MSFR-GCN: A Multi-Scale Feature Reconstruction Graph Convolutional Network for EEG Emotion and Cognition Recognition

Deng Pan, Haohao Zheng, Feifan Xu, Yu Ouyang, Zhe Jia, Chu Wang, and Hong Zeng, *Member, IEEE*

***Abstract*— Graph Convolutional Network (GCN) excels at EEG recognition by capturing brain connections, but previous studies neglect the important EEG feature itself. In this study, we propose MSFR-GCN, a multi-scale feature reconstruction GCN for recognizing emotion and cognition tasks. Specifically, MSFR-GCN includes the MSFR and feature-pool characteristically, with the MSFR consisting of two sub-modules, multi-scale Squeeze-and-Excitation (MSSE) and multi-scale sample re-weighting (MSSR). MSSE assigns weights to channels and frequency bands based on their separate statistical information, while MSSR assigns sample weights based on combined channel and frequency information. The feature-pool, which pools across the feature dimension, is applied after GCN to retain EEG channel information. The MSFR-GCN achieves excellent results in emotion recognition when first tested on two public datasets, SEED and SEED-IV. Than the MSFR-GCN is tested on our self-collected Emotion and Cognition EEG dataset (ECED) for both emotion and cognition classification tasks. The results show MSFR-GCN's good performance in emotion and cognition classification tasks and reveal the implicit relationship between the two, which may provide aid in the rehabilitation of people with cognitive impairments from an emotional perspective.**

***Index Terms*— EEG emotion recognition, EEG cognition recognition, graph convolutional network, SE attention, graph pool.**

Manuscript received 5 March 2023; revised 6 June 2023; accepted 6 August 2023. Date of publication 14 August 2023; date of current version 18 August 2023. This work was supported in part by the National Science Foundation of China under Grant 62076083, in part by the National Key Research and Development Program of China under Grant 2022YFE0199300, in part by the Hangzhou Artificial Intelligence Major Technological Innovation Project under Grant 2022AIZD0159, and in part by the Leading Goose Research and Development Program of Zhejiang under Grant 2023C03026. *(Corresponding author: Hong Zeng.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Zhejiang Chinese Medical University's Ethical Committee under Approval No. 2021-KL-021-01.

Deng Pan, Haohao Zheng, Feifan Xu, Yu Ouyang, Zhe Jia, and Hong Zeng are with the School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China (e-mail: dpan@hdu.edu.cn; zhenghaohao@hdu.edu.cn; 212050125@hdu.edu.cn; 222050269@hdu.edu.cn; 222050129@hdu.edu.cn; jivon@hdu.edu.cn).

Chu Wang is with the Second School of Clinical Medicine, Zhejiang Chinese Medical University, Hangzhou 310053, China (e-mail: 21718608@zju.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2023.3304660

I. Introduction

EMOTION is vital in our lives, but aging and declining health can reduce cognitive ability and sensitivity to emotion. Generally, emotional behavior patterns partly reflect cognitive ability, which also affects emotional response [1]. The cognitive theory suggests that emotions have cognitive components and a lack of emotional balance can cause various mental disorders, from neuroses to psychoses [2]. Recognizing emotion correctly and understanding the emotional-cognitive interaction are crucial not only to understanding the mind but also to revealing the causes of cognitive disorders [3].

Neuroscience researches argue that physiological signals, which are difficult to disguise and hide [4], are more representative of describing emotion than behavioral signals [5]. EEG is a physiological signal that directly and accurately reflects human brain activity, and with the rapid development of non-invasive and inexpensive EEG recording devices, EEG-based emotion recognition has received increasing attention [6].

Research on EEG-based emotion recognition involves two main aspects, namely, EEG discriminant feature extraction and emotion classification. Basically, the extracted EEG features can be classified into two types: time domain feature type (e.g., Hjorth feature [7], fractal dimension feature [8]) and frequency domain feature type (e.g., difference entropy [9]). Several researches [10] have indicated that distinct frequency bands play an vital role in human emotional behavior. Furthermore, certain EEG-based emotion detection study [11] have demonstrated the usefulness of frequency domain features in emotion recognition. To obtain frequency domain information, most researchers divide EEG signals into multiple frequency bands, e.g., δ band (1)-3 Hz), θ band (4)-7 Hz), α band (8)-13 Hz), β band (14-30 Hz) and γ band (31-50 Hz) [12], and then extract distinct features separately.

CNN-based deep learning methods have become popular in many fields in recent years and can extract local features on regular grid data using convolutional kernels. However, EEG signals are irregular due to electrode placement on an irregular grid. Some studies [13], [14] that use CNNs and RNNs have effectively transformed continuous EEG signals into a regular grid for convolution, but their methods require

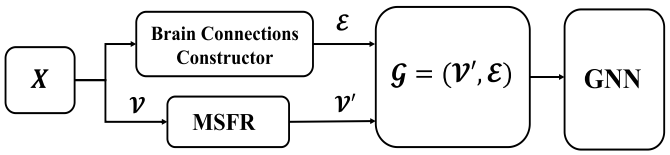


Fig. 1. Our model uses MSFR to reconstruct EEG features (V) and demonstrates the simultaneous use of E and V for better results, while others rely on E.

a 2D representation of the EEG channels on the scalp, which leads to a loss of spatial information in these channels. Representing a collection of EEG channels ($\mathcal{V}$) and brain connections ($\mathcal{E}$) through graphs (denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$) is an effective method. We characterize $\mathcal{V}$ as structure-independent information, while $\mathcal{E}$ are seen as structure-dependent information. Due to GNNs' superiority in irregular data, further EEG signal topology studies using GNNs are underway in EEG emotion recognition. Zhang et al. [15] designed Graph Convolutional Broad Network (GCB-net) to explore the deeper-level information of graph-structured data. Regularized graph neural network (RGNN) was introduced by Zhong et al. [16] to capture local and global interactions between EEG channels. Song et al. [17] proposed an instance-adaptive graph neural network (IAG) to model the relationships between EEG channels more reasonable. Ye et al. [18] proposed hierarchical dynamic GCN (HD-GCN) to explore dynamic multilevel spatial information among EEG channels.

Despite GNN's superiority over CNN for EEG emotion recognition, it still has two significant limitations. Firstly, most GNN-based emotion models construct brain connections from EEG channel associations, then use Message-Passing [19] in GNN to learn more about the connection and EEG feature. However, it appears that prior studies are dependent on brain connections overly in EEG emotion recognition, lacking proper utilization of EEG features ($\mathcal{V}$), as shown in Fig. 1. According to certain study [20], each frequency band and channel has a crucial role in emotional responses and should be assigned varying levels of importance. The aforementioned GNN-based emotion detection models often use pre-processed EEG features to generate graphs directly; all channels and frequency bands share the same importance for EEG emotion recognition, which results in an inappropriate graph representation. In brief, previous approaches undervalued EEG feature and treated all channels/frequency bands equally. Secondly, past researches have used simple pool methods (average or sum) that are applied across the node dimension after GCN, which is not ideal for EEG applications. GCN's output (i.e., $Z \in \mathbb{R}^{n \times d'}$) has channel-wise and feature-wise dimensions, but the latter is artificially designated and doesn't directly correspond to realistic meaning. Pooling across the node dimension leads to loss of specific and crucial channel information for EEG task.

