Spatial-Energy-Aware Dynamic Filtering with Sparse Graph Convolutions for EEG Emotion Recognition

Jingjing Hu[1,4], **Shuaiqi Fu**[2,4], **Zhan Si**[3,4], **Junran Chen**[4]
1.School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China
2.School of Software and IoT Engineering, Jiangxi University of Finance and Economics, Nanchang, China
3.Department of Chemistry and Centre for Atomic Engineering of Advanced Materials, Anhui University, Hefei, China
4.Wuhu Brain Media Information Technology Co., Ltd., Wuhu, China
xianhjj623@gmail.com, fushuaiqi8@gmail.com, naa0528@163.com, junranchen412@gmail.com

Abstract

Accurate recognition of human emotions from EEG signals plays a critical role in affective computing and human-computer interaction. However, existing methods face significant challenges in effectively capturing the sparse, dynamic, and energy-dependent characteristics of brain activity during emotional experiences. To address these challenges, we propose a novel framework, **Spatial-Energy-Aware Dynamic Filtering with Sparse Graph Convolutions (SEASGC)**, which rethinks EEG graph modeling from three perspectives: (1) sparse graph construction to adaptively capture the essential functional relationships between brain regions, (2) dynamic and location-dependent filtering to model nonlinear interactions between EEG nodes, and (3) energy-aware feature aggregation to leverage energy changes as critical indicators of emotional intensity. By explicitly integrating these principles, SEASGC provides a more comprehensive representation of EEG signals for emotion recognition. Extensive experiments on benchmark EEG emotion datasets demonstrate that SEASGC achieves state-of-the-art performance, highlighting its effectiveness and generalizability in modeling the complex spatial-spectral dynamics of EEG signals.

Keywords:electroencephalogram (EEG); emotion recognition

Introduction

Emotion is a complex human experience associated with specific patterns of physiological activity (Scherer, 1987). Accurate recognition of human emotional states plays a critical role in affective computing and human-computer interaction. Neural signals, as direct indicators of brain activity (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012), provide objective and reliable evidence for emotion recognition. Among these, **electroencephalography (EEG)** is a widely used tool for recording neural activity and has become a cornerstone in emotion research. The constructionist theory offers a crucial framework for understanding the neural representation of emotions (Lindquist et al., 2012), suggesting that emotional experiences are products of psychological events formed by coordinated interactions between cortical and subcortical neural networks. These interactions manifest in EEG signals as fluctuations across different frequency bands (*e.g.*, δ, θ, α, β, γ) and spatial distribution patterns. Furthermore, the intensity of emotional experiences has been shown to correlate closely with changes in the coordinated interactions of brain networks (Du, Fu, Wen, & He, 2023), emphasizing the importance of modeling brain activity from a complex network perspective.

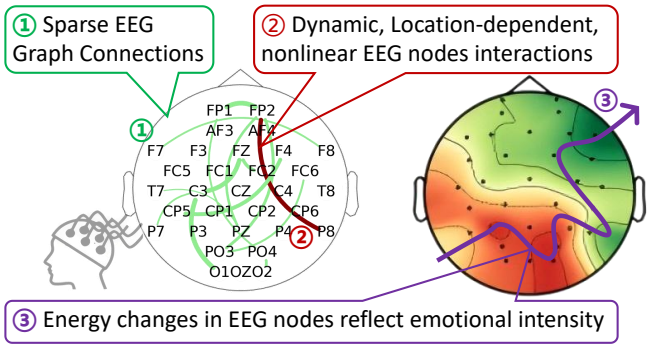

Figure 1: Rethinking EEG Graph Modeling from three key aspects: (1) Sparse EEG graph connections capture the essential functional relationships between brain regions while avoiding overfitting. (2) Dynamic, location-dependent, nonlinear interactions between EEG nodes reflect the complex and flexible nature of brain activity during emotional experiences. (3) Energy changes in EEG nodes serve as crucial indicators of emotional intensity, providing additional insights into emotion representation.

However, EEG-based emotion recognition faces significant challenges. Beyond the inherent complexity in linking the hidden neural activity with overt emotional expressions, there remain critical questions about the specific mechanisms underlying the interactions between different brain regions. Specifically, *how these interactions dynamically adapt to spatial and spectral changes in neural signals, and how these dynamics contribute to emotional experiences, remain inadequately understood.* Addressing these challenges necessitates advanced methods to capture the spatial-spectral complexity and dynamic nature of brain networks during emotional experiences.

Existing methods for EEG-based emotion recognition have made substantial progress, yet several limitations persist. Traditional approaches often rely on handcrafted features, which may fail to capture the complex and dynamic patterns of brain activity (Zeng, Pantic, Roisman, & Huang, 2007; Hu, Guo, Li, et al., 2024). More recent methods, particularly those employing graph-based models, have shown promise in capturing the spatial and spectral properties of EEG signals (Song, Zheng, Song, & Cui, 2018; Zhang et al.,

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

2021; Hu et al., 2024; Wu et al., 2020; Zhong, Wang, & Miao, 2022; Liu et al., 2024; Hu, Guo, Si, et al., 2024). For example, DGCNN (Song et al., 2018) effectively models spatial dependencies, while SparseDGCNN (Zhang et al., 2021) incorporates sparsity constraints to reduce redundancy in connections. Methods like MTGNN (Wu et al., 2020) and RGNN (Zhong et al., 2022) further integrate spectral information into graph representations.

