mathbf{S}^k)\right)\right) \in \mathbb{R}^{N\times N}, \qquad k = 1, \ldots, K. \tag{9}$$

Thus, the K attention heads generate K head-specific local graphs, with each head capturing pairwise electrode connectivity in a distinct representation subspace. Together with the functional diversity regularization, these graphs are encouraged to learn complementary and spatially specialized connectivity patterns.

To extract features from these learned graphs while preserving their distinct topological characteristics, we employ **Locally Constrained Parallel GCN Layers**. As illustrated in Fig. 1, K independent GCN modules operate in parallel, with each GCN associated with one learned local graph. For the k-th graph, the feature update is defined as:

$$\mathbf{H}_k = \mathrm{GELU}\left(\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}^k_{local}\tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{X}_{DE}\mathbf{W}_k\right) \in \mathbb{R}^{N\times F_{out}}, \tag{10}$$

where $\tilde{\mathbf{A}}^k_{local} = \mathbf{A}^k_{local} + \mathbf{I}$, $\mathbf{W}_k \in \mathbb{R}^{d\times F_{out}}$ is the learnable projection matrix of the k-th GCN, and F_{out} denotes its output feature dimension.

Finally, the representations obtained from all K local graphs are concatenated to form the local feature representation:

$$\mathbf{H}_{local} = \mathrm{Concat}(\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_K) \in \mathbb{R}^{N\times KF_{out}}. \tag{11}$$

The resulting $\mathbf{H}_{local}$ integrates complementary topological features captured by the head-specific local graphs and serves as the input token sequence for the subsequent iTransformer encoder.

2.3.2 Inter-Regional Integration via iTransformer

While parallel GCNs capture short-range dependencies, emotion regulation heavily relies on the long-range coordination between varying brain regions. Standard Transformers typically treat temporal steps as tokens, which is suboptimal for multivariate EEG data where spatial channel correlations are paramount.

To address this, we adopt the **iTransformer** encoder with a dimension-as-token strategy. Specifically, each electrode representation, enriched with region-specific features from the parallel GCNs, is treated as an independent token, enabling the self-attention mechanism to capture inter-regional dependencies. The aggregated local features $\mathbf{H}_{local}$ are used as the input to the encoder, denoted by $\mathbf{Z}^{(0)} = \mathbf{H}_{local}$. The encoder contains L stacked layers, each consisting of a Multi-Head Self-Attention (MSA) module and a Feed-Forward Network (FFN). The update process at the l-th layer is formulated as follows:

$$\mathbf{Z}' = \mathrm{LayerNorm}(\mathbf{Z}^{(l-1)} + \mathrm{MSA}(\mathbf{Z}^{(l-1)})), \tag{12}$$

$$\mathbf{Z}^{(l)} = \mathrm{LayerNorm}(\mathbf{Z}' + \mathrm{FFN}(\mathbf{Z}')), \tag{13}$$

where $\mathbf{Z}^{(L)} \in \mathbb{R}^{N \times d_{model}}$ denotes the output of the final encoder layer. In this formulation, the self-attention mechanism acts as a dynamic, fully connected graph that integrates information across the entire brain topology, effectively capturing high-level inter-regional dependencies.

To map these high-level inter-regional representations into the emotion space, we employ an MLP Projection Head. The unnormalized local logits $\mathbf{y}_{local} \in \mathbb{R}^C$ are computed by projecting the hidden features, applying GELU activation, flattening the resulting representations, and applying a final linear layer:

$$\mathbf{y}_{local} = \text{Flatten}(\text{GELU}(\mathbf{Z}^{(L)}\mathbf{W}_1 + \mathbf{b}_1))\mathbf{W}_2 + \mathbf{b}_2. \tag{14}$$

where $\mathbf{W}_1, \mathbf{b}_1, \mathbf{W}_2$, and $\mathbf{b}_2$ are learnable parameters, and $\mathbf{y}_{local} \in \mathbb{R}^C$ represents the unnormalized local-view logits, which provide complementary information to the global stream and will be jointly optimized.

2.4 Multi-Objective Joint Optimization

To jointly learn discriminative representations and meaningful graph structures, Neuro-HGLN is trained in an end-to-end manner with multiple complementary objectives. Specifically, the global and local streams are independently supervised by classification losses, while the learnable local graphs are regularized by geometric and functional constraints. Accordingly, the overall training objective comprises two classification losses, a geometric constraint, and functional diversity regularization.

