

Instance-Adaptive Graph for EEG Emotion Recognition

Tengfei Song,^{1,2} Suyuan Liu,¹ Wenming Zheng,^{1*} Yuan Zong,¹ Zhen Cui³

¹Key Laboratory of Child Development and Learning Science, Ministry of Education, Southeast University, Nanjing, China

²School of Information Science and Engineering, Southeast University, Nanjing, China

³School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
{songtf, syl, wenming_zheng, xhzyuan}@seu.edu.cn; zhen.cui@njust.edu.cn

Abstract

To tackle the individual differences and characterize the dynamic relationships among different EEG regions for EEG emotion recognition, in this paper, we propose a novel instance-adaptive graph method (IAG), which employs a more flexible way to construct graphic connections so as to present different graphic representations determined by different input instances. To fit the different EEG pattern, we employ an additional branch to characterize the intrinsic dynamic relationships between different EEG channels. To give a more precise graphic representation, we design the multi-level and multi-graph convolutional operation and the graph coarsening. Furthermore, we present a type of sparse graphic representation to extract more discriminative features. Experiments on two widely-used EEG emotion recognition datasets are conducted to evaluate the proposed model and the experimental results show that our method achieves the state-of-the-art performance.

Introduction

Emotion recognition makes machine capture human emotional states, which is a crucial part in the research fields of man-machine interaction and artificial intelligence. As a hot topic in affective computing area, emotion recognition has recently caught more attentions. The responses of emotion can be characterized by behavioral and physiological signals. Compared to behavioral signals, such as facial expression and speech, physiological signals provide a more reliable way to identify different emotion states since they are difficult to be disguised (Liu, Sourina, and Nguyen 2011). Among various non-invasive physiological signals such as galvanic skin response, electrocardiogram, respiratory, blood pressure and electromyogram et al., electroencephalograph (EEG) can directly measure brain electrical activities, which contain richer information related to emotions. Therefore, the study on EEG signals can contribute to the revelation of human emotions.

For EEG signals, the data are distributed on irregular grid. Graph provides an effective data structure to characterize a

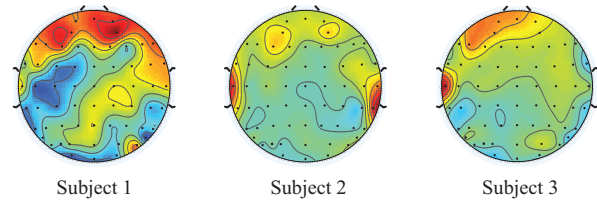


Figure 1: The individual energy distributions of three subjects in resting state.

set of objects and their relationships. Commonly, the graph connections, used in the classical graph CNNs (Bruna et al. 2013) and GCNN (Defferrard 2016), are mostly based on some natural connections or spatial positions between nodes. Apparently, these methods to construct graph connections are not suitable to model EEG signals, which contain complicated dynamic functional connections between different EEG regions. These functional connections of brain regions and their functional organizations are particularly important in the cognitive generation of emotions (Kober et al. 2008). Moreover, the imperceptible neuromechanism makes it hard to directly predefine the functional connections related to emotions. Song et.al (Song et al. 2018) proposed the dynamical graph convolutional neural network (DGCNN) to learn an optimal connection based on the distribution of EEG data. Li et.al (Li et al. 2018a) proposed the adaptive graph convolution neural network (AGCNN) to modify the graph connections in a small range. Nevertheless, these methods just provide the common graph connection or weakly modified graph connections. The studies (Lee and Hsieh 2014; Davidson et al. 1999) on human emotions have investigated the significance of individual differences, like the energy distributions shown in Figure 1, and different patterns of functional connectivity in terms of different emotion states, which should also be considered for EEG emotion recognition.