Despite these advancements, several challenges persist. First, most existing graph-based methods assume a fixed, fully-connected graph structure or predefine connections between EEG electrodes (Song et al., 2018; Liu et al., 2024), which may fail to reflect the sparse and dynamic nature of neural interactions (Lindquist et al., 2012; Zhang et al., 2021). Second, while some methods focus on spatial and spectral properties (Zhong et al., 2022; Zhang et al., 2021), they often overlook the energy fluctuations in EEG signals that are closely related to emotional intensity (Du et al., 2023). These energy changes provide crucial insights into the dynamics of emotional experiences but are underexplored in existing approaches. Additionally, the nonlinear and location-dependent spatial interactions between EEG nodes, which play an essential role in encoding complex brain dynamics, remain inadequately modeled.

To address these challenges, we propose a novel framework, **Spatial-Energy-Aware Dynamic Filtering with Sparse Graph Convolutions (SEASGC)**. As illustrated in Figure 1, our approach rethinks EEG graph modeling from three key aspects: **(1) incorporating sparse EEG graph connections to better capture the essential functional relationships between brain regions, (2) modeling dynamic, location-dependent, nonlinear interactions between EEG nodes, and (3) leveraging energy changes in EEG nodes as critical indicators of emotional intensity.** By explicitly integrating these principles, our method aims to provide a more comprehensive representation of EEG signals for emotion recognition.

Specifically, SEASGC consists of the following key components: **(1) Sparse EEG Connection Graph Construction.** Recognizing the sparse and dynamic nature of functional brain interactions, we impose a sparsity constraint on the graph adjacency matrix to adaptively learn the essential connections between EEG electrodes. This avoids the over-simplified assumption of fully-connected graphs and ensures that only the most relevant connections are retained for emotion representation. A residual mechanism is further integrated to enhance the robustness of the learned EEG representations. **(2) Multi-scale Gaussian Spatial Filtering.** To model the nonlinear and location-dependent interactions between EEG nodes, we introduce a set of multi-scale Gaussian basis functions that dynamically generate location-dependent filters. These filters are learned through an MLP and enable the model to capture diverse spatial interactions across EEG nodes, reflecting the flexible nature of brain activity during emotional experiences. **(3) Energy-Aware Gradient Aggregation.** As energy changes in EEG signals closely correlate with emotional intensity, we propose an energy-aware gradient aggregation mechanism to explicitly capture these variations. By computing gradient-based features that encode energy changes between EEG nodes, this module enriches the representation of EEG signals and enhances the model's ability to distinguish emotional states. By combining these components, SEASGC effectively addresses the limitations of existing EEG-based emotion recognition methods, providing a more comprehensive approach to capturing the spatial-spectral complexity, dynamic nature, and energy variations of EEG signals.

The main contributions of this work are as follows:

- We propose SEASGC, a novel framework for EEG-based emotion recognition that integrates sparse graph learning, dynamic spatial filtering, and energy-aware modeling to comprehensively capture the spatial, spectral, and dynamic characteristics of EEG signals.
- SEASGC introduces three key innovations: a sparsity-constrained graph construction method to adaptively learn EEG graph structures, a multi-scale Gaussian spatial filtering module to model nonlinear and location-dependent interactions, and an energy-aware gradient aggregation mechanism to leverage energy changes as critical indicators of emotional intensity.
- Extensive experiments on benchmark EEG emotion datasets demonstrate that SEASGC achieves state-of-the-art performance, validating its effectiveness and generalizability.

Related Work

EEG-based emotion recognition has been extensively studied due to its critical role in affective computing and human-computer interaction. Early approaches primarily relied on handcrafted features or shallow learning models, which often failed to capture the complex and dynamic patterns of brain activity (Zeng et al., 2007; Schirrmeister et al., 2017). Recent advancements have shifted towards graph-based methods that leverage the structural and functional connectivity of EEG signals. For instance, DGCNN (Song et al., 2018) and SparseDGCNN (Zhang et al., 2021) introduced graph convolutional networks to model spatial dependencies between EEG nodes, with the latter incorporating sparsity constraints to reduce redundancy. Similarly, MTGNN (Wu et al., 2020) and RGNN (Zhong et al., 2022) integrated spectral information into graph modeling, further enhancing the representation of EEG signals.

