



## OPEN ACCESS

## EDITED BY

Redha Taiar,  
Université de Reims Champagne-Ardenne,  
France

## REVIEWED BY

Francesca Gasparini,  
University of Milano-Bicocca, Milan, Italy  
Guojun Dai,  
Hangzhou Dianzi University, China  
Ziyu Jia,  
Chinese Academy of Sciences (CAS), China

## \*CORRESPONDENCE

Liya Huang  
✉ huangly@njupt.edu.cn

RECEIVED 15 April 2024

ACCEPTED 26 June 2024

PUBLISHED 19 July 2024

## CITATION

Chen W, Liao Y, Dai R, Dong Y and  
Huang L (2024) EEG-based emotion  
recognition using graph convolutional neural  
network with dual attention mechanism.  
*Front. Comput. Neurosci.* 18:1416494.  
doi: 10.3389/fncom.2024.1416494

## COPYRIGHT

© 2024 Chen, Liao, Dai, Dong and Huang.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# EEG-based emotion recognition using graph convolutional neural network with dual attention mechanism

Wei Chen, Yuan Liao, Rui Dai, Yuanlin Dong and Liya Huang\*

College of Electronic and Optical Engineering & College of Flexible Electronics (Future Technology), Nanjing University of Posts and Telecommunications, Nanjing, China

EEG-based emotion recognition is becoming crucial in brain-computer interfaces (BCI). Currently, most researches focus on improving accuracy, while neglecting further research on the interpretability of models, we are committed to analyzing the impact of different brain regions and signal frequency bands on emotion generation based on graph structure. Therefore, this paper proposes a method named Dual Attention Mechanism Graph Convolutional Neural Network (DAMGCN). Specifically, we utilize graph convolutional neural networks to model the brain network as a graph to extract representative spatial features. Furthermore, we employ the self-attention mechanism of the Transformer model which allocates more electrode channel weights and signal frequency band weights to important brain regions and frequency bands. The visualization of attention mechanism clearly demonstrates the weight allocation learned by DAMGCN. During the performance evaluation of our model on the DEAP, SEED, and SEED-IV datasets, we achieved the best results on the SEED dataset, showing subject-dependent experiments' accuracy of 99.42% and subject-independent experiments' accuracy of 73.21%. The results are demonstrably superior to the accuracies of most existing models in the realm of EEG-based emotion recognition.

## KEYWORDS

attention mechanism, EEG, electrode channels, emotion recognition, frequency bands, graph convolutional neural network, transformer

## 1 Introduction

EMOTION is the subjective emotional response that humans experience in specific moments or situations. It plays a vital role in correctly interpreting behavior and facilitating effective communication (Jenke et al., 2014). In the development of brain-computer interface (BCI) systems, the urgency to empower machines with the ability to assist in analyzing human emotions is significant (Gu et al., 2023). Consequently, emotion recognition has emerged as one of the crucial research directions in affective computing (Hu et al., 2019). Through numerous studies, it has been discovered that the generation of human emotions is highly correlated with electrical signals in the cerebral cortex of the brain (She et al., 2023). Additionally, humans may involuntarily or intentionally conceal their real emotions through facial expressions and language except EEG signals (Zhang et al., 2020). As a result, researchers prefer emotion recognition methods based on EEG signals as they are more reliable and objective in capturing an individual's emotional state (Lu et al., 2023).

In earlier studies on emotion recognition, traditional machine learning methods were predominantly relied upon, such as Support Vector Machines (SVM) (Kumar and Nataraj, 2019), which were extensively used due to their effectiveness in handling high-dimensional feature spaces and their ability to perform well with a limited amount of training data. However, as deep learning continues to progress, we are now witnessing a shift in the landscape. It has not only demonstrated significant performance in the field of computer vision (Yuan et al., 2018; Zhang and Zheng, 2022) and natural language processing (Lauriola and Aioli, 2022), but has also gained widespread popularity in biomedical signal processing (Rahman et al., 2021). Initially, Wang et al. (2022) utilized convolutional neural network (CNN) to classify positive, neutral, and negative emotions. Building on the premise that CNN plays a vital role in emotion detection, Yang et al. (2018) designed a parallel convolutional recurrent neural network model for emotion recognition, which yielded promising results. To further investigate the temporal and spatial aspects of brain networks, Cui et al. (2022) employed a CNN-BiLSTM architecture to investigate the temporal complexity and spatial location of brain networks. At the same time, some researchers (Jia et al., 2021; Yang et al., 2024) recognized that the brain exhibits a complex graph structure in three-dimensional space, leading to investigations from spatial perspective. Liu et al. (2024) provide a comprehensive and systematic review of existing graph neural networks in EEG-based emotion recognition. For example, Liu et al. (2023) and Ding et al. (2022) effectively leveraged Graph Convolutional Neural Networks for the efficient feature extraction through both global information aggregation between brain regions and local information integration within brain regions under emotional states. This represents a promising start in EEG-based emotion recognition research, yet they did not proceed to further explore critical factors for classifying emotions. Liu et al. (2023) utilized transformer model for multimodal knowledge extraction, thereby enhancing recognition performance. Gong et al. (2023) designed an attention-based feature extraction and fusion module, which can selectively obtain key features based on their spatial and temporal significances. Guo et al. (2022) delved into understanding the dependence of emotion recognition building on the transformer model on each EEG channel and visualized the captured features. While they succeeded in extracting vital information from time segments or channels, their approach overlooked the foundational graph structure. This oversight led to the loss of significant information, consequently capping the potential of their model's classification capability.

To address these challenges, our study introduces a brain decoding approach that primarily relies on graph convolutional neural network and attention mechanism of transformer. To be precise, we construct a three-dimensional spatial adjacency matrix and employ graph convolutional neural networks to aggregate information from multiple channels, extracting representative spatiotemporal features. Additionally, we utilize two attention mechanisms: electrode channel attention and signal frequency band attention. These mechanisms reveal the contributions of individual electrode channels to emotional responses in different brain regions and the relative impact of various frequency bands on emotions. By employing these attention mechanisms, we effectively leverage the information embedded in EEG signals, leading to improved overall decoding performance. The main contributions can be summarized as follows:

- 1) A graph convolutional neural network framework with a dual attention mechanism is proposed by combining GCN and Transformer. Experiments were conducted on DEAP, SEED, and SEED-IV datasets, covering binary, ternary, and quaternary classification tasks. The subject-dependent and subject-independent experimental results demonstrate that our model outperforms most existing models.
- 2) The graph convolution operation aggregates brain channel features based on the adjacency matrix composed of three-dimensional distances, effectively utilizing spatial information. After that, dual attention mechanism operates on both electrode channels and signal frequency bands, allocates weights and allows for better extraction of crucial information from temporal and spectral EEG data sequences.
- 3) After model training, we visually analyze the dual attention mechanism through model parameters, which can better analyze the roles of different electrode channels and signal frequency bands in emotion processing. It provides valuable insights for further research on emotion regulation and cognitive processes, offering important clues for exploring these domains.

The remaining sections of this paper are organized as follows: Section 2 outlines the related work, providing context and background for our study. Section 3 describes the methodology, including the development and implementation of our model. Section 4 details the experimental setup, dataset, evaluation metrics and presents the results. Finally, Section 5 concludes the paper with a discussion of the implications, limitations, and future directions for research in brain science.

## 2 Related work

In this section, we begin by showcasing several notable emotional recognition features for EEG signals. After that, we provide a concise overview of GCN and attention mechanism, fundamental components of the proposed model.

### 2.1 Emotional recognition features for EEG signals

In general, most EEG-based emotion recognition methods begin by extracting features from processed EEG signals. Subsequently, these extracted features are then employed as input for classification algorithms to achieve accurate classification of emotional states (Cai et al., 2021).

Zhang et al. (2020) has indicated that the EEG features employed in emotion recognition can be broadly classified into time-domain, frequency-domain and time-frequency domain features. Recent studies (García-Martínez et al., 2021; Yan et al., 2023) shows that the human brain functions as a nonlinear dynamic system, with EEG signals analyzable via nonlinear methods and feature extraction. Commonly used nonlinear features for EEG signals include differential entropy (DE) (Duan et al., 2013), permutation entropy (Nicolau and Georgiou, 2012), discrete wavelet transform (DWT) (Chen et al., 2017), power spectral

density (PSD) (Alam et al., 2020) and various other entropy measures. Among them, DE was initially proposed by Duan et al. (2013) and validated to be effective in the field of emotion recognition. As a result, DE has gained significant popularity as a widely used and effective feature extraction technique in the domain of EEG-based emotion recognition.

## 2.2 Graph convolutional neural network

Graph Convolutional Neural Network is a deep learning model specifically designed for graph data (Kipf and Welling, 2017). It extends the idea of convolutional operations to the graph domain and can encode graph structures and node features in a useful way for semi-supervised classification. GCN has shown excellent performance in tasks such as social networks (Zhong et al., 2020), machine fault diagnosis (Li et al., 2021), and recommendation system (He et al., 2020). Since the brain can be considered as a complex graph network, GCN is capable of effectively capturing both local and global information in brain networks. This enables it to enhance the performance of EEG signal analysis and facilitate research and applications in neuroscience and neurology. Song et al. (2020) first applied GCN for EEG emotion recognition, using a Dynamic Graph Convolutional Neural Network (DGCNN) that operates on multi-channel EEG data. Qiu et al. (2023) introduced multi head attention mechanism and residual network, proposing the Multi-head Residual Graph Convolutional Neural Network (MRGCN) model which combines short-range and long-range connections for EEG-based emotion recognition.

## 2.3 Attention mechanism

Graph Attention Network (GAT) (Veličković et al., 2018), as a graph neural network based on the attention mechanism, has shown excellent performance in processing graph data. However, in brain network research, GAT has a limitation in capturing global information. The attention mechanism of GAT is based on the interaction between nodes for weighted aggregation, which may result in insufficient capture of global information in the entire brain network, especially in the presence of long-range dependencies. The Transformer model (Vaswani et al., 2023) can effectively address this limitation. The Transformer model initially caused a great sensation in natural language processing. Its attention mechanism can adaptively focus on different positions of information according to the task requirements. This adaptability enables the model to capture global relationships and better differentiate important information stored in multiple channels of EEG signals. As a result, many researchers have applied the Transformer model in EEG studies. Wang et al. (2022) utilized the Transformer encoder to capture spatial dependencies between brain regions. Liu et al. (2023) made full use of the relative spatial information in EEG data and constructed a dual-layer capsule network for emotion recognition. (Song et al. (2021) employed attention along the feature channel dimension to weight the preprocessed and spatially filtered data, while also considering the global dependencies along the temporal dimension for emotion recognition.