2.4.1 Deep Supervision Classification Loss

To ensure that both the global and local streams learn discriminative representations, we apply deep supervision by imposing classification objectives on the two streams independently. Given the ground-truth label y, the classification losses are defined as

$$\mathcal{L}_{global} = \mathcal{L}_{CE}(\mathbf{y}_{global}, y), \mathcal{L}_{local} = \mathcal{L}_{CE}(\mathbf{y}_{local}, y), \tag{15}$$

where $\mathbf{y}_{global}$ and $\mathbf{y}_{local}$ denote the logits produced by the global and local streams, respectively. These two supervised objectives provide direct gradient signals to the corresponding streams, encouraging each stream to learn discriminative representations rather than relying solely on the final fusion.

2.4.2 Geometric Constraint (KL Divergence)

Although the region-level local graphs are derived from the shared spatial prior, the learnable topology transformation and multi-head graph generation allow their connectivity patterns to adapt during training. To prevent the learned graphs from deviating excessively from the underlying anatomical structure, we introduce a **Geometric Constraint Loss** based on the Kullback-Leibler (KL) divergence (Naidji et al., 2026). Specifically, each learned local adjacency matrix $\mathbf{A}^k_{local}$ is encouraged to preserve the connectivity distribution encoded by the spatial prior $\mathbf{A}_{prior}$:

$$\mathcal{L}_{dist} = \sum_{k=1}^{K} \text{KL}\left(\mathbf{A}_{prior} \parallel \mathbf{A}^k_{local}\right). \tag{16}$$

Minimizing $\mathcal{L}_{dist}$ provides a soft geometric constraint that prevents excessive deviation from the spatial prior while retaining the flexibility to learn complementary connectivity patterns.

2.4.3 Functional Diversity Regularization

While the geometric constraint preserves anatomical consistency, it alone does not prevent different local graphs from learning highly similar connectivity patterns. To encourage the head-specific local graphs to capture complementary connectivity patterns, we introduce a **Functional Diversity Loss** that penalizes overlapping connections:

$$\mathcal{L}_{div} = \sum_{i=1}^{K} \sum_{j=i+1}^{K} \left\| \mathbf{A}^i_{local} \odot \mathbf{A}^j_{local} \right\|_F^2, \tag{17}$$

where $\odot$ denotes the element-wise product and $\|\cdot\|_F$ represents the Frobenius norm. Minimizing $\mathcal{L}_{div}$ discourages redundant connectivity across the head-specific local graphs and promotes complementary specialization.

2.4.4 Overall Training Objective

The four loss terms are jointly optimized to balance classification performance, anatomical consistency, and functional diversity. The total training objective is formulated as

$$\mathcal{L}_{total} = \alpha\mathcal{L}_{dist} + \beta\mathcal{L}_{div} + \gamma\mathcal{L}_{global} + \delta\mathcal{L}_{local}, \tag{18}$$

where α and β control the contributions of the geometric constraint and functional diversity regularization terms, respectively, while γ and δ weight the classification objectives of the global and local streams. The entire Neuro-HGLN framework is optimized by minimizing $\mathcal{L}_{total}$, allowing the representation learning and graph structure learning modules to be jointly optimized in an end-to-end manner.

2.4.5 Decision-Level Fusion during Inference

After training, the predictions of the global and local streams are combined through decision-level fusion. Specifically, their logits are equally averaged as $\mathbf{y}_{final} = 0.5(\mathbf{y}_{global} + \mathbf{y}_{local})$, and the final prediction is obtained by $\hat{y} = \arg\max(\mathbf{y}_{final})$. This late-fusion strategy integrates the complementary information captured by the two streams for final classification.

The complete training procedure of Neuro-HGLN is summarized in Algorithm 1 in Appendix B.