To address aforementioned issues, in this paper, we propose a novel instance-adaptive graph neural network (IAG) framework to model the relationships of different EEG channels in a more flexible way. Neuroscience research (LeDoux

*Corresponding author

Copyright © 2020, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

2000) has indicated emotion processing is the directed pattern. Specifically, the IAG model provides a directed graphic representation to construct the graph connections so as to reveal the intrinsic relationships between EEG regions. To alleviate the influence of individual differences and different patterns among various emotions, we try to design an instance-level graphic construction method. Motivated by attention mechanism, an additional branch is employed to adaptive generate adjacency matrices, i.e., the directed graphic connections. The generation of graphic connections is determined by input data, i.e., input instances, such that graphic connections can be adjusted dynamically according to different input instances. For this additional branch, we employ left multiplication matrix and right multiplication matrix to fuse the spatial and frequency information, respectively, which makes this structure more effective to represent the relationships between EEG channels. Considering that EEG signals from different frequencies may be represented by different topological relationships, we define a novel multi-level and multi-graph convolution method to model EEG signals. In terms of graph clustering, we define a type of pooling method to abstract EEG features according to spatial information rather than the distance between features. Besides, we also analyze the performance of sparse graphic connections on EEG emotion recognition.

In experiments, we evaluate the proposed method on two widely used EEG emotion datasets, i.e., SEED (Zheng and Lu 2015) and MPED (Song et al. 2019), and experimental results demonstrate that our method achieves the state-of-the-art performance. The contributions of the proposed IAG are summarized as below:

- **Instance-adaptive graphic connections.** A new flexible way to adaptive generate graphic connections is proposed so as to fit every instance and characterize the individual differences for EEG emotion recognition.
- **Graph convolution and graph coarsening.** A novel multi-level and multi-graph convolution method is proposed to model EEG features from different frequency bands and a spatial graph coarsening method is designed to abstract EEG signals.
- **Sparse graphic connections.** The l_1 -norm is employed to constrain graphic connections so as to evaluate the sparse graphic representation.

Related Work

Graph Convolutional Neural Network

Graph convolutional neural network is an extension of CNN to deal with irregular data. Particularly, graph convolution methods can be divided into two categories, i.e., spatial methods and spectral methods (Bruna et al. 2013). For spatial methods, sorting or aggregating neighbor nodes are generally presented to model spatial relationships. For instance, Niepert et.al proposed PSCN that sorted neighbor nodes and then convolution was employed to deal with the sorted nodes (Niepert, Ahmed, and Kutzkov 2016). Luo et.al transformed raw graphs into sequences by introducing the concept of n-gram block (Luo et al. 2017). In contrast, spectral methods

transform the signal into spectral domain through spectral graph theory (Sandryhaila and Moura 2014). Spectral methods often suffer high computational complexity due to the eigenvalue decomposition. The polynomial approximation may deal with this issue to some extent (Defferrard 2016). Many spectral methods aim to extract the local stationarity property of the signals around the central nodes, which is similar with CNNs. Li et.al proposed an adaptive graph convolutional neural networks so as to adjust graphic connections in a short distance (Li et al. 2018a). All these methods need to define the graph connections to extract the localized features. However, for some tasks, like EEG emotion recognition, it is hard to predefine the graphic connections according to the prior information and global relationships are crucial to be considered. How to design a more flexible graph convolution method is still an open issue.

EEG Emotion Recognition

Generally, EEG emotion recognition can be divided into two steps, i.e., feature extraction and classification. In feature extraction stage, we commonly distinguish features in time domain or frequency domain features. Especially, energy features from different frequency bands are the most popular for EEG emotion recognition. And then these features are processed by classification algorithms so as to predict various emotion states (Jenke, Peer, and Buss 2014).

With processed EEG features, many classification algorithms are introduced to effectively distinguish different emotion states. In (Zheng 2017), Zheng et.al proposed a group sparse canonical correlation analysis (GSCCA) method for EEG channel selection. Li et.al (Li et al. 2018c) proposed a graph regularized sparse linear regression (GRSLR) algorithm to deal with sparse transform matrix learning so as to improve EEG emotion classification accuracies. In (Zheng and Lu 2016), Zheng et.al proposed a transfer learning method to tackle personalizing EEG emotion recognition. Recently, more deep learning methods have been successfully applied for EEG emotion recognition with prominent performance. In (Zheng and Lu 2015), Zheng et.al employed Deep Belief Network (DBN) to extract the high-level information from EEG signals. In (Zhang et al. 2017), Zhang et.al proposed spatial-temporal recurrent neural network (STRNN) to investigate both spatial and temporal dependencies of EEG signals. In (Song et al. 2018), Song et.al proposed dynamical graph convolutional neural networks (DGCNN) for EEG emotion recognition, which aims to learn graphic connections between multiple EEG channels based on the training data so as to extract more discriminative features. In (Li et al. 2018b), Li et.al proposed a bi-hemispheres domain adversarial neural network (BiDANN) to narrow the distribution gap between training and testing data. In (Li et al. 2019), a hierarchical spatial-temporal neural network (R2G-STNN) is proposed to learn both regional and global spatial-temporal features for EEG emotion recognition. In contrast, in this paper, we most focus on a type of instance-adaptive global connection to model EEG signals.