## 3 Method

In section 3, we provide a comprehensive explanation of the DAMGCN model, as illustrated in Figure 1. The model comprises four main blocks: (A) Feature Extraction Block, (B) Graph Convolution Block, (C) Dual Attention Mechanism Block, and (D) Classifier Block.

### 3.1 Feature extraction block

As shown in Figure 1, we extract DE features from EEG signals and the three-dimensional spatial positions of electrodes for emotion classification. DE is a measure in information theory that describes the uncertainty of continuous random variables and defined as  $H(X)$  (Feutrill and Roughan, 2021):

$$H(X) = - \int_{-\infty}^{+\infty} p(x) \log(p(x)) dx \quad (1)$$

where  $X$  is a time series,  $p(x)$  represents the probability density function of the continuous information. Assuming  $X$  as the EEG signal and following a Gaussian distribution  $N(\mu, \sigma^2)$ ,  $\mu$  and  $\sigma$  are the mean and variance of  $X$ , then  $p(x)$  can be expressed as shown in Eq. (2):

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{x-\mu}{2\sigma^2}\right)} \quad (2)$$

As a result, Eq. (1) can be expressed as:

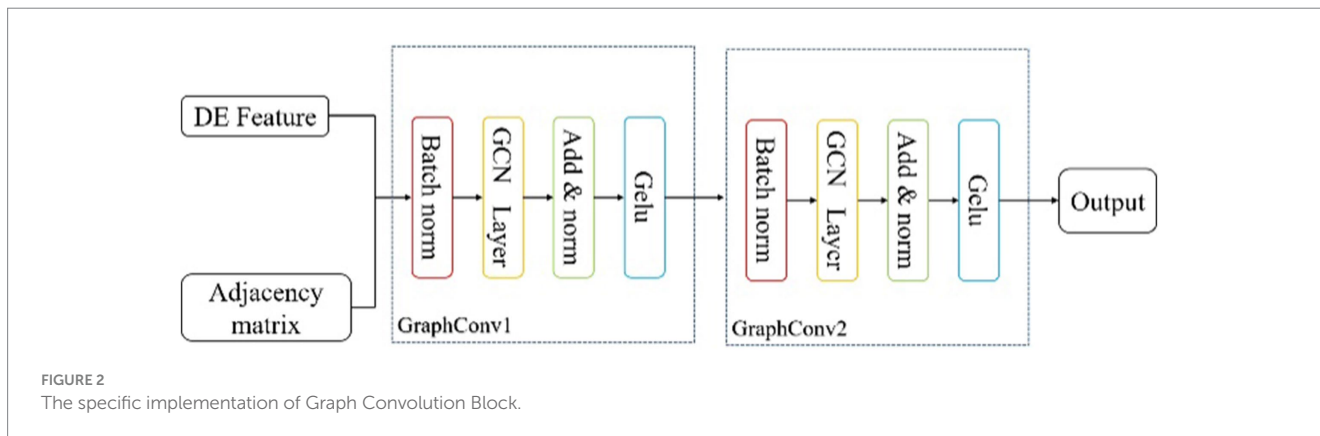
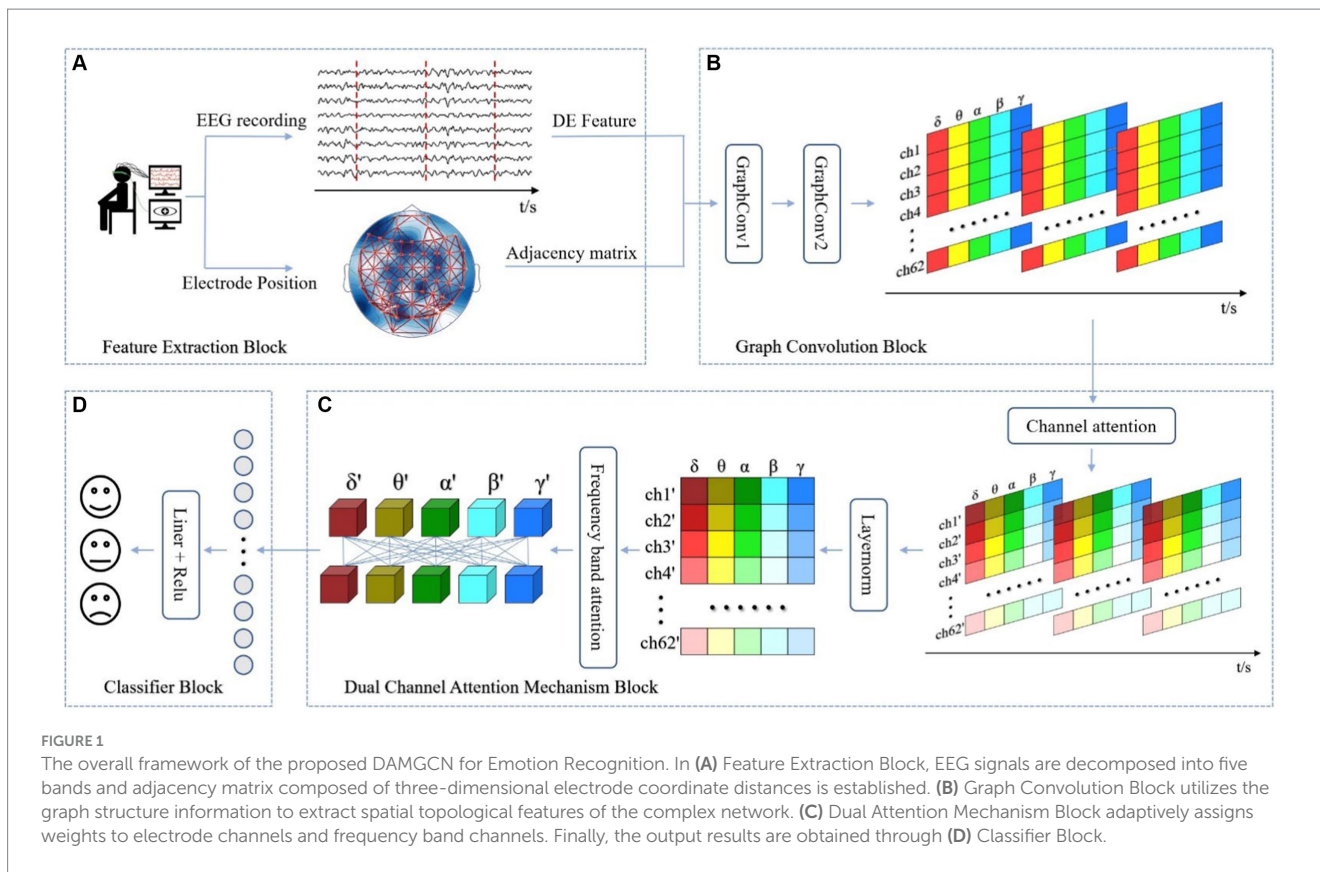
$$H(X) = \frac{1}{2} \log(2\pi\sigma^2) \quad (3)$$

We divide the EEG signals into five frequency bands using the Short-Time Fourier Transform (STFT):  $\delta$  wave (1–4 Hz),  $\theta$  wave (4–8 Hz),  $\alpha$  wave (8–13 Hz),  $\beta$  wave (13–30 Hz), and  $\gamma$  wave (>30 Hz). After that, we utilize the EEG data of the five wave bands as inputs to Eq. (3) to calculate DE.

After EEG signal processing, another feature that we need to extract is the adjacency matrix based on brain network nodes. 3D electrode coordinates are employed to compute the connectivity matrices of electrode channels to construct a three-dimensional spatial adjacency matrix, as shown in Eq. (4):

$$\begin{cases} d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \\ w_{ij} = \frac{1}{d_{ij}}; \text{ if } i \neq j \\ w_{ij} = 1; \text{ if } i = j \end{cases} \quad (4)$$

This matrix describes the structure and connectivity patterns of the brain network, allowing us to analyze functional connections between nodes, study the complexity of the brain network. The



adjacency matrix plays a crucial role in understanding the organization, functionality, and characteristics of the brain network (Gómez-Tapia and Longo, 2022).

### 3.2 Graph convolution block

As shown in the Figure 2, this block consists of two graph convolution layers. The DE features and the adjacency matrix obtained from last block are input in batches into the Graph Convolution Block. In the GraphConv Model, the input is normalized using a batch normalization layer (Ioffe and Szegedy, 2015) to reduce the absolute differences between the data, thereby accelerating convergence speed and improving stability. In a batch of data  $X^B = \{X_1, X_2, X_3, \dots, X_i\}$ ,

$X_i \in R^{C \times F}$ , where  $X^B$  has three dimensions called batch-size(B), channel(C), frequency(F) band, the normalization formula is as follows:

$$\hat{y}_i = \gamma \left( \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} \right) + \beta \tag{5}$$

$x_i$  denotes the dimension of channel,  $\mu$  and  $\sigma$  represent the mean and standard deviation,  $\epsilon$  is a small constant added to the batch variance for numerical stability,  $\gamma$  is the scaling factor,  $\beta$  is the shifting factor, and  $\hat{y}_i$  represents the data after normalization.