3 Experiments

3.1 Datasets

To comprehensively evaluate Neuro-HGLN, we employ four widely recognized benchmarks covering coarse-to-fine emotion granularities: **SEED** (Zheng & Lu, 2015), **SEED-IV** (Zheng et al., 2018), **SEED-V** (Liu et al., 2021), and **MPED** (Song et al., 2019). The **SEED series** share a unified acquisition protocol using a 62-channel ESI NeuroScan system. Following standard procedures, we utilize Differential Entropy (DE) features extracted from 1-second non-overlapping segments across five frequency bands ($\delta, \theta, \alpha, \beta, \gamma$). The three datasets differ primarily in complexity: SEED (15 subjects, 3 sessions) targets 3 emotions (*positive, neutral, negative*); SEED-IV (15 subjects) targets 4 emotions (*happy, sad, fear, neutral*); and SEED-V (20 subjects) targets 5 emotions (*happy, sad, fear, disgust, neutral*). Additionally, we utilize the large-scale **MPED** dataset (Song et al., 2019) (30 subjects) for fine-grained 7-class classification (*joy, funny, anger, fear, disgust, sadness, neutral*). Distinct from the SEED series, we extract 256-point Short-Time Fourier Transform (STFT) features for MPED to better capture fine-grained time-frequency characteristics, while maintaining the same 62-channel and five-band configuration. We report implementation details and hyperparameter settings in Appendix C.

3.2 Subject-dependent Evaluation

Table 1: Subject-dependent classification performance (Accuracy / STD %) on SEED, SEED-IV, SEED-V, and MPED.

Method	ACC / STD (%)									
	Session 1 → Session 2			Session 1 → Session 3			Session 2 → Session 3			–
	SEED	**SEED-IV**	**SEED-V**	**SEED**	**SEED-IV**	**SEED-V**	**SEED**	**SEED-IV**	**SEED-V**	**MPED**
SVM	51.62/11.80	59.81/16.61	41.56/15.36	50.23/11.99	59.50/17.07	41.87/15.18	44.65/11.90	67.77/10.94	43.40/16.76	32.39/09.53
DGCNN	86.96/09.24	67.47/13.31	69.93/11.16	87.42/08.59	66.38/12.02	63.50/09.92	90.96/07.95	74.19/14.07	71.54/14.89	32.37/06.08
PGCN	91.24/03.75	79.87/11.60	75.07/10.13	90.97/08.59	80.70/11.60	70.23/10.72	93.43/07.40	82.22/08.50	75.05/13.38	43.56/08.21
RGNN†	89.77/08.85	70.08/10.54	72.56/10.52	88.59/07.14	70.07/14.59	69.14/10.61	92.33/11.56	74.59/11.31	69.28/10.11	38.77/08.14
TANN†	90.91/12.47	73.85/09.46	74.78/08.66	90.40/12.57	73.38/15.29	69.08/09.43	91.64/17.10	80.02/08.49	77.63/09.20	39.82/07.98
Neuro-HGLN	**92.57/07.80**	**82.23/12.45**	**76.80/11.75**	**94.64/09.37**	**84.74/13.30**	**73.72/12.83**	**95.21/09.21**	**85.45/15.11**	**78.79/08.94**	**44.96/08.25**

† Methods that utilize unlabeled test data for training (Domain Adaptation).

We evaluate Neuro-HGLN using a strict cross-session protocol on the SEED series and a standard 3:1 train-test split on MPED. Table 1 summarizes the comparison against state-of-the-art baselines,

including dynamic graph methods (DGCNN (Song et al., 2018), PGCN (Zhou et al., 2023)) and domain adaptation techniques (RGNN (Zhong et al., 2020), TANN (Li et al., 2021)).

Neuro-HGLN consistently outperforms all competing methods. On the **SEED series**, it achieves consistently strong performance; specifically, in the challenging *Session 1 → 3* task, it achieves **94.64%** (SEED) and **84.74%** (SEED-IV), surpassing the strong baseline PGCN by over 3.5%. This robustness may benefit from the **Geometric Constraint**, which anchors the learnable graphs to the spatial prior and helps reduce sensitivity to session-specific noise.

Crucially, Neuro-HGLN outperforms domain adaptation methods (e.g., +5.43% over TANN on SEED-IV) without accessing unlabeled test data. This suggests that the **Functional Diversity Regularization** encourages complementary functional patterns that generalize across sessions. On the fine-grained **MPED** (7 emotions), our model achieves **44.96%**, suggesting that cross-region integration via the iTransformer is beneficial for modeling complex multivariate dependencies.

3.3 Subject-independent Evaluation

Table 2: Subject-independent emotion recognition results (Accuracy / STD %) on SEED, SEED-IV, SEED-V, and MPED.