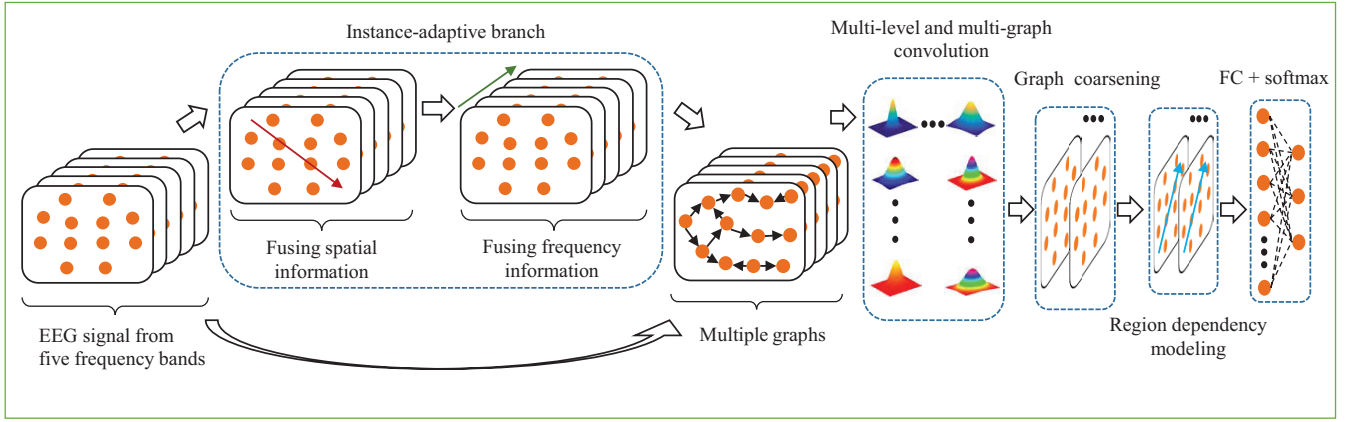


Figure 2: The framework of the proposed IAG. An instance-adaptive branch is provided to achieve the dynamic graphic connections. EEG signals are processed by multi-level and multi-graph convolution, graph coarsening, region dependency modeling, full connection layer (FC) and softmax layer.

Method

Overview

To deal with EEG emotion recognition, we need to model the dynamic relationships between different EEG channels and distinguish different emotion states. The framework of the proposed IAG is shown in Fig. 2. To better characterize the dynamic graphic connections of EEG signals, we propose an additional branch to generate graphic connections, which are adaptively changed along with input instance. With the generated graphic connections, EEG data is processed by multi-level and multi-graph convolution to diffuse the information among different channels. To achieve more robust features and reduce computational complexity, we conduct the graph coarsening to abstract them, which generates the features related to specific regions. To building the dependency of these regions, an iterative model, i.e., long short-term memory networks (LSTM), is applied to model the nodes after graph coarsening. The role of LSTM is to model dependencies of the features having sequential structure in graph coarsening layer to an emotion-discriminative feature vector, which is helpful to provide a good performance on EEG emotion recognition. Finally, all the hidden states of LSTM are connected to a full connection layer and a softmax layer is applied to output the predicted labels.