Despite these advancements, existing methods face several limitations. Many assume fixed or predefined graph structures that fail to adapt to the dynamic and sparse nature of neural interactions (Song et al., 2018; Zhang et

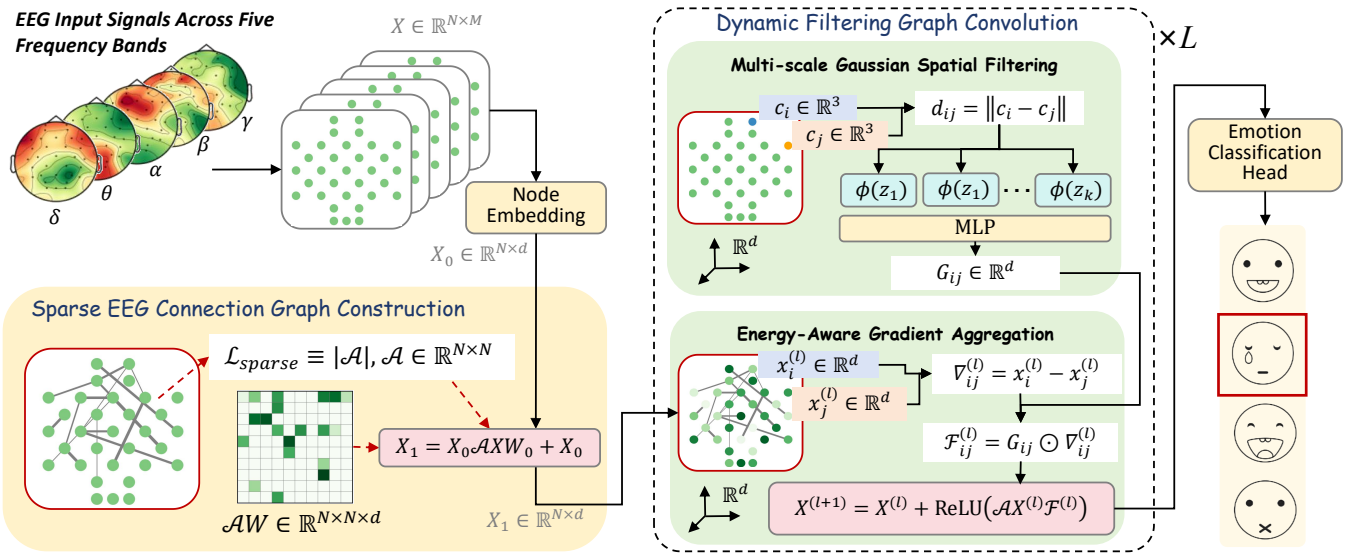

Figure 2: The overall framework of the proposed Spatial-Energy-Aware Dynamic Filtering with Sparse Graph Convolutions (SEASGC) for EEG Emotion Recognition. The framework is composed of two core modules: **(1) Sparse EEG Connection Graph Construction.** EEG signals across five frequency bands (δ, θ, α, β, γ) are treated as inputs to construct a sparse EEG graph. We impose sparsity constraints on the adjacency matrix $\mathcal{A} \in \mathbb{R}^{N\times N}$ using a loss function $\mathcal{L}_{sparse}$, and perform a graph convolution $X_0\mathcal{A}W$ with a learnable weight matrix $W \in \mathbb{R}^{N\times N\times d}$. This is combined with a residual structure to refine the representation and produce X_1. **(2) Dynamic Filtering Graph Convolution.** This module consists of two subcomponents: *Multi-scale Gaussian Spatial Filtering*, which introduces K Gaussian basis functions and an MLP layer to generate location-dependent filters $\mathcal{G}_{ij} \in \mathbb{R}^d$; and *Energy-Aware Gradient Aggregation*, which computes gradient-based features $\nabla_{ij}^{(1)}$ to reflect energy changes. The dynamic filter $\mathcal{F}_{ij}^{(l)}$, combining $\mathcal{G}_{ij}$ and $\nabla_{ij}^{(l)}$, is applied to perform graph convolution, resulting in the final representation $X^{(L)}$. Finally, an emotion classification head predicts the emotional states based on $X^{(L)}$.

al., 2021). Furthermore, while some models incorporate spatial and spectral features (Wu et al., 2020; Zhong et al., 2022), they often neglect energy fluctuations in EEG signals, which are closely related to emotional intensity (Du et al., 2023). Additionally, the nonlinear and location-dependent interactions between EEG nodes, which are critical for encoding complex brain dynamics, remain underexplored (Ding, Robinson, Zhang, Zeng, & Guan, 2022; Ding, Robinson, Tong, Zeng, & Guan, 2024). To address these gaps, SEASGC introduces several innovations, including adaptive sparse graph construction, multi-scale Gaussian spatial filtering, and energy-aware gradient aggregation, which explicitly capture the spatial, spectral, and energy-dependent dynamics of EEG signals. These contributions align with recent findings on neural connectivity dynamics (Lindquist et al., 2012; Lettieri et al., 2022), demonstrating the importance of spatial and spectral features, as well as energy variations, in modeling emotional experiences.

Methodology

To address the challenges in EEG-based emotion recognition, we propose the **Spatial-Energy-Aware Dynamic Filtering with Sparse Graph Convolutions (SEASGC)** framework. This method rethinks EEG graph modeling from three key aspects: **(1) Sparse EEG Connection Graph Construction:** capturing essential functional relationships between brain regions, **(2) Multi-scale Gaussian Spatial Filtering:** modeling nonlinear and location-dependent interactions, **(3) Energy-Aware Gradient Aggregation:** leveraging energy changes to reflect emotional intensity. The overall framework is shown in Figure 2. Below, we detail the two main modules of SEASGC.