The GCN layer convolves and aggregates the feature information of nodes using the adjacency matrix. The calculation formula can be concluded as follows (Kipf and Welling, 2017):

$$H^{(l+1)} = \sigma \left( \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right) \quad (6)$$

here,  $H^{(l)}$  represents the node features at layer  $l$ .  $W^{(l)}$  denotes the trainable weight matrix.  $\tilde{A} = A + I_N$  is the self-connected adjacency matrix of the graph  $G$ , where  $A$  is the original adjacency matrix and  $I_N$

is the identity matrix,  $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$  is the degree of matrix of  $\tilde{A}$ .  $\sigma(\cdot)$  is

the activation function, typically a non-linear function like  $\text{ReLU}(\cdot)$ . We also cited the advantages of residual networks (He et al., 2015), including their facilitation of model training, alleviation of overfitting, and increased network depth. Thus, Eq. (6) can be further expressed as shown in Eq. (7):

$$H_{res}^{(l+1)} = H^{(l+1)} + H^{(l)} \quad (7)$$

After passing through the residual network, the data is activated using the GELU activation function (Hendrycks and Gimpel, 2023), which is defined by Eq. (8):

$$\text{GELU}(x) = x * \sigma(1.702x) = x * \frac{1}{1 + e^{-1.702x}} \quad (8)$$

### 3.3 Dual channel attention mechanism block

The dual attention mechanism block employs the attention mechanism of Transformer to separately allocate attention weights to the EEG channels and the frequency bands of the DE data. In order to fully utilize the graph structure information obtained by the graph convolution block, we first implement the electrode channel mechanism to enhance the emotional relevance of certain electrodes within the channels while suppressing the irrelevant ones. Frequency adaptation mechanism amplifies the impact of relevant frequency band signals on emotions while attenuating the influence of irrelevant frequency band signals.

As shown in the Figure 3, the data after Graph Convolution Block is first mapped to a high-dimensional embedding vector  $Y^B$  ( $Y^B = \{Y_1, Y_2, Y_3, \dots, Y_i\}$ ,  $Y_i \in R^{C \times E}$ , where  $Y^B$  has three dimensions called batch-size, channel, embedding-vector) through Input embedding and then normalized using LayerNorm (Ba et al., 2016). The formula is the same as Eq. (5), while  $x_i$  represents the dimension of embedding vector here. Next, we transform the embedding vector into Query, Keys, and Values vectors using three weight matrices  $W^Q$ ,  $W^K$ , and  $W^V$  to compute the attention, as shown in Eq. (9):

$$\begin{cases} Q = X \cdot W^Q \\ K = X \cdot W^K \\ V = X \cdot W^V \\ \text{Attention} = \text{softmax} \left( \frac{Q \cdot K^T}{\sqrt{d_k}} \right) \end{cases} \quad (9)$$

To enhance the robustness and stability of the model, we employ Multi-Head Attention to obtain multiple sets of Query, Keys, and Values. Each set is used to calculate a Z matrix separately, and the resulting Z matrices can be concatenated together, as shown in Eq. (10):

$$Z = \text{concat} \left\{ \text{softmax} \left( \left( \frac{Q_i \cdot K_i^T}{\sqrt{d_k}} \right) \cdot V_i \right), i = 1, 2, 3, \dots \right\} \quad (10)$$

During the computation of the attention mechanism, we also employ residual connections and LayerNorm to prevent training degradation and other issues. Finally, the output is obtained through the forward propagation network based on residual connections, as shown in Eq. (11):

$$\text{output} = (Z \cdot W^T + b) + Z \quad (11)$$

The frequency band attention mechanism is similar to the channel attention mechanism, and the flowchart is shown in the Figure 3.

### 3.4 Classifier block

During the model training phase, the feature vector obtained from the forward propagation is passed through a fully connected layer to achieve dimensionality reduction. This is followed by generating predicted labels, resulting in the final classification results. The cross-entropy loss function is then employed to calculate the loss between the true emotion labels  $y$  and the predicted emotion  $\hat{y}$  labels, as shown in Eq. (12):

$$\begin{cases} \hat{y} = \text{argmax}(\text{output}) \\ \text{Loss}(\theta) = - \sum_{i=1}^N y_i \log \hat{y}_i \end{cases} \quad (12)$$

where  $\theta$  represents all the parameters in the DAMGCN model. To evaluate the classification results of the DAMGCN model, we use accuracy as performance metrics, as shown in Eq. (13):

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (13)$$

The formula is an example of binary tasks. Total samples are the sum of true positive (TP) predictions, true negative (TN) predictions, false positive (FP) predictions, false negative (FN) predictions, with

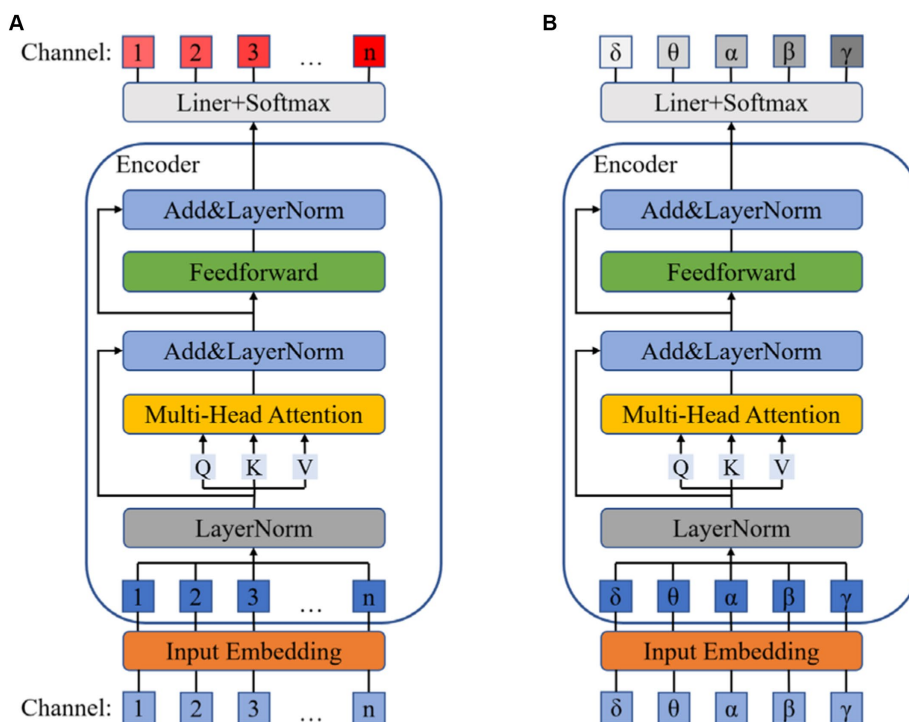


FIGURE 3 The structure of electrode channel adaptation (A) and frequency band adaptation (B).

the sum of TP and TN representing the count of samples predicted correctly.

### 4 Experiment setting and results

In this section, we first introduce three different types of datasets and describe the experimental setup and preprocessing steps. Based on this, we mainly conduct subject-dependent experiments to demonstrate the EEG emotion classification performance of the proposed DAMGCN model and promote it in subject-independent experiments. Subsequently, we visualize the experimental results through subject-dependent experimental data for mechanism analysis and conducted ablation experiments.

#### 4.1 Datasets

The DEAP dataset (Koelstra et al., 2012) collected physiological signals and emotion label data from 32 participants. Each participant watched 40 segments of audio-visual stimuli. EEG signals were captured using a 40-electrode EEG cap distributed according to the 10–20 system. The duration of each trial was 63 s, consisting of a 3-s baseline data at the start followed by 60 s of test data. The data was downsampled to 128 Hz and bandpass frequency filtering was applied in the range of 4.0–45.0 Hz. The labels were provided through questionnaire surveys to assess the emotional evaluation of the stimulus videos. The participants were instructed to provide subjective ratings for the stimulus videos across four dimensions: valence, arousal, dominance, and liking. The

points ranged from 1 to 9 to express self-states, so we compromise by selecting a threshold of 5 to binarize the labels.

The SEED dataset (Zheng and Bao-Liang, 2015) contains EEG signal data from 15 subjects collected using the 62-channel ESI NeuroScan System. The database comprises three sessions, and within each session, participants were instructed to choose 15 segments for emotion elicitation. The data was downsampled to 200 Hz, and a bandpass frequency filter ranging from 0 to 75.0 Hz was applied. The emotion labels include three emotional states: positive, neutral, and negative.

The SEED-IV dataset (Zheng et al., 2019) is an extension of the SEED dataset, with the main difference being the videos viewed by the participants. Each session in SEED-IV consists of 24 trials, and the

TABLE 1 Datasets introduction.

Dataset	DEAP	SEED	SEED-IV
Channels	40	62	62
Sampling rate	128 Hz	200 Hz	200 Hz
Subjects	32	15	15
Sessions	1	3	3
Trials	40	15	72
Trial duration	63 s	120 ~ 240 s	120 ~ 240 s
Frequency range	4.0 ~ 45.0 Hz	0 ~ 75.0 Hz	1.0 ~ 75.0 Hz
Frequency bands	4	5	5
Label types	4	1	1
Classification	2	3	4

emotion classification labels are categorized into four classes: happy, sad, neutral, and fear.

In summary, the similarities and differences among the three datasets are shown in the Table 1. It should be noted that the duration of trials in the SEED and SEED-IV dataset are inconsistent.

## 4.2 Data preprocessing

The approach to data preprocessing and feature extraction in emotion recognition tasks is critical for optimal model performance. Our methodology for processing the DEAP, SEED, and SEED-IV datasets is outlined as follows:

For the DEAP dataset, we selected 32 channels related to EEG and set the non-overlapping duration of each segment to 0.5 s to obtain 120 samples every trial. Since the provided data from the official source has already been filtered using a bandpass filter in the range of 4.0–45.0 Hz, according to the section 3.1 mentioned, we further filtered the raw data into four frequency bands ( $\theta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ) and extracted DE features. Data format for each subject is  $4800 \times 32 \times 4$  (*sample*  $\times$  *channel*  $\times$  *frequency band*).

For the SEED and SEED-IV datasets, EEG data from each channel is segmented into temporal window of 1 s each, with no overlap between them. Unlike the DEAP dataset, SEED and SEED-IV do not filter out signals in the  $\delta$  (1–4 Hz) frequency band. We have summarized all the trials of one subject because of different sample sizes for each trial. In each session, the data format for a single subject in the SEED dataset is  $3394 \times 62 \times 5$  (*sample*  $\times$  *channel*  $\times$  *frequency band*). However, what sets it apart from the SEED dataset is that trials for each session in the SEED-IV dataset are inconsistent, resulting in 851, 832, 822 samples in 3 sessions. The data format for a single subject in the SEED-IV dataset is  $851 / 832 / 822 \times 62 \times 5$  (*sample*  $\times$  *channel*  $\times$  *frequency band*).

## 4.3 Evaluation strategy

In this article, the strategy for model training includes both subject-dependent and subject-independent experiments.