Aside from limitations in model design, the previously used emotion datasets also have areas for improvement. Previous datasets explored the task of recognizing emotions across subjects and sessions, producing findings to improve our understanding of brain activity patterns. However, these datasets were limited to young, healthy subjects. Cognitive ability remains stable in a short session but can change with age and disease (e.g., Alzheimer's disease). Exploring the emotional responsiveness of the elderly and the interaction between emotion and cognition is difficult but rewarding, and it may aid in the rehabilitation of people with cognitive impairment.

Previous GNN-based models' shortcomings are addressed by our multi-scale feature reconstruction GCN (MSFR-GCN) for EEG emotion recognition. The MSFR combines with a feature-adaptive graph connections module [21] to enhance the utilization of structure-independent and structural information, respectively. Channel-SE and frequency-SE of MSSE, based on designed variant of SE attention, explores what channel and band weight reconstruction achieves the overall optimum on the dataset. And the MSSR module integrates channel and frequency domain information to assign variable importance to samples. When GCN has enough structure and structure-independent information, we use a reasonable feature-pool across feature dimension to extract powerful representation. Finally, we conduct subject-dependent and subject-independent experiments on two widely used EEG emotion datasets (SEED [11] and SEED-IV [22]). Additionally, we introduce a dataset to explore the link between emotion and cognition. Specifically, a group of elderly individuals, who experience cognitive problems, in different cognitive stages were recruited from hospital and subjected to experiments that induced joyful, neutral, and sad emotions. On that emotion and cognition EEG dataset (ECED), we perform subject-independent emotion and cognition classification experiments and analyze the results.

In summary, the contributions of this work can be outlined as:

- We propose the MSFR module to enhance GCN's learning of structure-independent information by assigning weights to channels, frequency bands, and samples, resulting in a more integrated and focused representation. After GCN, we use the feature-pool to extract powerful representation, and we get our MSFR-GCN.
- On two public EEG datasets, we conducte comprehensive tests in both subject-dependent and subject-independent classification conditions. The results show that our proposed model is effective, and structure-independent information is valuable. Moreover, our MSFR-GCN model outperforms the competitive models in most experimental conditions.
- We design an experimental paradigm that can simultaneously reflect the emotional and cognitive state of the subjects and then perform subject-independent emotion classification and cognition identification tests using our self-collected EEG dataset. The findings suggest that strong emotion can help with cognition recognition, whereas low cognitive competence leads to decreased sensitivity to emotion stimuli.

The remainder of this article is organized as follows. Section II describes the proposed method in detail. Section III introduces the details of our ECED and two public datasets,

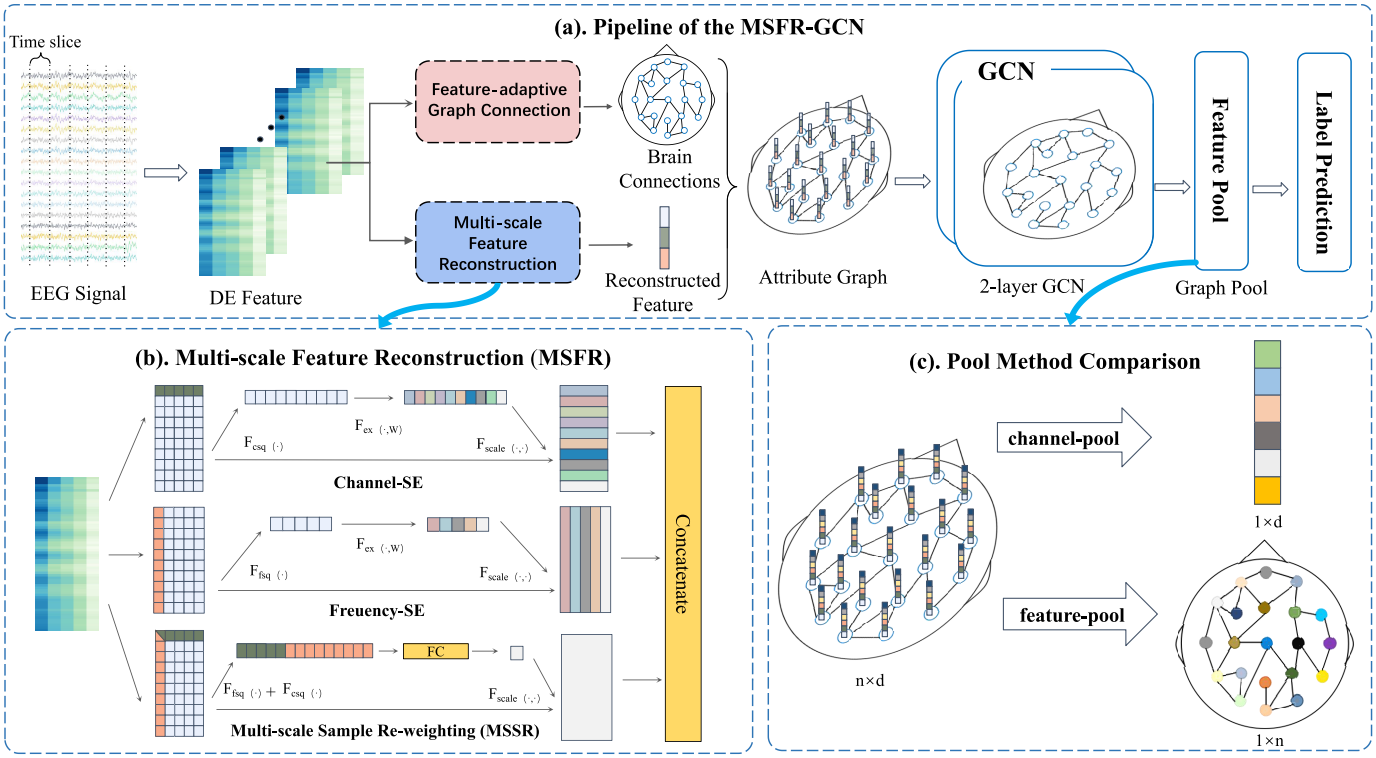


Fig. 2. (a) The pipeline of the proposed MSFR-GCN. (b) The MSFR branch is provided to learn more structure-independent information. The channel-SE, frequency-SE, and MSSR modules calculate the weights of various channels, frequency bands, and samples based on statistical information from different perspectives, and then reconstruct the DE features individually. Finally, we concatenate these three modules to obtain the final output of MSFR. (c) By comparing and visualizing the sum pool across different dimensions, we find that the feature-pool (sum across feature dimensions) better maintains channel information. The detailed MSFR-GCN are described in Section II.

and presents experimental results and analysis of all three datasets. Section IV does ablation experiments, visualization of the MSSE, and joint parameter sensitivity analysis for the MSSE. Finally, Section V summarizes this study.

II. Method

A. Overview

In this section, we first introduce the attribute graph for EEG features. Then we present the details of the proposed MSFR-GCN model.