Sparse EEG Connection Graph Construction

To effectively model the neural signal interaction patterns in EEG-based emotion recognition, we first construct a graph structure $\mathcal{G} = \{X, \mathcal{A}, \mathcal{C}\}$ that can capture both spatial and spectral properties of EEG signals. Where nodes $X \in \mathbb{R}^{N\times M}$ represent EEG electrodes with features $x_i \in \mathbb{R}^M$ being the M frequency band values of electrode i. The adjacency matrix $\mathcal{A} \in \mathbb{R}^{N\times N}$ denotes potential neural connections between electrodes. The geometric characteristics $\mathcal{C} \in \mathbb{R}^{N\times 3}$ contain the electrode 3D coordinates from the standard 10-20 system (Koelstra et al., 2012; Soleymani, Lichtenauer, Pun, & Pantic, 2012) or 10-05 system (Duan, Zhu, & Lu, 2013) placement, which reflect the spatial location of the EEG signals. To enhance the representation of EEG signals, the input X is projected to a latent space using a learnable embedding matrix $E \in \mathbb{R}^{M\times d}$, resulting in $X_0 = X \cdot E$, where

$X_0 \in \mathbb{R}^{N\times d}$.

Sparse Edge Optimization. Since the precise graph structure is often unknown (Du et al., 2023), we initialize a fully-connected graph structure where edge weights $\mathcal{W} \in \mathbb{R}^{N\times N\times d}$ can be adaptively learned to capture dynamic neural interactions. Based on tomography studies suggesting functional relationships between electrodes are typically sparse (Zhang et al., 2021), we introduce sparsity constraints to the adjacency matrix $\mathcal{A}$ during training. Specifically, we first introduce W_A as a learnable parameter of size $N \times N$, and use a hard activation function with a threshold of 0.5 that maps elements of W_A to binary values (0 or 1), to represent the current adjacency matrix $\mathcal{A}$. This optimization is:

$$\mathcal{L}_{sparse} = \frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}|\mathcal{A}| - \alpha N^2\right), \tag{1}$$

where α (*i.e.*, $\alpha = 0.25$) controls the proportion of retained edges. This approach ensures that the number of edges in the graph is constrained within a desirable range, reflecting realistic sparsity patterns.

Preliminary Signal Convolution. To enhance the initial representation of EEG signals while preserving the original feature information, we perform a preliminary signal convolution using a learnable weight matrix $W \in \mathbb{R}^{N\times N\times d}$, followed by a residual connection:

$$X_1 = X_0\mathcal{A}W + X_0, \quad X_1 \in \mathbb{R}^{N\times d}. \tag{2}$$

This step captures the initial spatial dependencies and feature interactions, providing a foundation for subsequent dynamic filtering while maintaining the original feature information through the residual connection.

Dynamic Filtering Graph Convolution

After constructing the sparse EEG graph structure, we designed the **Dynamic Filtering Graph Convolution** module to better capture the dynamic, nonlinear, and spatially dependent interactions in EEG signals. This module consists of two subcomponents: **Multi-scale Gaussian Spatial Filtering** and **Energy-Aware Gradient Aggregation**, which address the spatial complexity and energy variations of EEG signals associated with emotional representation.

Multi-scale Gaussian Spatial Filtering. To effectively model the nonlinear, dynamic, and location-dependent spatial dependencies between EEG electrodes, we propose a multi-scale Gaussian filtering mechanism that adapts to the spatial configuration of electrode locations. Given the 3D coordinates $c_i, c_j \in \mathbb{R}^3$, we compute pairwise distances $d_{ij} = \|c_i - c_j\|_2$, which are then projected through K Gaussian basis functions parameterized by learnable bandwidth γ and spatial anchors $\{z_k\}_{k=1}^{K}$:

$$\mathcal{G}_{ij} = \mathrm{MLP}\left(\bigoplus_{k=1}^{K}\exp(-\gamma(d_{ij} - z_k)^2)\right) \in \mathbb{R}^d. \tag{3}$$

This design captures three key properties of EEG spatial interactions. First, the **dynamicity** is achieved by allowing the Gaussian basis functions to adjust their influence based on the learned γ and z_k, which adapt to varying spatial configurations. Second, the **nonlinearity** is introduced through the exponential transformation and the MLP, enabling the model to represent complex spatial relationships. Third, the **location-dependence** is ensured by explicitly incorporating the pairwise distances d_{ij} and generating filters $\mathcal{G}_{ij}$ that are unique to each electrode pair. The multi-scale design ($K = 5$ anchors spanning $[0, 0.4]$ m) ensures the model captures both local interactions (e.g., $d_{ij} \leq 0.1$ m, nearby electrodes) and global dependencies (e.g., $d_{ij} \geq 0.3$ m, distant electrodes), while Gaussian kernels provide a smooth and biologically plausible transition for modeling neural signal propagation. The MLP layer further refines these filters, generating $\mathcal{G}_{ij}$ with diverse spatial patterns tailored to the unique characteristics of EEG signals.