In the subject-dependent experiment, we use ten-fold cross validation and leave-one-trial-out strategy to analyze each subject. For ten-fold cross validation strategy, the entire dataset is randomly divided into 10 equally sized subsets, each of which strives to maintain the overall distribution of the data. The model is trained using merged data from 9 subsets, and then evaluated on the reserved test set to obtain performance metrics such as accuracy. This process is repeated ten times to ensure that each subset has a chance to be used as a test set, resulting in 10 independent training and validation processes. Leave-one-trial-out strategy aims to evaluate the model's generalization ability to new experiments. If there are N experiments of one subject, in each validation process, select one experiment as the test set and the remaining N-1 experiments as the training set. This process will be repeated N times, each time selecting a different experiment as the test set to ensure that each experiment has the opportunity to be used as data to validate the performance of the model.

In the subject-independent experiment, leave-one-subject-out cross validation strategy was adopted. This strategy is used to evaluate

the model's generalization ability to new individual data. Assuming there are N subjects, data from N-1 subjects is selected as the training set for each experiment, leaving one subject as the testing set until each subject's data is tested once.

In terms of data label selection, we use the valence, arousal, dominance labels of the DEAP dataset, and the emotional state labels of the SEED and SEED-IV datasets.

## 4.4 Model training details

In the development of our DAMGCN model, it is necessary to quantify these model parameters in Dual Channel Attention Mechanism Block: Encoder, embedding vector, Multi-Head Attention. EEG data may contain more direct emotional signals compared to natural language processing tasks. The number of Encoders can start with fewer layers to avoid overfitting and maintain computational efficiency. The size of the embedding vector is determined based on the size and complexity of the dataset. In EEG emotion recognition tasks, it is possible to consider setting it between 32 and 64 because of the number electrode channels. The number of Multi-Head Attention can start with 4, which means it is possible to simultaneously focus on multiple aspects of the signal. For the Classifier Block, we use GELU activation function and two linear layers to gradually map high-dimensional features to the output dimension of emotion categories to increase the non-linear ability of the model. However, excessively large parameter settings may not result in significant performance improvements but could instead increase computational complexity. Based on experimental results and computational resources, we have made appropriate adjustments, with the parameter values shown in Table 2. Other parameters are adjusted through the first experiment of ten-fold cross validation. The number of epochs was set by early stopping strategy. After 200 epochs, the classification performance of the model did not show significant improvement and the training would be stopped. Batch size is conventionally established as a power of 2. Through our rigorous experimentation, it has been determined that a batch size of 64 facilitates more stable gradient descent. After experimenting with dropout rates between 0.1 and 0.5, we selected 0.5 for optimal performance. We tested multiple learning rates from  $1e-6$  to  $1e-1$  and found that when the learning rate was 0.001, the model was able to perform better. The loss function and optimizer we used are Cross entropy and Adam. Our experimental platform relies on the hardware condition of NVIDIA GeForce RTX 3080 Ti and deep

TABLE 2 Parameters settings of DAMGCN.

Hyper-parameter of DAMGCN	Value or type
Number of Encoders	2
Size of the embedding vector	64
Number of multi-head attention	6
Linear layers in classifier block	2
Max number of epochs	200
Batch size	64
Dropout	0.5
Learning rate	0.001

learning framework used was PyTorch 1.11.0. The parameter settings are shown in the Table 2.

## 4.5 Results and comparison

As shown in the Figure 4, we obtained ten-fold cross validation experiment's average accuracies of 96.96, 97.17, 97.50% for the two-class dimension labels of valence, arousal, dominance in DEAP dataset. For the three-class labels in SEED dataset, we achieved an average accuracy of 99.42% in three sessions. And in the case of the four-class labels in the SEED-IV dataset, the average accuracy obtained was 96.86%. The data indicates that our model has achieved an accuracy of over 96% on various types of datasets, demonstrating its wide applicability across different datasets.

Based on the results of ten-fold cross validation experiments, we selected 5% DE features of all subjects and plotted t-SNE in Figure 5. It can be observed that after DAMGCN training, the sample distribution becomes more distinct, and the level of disorder decreases. To analyze the performance of each dataset in different emotion categories more comprehensively, we have calculated confusion matrices using the proposed DAMGCN model in Figure 6. In the confusion matrix, the row sum represents the total number of samples, the diagonal elements represent the percentage of correctly classified samples for each emotion, and the remaining elements indicate the percentage of misclassified samples. Our findings reveal that the accuracy of classifying positive emotions consistently exceeds that of negative emotions. This suggests that the proposed method exhibits higher discriminative capability for positive emotions, which aligns with similar observations in other related works (Koelstra et al., 2012; Li et al., 2021; Guo et al., 2022).

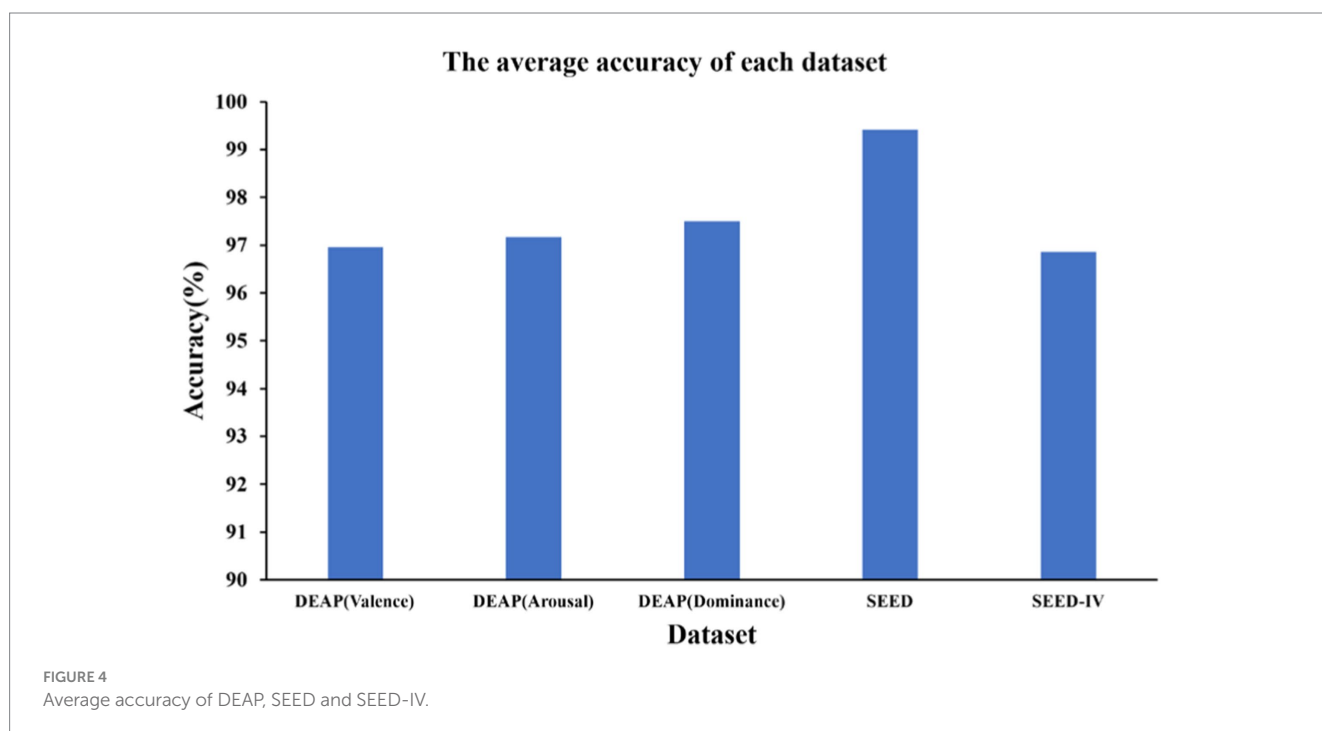
To evaluate the performance of our method, we conducted comparative studies with relevant literature that employed the same

experimental methods and datasets in Tables 3–5, which including traditional machine learning models as well as some state-of-the-art neural network models. Session-average represents the average accuracy across three sessions on SEED and SEED-IV datasets. Based on the comparison, it has been demonstrated that our proposed DAMGCN model outperforms existing algorithms in terms of accuracy and stability on the DEAP, SEED, and SEED-IV datasets. In addition, we extended our investigation through leave-one-trial-out strategy and subject-independent experiment, the outcomes delineated in Tables 6, 7 reveal that DAMGCN continues to exhibit comparative superiority in relation to existing methods. Comparing our results with existing graph neural network methods (Kipf and Welling, 2017), it can be concluded that DAMGCN designed with EEG signal characteristics performs better in emotion recognition tasks. Furthermore, Transformer's attention mechanism can selectively focus on electrical signals in certain regions or frequency bands of EEG data that are more relevant to emotion recognition tasks, dynamically assigning weights to different features, thereby enhancing the influence of informative features while reducing less useful ones.