The framework of the proposed MSFR-GCN for EEG emotion recognition is shown in Fig. 2. In MSFR-GCN, the feature-adaptive branch is used to learn dynamic brain connections. The MSFR branch examines the impact of assigning varying weights to channels, frequency bands, and samples on emotion detection and concatenates the scaled EEG features to produce the final node features ($\mathcal{V}$). And those two branches provide graph input to GCN, and after two layers, we sum the graph via feature-pool to create the final EEG graph representation in MSFR-GCN. The final EEG graph representation is then connected to a fully connected layer, and the predicted labels are obtained through a softmax layer.

B. Attribute Graph for EEG Signals

Generally, we use EEG energy features $X \in \mathbb{R}^{n\times d}$ (n and d are the number of EEG channels and frequency bands, respectively) from five frequency bands (δ: 1-4 Hz, θ: 4-8 Hz, α: 8-14 Hz, β: 14-30 Hz, and γ: 30-50 Hz) as input. A graph is defined as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_i\}_{i=1}^{n}$ is the set of nodes, $\mathcal{E} \in \mathbb{R}^{n\times n}$ is the set of edges. Generally, we describe $\mathcal{E}$ and $\mathcal{V}$ in terms of the adjacency matrix A and X. $A_{i,j}$ represents the connection between source node v_i and destination node v_j. v_i and v_j are not connected if $A_{i,j} = 0$, otherwise $A_{i,j} \neq 0$.

The graph Fourier transform (GFT) [23] is a common example of spectral graph filtering, or graph convolution. To describe node connections, GFT uses the Laplacian matrix $L = D - A$ or the normalized Laplacian matrix:

$$\hat{L} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \tag{1}$$

where I is the n-th order identity matrix and D is the diagonal degree matrix of A. Given that $\hat{L}$ is a symmetric positive semi-definite matrix, it can be decomposed into $\hat{L} = U\Lambda U^{\mathrm{T}}$, where U is the orthonormal eigenvector matrix of $\hat{L}$ and $\Lambda = \mathrm{diag}(\lambda_1, \hat{\mathrm{A}}\cdot\ \hat{\mathrm{A}}\cdot\ \hat{\mathrm{A}}\cdot, \lambda_n)$ is a diagonal matrix.

For a given feature matrix X, its graph fourier transform is $\hat{X} = U^T X$, and the corresponding inverse fourier transform is $X = U\hat{X}$. The graph convolution operation $*_{\mathcal{G}}$ between the input X and the filter G can be expressed as:

$$Y = X *_{\mathcal{G}} G = U\left(\left(U^{\mathrm{T}}X\right) \odot \left(U^{\mathrm{T}}G\right)\right) = U\hat{G}U^{\mathrm{T}}X \tag{2}$$

where $\odot$ is the element-wise Hadamard product, and $\hat{G}$ is the convolution kernel in the spectral domain. Finally, the convolution theorem from classical signal processing has been extended to graphic frequency representation.

C. Feature-Adaptive Graph Connection

The adjacency matrix $A \in \mathbb{R}^{n\times n}$ describes EEG channels' topology and is crucial for learning graph representation. Our model doesn't aim to find a better method for building channel connections, but an efficient strategy is still necessary. Jia et al. [21] provided an adaptive graph learning approach that learns the graph structure dynamically rather than being formed by prior information or artificial means. $A_{i,j} = g\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) (i, j \in \{1, 2, \cdots, N\})$ represents the connection between node $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ based on the input feature matrix $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, x_n)^T \in \mathbb{R}^{n\times d}$, and $g(\cdot)$ is defined as:

$$g\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \frac{\exp\left(\mathrm{ReLU}\left(\boldsymbol{w}^T \left|\boldsymbol{x}_i - \boldsymbol{x}_j\right|\right)\right)}{\sum_{j=1}^{N} \exp\left(\mathrm{ReLU}\left(\boldsymbol{w}^T \left|\boldsymbol{x}_i - \boldsymbol{x}_j\right|\right)\right)} \tag{3}$$

where w is a learnable vector. The approach described above subtracts $\boldsymbol{x}_i$ from $\boldsymbol{x}_j$, then normalizes them with the softmax operation to obtain the connection weights $A_{i,j}$.

D. Multi-Scale Feature Reconstruction

1) Multi-Scale SE: Since different channels and frequence bands contribute unequally for emotion recognition, so adaptive assignment of varying importance to them is necessary. The Squeeze-and-Excitation(SE) [24] attention is able to recalibrate channel-wise feature responses by explicitly modeling interdependencies between channels. However, original SE attention is biased towards image-like data (i.e., $U \in \mathbb{R}^{H\times W\times C}$). For a graph feature map $X \in \mathbb{R}^{n\times d}$ with only two dimensions, we cannot use SE directly because it only evaluates channel-wise descriptors for image-like data. Due to the importance of spatial and frequency information in human emotion recognition, we propose a multi-scale SE (MSSE) unit for graph-like data to highlight the most informative features for EEG tasks. The MSSE is composed of two variants of the SE attention, channel-SE and frequency-SE, which obtain weights based on the spatial and frequency perspectives, respectively.

Fig. 2b depicts the structure of the MSFR block, with the first two branches representing the MSSE. EEG features are passed through channel-squeeze and frequency-squeeze operations to obtain the channel and frequency descriptors, respectively, by calculating the mean across their spatial (d) and frequency dimensions (n). The two descriptors enable the embedding of global channel-wise and frequency-wise feature responses, giving all network layers access to global receptive field information. That aggregation is followed by an excitation operation, taking the form of a self-gating mechanism, and outputting per-channel and per-frequency modulation weights.

Specifically, both channel-squeeze and frequency-squeeze are done via 1D global-average pooling to generate channel-wise and frequency-wise statistic features. Formally, the statistic output calculated by global average pooling is defined as:

$$z_{csq} = F_{csq}(X) = \frac{1}{d}\sum_{t=1}^{d} X(n, t) \tag{4}$$

$$z_{fsq} = F_{fsq}(X) = \frac{1}{n}\sum_{t=1}^{n} X(t, d) \tag{5}$$

where X is the EEG feature. The squeeze operation (i.e., z_{csq} and z_{fsq}) generates global statistical features for both channel and frequency information, followed by an excitation operation (F_{ex}) that fully utilizes their dependencies through two fully connected (FC) layers. Excitation operation generates weights for channels and frequency bands, which are multiplied by input X to generate reconstructed features from various scales. The outputs of the channel-SE and frequency-SE operations are denoted by:

$$Z_c = F_{ex}(z_{csq})X = \sigma\left(W_2\delta\left(W_1 z_{csq}\right)\right)X \tag{6}$$

$$Z_f = F_{ex}(z_{fsq})X = \sigma\left(W_4\delta\left(W_3 z_{fsq}\right)\right)X \tag{7}$$

where W_i refers to the weight of FC layers, $\delta(\cdot)$ is the RELU function, $\sigma(\cdot)$ is the Sigmoid function, Z_c and Z_f are the results of the channel-SE and frequency-SE operation. Among the F_{ex} operation, there is a reduced dimension r (i.e., $W_1 \in \mathbb{R}^{r\times c}$ and $W_2 \in \mathbb{R}^{c\times r}$) and that hyperparameter settings will be discussed in Section IV.