Method	ACC / STD (%)			
	SEED	**SEED-IV**	**SEED-V**	**MPED**
SVM	56.73/16.29	37.99/12.52	23.71/08.25	19.66/03.96
DGCNN	79.95/09.02	52.82/09.23	41.92/06.67	25.12/04.20
PGCN	83.28/11.09	76.96/07.19	71.40/09.43	28.39/05.02
RGNN†	85.30/06.72	73.84/08.02	66.28/16.71	27.95/03.05
TANN†	84.41/08.75	68.00/08.35	67.36/08.22	28.32/05.11
Neuro-HGLN	**90.92/09.49**	**79.30/09.11**	**78.34/10.68**	**28.56/04.67**

† Methods using unlabeled test data (Domain Adaptation).

We employ a Leave-One-Subject-Out (LOSO) cross-validation strategy to evaluate robustness against physiological variance. As shown in Table 2, Neuro-HGLN achieves state-of-the-art performance across all datasets.

Under LOSO evaluation, Neuro-HGLN achieves **90.92%**, **79.30%**, **78.34%**, and **28.56%** on SEED, SEED-IV, SEED-V, and MPED, respectively. It consistently outperforms graph-based and domain-adaptation baselines without accessing unlabeled test data, demonstrating strong cross-subject generalization, particularly on SEED-V.

3.4 Ablation Study

Table 3: Ablation study results (Accuracy %) on SEED-IV and MPED.

Model Variant	**SEED-IV (Avg.)**	**MPED**
Full Model (Neuro-HGLN)	**85.14**	**44.96**
w/o Global Stream	82.10	41.50
w/o Local Stream	80.35	38.20
w/o Geo. Constraint ($\mathcal{L}_{dist}$)	83.90	43.10
w/o Funct. Diversity ($\mathcal{L}_{div}$)	84.15	42.80

To assess the contributions of individual components, we conduct ablation studies on SEED-IV and MPED using four model variants, as summarized in Table 3.

Impact of Dual-Stream Architecture: Removing either the Global (**w/o Global**) or Local (**w/o Local**) stream degrades performance. Notably, removing the Local Stream causes a larger drop on the fine-grained MPED dataset, supporting its importance for capturing subtle emotional patterns, while the Global Stream provides complementary holistic context.

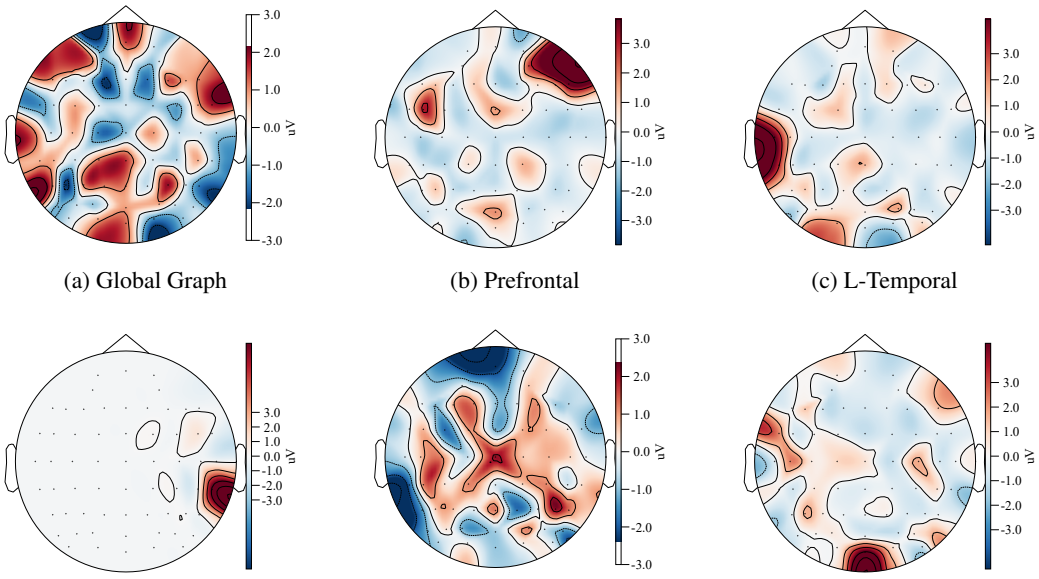
Impact of Regularization: Removing either the Geometric Constraint (**w/o Geo**) or Functional Diversity regularization (**w/o Div**) consistently reduces performance. These results support their com-

plementary roles in preserving anatomically grounded topology and encouraging diverse regional representations.