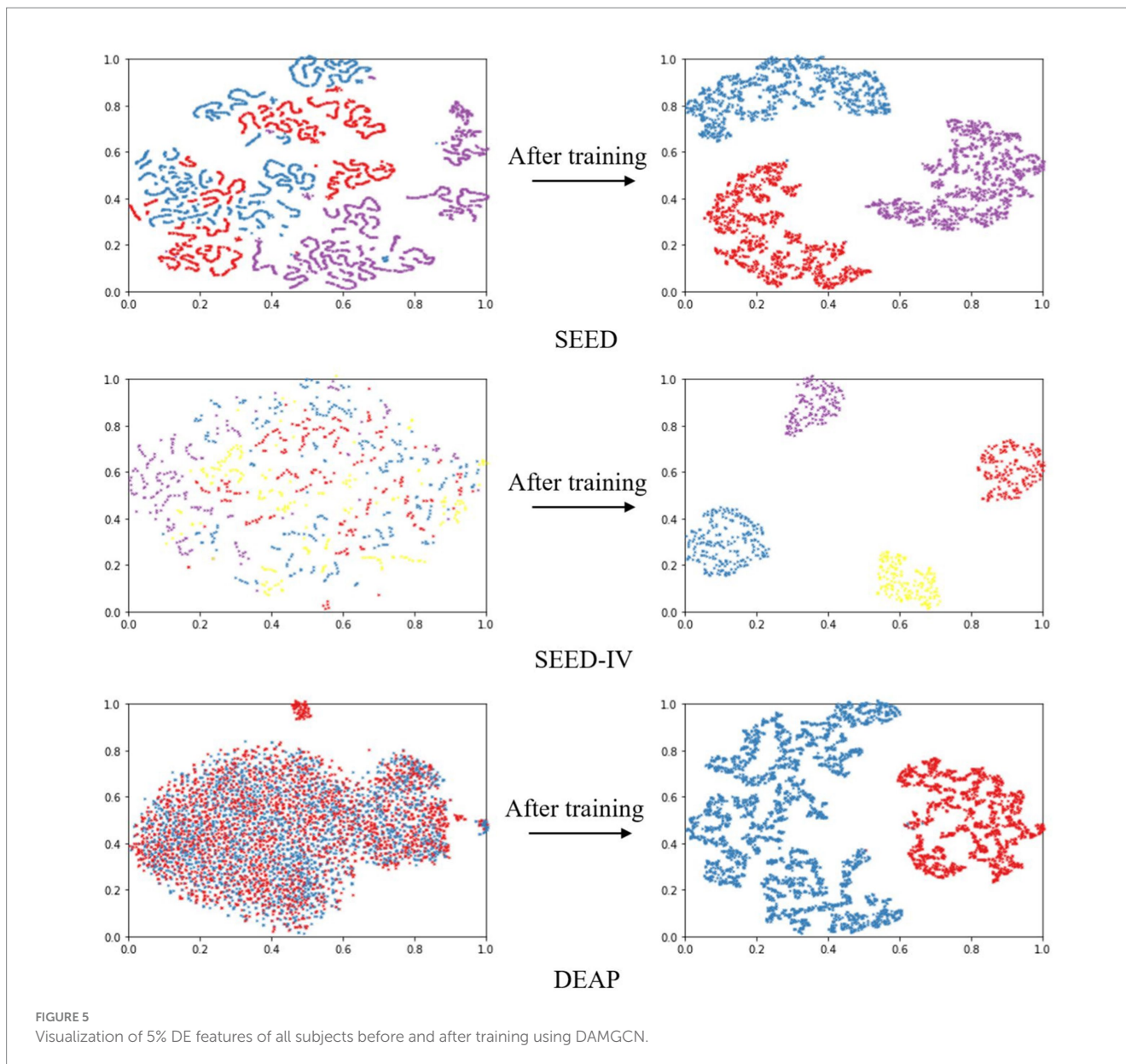
## 4.6 Interpretable analysis

Benefiting from the combination of the dual attention mechanism and GCN, our model has an advantage in interpretability. In frequency band analysis, the DEAP dataset is not included due to the absence of data in the  $\alpha$  wave band (1–4Hz). Figure 7 shows the parameter visualization after training convergence on the SEED and SEED-IV datasets, obtaining the contribution of different frequency bands to emotion recognition tasks.

The initial weight coefficients for each band before training are 0.2. After training, the weight  $\delta$  coefficient of the band is the lowest and consistently below 0.2. This is similar to the conclusion from







previous literature (Duan et al., 2013; Zhang et al., 2020), which indicate that the  $\delta$  band is associated with unconscious states and often appears during deep dreamless sleep, while emotional responses typically occur during wakefulness, especially in  $\gamma$  frequency band that is more prominent. We speculate that the weight coefficients of  $\delta$  should gradually approach 0 as the epochs increase. However, in Figure 7, the coefficient remains around 0.15, suggesting that during the optimization process, the model parameters might have stagnated at a local minimum or maximum, causing the model to become trapped in a local optimum. We attempted to increase the learning rate to mitigate this situation in 4.3 section. Unfortunately, a large learning rate destabilized the optimization process, causing the loss function of the model to gradually increase instead of decreasing, preventing the model from converging to a suitable solution and ultimately resulting in ineffective training results. The issue of local optima is inevitable in deep learning. As a result, it's necessary to analyze from the trend of parameter changes rather than

the results. This approach can provide us with directions for exploration in unknown domains and serve as a reliable way to validate conclusions drawn by previous researchers in clinical settings.

On the other hand, we extracted the attention matrix of the electrode channels and used degree centrality to evaluate the importance of nodes. The formula is as follows in Eq. (14):

$$DC_i = \frac{k_i}{N-1} \quad (14)$$

$N$  represents the number of nodes,  $k_i = \sum_{j=1}^n d_{ij}$  represents the sum

of the weights connected to current node and all other nodes. By calculating the mean of the degree centrality weights for all participants, we generated a distribution map of node importance in the brain regions involved in emotional activity.

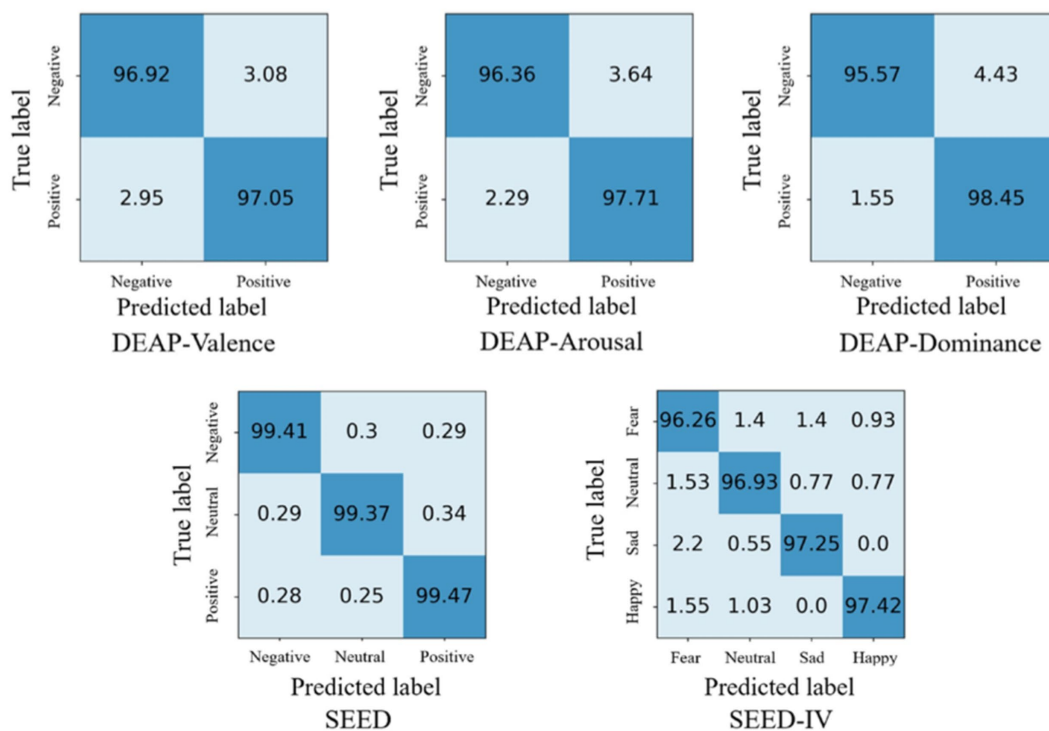


FIGURE 6 Confusion matrixes of the results on DEAP, SEED, SEED-IV.

TABLE 3 The average accuracy /standard deviation (%) of different methods on DEAP two-class.

Method	Feature	Valence (acc/std)	Arousal (acc/std)	Dominance (acc/std)
SVM (Kumar and Nataraj, 2019)	PSD/DE	74/–	86/–	72/–
CNN-BiLSTM (Cui et al., 2022)	DE	94/–	–/–	–/–
CNN (Li et al., 2019)	DE	89.43/4.72	90.40/4.05	–/–
LSTM-RNN (Xing et al., 2019)	PSD	81.1/–	74.38/–	–/–
GCN (Kipf and Welling, 2017)	DE	82.56/2.11	80.47/3.84	85.23/2.32
MRGCN (Qiu et al., 2023)	DE	94.97/3.8	95.72/2.5	–/–
GLFANet (Liu et al., 2023)	DE	94.53/1.02	94.91/1.05	95.35/0.90
AP-CapsNet (Liu et al., 2023)	MobileNet	93.89/–	95.04/–	95.08/–
TR&CA (Peng et al., 2023)	Raw signal	95.18/2.46	95.58/2.28	95.78/2.16
ST-CLSM (Feng et al., 2022)	Raw signal	95.52/–	95.04/–	–/–
EESCN (Xu et al., 2024)	DE	94.56/4.18	94.81/3.62	94.73/4.12
DAMGCN (ours)	DE	96.96/2.24	97.17/2.35	97.50/2.03

The results in Figure 8 reveal that frontal lobe, temporal lobe, and occipital lobe regions exhibit higher node weight coefficients, indicating heightened emotional activity. Our finding aligns partially with the observations reported in reference (Liao et al., 2024; Nie et al., 2011; Li et al., 2024). In addition, we found that there is a greater difference in brain regions between the SEED and SEED-IV datasets through the comparison of the ab and c graphs. Based on the significant differences in the data in Tables 6, 7, it can be analyzed that as the number of electrode channels increases and the graph structure information becomes richer, the model can learn more common features, especially in subject-independent experiments. This can not only highlight the importance of

graph convolutional neural networks but also offer insights for neuroscientists to assess the reliability of emotion recognition results based on brain activity regions.

### 4.7 Ablation

We conducted subject-dependent ablation study by removing the GCN block and the DAM block to examine the importance of each block in our proposed model. The comparison of average

TABLE 4 The average accuracy /standard deviation (%) of different methods on seed three-class.

Method	Feature	Session-average (acc/std)
CNN-BiLSTM (Cui et al., 2022)	DE	94.82/–
GCN (Kipf and Welling, 2017)	DE	92.54/1.89
MRGCN (Qiu et al., 2023)	DE	98.98/1.5
DPGAT (Li et al., 2021)	Raw signal	95.76/5.77
GLEFNet (Liu et al., 2023)	DE	93.19/1.54
Bi-ViTNet (Lu et al., 2023)	PSD/DE	97.55/1.58
ACTNN (Gong et al., 2023)	DE	98.47/1.72
ST-CLSM (Feng et al., 2022)	Raw signal	96.20/–
DAMGCN (ours)	DE	99.42/0.24

TABLE 5 The average accuracy /standard deviation (%) of different methods on SEED-IV four-class.