Attribute Graph for EEG signals

After energy feature extraction from five frequency bands, i.e., δ band (1-4 Hz), θ band (4-8 Hz), α band (8-14 Hz), β band (14-30 Hz) and γ band (30-50 Hz), an EEG sample is represented by $\mathbf{X} \in \mathbb{R}^{n \times d}$, where n is the number of EEG channels (nodes) and d is the number of frequency bands. To model this EEG signal, we employ a directed attribute graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{A}\}$ of n nodes, in which $\mathcal{V} = \{v_i\}_{i=1}^n$ represents the set of nodes, \mathcal{E} denotes the set of edges between these nodes and $\mathbf{A} \in \mathbb{R}^{n \times n}$ is an adjacency matrix. The adjacency matrix \mathbf{A} characterizes the connections between nodes. If source node v_i and destination node v_j are not connected, then $A(i, j) = 0$, otherwise $A(i, j) \neq 0$.

According to graphic signal processing theory (Sandryhaila and Moura 2014), the Jordan decomposition of the graph adjacency matrix \mathbf{A} can be defined as $\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}$, in which \mathbf{J} is a block-diagonal matrix and $\mathbf{F} = \mathbf{V}^{-1}$ is the graph Fourier Transform matrix. The EEG data can be transformed into frequency domain by $\hat{\mathbf{X}} = \mathbf{F}\mathbf{X}$ and the inverse graph Fourier transform is given by $\mathbf{X} = \mathbf{F}^{-1}\hat{\mathbf{X}}$. Given a filtering function $h(\cdot)$, the filtering process can be characterized as

$$\tilde{\mathbf{X}} = h(\mathbf{A})\mathbf{X} = \mathbf{F}^{-1}h(\mathbf{J})\mathbf{F}\mathbf{X}, \quad (1)$$

in which $h(\mathbf{J})$ is the graph frequency response of the filter $h(\mathbf{A})$, i.e., $h(\hat{\mathbf{A}}) = h(\mathbf{J})$. The convolution theorem from classical signal processing has been extended to graphic frequency representation.

Instance-Adaptive Graph Connections

To characterize the individual differences and the dynamic functional relationships between EEG regions plays an important role in distinguishing different emotion states. Motivated by design of attention framework, we employ an additional branch to adaptively construct multiple graphic connections such that the graphic connections are adaptively changed along with input instance during the training and classification processes. Especially, this instance-adaptive branch aims to get more effective graph connections by fusing spatial and frequency information, which is shown in Figure 2.

The t th column of the EEG data \mathbf{X} is denoted by \mathbf{X}_t , which represents EEG features in t th frequency band. First, we conduct the left multiplication projection, which can be expressed as this form:

$$\mathbf{O}_t = \mathbf{P}\mathbf{X}_t + \mathbf{B}_t, \quad (t = 1, 2, \dots, d) \quad (2)$$

in which $\mathbf{P} \in \mathbb{R}^{n \times n}$ is the left multiplication matrix to fuse the spatial information, i.e., various EEG channels,

$\mathbf{O} \in \mathbb{R}^{n \times d}$ is the output of the left multiplication projection and \mathbf{B} is the bias matrix.

The relationship between different frequency bands is also a key point to be considered. To fuse the emotional information between different frequency bands, we conduct the right multiplication projection, which can be expressed by following equation:

$$\mathbf{G} = \text{Relu}(\mathbf{O}\mathbf{Q}\Theta), \quad (3)$$

in which $\mathbf{Q} \in \mathbb{R}^{d \times d}$ is the right multiplication matrix to fuse frequency information, $\Theta \in \mathbb{R}^{d \times nd}$ is the projection matrix and $\mathbf{G} \in \mathbb{R}^{n \times nd}$ is the output. *Relu* function is applied to guarantee elements non-negative and \mathbf{G} is reshaped into d adjacency matrices, i.e., $[\mathbf{A}_1^*, \dots, \mathbf{A}_d^*]$, to represent graphs in d frequency bands.

The whole projection process can be expressed as

$$\mathbf{G} = f(\mathbf{X}, \mathbf{P}, \mathbf{B}, \mathbf{Q}, \Theta), \quad (4)$$

in which $\mathbf{P}, \mathbf{B}, \mathbf{Q}$ and Θ are the parameters to be solved. The output \mathbf{G} of left and right multiplication projection is adaptively changed along with the input instance \mathbf{X} . That is, this branch will adaptively adjust graph connections to fit every input instance, which provides a flexible way to represent the intrinsic relationships between EEG regions. To achieve the normalized version, each element of adjacency matrix A_{ij} is multiplied by $\frac{1}{\sqrt{D_{ii}D_{jj}}}$, formally, $\mathbf{A}^{norm} = \mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}$, in which \mathbf{D} is a diagonal matrix calculated by $D_{ii} = \sum_j A_{ij}$. Unless otherwise specified, we use the normalized \mathbf{A} below.