3.5 Visualization of Learned Topology

To examine the learned connectivity patterns, we visualize the topological structures learned by the Global Stream and the distinct Local Streams in Fig. 2.

As illustrated in Fig. 2 (a), the **Global Graph** exhibits a widespread, complex activation pattern that extends across both hemispheres. This distributed topology indicates that the global stream captures long-range dependencies and holistic connectivity patterns.



(a) Global Graph (b) Prefrontal (c) L-Temporal

(d) R-Temporal (e) Parietal (f) Occipital

Figure 2: **Visualization of learned topological structures.** (a) The Global Graph captures holistic, long-range dependencies. (b-f) The Local Graphs exhibit distinct, region-specific activation patterns, supporting the role of functional diversity regularization in promoting complementary representations.

In sharp contrast, the **Local Graphs** (Fig. 2 (b-f)) demonstrate highly localized and region-specific patterns. Specifically, the **Prefrontal** map shows concentrated activation in the anterior region, which has been associated with emotion regulation and higher-order cognition (Ochsner & Gross, 2005). The **Temporal** maps (Left and Right) show distinct focal activations on the lateral sides, while the **Occipital** map exhibits a concentrated response in the posterior visual cortex (Olson et al., 2007; Bradley et al., 2001).

To further examine the discriminability of the learned representations, additional t-SNE visualizations of feature distributions are provided in Appendix D.

4 Conclusion

We propose Neuro-HGLN, a hierarchical graph–Transformer framework that integrates global connectivity modeling with region-specific representations under a spatial anatomical prior. Experiments on SEED, SEED-IV, SEED-V, and MPED demonstrate strong performance across both cross-session and cross-subject settings without using unlabeled test data. Ablation and topology visualizations further support the complementary roles of the global and local streams and the proposed regularization objectives.

AI USE STATEMENT

In this work, we have not used generative AI tools for any of the required disclosure tasks listed in the ICLR 2027 AI use policy. Additionally, we used generative AI tools for editing the manuscript to improve readability and clarity, rewriting sentences, discussing manuscript structure and presentation, and adjusting LaTeX formatting. We have reviewed and verified all AI-assisted work. We take responsibility for the final content of this work, including all text, claims, analyses, and other artifacts produced with the aid of generative AI.

ETHICS STATEMENT

There are no ethical issues.

REFERENCES

Arjun Arjun, Aniket Singh Rajpoot, and Mahesh Raveendranatha Panicker. Introducing attention mechanism for eeg signals: Emotion recognition with vision transformers. In *2021 43rd annual international conference of the IEEE engineering in medicine & biology society (EMBC)*, pp. 5723–5726. IEEE, 2021.

Margaret M Bradley, Maurizio Codispoti, Bruce N Cuthbert, and Peter J Lang. Emotion and motivation i: defensive and appetitive reactions in picture processing. *Emotion*, 1(3):276, 2001.

Dong-Wei Chen, Rui Miao, Wei-Qi Yang, Yong Liang, Hao-Heng Chen, Lan Huang, Chun-Jian Deng, and Na Han. A feature extraction method based on differential entropy and linear discriminant analysis for emotion recognition. *Sensors*, 19(7):1631, 2019.

Ruo-Nan Duan, Jia-Yi Zhu, and Bao-Liang Lu. Differential entropy feature for eeg-based emotion classification. In *2013 6th international IEEE/EMBS conference on neural engineering (NER)*, pp. 81–84. IEEE, 2013.

Soobeom Jang, Seong-Eun Moon, and Jong-Seok Lee. Eeg-based emotional video classification via learning connectivity structure. *IEEE Transactions on Affective Computing*, 14(2):1586–1597, 2021.

Berke Kılıç and Serap Aydın. Classification of contrasting discrete emotional states indicated by eeg based graph theoretical network measures. *Neuroinformatics*, 20(4):863–877, 2022.

Yang Li, Boxun Fu, Fu Li, Guangming Shi, and Wenming Zheng. A novel transferability attention neural network model for eeg emotion recognition. *Neurocomputing*, 447:92–101, 2021. ISSN 0925-2312. URL https://www.sciencedirect.com/science/article/pii/S0925231221003027.

Wei Liu, Jie-Lin Qiu, Wei-Long Zheng, and Bao-Liang Lu. Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition. *IEEE Transactions on Cognitive and Developmental Systems*, 2021.