Method	Feature	Session-average (acc/std)
SVM (Kumar and Nataraj, 2019)	PSD/DE	77.33/–
GCN (Kipf and Welling, 2017)	DE	82.86/3.64
Bi-ViTNet (Lu et al., 2023)	PSD/DE	88.08/6.32
ACTNN (Gong et al., 2023)	DE	91.90/5.43
ST-CLSM (Feng et al., 2022)	Raw signal	93.86/–
EESCN (Xu et al., 2024)	DE	79.65/8.22
DAMGCN (ours)	DE	96.86/1.33

TABLE 6 The average accuracy /standard deviation (%) of different methods (leave-one-trial-out strategy).

Method	Feature	DEAP (acc/std)	SEED (acc/std)	SEED-IV (acc/std)
SVM (Kumar and Nataraj, 2019)	PSD/DE	66.25/9.04	–/–	–/–
DBN-CRF (Chao and Liu, 2020)	Power features	76.13/–	83.46/–	–/–
GAT (Veličković et al., 2018)	DE	72.65/6.88	80.15/5.56	78.55/6.32
GCN (Kipf and Welling, 2017)	DE	75.86/4.35	83.56/2.98	81.05/4.73
DAMGCN (ours)	DE	78.21/5.64	90.4/2.61	83.55/3.88

accuracies of the DAMGCN, DAM, and GCN models on the DEAP, SEED, and SEED-IV datasets is shown in the Figure 9. When only the GCN block was involved in emotion recognition, the accuracy was 96.13, 95.1, 96.2, 98.98, 92.2%, resulting in a decrease of 0.73, 2.07, 1.3, 0.44, 4.66% in accuracy, respectively. This indicates that there might be channels unrelated to emotions among all electrode channels, and the GCN module was not effective in distinguishing them. DAM is the block proposed in this paper for effectively allocating channel weights. Without the

TABLE 7 The average accuracy /standard deviation (%) of different methods (subject-independent).

Method	Feature	DEAP (acc/std)	SEED (acc/std)	SEED-IV (acc/std)
SVM (Kumar and Nataraj, 2019)	PSD/DE	59.6/–	56.73/16.29	51.78/12.85
VDM-DNN (Pandey and Seeja, 2022)	VMD	61.88/–	–/–	–/–
GCN (Kipf and Welling, 2017)	DE	58.4/13.56	62.4/10.14	59.13/14.55
LGG (Ding et al., 2022)	Raw signal	63.07/–	–/–	–/–
AP-CapsNet (Liu et al., 2023)	MobileNet	63.41/–	–/–	–/–
SSL-EEG (Xie et al., 2021)	Raw signal	–/–	67.52/12.73	53.62/8.47
MetaEmotionNet (Ning et al., 2024)	DE		77.5/0.088	61.2/0.083
A-LSTM (Li et al., 2022)	Raw signal	–/–	72.18/10.85	69.5/15.65
DAMGCN (ours)	DE	64.61/10.54	73.21/8.35	68.22/7.03

GCN module, the accuracy on the DEAP dataset decreased by 1.46, 1.97, and 1% in the three different labels, while the accuracy on the SEED and SEED-IV datasets decreased by 0.92 and 3.46%. In addition, we removed the DAM and GCN modules, and compared the results obtained from only the linear layer model in Figure 9 (labeled as None), it can be seen that the classification accuracy the accuracy was 70.62, 72.13, 74.26, 80.46, 72.54%. We can analyze the significant classification performance of DAM and GCN modules on EEG signals through data ablation experiments. We believe that the reason why the method proposed in this article performs well on these datasets can be attributed to the adaptability of the proposed model to EEG signals: EEG signals are electrical signals collected from multiple positions on the surface of the scalp, with a fixed spatial structure. The connections between different brain regions form a dynamic network, and different brain regions interact with each other through neural networks to influence emotional states. GCN naturally matches the spatial structure and dynamic network characteristics of EEG data through its ability to process graph structured data, enabling it to extract more complex and in-depth features from EEG signals. Transformer's attention mechanism can selectively focus on electrical signals in certain regions or frequency bands of EEG data that are more relevant to emotion recognition tasks, dynamically assigning weights to different features, thereby enhancing the influence of informative features while reducing less useful ones.

In conclusion, our research results demonstrate the complementary roles played by the information aggregation of the Graph Convolutional Neural Network and the weight allocation of the Dual Attention Mechanism in extracting significant information from brain networks and ensuring stability in channel selection for classification.

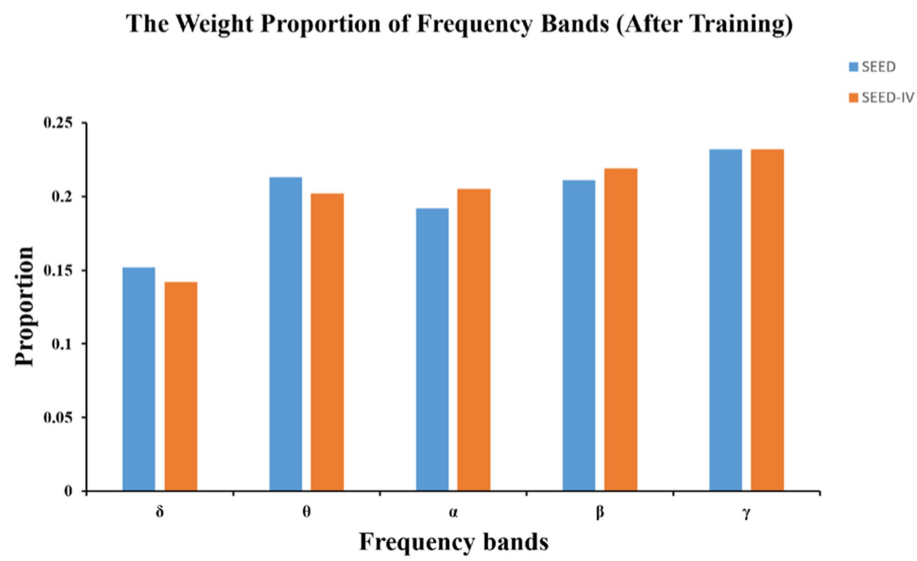


FIGURE 7  
The weight proportion of frequency before and after training using DAMGCN.

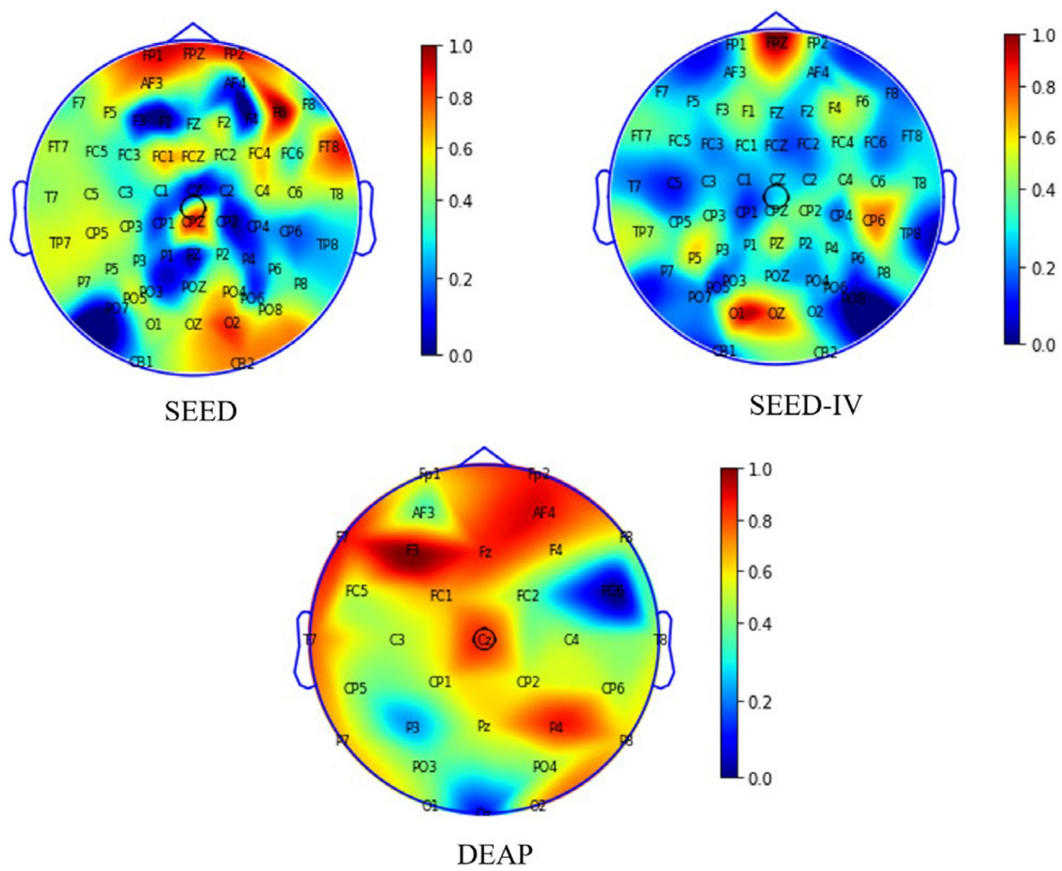
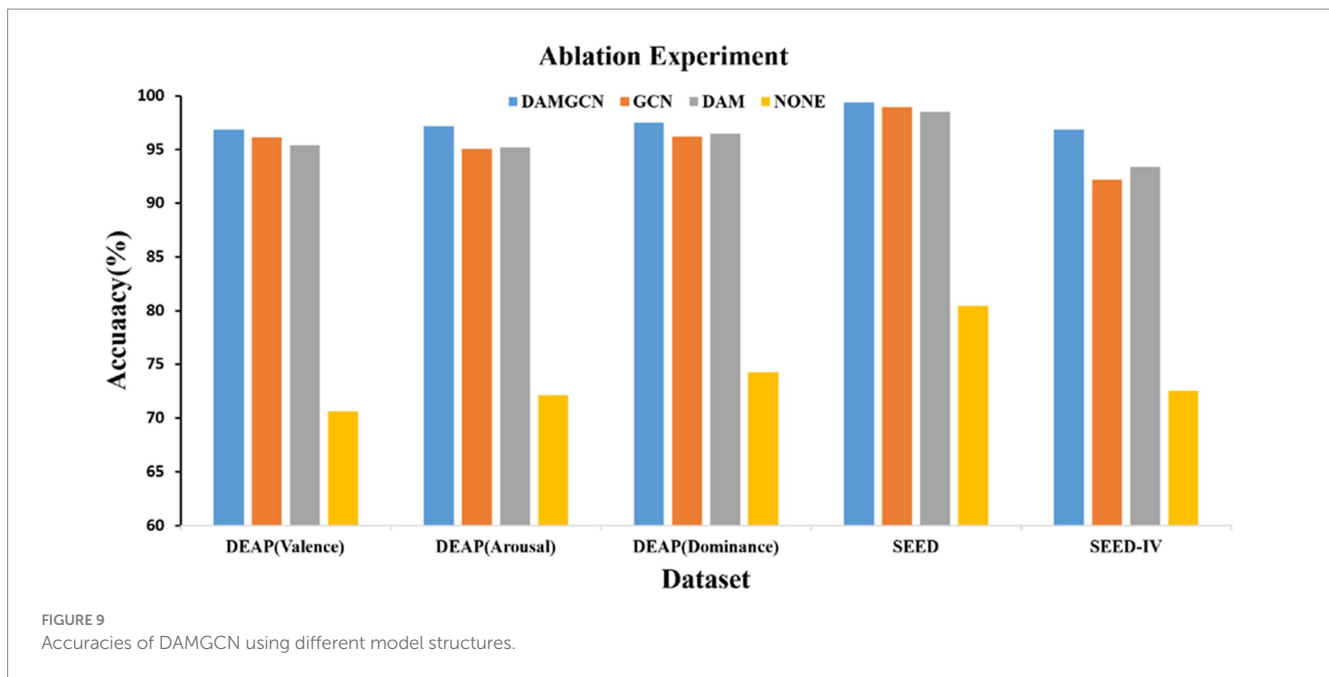