2) Multi-Scale Sample Re-Weighting: Although channel and frequency branches do not interflow information, they can be viewed as a set of local descriptors whose statistics are informative for the entire input EEG feature [24]. Therefore, feature reconstruction should be jointly recalibrated across multiple domains. Additionally, DNNs tend to overfit to biased training data with class imbalance [25], and sample re-weighting strategies are commonly used to mitigate this issue.

To address the concerns, we propose MSSR, a multi-scale sample re-weighting unit that learns an adaptive weighting function from both a spatial and a frequency perspective. The MSSR's input is a concatenation of channel descriptor (z_{csq}) and frequency descriptor (z_{fsq}). The weighting function is an MLP with one hidden layer, enabling the approach to fit a variety of weighting functions. The formula of MSSR is as follows:

$$Z_s = S_s X = \sigma\left(W_6\delta\left(W_5(z_{csq} \| z_{fsq})\right)\right)X \tag{8}$$

where W_i refers to the parameters of the MLP, S_s represents the weight of the sample, and $\|$ is the concatenating operator. Based on the descriptors from different scales and MLP, we finally get re-weighting samples as Z_s.

To make rational and efficient use of the reconstructed features (Z_c, Z_f and Z_s), the output of the MSFR module is designed as:

$$X_{re} = (Z_f \| Z_c \| Z_s) \tag{9}$$

E. Graph Construction and Graph Convolution Network

From the above two subsections, we obtain matrix A describing the brain connections and the reconstructed EEG features X_{re} as the input for GCN.

Kipf and Welling [26] presented an efficient variant of CNNs operating directly on graphs. They made a localized first-order approximation to ChebNet [27], and formulated a 2-layer GCN as follows:

$$Z = \mathrm{softmax}\left(\hat{L}\,\mathrm{ReLU}\left(\hat{L}X_{re}W^{(1)}\right)W^{(2)}\right) \tag{10}$$

The normalized laplacian matrix $\hat{L}$ prevents the value in the feature matrix from getting too large. Finally, the graph pool is used to resize the features to a fixed size before graph classification. It's worth noting that the global-add pool is performed across the feature dimension instead of the channel dimension [16], and we refer to it as feature-pool. And the final graph represention of MSFR-GCN is defined as:

$$g = \sum_{t=1}^{d} Z(n, t) \tag{11}$$

where Z is the output of GCN, and g is the input of classification net.

F. Algorithm For MSFR-GCN

We iteratively update network parameters using backpropagation (BP) until optimal results are achieved. We establish a cross-entropy loss function with the following form:

$$Loss = crossentropy\left(\mathbf{l}, \mathbf{1^p}\right) + \alpha\|\Theta\|_2 \tag{12}$$

where $\mathbf{l}$ and $\mathbf{1^p}$ denote the training data's actual and predicted label vectors, Θ denote all of the model's parameters, α denote the trade-off regularization weight, and $\|\cdot\|_2$ denote the $l_2 - norm$. We employ cross entropy loss to quantify the dissimilarity between predicted and real labels, as well as regularization $\alpha\|\Theta\|$ to prevent overfitting of model parameters learning and limit adjacency matrices A.

The Algorithm 1 summarizes the training procedures of the proposed MSFR-GCN on EEG-based emotion and cognition recognition.

Algorithm 1 The Training Procedure of MSFR-GCN

Require: X, Y: EEG features associated with multiple frequency bands and the corresponding labels; η: The learning rate for backward updating parameters; T: number of epochs; B: batch size; α: The hype-parameter of the L2 regularization; r: The reduced dimension of the MSSE unit

Ensure: The learned model parameters in MSFR-GCN;

Randomly initialize model parameters in MSFR-GCN using Xavier initialization [28];
for $i = 1 : T$ **do**
  repeat
    Draw one batch of training samples;
    Calculate the adjacency matrix A using (3);
    Calculate the output (X_{re}) of MSFR using (6), (7), (8) and (9);
    Calculate GCN's output (Z) using (10);
    Do feature-pool(11) after GCN;
    Calculate the result (g) of the classification net;
    Calculate the loss function using (12);
    Update parameters in the MSFR-GCN model with gradient descent;
  until the iterations satisfies the predefined algorithm convergence condition.
end for

III. Experiments

In this section, we verify our model's performance on one private dataset and two widely used EEG emotion datasets (ECED, SEED, and SEED IV) and analyze the results.

A. Protocol of the ECED

To better use emotional stimuli to help determine cognitive states, we chose the more established Oddball paradigm for better cognitive judgment [29]. Specifically, there are three emotion pictures used to elicit happy, neutral and sad emotion, and squares and circles (representing standard and target stimuli) are used to facilitate cognitive activities. The participants belonged to three main cognitive states: Alzheimer's disease (AD), mild cognitive impairment (MCI), and healthy control (HC). AD is a progressive neurodegenerative disorder characterized by cognitive, and functional changes and its prevalence escalates with age [30]. MCI subjects display memory impairment beyond their age expectations but do not meet the criteria for dementia diagnosis. HC refers to elderly subjects without cognitive impairment.

All participants were informed about the study protocol and the groups had a homogeneous age (mean: 66.24, std: 7.10) and gender (14 males and 19 females). EEG data were recorded from 33 subjects: 11 AD patients, 11 MCI patients, and 11 HC subjects. All subjects completed Montreal Cognitive Assessment (MoCA) [31], Cognitive Abilities Screening Instrument (CASI) [32]. The hospitalists assess the subject's overall cognition and complete the Clinical Dementia Rating (CDR) [33] based on the subject's questionnaire performance.

Subjects sat 100 cm from a computer screen in a dimly lit room. Each cognition group viewed three randomly ordered emotional scenarios, and the test flow is shown in Fig. 3. After each experiment, the participants rated their corresponding emotional response to the experiment based on their actual feelings (0-10, with 0 indicating no emotional response). Self-measurement scores for neutral, happy, and sad were as follows: AD (6.55/7.00/6.97), MCI (7.73/7.82/6.82), and HC (8.27/8.63/8.18). Individuals with more severe cognitive deterioration show lower sensitivity to emotional stimuli and record lower self-assessment scores.

According to the international 10-20 system, the 32-channel EEG signals were recorded using g.Nautilus Research EEG cap at a sampling rate of 250 Hz. Additionally, the signals were band-pass filtered between 0.1 and 30 Hz. After that, ICA [34] was adopted to remove eye blinking artifacts.

B. Brief Description of Datasets

To assist identification, we directly employ pre-computed DE features smoothed by linear dynamic systems [11] to establish a fair comparison with prior works.

1) SEED and SEED-IV: For SEED and SEED-IV, DE features are pre-computed over five frequency bands (delta, theta, alpha, beta, and gamma) for each second of EEG signals (without overlapping) in each channel.