Energy-Aware Gradient Aggregation. To encode energy variations in spectral features that reflect emotional intensity, we compute gradient features $\nabla_{ij}^{(l)} = x_i^{(l)} - x_j^{(l)} \in \mathbb{R}^d$ at layer l, representing the difference in spectral embeddings between connected EEG nodes i and j. These gradients capture localized energy changes across the EEG graph and are concatenated with the spatial filters $\mathcal{G}_{ij}$, which are shared across layers to ensure consistent modeling of spatial structure. Specifically, the gradients and spatial filters are processed through a gated aggregation mechanism to generate dynamic filters $\mathcal{F}_{ij}^{(l)}$:

$$\mathcal{F}_{ij}^{(l)} = \sigma(W_g[\mathcal{G}_{ij} \oplus \nabla_{ij}^{(l)}]) \odot (W_f\mathcal{G}_{ij}), \tag{4}$$

where $\sigma(\cdot)$ denotes the sigmoid gating function, $\oplus$ indicates concatenation, and $\odot$ represents element-wise multiplication. Here, W_g and W_f are learnable weight matrices. The shared spatial filters $\mathcal{G}_{ij}$ ensure that the dynamic filters $\mathcal{F}_{ij}^{(l)}$ are consistently grounded in the underlying spatial relationships, while the layer-specific gradients $\nabla_{ij}^{(l)}$ allow the model to adaptively emphasize energy-salient connections at each layer.

Dynamic Filtering Graph Convolution. The dynamic filtering graph convolution integrates spatial and spectral energy dynamics to propagate emotion-relevant features while suppressing irrelevant connections. At each layer l, the convolution operation is defined as:

$$X^{(l+1)} = \mathrm{LayerNorm}(X^{(l)} + \mathrm{ReLU}(\mathcal{A}X^{(l)}\mathcal{F}^{(l)})), \tag{5}$$

Table 1: Performance comparison of all baselines in terms of Accuracy for classification tasks (the greater, the better) on DEAP and DREAMER. The best results are bolded.

Methods	DEAP				DREAMER			
	Arousal		Valence		Arousal		Valence	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
DGCNN (Song et al., 2018)	64.89 ± 8.73	66.17 ± 10.16	68.14 ± 6.72	72.89 ± 6.44	68.56 ± 8.51	67.85 ± 9.55	68.00 ± 6.51	63.95 ± 7.11
MM-ResLSTM (Ma, Tang, Zheng, & Lu, 2019)	70.84 ± 9.48	68.07 ± 10.81	70.89 ± 7.64	64.21 ± 7.12	68.43 ± 7.25	63.77 ± 7.31	71.34 ± 6.72	70.63 ± 5.39
ACRNN (Tao et al., 2020)	69.71 ± 8.41	65.48 ± 8.04	68.82 ± 6.23	69.33 ± 7.09	70.24 ± 9.26	65.78 ± 9.54	65.23 ± 5.11	60.27 ± 6.99
MTGNN (Wu et al., 2020)	70.98 ± 7.37	68.45 ± 9.98	69.14 ± 6.47	68.97 ± 6.19	71.24 ± 7.26	69.87 ± 9.53	67.07 ± 5.93	68.07 ± 8.31
SST-EmotionNet (Jia et al., 2020)	73.49 ± 6.61	76.84 ± 10.72	70.94 ± 6.21	69.87 ± 6.99	71.54 ± 8.43	69.83 ± 8.71	70.39 ± 7.03	68.10 ± 5.27
HetEmotionNet (Jia et al., 2021)	74.36 ± 7.85	75.71 ± 10.82	72.08 ± 5.66	71.15 ± 6.1	70.58 ± 7.63	69.24 ± 4.92	70.03 ± 6.64	67.49 ± 6.38
ASTG-LSTM (Li, Zheng, Zong, Chang, & Lu, 2021)	72.55 ± 6.83	73.18 ± 11.01	71.96 ± 6.48	70.8 ± 7.26	70.23 ± 8.43	69.14 ± 5.34	68.76 ± 6.76	68.94 ± 5.49
RGNN (Zhong et al., 2022)	60.55 ± 5.31	62.49 ± 7.14	61.37 ± 6.93	60.75 ± 6.16	62.48 ± 5.03	61.07 ± 6.47	61.98 ± 5.87	61.11 ± 5.37
TSception (Ding et al., 2022)	74.38 ± 8.18	76.21 ± 9.84	68.94 ± 7.25	72.04 ± 6.94	69.77 ± 6.56	69.27 ± 9.43	68.16 ± 7.12	65.06 ± 6.71
ST-GCLSTM (Feng, Cheng, Zhao, Deng, & Zhang, 2022)	76.68 ± 7.37	71.34 ± 9.85	71.81 ± 6.72	70.36 ± 9.14	69.29 ± 6.21	69.43 ± 5.93	70.49 ± 5.98	70.23 ± 5.84
LGGNet (Ding et al., 2024)	72.31 ± 8.38	73.28 ± 8.78	68.53 ± 7.79	69.15 ± 7.59	70.49 ± 8.84	68.94 ± 8.47	68.69 ± 6.08	66.12 ± 8.14
VSGT (Liu et al., 2024)	78.95 ± 6.89	78.65 ± 8.24	74.15 ± 6.41	74.05 ± 6.43	75.19 ± 6.87	72.98 ± 7.54	73.41 ± 6.36	72.73 ± 5.71
SEASGC (Ours)	**80.03 ± 5.94**	**79.96 ± 7.35**	**77.12 ± 5.63**	**77.01 ± 5.77**	**76.97 ± 5.91**	**73.89 ± 6.42**	**75.22 ± 5.37**	**74.56 ± 5.12**