FIGURE 8  
Visualization of brain regions weight distribution map on SEED, SEED-IV and DEAP.





## 5 Conclusion and future work

In this manuscript, we present the DAMGCN emotion recognition model, which combines the synergistic power of GCN and Transformer. The proposed model leverages the inherent connections between brain channels and utilizes the graph structure information to extract spatial topological features of the complex neural network. Additionally, it assigns weight coefficients to individual information to enable effective emotion classification. We conducted an extensive array of experiments on the DEAP, SEED, and SEED-IV datasets, and the results indicate that the model is competitive compared to state-of-the-art methods. Additionally, through ablative experiments, we corroborated the substantial contributions made by both the GCN block and the DAM block of our model in augmenting the classification performance. We also employed attention mechanism to visualize the significance of each EEG channel and different frequency bands in emotion recognition. Through this analysis, we observed that the weight coefficients associated with the  $\delta$  frequency band were relatively low across most participants, suggesting a weak correlation between this particular EEG band and human emotions. Finally, our observations indicate a strong association between emotional activity and specific brain regions, notably the prefrontal and occipital lobes. This method we proposed offers a valuable framework for subsequent research endeavors in the field of emotion recognition.

In terms of models, our model requires more time to learn its parameters during the training phase, yet it remains susceptible to the challenge of getting trapped in local optima, a prevalent issue in many deep learning studies. Fortunately, we can mitigate this concern by focusing on the physical implications of parameter variations rather than solely relying on outcomes. In our subject-dependent and subject-independent experiments, the results of the subject-independent experiments were not very impressive. Therefore, our forthcoming study will attempt to introduce contrastive learning and

transfer learning methods to improve the model, so that the model can learn common features between subjects to achieve high classification results in subject-independent experiments.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/ participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

## Author contributions

WC: Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. YL: Data curation, Formal analysis, Software, Supervision, Validation, Visualization, Writing – review & editing. RD: Data curation, Formal analysis, Funding acquisition, Supervision, Validation, Writing – review & editing. YD: Investigation, Project administration, Resources, Software, Writing – review & editing. LH: Funding

acquisition, Project administration, Supervision, Validation, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (Grant No. 61977039), “New Infrastructure Development & University Informatization” research project of China Association for Educational Technology (Grant No. XJJ1202205007) and Open subject of cognitive EEG and transcranial, electrical stimulation regulation of neuracle (Grant No. BRKOT-NJUPT-20220630H).

## References

- Alam, R., Zhao, H., Goodwin, A., Kavehei, O., and McEwan, A. (2020). Differences in power spectral densities and phase quantities due to processing of EEG signals. *Sensors* 20:6285. doi: 10.3390/s20216285
- Ba, J. L., Kiros, J. R., and Hinton, G. E. (2016). Layer normalization. *arXiv*. Available at: <http://arxiv.org/abs/1607.06450>
- Cai, J., Xiao, R., Cui, W., Zhang, S., and Liu, G. (2021). Application of electroencephalography-based machine learning in emotion recognition: a review. *Front. Syst. Neurosci.* 15:729707. doi: 10.3389/fnsys.2021.729707
- Chao, H., and Liu, Y. (2020). Emotion recognition from Multi-Channel EEG signals by exploiting the deep belief-conditional random field framework. *IEEE Access* 8, 33002–33012. doi: 10.1109/ACCESS.2020.2974009
- Chen, D., Wan, S., Xiang, J., and Bao, F. S. (2017). A high-performance seizure detection algorithm based on discrete wavelet transform (DWT) and EEG. *PLoS One* 12:e0173138. doi: 10.1371/journal.pone.0173138
- Cui, F., Wang, R., Ding, W., Chen, Y., and Huang, L. (2022). A novel DE-CNN-BiLSTM multi-fusion model for EEG emotion recognition. *Mathematics* 10:582. doi: 10.3390/math10040582
- Ding, Y., Robinson, N., Tong, C., Zeng, Q., and Guan, C. (2022). LGGNet: learning from local-global-graph representations for brain-computer Interface. *arXiv*. Available at: <http://arxiv.org/abs/2105.02786>
- Duan, R.-N., Zhu, J.-Y., and Lu, B.-L., “Differential entropy feature for EEG-based emotion classification,” In 2013 6th international IEEE/EMBS conference on neural engineering (NER), San Diego, CA, USA: IEEE, (2013), pp. 81–84.
- Feng, L., Cheng, C., Zhao, M., Deng, H., and Zhang, Y. (2022). EEG-based emotion recognition using spatial-temporal graph convolutional LSTM with attention mechanism. *IEEE J. Biomed. Health Inform.* 26, 5406–5417. doi: 10.1109/JBHI.2022.3198688
- Feutrill, A., and Roughan, M. (2021). A review of Shannon and differential entropy rate estimation. *Entropy* 23:1046. doi: 10.3390/e23081046
- García-Martínez, B., Fernández-Caballero, A., Zunino, L., and Martínez-Rodrigo, A. (2021). Recognition of emotional states from EEG signals with nonlinear regularity- and predictability-based entropy metrics. *Cogn. Comput.* 13, 403–417. doi: 10.1007/s12559-020-09789-3
- Gómez-Tapia, B. B., and Longo, L. (2022). On the minimal amount of EEG data required for learning distinctive human features for task-dependent biometric applications. *Front. Neuroinform.* 16:844667. doi: 10.3389/fninf.2022.844667
- Gong, P., Jia, Z., Wang, P., Zhou, Y., and Zhang, D., “ASTDF-net: attention-based spatial-temporal dual-stream fusion network for EEG-based emotion recognition,” In Proceedings of the 31st ACM international conference on multimedia, Ottawa ON Canada: ACM, (2023), pp. 883–892.
- Gong, L., Li, M., Zhang, T., and Chen, W. (2023). EEG emotion recognition using attention-based convolutional transformer neural network. *Biomed. Signal Process. Control* 84:104835. doi: 10.1016/j.bspc.2023.104835
- Gu, Y., Zhong, X., Qu, C., Liu, C., and Chen, B. (2023). A domain generative graph network for EEG-based emotion recognition. *IEEE J. Biomed. Health Inform.* 27, 2377–2386. doi: 10.1109/JBHI.2023.3242090
- Guo, J.-Y., Cai, Q., An, J. P., Chen, P. Y., Ma, C., Wan, J. H., et al. (2022). A transformer based neural network for emotion recognition and visualizations of crucial EEG channels. *Physica: Stat. Mech. Appl.* 603:127700. doi: 10.1016/j.physa.2022.127700
- He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., and Wang, M., “LightGCN: simplifying and powering graph convolution network for recommendation,” In Proceedings of the

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

43rd international ACM SIGIR conference on Research and Development in information retrieval, China: ACM, (2020), pp. 639–648.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. *arXiv*. Available at: <http://arxiv.org/abs/1512.03385>

Hendrycks, D., and Gimpel, K. (2023). Gaussian error linear units (GELUs). *arXiv*. Available at: <http://arxiv.org/abs/1606.08415>

Hu, X., Chen, J., Wang, F., and Zhang, D. (2019). Ten challenges for EEG-based affective computing. *Brain Sci. Adv.* 5, 1–20. doi: 10.1177/2096595819896200

Ioffe, S., and Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv*. Available at: <http://arxiv.org/abs/1502.03167>

Jenke, R., Peer, A., and Buss, M. (2014). Feature extraction and selection for emotion recognition from EEG. *IEEE Trans. Affect. Comput.* 5, 327–339. doi: 10.1109/TAFFC.2014.2339834

Jia, Z., Lin, Y., Wang, J., Ning, X., He, Y., Zhou, R., et al. (2021). Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* 29, 1977–1986. doi: 10.1109/TNSRE.2021.3110665

Kipf, T. N., and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *arXiv*. Available at: <http://arxiv.org/abs/1609.02907>

Koelstra, S., Muhl, C., Soleymani, M., Jong-Seok Lee, , Yazdani, A., Ebrahimi, T., et al. (2012). DEAP: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15

Kumar, D. K., and Nataraj, J. L. (2019). Analysis of EEG based emotion detection of DEAP and SEED-IV databases using SVM. *SSRN J.* doi: 10.2139/ssrn.3509130

Lauriola, A. L., and Aioli, F. (2022). An introduction to deep learning in natural language processing: models, techniques, and tools. *Neurocomputing* 470, 443–456. doi: 10.1016/j.neucom.2021.05.103

Li, Y., Chen, J., Li, F., Fu, B., Wu, H., Ji, Y., et al. (2022). GMSS: graph-based multi-task self-supervised learning for EEG emotion recognition. *arXiv*. Available at: <http://arxiv.org/abs/2205.01030>

Li, X., Li, J., Zhang, Y., and Tiwari, P., “Emotion recognition from multi-channel EEG data through a dual-pipeline graph attention network,” In 2021 IEEE international conference on bioinformatics and biomedicine (BIBM), Houston, TX, USA: IEEE, (2021), pp. 3642–3647.