There are 15 subjects and 3 sessions per subject in the SEED dataset. Each session included 15 film clips, five for each mood, evoking three types of emotions: positive, neutral,

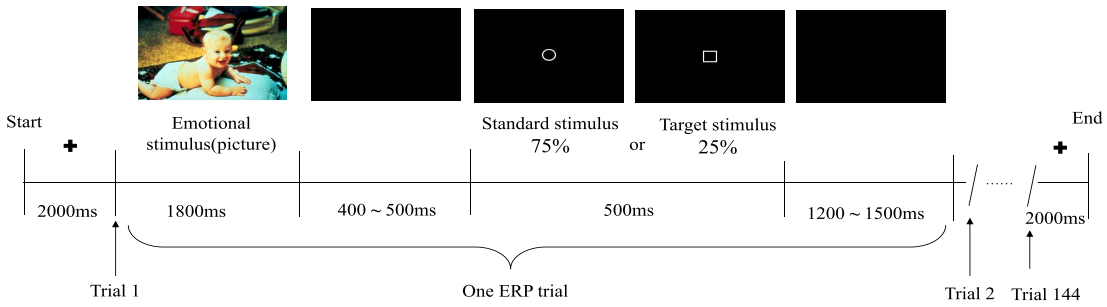


Fig. 3. Protocol of the EEG experiments on ECED.

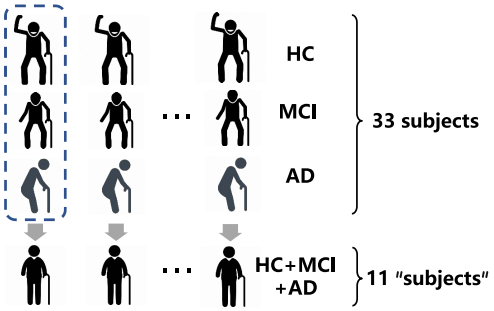


Fig. 4. Subject-independent cognition classification protocol for ECED. There are 33 different subjects in one emotional state.

and negative. Each person had 15 trials per session, with 185-238 samples per trial, resulting in a total of approximately 3400 samples. SEED-IV is an extension of the SEED dataset, which includes four emotions: happy, neutral, sad, and fear, each with six film clips. Each session has 24 trials with 12-64 samples, totaling approximately 830 samples.

2) ECED: One-second sliding window was applied to segment EEG signals, and the DE features were extracted from four frequency bands (δ, θ, α, and β band). Therefore, for 33 subjects, an average of 3600 samples could be obtained for each subject, and each sample had 32 dimensional features.

C. Experimental Protocol

To evaluate our model, two types of experiments are used: subject-dependent and subject-independent. The training and testing data for the subject-dependent task are obtained from different trials of the same individual, whereas for the subject-independent task, they are obtained from different subjects.

1) SEED and SEED-IV: The same protocol as in [16], [35], and [36] is used for the subject-dependent experiment. In particular, we train our model on each subject in the SEED dataset using the first 9 trials for training and the remaining 6 trials for testing in two sessions. For the SEED-IV dataset, we train our model on the first 16 trials and test it on the remaining 8 trials, which cover all emotions (two trials per emotion class) in three sessions.

For the subject-independent experiment, we use the leave-one-subject-out (LOSO) cross-validation strategy as in [37], where one subject's EEG emotion data is used for testing and the remaining subjects' data for training. Specifically, we evaluate the model's performance across all test subjects in one session for SEED and three sessions for SEED-IV.

2) ECED: We conduct subject-independent emotion and cognition classification experiments. For emotion classification on the ECED, we use the LOSO strategy across three cognitive groups (HC-group, MCI-group, and AD-group) and evaluate model performance by averaging accuracy across the three groups.

Direct subject-independent cognition classification is not possible for the ECED, as a person cannot be in all three cognitive states simultaneously. We categorize the experimental data into three emotion groups (happy group, neutral group, and sad group) based on their emotional states. Fig. 4 displays three groups in the same emotional state, but belonging to different cognitive states. To create a "subject" with three cognitive states in one emotion states, we choose one subject sequentially from each set of cognitive states. For cognition classification, we perform experiments using data from all three emotion groups.

D. Setup of the Experiments

In the experiments, the size of the input X is $n \times f$ and the output dimension of each electrode is 20. In particular, MSFR-GCN is implemented using PyTorch on a Nvidia 3070 laptop GPU and trained using the Adam optimizer with batch size 128, learning rate 0.005, and weight decay rate 2e-5. The mean accuracy (ACC) and standard deviation (STD) are employed as evaluation criteria in all datasets.

E. Experimental Resluts and Analysis for ECED

Table I show the results of two experiments (cognition task and cognition task) conducted on the ECED.

To verify the validity of MSFR-GCN on the ECED, we select the baseline model, which combines GCN and a feature-adaptive graph connections module, along with EEG-Net [38] as the comparative method. The former serves as the baseline for our model, while the latter represents a typical application of the convolutional model in the EEG domain. The comparison results with baseline show the significant effectiveness of our designed MSFR and feature-pool on two cross-subject tasks, and the comparison results with EEG-Net demonstrate the superior performance of our designed MSFR-GCN than convolutional network. Next, we reanalyzed the results in terms of group. Firstly, the results of all models show that low cognitive group(AD,MCI) will lead to worse emotion recognition, and happy emotion can obviously help to identify the cognitive state. Secondly, our model narrows

TABLE I
SUBJECT-INDEPENDENT EMOTION CLASSIFICATION ACCURACY (MEAN/STD) ON ECED

Model(Cognition Task)	Sad-group	Neutral-group	Happy-group	Average
EEGNet	45.89/12.60	46.22/15.67	48.48/12.81	46.86/14.03
baseline	48.76/11.01	48.53/13.48	51.97/11.21	49.75/12.06
MSFR-GCN	**63.87/08.29**	**63.65/11.35**	**64.77/07.53**	**64.10/08.92**
Model(Emotion Task)	HC-group	MCI-group	AD-group	Average
EEGNet	43.59/09.50	41.78/10.85	40.54/08.46	41.57/09.82
baseline	46.05/09.61	45.39/09.37	43.22/07.72	44.88/09.06
MSFR-GCN	**57.98/08.67**	**56.72/09.88**	**55.93/05.44**	**56.88/08.00**

TABLE II
SUBJECT-DEPENDENT AND SUBJECT-INDEPENDENT CLASSIFICATION ACCURACY (MEAN/STD) ON SEED AND SEED-IV

Model	SEED		SEED-IV	
	dependent	independent	dependent	independent
SVM	83.99/09.72	56.73/16.29	56.61/20.05	37.99/12.52
DGCNN	90.40/08.49	79.95/09.02	69.88/16.29	52.82/09.23
BiDANN	92.38/07.04	83.28/09.60	70.29/12.63	65.59/10.39
RGNN	94.24/05.95	85.30/06.72	79.37/10.54	73.84/08.02
IAG	95.44/05.48	86.30/06.91	—	—
GMSS	96.48/04.63	86.52/06.22	86.37/11.45	73.48/07.41
RGNN w/o DA	—	81.92/09.35	—	71.65/09.34
MSFR-GCN	**96.63/04.60**	**86.78/05.40**	**89.02/11.31**	**73.43/07.32**

"w/o DA" means removing node-wise domain adversarial training component.

the gap between groups and has better robustness. Specifically, the maximum difference of cognition task between groups decreased from 3.44%(baseline) to 1.12%(MSFR-GCN), and the maximum difference of emotion tasks decreased from 3.05%(EEGNet) to 2.05%(MSFR-CGN). Lastly, the cognition task results of all models showed that the neutral-group has a significantly larger standard deviation, and the AD-group has a significantly smaller standard deviation in the emotion task. This proves that neutral emotion do not have stable patterns that help cognitive recognition, and AD-group have consistently poor performance in emotional responses, reflecting less complex and more consistent brain activity.