Yong Liu, Tengge Hu, Haoran Zhang, Haixu Wu, Shiyu Wang, Lintao Ma, and Mingsheng Long. itransformer: Inverted transformers are effective for time series forecasting. *arXiv preprint arXiv:2310.06625*, 2023.

Ilyes Naidji, Ahmed Tibermacine, Imad Eddine Tibermacine, Samuele Russo, and Christian Napoli. Egdn-kl: dynamic graph–deviation network for eeg anomaly detection. *Biomedical Signal Processing and Control*, 112:108597, 2026.

Marc R. Nuwer. 10-10 electrode system for eeg recording. *Clinical Neurophysiology*, 129(5): 1103, 2018. ISSN 1388-2457. doi: https://doi.org/10.1016/j.clinph.2018.01.065. URL https://www.sciencedirect.com/science/article/pii/S1388245718300907.

Kevin N Ochsner and James J Gross. The cognitive control of emotion. *Trends in cognitive sciences*, 9(5):242–249, 2005.

Ingrid R Olson, Alan Plotzker, and Youssef Ezzyat. The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain*, 130(7):1718–1731, 2007.

Raymond Salvador, John Suckling, Martin R Coleman, John D Pickard, David Menon, and ED Bullmore. Neurophysiological architecture of functional magnetic resonance images of human brain. *Cerebral cortex*, 15(9):1332–1342, 2005.

Tengfei Song, Wenming Zheng, Peng Song, and Zhen Cui. Eeg emotion recognition using dynamical graph convolutional neural networks. *IEEE Transactions on Affective Computing*, 11(3): 532–541, 2018.

Tengfei Song, Wenming Zheng, Cheng Lu, Yuan Zong, Xilei Zhang, and Zhen Cui. Mped: A multi-modal physiological emotion database for discrete emotion recognition. *IEEE Access*, 7: 12177–12191, 2019.

Tengfei Song, Wenming Zheng, Suyuan Liu, Yuan Zong, Zhen Cui, and Yang Li. Graph-embedded convolutional neural network for image-based eeg emotion recognition. *IEEE Transactions on Emerging Topics in Computing*, 10(3):1399–1413, 2021.

Olaf Sporns. Network attributes for segregation and integration in the human brain. *Current opinion in neurobiology*, 23(2):162–171, 2013.

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

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Yiming Wang, Bin Zhang, and Lamei Di. Research progress of eeg-based emotion recognition: a survey. *ACM Computing Surveys*, 56(11):1–49, 2024.

Yongling Xu, Yang Du, Ling Li, Honghao Lai, Jing Zou, Tianying Zhou, Lushan Xiao, Li Liu, and Pengcheng Ma. Amdet: Attention based multiple dimensions eeg transformer for emotion recognition. *IEEE Transactions on Affective Computing*, 15(3):1067–1077, 2023.

Tong Zhang, Wenming Zheng, Zhen Cui, Yuan Zong, and Yang Li. Spatial–temporal recurrent neural network for emotion recognition. *IEEE transactions on cybernetics*, 49(3):839–847, 2018.

Yuxuan Zhao, Siqi Liu, Linjing Cao, Yidao Ji, Wei Wu, and Bo Wang. Electrode quantity outperforms spatial topology in eeg-based emotion recognition: A dual-cnn study on electrode configuration. *Biomedical Signal Processing and Control*, 113:109082, 2026.

W. Zheng, W. Liu, Y. Lu, B. Lu, and A. Cichocki. Emotionmeter: A multimodal framework for recognizing human emotions. *IEEE Transactions on Cybernetics*, pp. 1–13, 2018. ISSN 2168-2267.

Wei-Long Zheng and Bao-Liang Lu. Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Transactions on autonomous mental development*, 7(3):162–175, 2015.

Peixiang Zhong, Di Wang, and Chunyan Miao. Eeg-based emotion recognition using regularized graph neural networks. *IEEE Transactions on Affective Computing*, 13(3):1290–1301, 2020.

Yijin Zhou, Fu Li, Yang Li, Youshuo Ji, Guangming Shi, Wenming Zheng, Lijian Zhang, Yuanfang Chen, and Rui Cheng. Progressive graph convolution network for eeg emotion recognition. *Neurocomputing*, 544:126262, 2023.