Li, Y., Wong, C. M., Zheng, Y., Wan, F., Mak, P. U., Pun, S. H., et al., “EEG-based emotion recognition under convolutional neural network with differential entropy feature maps,” In 2019 IEEE international conference on computational intelligence and virtual environments for measurement systems and applications (CIVEMSA), Tianjin, China: IEEE, (2019), pp. 1–5.

Li, D., Xie, L., Wang, Z., and Yang, H. (2024). Brain emotion perception inspired EEG emotion recognition with deep reinforcement learning. *IEEE Trans. Neural Netw. Learn. Syst.*, 1–14. doi: 10.1109/TNNLS.2023.3265730

Li, T., Zhao, Z., Sun, C., Yan, R., and Chen, X. (2021). Multireceptive field graph convolutional networks for machine fault diagnosis. *IEEE Trans. Ind. Electron.* 68, 12739–12749. doi: 10.1109/TIE.2020.3040669

Li, Y., Zheng, W., Zong, Y., Cui, Z., Zhang, T., and Zhou, X. (2021). A bi-hemisphere domain adversarial neural network model for EEG emotion recognition. *IEEE Trans. Affect. Comput.* 12, 494–504. doi: 10.1109/TAFFC.2018.2885474

Liao, Y., Zhang, Y., Wang, S., Zhang, X., Zhang, Y., Chen, W., et al. (2024). CLDTA: Contrastive Learning based on Diagonal Transformer Autoencoder for Cross-Dataset EEG Emotion Recognition. Available at: doi: 10.48550/arXiv.2406.08081

- Liu, C., Feng, H., Taylor, P. A., Kang, M., Shen, J., Saini, J., et al. (2024). Graph neural networks in EEG-based emotion recognition: a survey. *arXiv*. doi: 10.48550/arXiv.2402.01138
- Liu, Y., Jia, Z., and Wang, H., “EmotionKD: a cross-modal knowledge distillation framework for emotion recognition based on physiological signals,” In Proceedings of the 31st ACM international conference on multimedia, Ottawa ON Canada: ACM, (2023), pp. 6122–6131.
- Liu, S., Wang, Z., An, Y., Zhao, J., Zhao, Y., and Zhang, Y.-D. (2023). EEG emotion recognition based on the attention mechanism and pre-trained convolution capsule network. *Knowl.-Based Syst.* 265:110372. doi: 10.1016/j.knsys.2023.110372
- Liu, S., Zhao, Y., An, Y., Zhao, J., Wang, S.-H., and Yan, J. (2023). GLFANet: a global to local feature aggregation network for EEG emotion recognition. *Biomed. Signal Process. Control* 85:104799. doi: 10.1016/j.bspc.2023.104799
- Lu, W., Tan, T.-P., and Ma, H. (2023). Bi-branch vision transformer network for EEG emotion recognition. *IEEE Access* 11, 36233–36243. doi: 10.1109/ACCESS.2023.3266117
- Nicolaou, N., and Georgiou, J. (2012). Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. *Expert Syst. Appl.* 39, 202–209. doi: 10.1016/j.eswa.2011.07.008
- Nie, D., Wang, X.-W., Shi, L.-C., and Lu, B.-L., “EEG-based emotion recognition during watching movies,” In 2011 5th international IEEE/EMBS conference on neural engineering, Cancun: IEEE, (2011), pp. 667–670.
- Ning, X., Wang, J., Lin, Y., Cai, X., Chen, H., Gou, H., et al. (2024). MetaEmotionNet: spatial-spectral-temporal-based attention 3-D dense network with Meta-learning for EEG emotion recognition. *IEEE Trans. Instrum. Meas.* 73, 1–13. doi: 10.1109/TIM.2023.3338676
- Pandey, P., and Seeja, K. R. (2022). Subject independent emotion recognition from EEG using VMD and deep learning. *J. King Saud Univ. Comput. Inf. Sci.* 34, 1730–1738. doi: 10.1016/j.jksuci.2019.11.003
- Peng, G., Zhao, K., Zhang, H., Xu, D., and Kong, X. (2023). Temporal relative transformer encoding cooperating with channel attention for EEG emotion analysis. *Comput. Biol. Med.* 154:106537. doi: 10.1016/j.compbiomed.2023.106537
- Qiu, X., Wang, S., Wang, R., Zhang, Y., and Huang, L. (2023). A multi-head residual connection GCN for EEG emotion recognition. *Comput. Biol. Med.* 163:107126. doi: 10.1016/j.compbiomed.2023.107126
- Rahman, M. M., Sarkar, A. K., Hossain, M. A., Hossain, M. S., Islam, M. R., Hossain, M. B., et al. (2021). Recognition of human emotions using EEG signals: a review. *Comput. Biol. Med.* 136:104696. doi: 10.1016/j.compbiomed.2021.104696
- She, Q., Zhang, C., Fang, F., Ma, Y., and Zhang, Y. (2023). Multisource associate domain adaptation for cross-subject and cross-session EEG emotion recognition. *IEEE Trans. Instrum. Meas.* 72, 1–12. doi: 10.1109/TIM.2023.3277985
- Song, Y., Jia, X., Yang, L., and Xie, L. (2021). Transformer-based spatial-temporal feature learning for EEG decoding. *arXiv*. Available at: <http://arxiv.org/abs/2106.11170>
- Song, T., Zheng, W., Song, P., and Cui, Z. (2020). EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans. Affect. Comput.* 11, 532–541. doi: 10.1109/TAFFC.2018.2817622
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2023). Attention is all you need. *arXiv*. Available at: <http://arxiv.org/abs/1706.03762>
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y. (2018). Graph attention networks. *arXiv*. Available at: <http://arxiv.org/abs/1710.10903>
- Wang, J.-G., Shao, H.-M., Yao, Y., Liu, J.-L., Sun, H.-P., and Ma, S.-W. (2022). Electroencephalograph-based emotion recognition using convolutional neural network without manual feature extraction. *Appl. Soft Comput.* 128:109534. doi: 10.1016/j.asoc.2022.109534
- Wang, Z., Wang, Y., Hu, C., Yin, Z., and Song, Y. (2022). Transformers for EEG-based emotion recognition: a hierarchical spatial information learning model. *IEEE Sensors J.* 22, 4359–4368. doi: 10.1109/JSEN.2022.3144317
- Xie, Z., Zhou, M., and Sun, H., “A novel solution for EEG-based emotion recognition,” In 2021 IEEE 21st international conference on communication technology (ICCT), Tianjin, China: IEEE, (2021), pp. 1134–1138.
- Xing, X., Li, Z., Xu, T., Shu, L., Hu, B., and Xu, X. (2019). SAE+LSTM: a new framework for emotion recognition from Multi-Channel EEG. *Front. Neurobot.* 13:37. doi: 10.3389/fnbot.2019.00037
- Xu, F., Pan, D., Zheng, H., Ouyang, Y., Jia, Z., and Zeng, H. (2024). EESCN: a novel spiking neural network method for EEG-based emotion recognition. *Comput. Methods Prog. Biomed.* 243:107927. doi: 10.1016/j.cmpb.2023.107927
- Yan, Y., Wu, X., Li, C., He, Y., Zhang, Z., Li, H., et al. (2023). Topological EEG nonlinear dynamics analysis for emotion recognition. *IEEE Trans. Cogn. Dev. Syst.* 15, 625–638. doi: 10.1109/TCDS.2022.3174209
- Yang, Y., Wang, Z., Tao, W., Liu, X., Jia, Z., Wang, B., et al. (2024). Spectral-spatial attention alignment for multi-source domain adaptation in EEG-based emotion recognition. *IEEE Trans. Affect. Comput.*, 1–13. doi: 10.1109/TAFFC.2024.3394436
- Yang, Y., Wu, Q., Qiu, M., Wang, Y., and Chen, X., “Emotion recognition from Multi-Channel EEG through parallel convolutional recurrent neural network,” In 2018 international joint conference on neural networks (IJCNN), Rio de Janeiro: IEEE, (2018), pp. 1–7.
- Yuan, N., Kang, B. H., Xu, S., Yang, W., and Ji, R., “Research on image target detection and recognition based on deep learning,” In 2018 international conference on information systems and computer aided education (ICISCAE), Changchun, China: IEEE, (2018), pp. 158–163.
- Zhang, J., Yin, Z., Chen, P., and Nichele, S. (2020). Emotion recognition using multimodal data and machine learning techniques: a tutorial and review. *Inf. Fusion* 59, 103–126. doi: 10.1016/j.inffus.2020.01.011
- Zhang, Y., and Zheng, X., “Development of image processing based on deep learning algorithm,” In 2022 IEEE Asia-Pacific conference on image processing, electronics and computers (IPEC), Dalian, China: IEEE, (2022), pp. 1226–1228.
- Zheng, W.-L., and Bao-Liang, L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Mental Dev.* 7, 162–175. doi: 10.1109/TAMD.2015.2431497
- Zheng, W.-L., Liu, W., Lu, Y., Lu, B.-L., and Cichocki, A. (2019). Emotion meter: a multimodal framework for recognizing human emotions. *IEEE Trans. Cybern.* 49, 1110–1122. doi: 10.1109/TCYB.2018.2797176
- Zhong, T., Wang, T., Wang, J., Wu, J., and Zhou, F. (2020). Multiple-aspect attentional graph neural networks for online social network user localization. *IEEE Access* 8, 95223–95234. doi: 10.1109/ACCESS.2020.2993876