where $\mathcal{A}$ is the sparse adjacency matrix representing the graph structure, and $\mathcal{F}^{(l)} \in \mathbb{R}^{N\times N\times d}$ is the dynamic filter combining spatial filters $\mathcal{G}$ and energy-aware gradients $\nabla^{(l)}$. The sparse structure of $\mathcal{A}$ (constrained by $\mathcal{L}_{sparse}$) ensures that only neurologically plausible connections are retained.

The graph convolution iteratively refines the node embeddings through L (*i.e.*, $L = 2$) layers, where the residual connection $(X^{(l)} + \cdot)$ preserves essential features from the previous layer, ReLU introduces nonlinearity, and LayerNorm stabilizes training. By leveraging both geometric constraints (via $\mathcal{A}$) and spectral energy dynamics (via $\mathcal{F}$), this architecture captures the complex neural interaction patterns critical for emotion recognition, as illustrated in Figure 2.

EEG Emotion Classification

The final EEG node representations $X^{(L)} \in \mathbb{R}^{N\times d}$, obtained from the dynamic filtering graph convolution module, are mapped to the emotion category space through a fully connected layer, producing predicted probabilities $p_{i,c}$ for each sample i and category c. The model is trained end-to-end with the following loss function:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log(p_{i,c}) + \lambda\mathcal{L}_{sparse}, \tag{6}$$

where the first term is the cross-entropy loss for classification and the second term $\mathcal{L}_{sparse}$ enforces sparsity in the graph structure. Here, $y_{i,c}$ is the one-hot encoded ground truth, $p_{i,c}$ is the predicted probability, and λ (set to $\lambda = 1.0$) controls the trade-off. This loss ensures high classification accuracy while maintaining a sparse and interpretable graph structure.

Experiments

We evaluate our proposed SEASGC framework on two widely used EEG emotion recognition datasets: DEAP (Koelstra et al., 2012) and DREAMER (Soleymani et al., 2012).

Both datasets are designed for dimensional emotion analysis, capturing affective states along valence (pleasure) and arousal (excitement) dimensions. DEAP includes 32-channel EEG recordings, while DREAMER features 14-channel recordings. Both are downsampled to 128 Hz, and continuous emotional ratings are converted into binary classifications (high/low) using thresholds of 5 and 3, respectively, following standard protocols (Koelstra et al., 2012; Soleymani et al., 2012).

For fair comparisons, we apply the same preprocessing and training pipeline to SEASGC and all baselines. We use 10-fold cross-validation, where one fold serves as the test set and the rest is split into training and validation sets (8:1). EEG signals are segmented into 4-second non-overlapping samples for classification. Accuracy and macro-F1 score are reported as evaluation metrics, with macro-F1 addressing class imbalance.

Overall Results

Table 1 presents the performance comparison of our proposed SEASGC framework with state-of-the-art methods across three benchmark datasets: DEAP, DREAMER, and SEED. SEASGC consistently outperforms all baselines in terms of both accuracy and F1-score, demonstrating its superior ability to model the complex spatial, spectral, and energy-dependent dynamics of EEG signals. Notably, SEASGC achieves the highest accuracy of 80.03% on DEAP (Arousal), 76.97% on DREAMER (Arousal), and 77.63% on SEED, surpassing traditional and graph-based approaches such as DGCNN (Song et al., 2018), SparseDGCNN (Zhang et al., 2021), and VSGT (Liu et al., 2024). These results highlight the effectiveness of our sparse graph learning, dynamic filtering, and energy-aware modeling in capturing emotion-relevant features while maintaining a compact and interpretable graph structure.

Key Components Analysis

Table 2 shows the ablation study results on the DEAP dataset, highlighting the contributions of SEASGC's components. The full model achieves the highest performance (Arousal: 80.03% accuracy, 79.96% F1; Valence: 77.12% accuracy, 77.01% F1). Removing Sparse Graph Construction (SGC) reduces accuracy to 77.15% (Arousal) and 74.65% (Valence), showing its importance for capturing sparse connectivity.

Table 2: Ablation study results on the DEAP dataset. The table shows the impact of removing key components from SEASGC on classification accuracy (%) and macro-F1 score (%) with standard deviation. The best results are highlighted in bold.