The finding is consistent with the notion that cognitive abilities can influence emotional reactions and strong emotion help to distinguish cognition [1], [3]. Though emotion task are challenging due to issues with the Oddball paradigm and single emotional picture stimuli, our model has great performance improvement and stability improvement on both cognitive and emotional tasks.

F. Experimental Resluts and Analysis for SEED and SEED-IV

To compare MSFR-GCN's benefits, similar experiments were carried out as with other methods such as linear support vector machine (SVM) [39], dynamical graph convolutional neural network (DGCNN) [40], bi-hemisphere domain adversarial neural network (BiDANN) [36], regularized graph neural network (RGNN) [16], instance-adaptive graph method (IAG) [17], and graph-based multi-task self-supervised(GMSS) [41]. These approaches represent earlier studies on emotion recognition, and their outcomes are directly cited from the literature

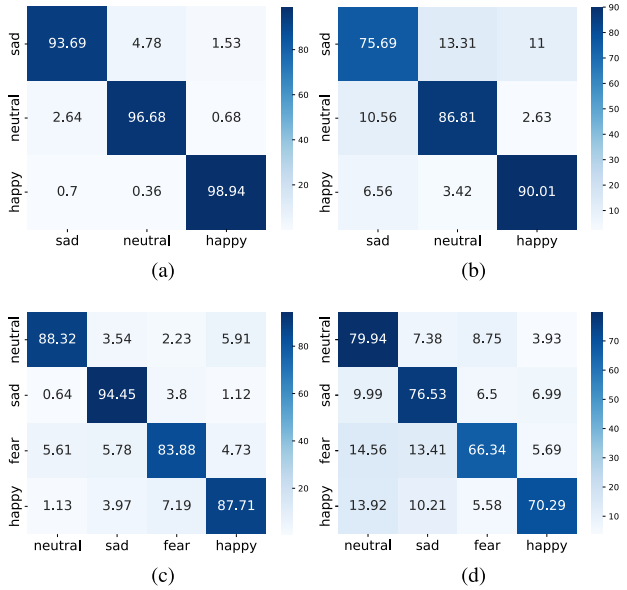


Fig. 5. Confusion matrices of MSFR-GCN. (a) Subject-dependent classification on SEED (b) subject-independent classification on SEED (c) Subject-dependent classification for SEED-IV (d) Subject-independent classification on SEED-IV.

to ensure a compelling comparison with the proposed method. The experiments results are summarized in Table II.

In subject-dependent experiments on two public EEG emotional datasets, MSFR-GCN outperforms all other mentioned methods, including four GNN-based models. MSFR-GCN performs 2.65% better than the most advanced approach, GMSS, on SEED-IV, indicating its superior discriminating ability in subject-dependent experiments.

Subject-independent experiments were conducted, and MSFR-GCN achieved the best performance on SEED, surpassing the previous best method, GMSS, by 0.26%. Our model also outperforms RGNN and BiDANN on the SEED dataset, which incorporate domain adversarial structure, by 1.48% and 3.50%. Meanwhile, MSFR-GCN performs similarly to RGNN and GMSS on SEED-IV, achieving 73.43%, 73.84%, and 73.48% accuracy, respectively. Removing RGNN's domain adversarial training component (RGNN w/o DA) designed for subject-independent tasks resulted in a 1.83% decrease in accuracy compared to MSFR-GCN. Compared to GMSS, MSFR-GCN exhibited competitive performance and improved stability, achieving slightly lower result but with a lower standard deviation. Additionally, MSFR-GCN achieves the lowest standard deviation when compared to all tasks, confirming its superior discriminating and generalization abilities.

Using our designed MSFR and feature pool, MSFR-GCN reconstructs EEG features by incorporating information from multiple scales, generating a more suitable graph representation for EEG tasks, resulting in positive outcomes on diverse datasets and tasks.

G. Confusion Matrix for SEED and SEED-IV

In Fig. 5, we present the confusion matrices of the EEG emotion recognition results using the proposed MSFR-GCN model on SEED and SEED-IV. In both subject-dependent and

TABLE III
ABLATION STUDY FOR SUBJECT-DEPENDENT AND SUBJECT-INDEPENDENT CLASSIFICATION ACCURACY (MEAN/STD) ON SEED AND SEED-IV

Model	SEED		SEED-IV	
	dependent	independent	dependent	independent
MSFR-GCN	96.63/04.60	86.78/05.40	89.02/11.31	73.43/07.32
w/o MSSE	95.13/05.38	83.73/05.87	86.41/11.50	69.75/08.37
w/o MSSR	95.97/05.18	83.98/06.92	87.63/10.70	72.30/09.32
w/o feature-pool	93.05/06.20	83.08/06.51	84.18/11.15	70.52/09.17
w/ MSSE	92.67/06.73	82.58/05.64	83.31/12.20	71.38/08.64
w/ feature-pool	91.34/07.94	81.86/05.30	80.81/12.24	64.22/09.56
baseline	89.91/07.60	71.70/13.10	74.35/14.09	58.62/07.23

subject-independent experiments on SEED, happy emotions are easiest to distinguish, while sad emotions are the hardest, consistent with previous works [41]. And also, we find that sad and neutral emotions are frequently misidentified to each other by the model, suggesting that they share some similarities.

For the SEED-IV dataset, neutral and sad emotions are easier to recognize than happy and fear emotions in both subject-dependent and subject-independent experiments. Notably, sad and neutral emotions have the highest accuracy (94.45% and 79.94%, respectively) in both subject-dependent and subject-independent experiments. This suggests that EEG patterns generated from neutral movies may share more similarities among people, while fear is the most challenging emotion to identify in both tasks. In subject-independent experiments, sad, fearful, and happy emotions are commonly misidentified as neutral emotions.

Due to the individual differences, subject-dependent experiments achieve higher classification accuracy than subject-independent experiments. While MSFR-GCN proves effective in emotion recognition, emotional activity patterns can vary significantly among individuals, so a fully general paradigm for frequency and channel perspectives still needs to be explored, and our model requires further improvement.

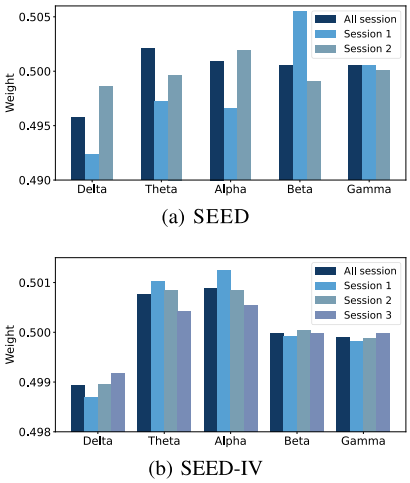
IV. DISCUSSION

This section investigates the proposed method and includes the ablation study of MSFR-GCN, the visualization of MSSE, and a joint parameter sensitivity analysis for MSSE.

A. Ablation Study

In this section, we perform an ablation study on subject-dependent and subject-independent experiments on SEED and SEED-IV to validate the effectiveness of each module in our method. As depicted in Table III, the notation "w/" indicates the combination of the baseline model with the subsequent component, while "w/o" indicates the removal of the component.

Compared to the baseline model, all other configurations show better performance, demonstrating the efficacy of the proposed modules (MSSE, MSSR, and feature-pool). Furthermore, the highest accuracy is attained when all three modules are combined for all tasks. A simple feature-pool design that only adjusts the dimension of the pool sum leads to improved results on all tests, highlighting the efficacy and rationality of the feature-pool approach for emotion recognition tasks. Although our module achieved good results on the subject-dependent task in the SEED dataset, the improvement in performance was less significant compared to the subject-dependent condition. Results suggest that MSFR-GCN excels at challenging tasks such as subject-independent classification and recognition of more emotions. Removing the MSSR module has a small effect on performance compared to removing the other two modules. When comparing the two experimental setups, w/ MSSE and w/ feature-pool, it is apparent that the more complicated and targeted design, w/ MSSE, performs better on all tasks.



(a) SEED

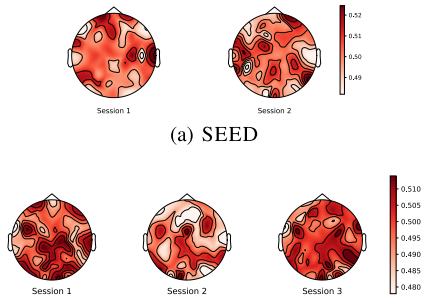
(b) SEED-IV

Fig. 6. Weights of frequency bands on SEED and SEED-IV.

B. The Visualization of the Learned Weight in Multi-Scale SE

To explore suitable weight reconstruction, we visualize the weights of channels and frequency bands for SEED and SEED-IV datasets (Fig. 6 and Fig. 7). The weights were collected from test samples of all participants. From Fig. 6a, we observe consistent results with previous studies [11], [20], showing that the β and γ bands have higher weights, while the α band has intermediate weights. This aligns with the higher classification accuracy observed for happy and neutral emotions in the confusion matrix.

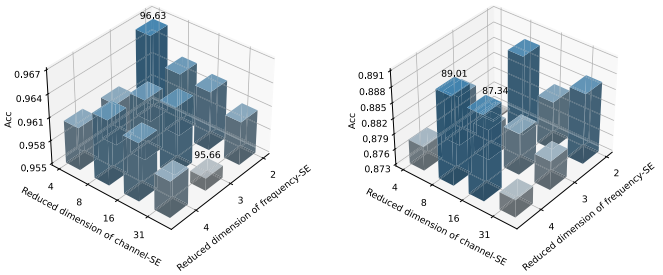
SEED-IV includes fear emotion, and negative and neutral emotion account for about a quarter of the data. According to research [11], [20], the β and γ bands are associated with good emotions, while the α band is associated with neutral emotions. Therefore the overall effect will be bad if the high frequency band is overemphasized. Knyazev et al. [42] found that explicit anxiety is associated with reduced theta band synchronization and higher alpha band de-synchronization in subjects experiencing high anxiety. Ertl et al. [43] found a positive correlation between EEG oscillations in the theta range and successful use of cognitive reappraisal in order to decrease negative emotions. Lesting et al. [44] claimed that theta activity thus seems to constitute an integrative mechanism for coordination of activity in the fear memory network. Tortella-Feliu et al. [45] thought higher spontaneous



(a) SEED

(b) SEED-IV

Fig. 7. Weights of channels on SEED and SEED-IV.



(a) SEED (b) SEED-IV

Fig. 8. Parameter sensitivity joint analysis of MSSE on SEED and SEED-IV.

frontal alpha activity was the predictor of faster recovery from discomfort elicited by normative neutral pictures. We argue that Alpha responses corresponded to participants' perception of neutral stimuli. Theta activity was observed during the processing of sad and fear stimuli, which can trigger fear memories and generate anxiety. Positive emotions were linked to enhanced beta and gamma responses. In order to achieve good global results, our model assigns more weight to theta and alpha in Fig. 6b, which is consistent with previous physiological studies.

As shown in Fig. 7a, channel weights on the SEED have high and asymmetric activation in the frontal and temporal lobes, which is consistent with studies on SEED [11], [20]. The results on SEED-IV datasets highlight the importance of the prefrontal and temporal lobes, as well as the involvement of the occipital and parietal lobes. Zheng et al. [20] reported that positive emotions activate more lateral temporal areas than negative emotions, neutral emotions elicit stronger responses at parietal and occipital sites, and negative emotions result in higher responses at parietal and occipital sites and prefrontal sites. Our findings in SEED-IV is consistent with above conclusion. Moreover, Zhong's [16] analysis of important brain regions on SEED-IV was consistent with ours. When there are more kinds of emotions, the brain's workings become more complex.

C. Sensitivity Analysis for Multi-Scale SE

Hu et al. [24] proposed the SE module, which utilizes an excitation operation with a reduction ratio that affects the model's performance and complexity. Since the number of EEG channels (62) and frequency bands (5) are not easily divisible by an integer, we introduce the reduced dimension without changing the core idea of the "excitation operation". Then we conduct a combined sensitivity analysis of the MSSE. Actually, we perform subject-dependent sensitivity analysis of reduced dimension on SEED and SEED-IV, and the results are shown in Fig. 8.

The model achieved the highest accuracy (96.63% and 89.01%) for subject-dependent experiments on SEED and SEED-IV with 4×2 and 8×4 reduced dimensions for channel-SE and frequency-SE, respectively. Furthermore, the model achieves the lowest accuracy (95.66% and 87.34%) when the reduction dimension is 31×3 and 4×2. In general, the results in Fig. 8a and Fig. 8b indicate that performance is robust for a range of reduced dimension combinations, and increasing complexity does not consistently improve performance.

Notably, given the robust performance of the SE mechanism, we did not explore all parameter combinations to achieve the highest accuracy in subsequent subject-independent experiments on SEED, SEED-IV, and ECED. We just chose suitable parameter combinations for multi-scale SE, using 8×3 for two public datasets and 8×2 for ECED. And the results also tell that the parameters used above can achieve competitive performance across the subject-independent task.

V. Conclusion

In this paper, a MSFR-GCN model is proposed for EEG emotion and cognition recognition. Different from earlier research that has focused on brain connections, we propose the MSFR to emphasize the value of structure-independent information in EEG data. Meanwhile, we adopt a more suitable feature-pool to get EEG graph representation. Extensive experiments on two public datasets demonstrate the superior performance of our model over the state-of-the-art GMSS in most experimental settings, and the visualization of channel weights and frequency band weights has proven that our MSFR can assign different and valid weights in most situations. We also do joint parameter analysis for the MSSE module and verify its robust performance. To address the limitation of the previous dataset, we introduce ECED and perform two subject-independent experiments on it. The results of ECED show that emotion and cognition can impact one another to some extent. In future work, better joint use of structural and unstructured information will be further investigated to explore how to further improve EEG emotion recognition.