Model	DEAP (Arousal)		DEAP (Valence)	
	Accuracy	F1-score	Accuracy	F1-score
Full SEASGC	**80.03**	**79.96**	**77.12**	**77.01**
w/o SGC	77.15	76.88	74.65	74.21
w/o MSGF	77.13	77.09	75.15	75.04
w/o EAGA	78.04	77.77	75.88	75.63
w/o MSGF & EAGA	75.98	75.34	73.42	73.01

Excluding Multi-scale Gaussian Spatial Filtering (MSGF) or Energy-Aware Gradient Aggregation (EAGA) also results in notable performance drops, with the most significant degradation (Arousal: 75.98%, Valence: 73.42%) when both are removed. These results confirm that each component is essential and their integration allows SEASGC to effectively capture spatial, spectral, and energy-related features for EEG emotion recognition.

Parameter Sensitivity Analysis

We conducted a sensitivity analysis to evaluate the impact of two key hyperparameters in SEASGC: the sparse ratio (α) and the number of Gaussian filters (K). For the sparse ratio, which controls the number of edges in the graph, performance improves as α increases from 0.1 to 0.25, reaching an accuracy of 80.03%. However, further increasing α reduces accuracy, indicating that moderate sparsity effectively balances capturing critical connections and avoiding noise. For K, which determines the number of multi-scale Gaussian filters, accuracy improves significantly from $K = 1$ to $K = 5$, reflecting the filters' ability to capture both localized and global spatial dependencies. Beyond $K = 5$, performance saturates or slightly declines due to redundancy. These results suggest that $\alpha = 0.25$ and $K = 5$ are optimal, highlighting the importance of tuning hyperparameters to maximize SEASGC's performance.

Visualization Analysis

Figure 4 illustrates the key brain region connections for Valence Label 1 (positive emotions) and Valence Label 0 (negative emotions). For Valence Label 1, the strongest connections are primarily located in the frontal lobe (*e.g.*, FP1-AF3, FP2-AF4), central axis (FZ-CZ), and parietal-occipital regions (P3-O1, P4-O2). These connections are associated with emotional regulation, reward mechanisms, and enhanced cognitive control under positive emotional states. In contrast, Valence Label 0 emphasizes connections in the frontal lobe (*e.g.*, FP1-FP2), frontal-central axis (AF3-FZ), and central-parietal regions (CZ-PZ). These regions are linked to emotional suppression, stress responses, and heightened threat sensitivity during negative emotions. The higher weights of frontal-central connections in Valence Label 0 highlight the brain's prioritization of internal emotional processing under stress or negative states.

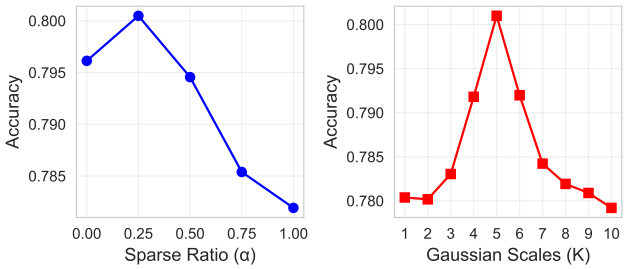

Figure 3: Sensitivity analysis of sparse ratio (α) and Gaussian filter scale (K) on DEAP (Arousal). The model achieves the highest accuracy at $\alpha = 0.25$ and $K = 5$.

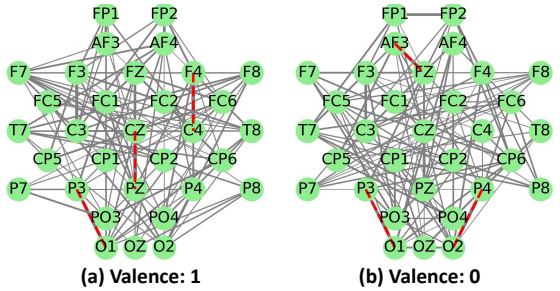

Figure 4: Comparison of brain connectivity patterns for Valence Label 1 (positive emotions) and Valence Label 0 (negative emotions). Red dashed edges highlight the top 3 strongest connections based on weight.

Conclusion

This paper proposed SEASGC, a novel framework for EEG-based emotion recognition that integrates sparse graph construction, dynamic spatial filtering, and energy-aware feature aggregation. Sparse graph construction captures key neural connections while avoiding redundancy, multi-scale Gaussian filtering models nonlinear spatial interactions, and energy-aware aggregation reflects emotional intensity. Experiments on DEAP, DREAMER, and SEED datasets show SEASGC achieves state-of-the-art accuracy and F1-score, outperforming existing methods. Ablation and sensitivity analyses confirm the effectiveness of all components and provide guidance for parameter tuning. SEASGC offers a robust and interpretable solution for EEG emotion recognition, with potential for applications in affective computing and brain-computer interfaces.

Acknowledgments

We would like to thank Wuhu Brain Media Information Technology Co., Ltd., Wuhu, China. and Shandong Tianruixiang Information Technology Co., Ltd., China. Additionally, this work was supported by the Anhui Provincial Graduate Quality Engineering Program.

References

Ding, Y., Robinson, N., Tong, C., Zeng, Q., & Guan, C. (2024). Lggnet: Learning from local-global-graph representations for brain-computer interface. *IEEE Trans. Neural Networks Learn. Syst.*, *35*(7), 9773–9786.