References

[1] F. Dolcos, A. D. Iordan, and S. Dolcos, "Neural correlates of emotion-cognition interactions: A review of evidence from brain imaging investigations," *J. Cognit. Psychol.*, vol. 23, no. 6, pp. 669–694, Sep. 2011.

[2] R. J. Dolan, "Emotion, cognition, and behavior," *Science*, vol. 298, no. 5596, pp. 1191–1194, 2002.

[3] H. Okon-Singer, T. Hendler, L. Pessoa, and A. J. Shackman, "The neurobiology of emotion-cognition interactions: Fundamental questions and strategies for future research," *Frontiers Hum. Neurosci.*, vol. 9, p. 58, Feb. 2015.

[4] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time EEG-based emotion recognition and its applications," in *Transactions on Computational Science XII*. Berlin, Germany: Springer, 2011, pp. 256–277.

[5] J. C. Britton, K. L. Phan, S. F. Taylor, R. C. Welsh, K. C. Berridge, and I. Liberzon, "Neural correlates of social and nonsocial emotions: An fMRI study," *NeuroImage*, vol. 31, no. 1, pp. 397–409, May 2006.

[6] S. M. Alarcão and M. J. Fonseca, "Emotions recognition using EEG signals: A survey," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 374–393, Jul. 2019.

[7] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, Sep. 1970.

[8] Y. Liu and O. Sourina, "Real-time fractal-based valence level recognition from EEG," in *Transactions on Computational Science XVIII*. Berlin, Germany: Springer, 2013, pp. 101–120.

[9] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification," in *Proc. 6th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, Nov. 2013, pp. 81–84.

[10] D. Nie, X.-W. Wang, L.-C. Shi, and B.-L. Lu, "EEG-based emotion recognition during watching movies," in *Proc. 5th Int. IEEE/EMBS Conf. Neural Eng.*, Apr. 2011, pp. 667–670.

[11] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Trans. Auto. Mental Develop.*, vol. 7, no. 3, pp. 162–175, Sep. 2015.

[12] M. Li and B.-L. Lu, "Emotion classification based on gamma-band EEG," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 1223–1226.

[13] D. Zhang, L. Yao, X. Zhang, S. Wang, W. Chen, and R. Boots, "Cascade and parallel convolutional recurrent neural networks on EEG-based intention recognition for brain computer interface," in *Proc. 32nd AAAI Conf. Artif. Intell. (AAAI)*. AAAI Press, 2018, pp. 1703–1710.

[14] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial–temporal recurrent neural network for emotion recognition," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 839–847, Mar. 2019.

[15] T. Zhang, X. Wang, X. Xu, and C. L. P. Chen, "GCB-Net: Graph convolutional broad network and its application in emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 379–388, Jan. 2022.

[16] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1290–1301, Jul. 2022.

[17] T. Song, S. Liu, W. Zheng, Y. Zong, and Z. Cui, "Instance-adaptive graph for EEG emotion recognition," in *Proc. AAAI*, 2020, pp. 2701–2708.

[18] M. Ye, C. L. P. Chen, and T. Zhang, "Hierarchical dynamic graph convolutional network with interpretability for EEG-based emotion recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Dec. 9, 2022, doi: 10.1109/TNNLS.2022.3225855.

[19] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," 2017, *arXiv:1704.01212*.

[20] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, "Identifying stable patterns over time for emotion recognition from EEG," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 417–429, Jul. 2019.

[21] Z. Jia et al., "GraphSleepNet: Adaptive spatial–temporal graph convolutional networks for sleep stage classification," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, Jul. 2020, pp. 1324–1330.

[22] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "EmotionMeter: A multimodal framework for recognizing human emotions," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1110–1122, Mar. 2019.

[23] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.

[24] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

[25] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.

[26] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.

[27] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. Adv. Neural Inf. Process. Syst.*, 29, 2016, pp. 1–9.

[28] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.

[29] E. Halgren, K. Marinkovic, and P. Chauvel, "Generators of the late cognitive potentials in auditory and visual oddball tasks," *Electroencephalogr. Clin. Neurophysiol.*, vol. 106, no. 2, pp. 156–164, Feb. 1998.

[30] C. Grady, "The cognitive neuroscience of ageing," *Nature Rev. Neurosci.*, vol. 13, no. 7, pp. 491–505, Jul. 2012.

[31] D. H. J. Davis, S. T. Creavin, J. L. Y. Yip, A. Noel-Storr, C. Brayne, and S. Cullum, "Montreal cognitive assessment for the diagnosis of Alzheimer's disease and other dementias," *Cochrane Database Systematic Rev.*, vol. 10, 2015, Art. no. CD010775. [Online]. Available: https://api.semanticscholar.org/CorpusID:27232007

[32] E. L. Teng et al., "The cognitive abilities screening instrument (CASI): A practical test for cross-cultural epidemiological studies of dementia," *Int. Psychogeriatrics*, vol. 6, no. 1, pp. 45–58, Mar. 1994.

[33] J. C. Morris, "Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type," *Int. Psychogeriatrics*, vol. 9, no. S1, pp. 173–176, Dec. 1997.

[34] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Netw.*, vol. 13, nos. 4–5, pp. 411–430, Jun. 2000.

[35] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 327–339, Jul. 2014.

[36] Y. Li, W. Zheng, Y. Zong, Z. Cui, T. Zhang, and X. Zhou, "A bi-hemisphere domain adversarial neural network model for EEG emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 12, no. 2, pp. 494–504, Apr. 2021.

[37] W. L. Zheng and B. L. Lu, "Personalizing EEG-based affective models with transfer learning," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, 2016, pp. 2732–2738.

[38] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013.

[39] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, Jun. 1999.

[40] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Trans. Affect. Comput.*, vol. 11, no. 3, pp. 532–541, Jul. 2020.

[41] Y. Li et al., "GMSS: Graph-based multi-task self-supervised learning for EEG emotion recognition," *IEEE Trans. Affect. Comput.*, early access, Apr. 28, 2022, doi: 10.1109/TAFFC.2022.3170428.

[42] G. G. Knyazev, A. V. Bocharov, E. A. Levin, A. N. Savostyanov, and J. Y. Slobodskoj-Plusnin, "Anxiety and oscillatory responses to emotional facial expressions," *Brain Res.*, vol. 1227, pp. 174–188, Aug. 2008.

[43] M. Ertl, M. Hildebrandt, K. Ourina, G. Leicht, and C. Mulert, "Emotion regulation by cognitive reappraisal—The role of frontal theta oscillations," *NeuroImage*, vol. 81, pp. 412–421, Nov. 2013.

[44] J. Lesting, R. T. Narayanan, C. Kluge, S. Sangha, T. Seidenbecher, and H.-C. Pape, "Patterns of coupled theta activity in amygdala-hippocampal-prefrontal cortical circuits during fear extinction," *PLoS ONE*, vol. 6, no. 6, Jun. 2011, Art. no. e21714.

[45] M. Tortella-Feliu, A. Morillas-Romero, M. Balle, J. Llabrés, X. Bornas, and P. Putman, "Spontaneous EEG activity and spontaneous emotion regulation," *Int. J. Psychophysiol.*, vol. 94, no. 3, pp. 365–372, Dec. 2014.