Ding, Y., Robinson, N., Zhang, S., Zeng, Q., & Guan, C. (2022). Tsception: Capturing temporal dynamics and spatial asymmetry from eeg for emotion recognition. *IEEE Trans. Affect. Comput.*, *14*(3), 2238–2250.

Du, C., Fu, K., Wen, B., & He, H. (2023). Topographic representation of visually evoked emotional experiences in the human cerebral cortex. *Iscience*, *26*(9).

Duan, R.-N., Zhu, J.-Y., & Lu, B.-L. (2013). Differential entropy feature for eeg-based emotion classification. In *2013 6th international ieee/embs conference on neural engineering (ner)* (p. 81-84).

Feng, L., Cheng, C., Zhao, M., Deng, H., & Zhang, Y. (2022). Eeg-based emotion recognition using spatial-temporal graph convolutional LSTM with attention mechanism. *IEEE J. Biomed. Health Informatics*, *26*(11), 5406–5417.

Hu, J., Guo, D., Li, K., Si, Z., Yang, X., Chang, X., & Wang, M. (2024). Unified static and dynamic network: Efficient temporal filtering for video grounding. *arXiv preprint arXiv:2403.14174*.

Hu, J., Guo, D., Li, K., Si, Z., Yang, X., & Wang, M. (2024). Maskable retentive network for video moment retrieval. In *Proceedings of the 32nd acm international conference on multimedia* (pp. 1476–1485).

Hu, J., Guo, D., Si, Z., Liu, D., Diao, Y., Zhang, J., ... Wang, M. (2024). Mol-mamba: Enhancing molecular representation with structural & electronic insights. *arXiv preprint arXiv:2412.16483*.

Jia, Z., Lin, Y., Cai, X., Chen, H., Gou, H., & Wang, J. (2020). Sst-emotionnet: Spatial-spectral-temporal based attention 3d dense network for EEG emotion recognition. In *Acm mm* (pp. 2909–2917). ACM.

Jia, Z., Lin, Y., Wang, J., Feng, Z., Xie, X., & Chen, C. (2021). Hetemotionnet: Two-stream heterogeneous graph recurrent neural network for multi-modal emotion recognition. In *Acm mm* (pp. 1047–1056). ACM.

Koelstra, S., Mühl, C., Soleymani, M., Lee, J., Yazdani, A., Ebrahimi, T., ... Patras, I. (2012). DEAP: A database for emotion analysis ;using physiological signals. *IEEE Trans. Affect. Comput.*, *3*(1), 18–31.

Lettieri, G., Handjaras, G., Setti, F., Cappello, E. M., Bruno, V., Diano, M., ... Cecchetti, L. (2022). Default and control network connectivity dynamics track the stream of affect at multiple timescales. *Social cognitive and affective neuroscience*, *17*(5), 461–469.

Li, X., Zheng, W., Zong, Y., Chang, H., & Lu, C. (2021). Attention-based spatio-temporal graphic LSTM for EEG emotion recognition. In *Ijcnn* (pp. 1–8). IEEE.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behavioral and brain sciences*, *35*(3), 121–143.

Liu, C., Zhou, X., Xiao, J., Zhu, Z., Zhai, L., Jia, Z., & Liu, Y. (2024). VSGT: variational spatial and gaussian temporal graph models for eeg-based emotion recognition. In *Ijcai* (pp. 3078–3086). ijcai.org.

Ma, J., Tang, H., Zheng, W.-L., & Lu, B.-L. (2019). Emotion recognition using multimodal residual lstm network. In *Acm mm* (pp. 176–183).

Scherer, K. R. (1987). Toward a dynamic theory of emotion: The component process model of affective states. *Geneva studies in Emotion and Communication*, *1*, 1–98.

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., ... Ball, T. (2017). Deep learning with convolutional neural networks for eeg decoding and visualization. *Human brain mapping*, *38*(11), 5391–5420.

Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2012). A multimodal database for affect recognition and implicit tagging. *IEEE Trans. Affect. Comput.*, *3*(1), 42–55.

Song, T., Zheng, W., Song, P., & Cui, Z. (2018). Eeg emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans. Affect. Comput.*, *11*(3), 532–541.

Tao, W., Li, C., Song, R., Cheng, J., Liu, Y., Wan, F., & Chen, X. (2020). Eeg-based emotion recognition via channel-wise attention and self attention. *IEEE Trans. Affect. Comput.*, *14*(1), 382–393.

Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. In *Kdd* (pp. 753–763). ACM.

Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2007). A survey of affect recognition methods: audio, visual and spontaneous expressions. In *Proceedings of the 9th international conference on multimodal interfaces* (pp. 126–133).

Zhang, G., Yu, M., Liu, Y.-J., Zhao, G., Zhang, D., & Zheng, W. (2021). Sparsedgcnn: Recognizing emotion from multichannel eeg signals. *IEEE Transactions on Affective Computing*, *14*(1), 537–548.

Zhong, P., Wang, D., & Miao, C. (2022). Eeg-based emotion recognition using regularized graph neural networks. *IEEE Trans. Affect. Comput.*, *13*(3), 1290–1301.