Multi-Level and Multi-Graph Convolutional Kernels

In the standard CNN, convolutional operation is repeated to extract high-level features with a local square spatial kernels. For graphs, it is hard to construct the convolutional kernels, due to the irregular structure. Besides, the graph filtering responses need homogeneous graph structure. To deal with this issues, we employ adjacency matrix \mathbf{A} to model the connections between different nodes, which is based on the graph filtering theories (Sandryhaila and Moura 2014). Similarly, graph convolutional operation can be repeated to model a high-level connections between nodes. \mathbf{A}^k expresses the connections after k -step graph convolution operations. To consider the graphs with different levels, we can model these graphs by a polynomial of \mathbf{A} . Here, we denote the k -order polynomial as $\varphi_k(\mathbf{A}) = \mathbf{A}^k$. Thus, we define the multi-level graph convolution as

$$\mathbf{Y} = \mathcal{G} * \mathcal{F} = \sum_{k=0}^{K-1} \varphi_k(\mathbf{A})\mathbf{X}, \quad (5)$$

in which $\varphi_k(\mathbf{A})$ is the k -th level graph. This process provides a potential way to consider the information from different levels.

Particularly, there are different relationships between human emotions and EEG signals from different frequency bands. Inspired by this consideration, we employ multiple graphs to model EEG features from different frequency

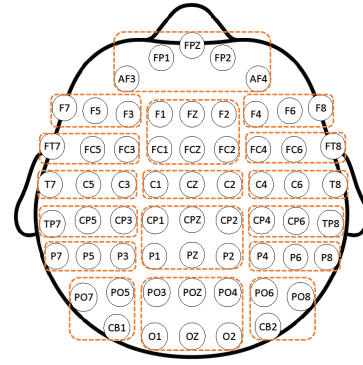


Figure 3: The region partition for spatial graph coarsening. The 62 EEG channels in the international 10-20 system are divided into 17 groups and the EEG nodes of each group are clustered into one node.

bands respectively, via d adjacency matrices $[\mathbf{A}_1^*, \dots, \mathbf{A}_d^*]$, which give a more specific graphic representation. Therefore, we model EEG data from different frequency bands as

$$\mathbf{Y} = \text{Cat}[\sum_{k=0}^{K-1} \varphi_k(\mathbf{A}_1^*)\mathbf{X}_1, \dots, \sum_{k=0}^{K-1} \varphi_k(\mathbf{A}_d^*)\mathbf{X}_d]\mathbf{U}, \quad (6)$$

in which $\text{Cat}[\]$ is a concatenation operator, $\mathbf{U} \in \mathbb{R}^{d \times d'}$ is the dimension transformation matrix to be solved and $\mathbf{Y} \in \mathbb{R}^{n \times d'}$ is the output. This convolution operation provides a multi-level and multi-graph representation, which diffuses the information from different channels so as to extract more discriminative features.

Graph Coarsening

Similar with the standard pooling operation in CNNs, down-sampling graphs is helpful to abstract them and reduce computational complexity. For graphs, the pooling operation relies on the meaningful relationships between nodes, which can be considered according to spatial locations or the distance of features between different nodes. For EEG signals, clustering these nodes with close features may destroy dynamic characteristics, which will lose useful information related to emotions. Therefore, we conduct the pooling operation based on spatial locations of EEG electrodes in case to lose important dynamic information.

With an EEG sample as example, the graph after convolution operation is divided in to n' groups and clustered into n' nodes. The neighbor nodes are fused and the pooling operation on graphs can be formulated as

$$\mathbf{Z}_p = \frac{1}{n_{te} - n_{ts}} \sum_{l=n_{ts}}^{n_{te}} \mathbf{Y}_l, \quad (p = 1, \dots, n') \quad (7)$$

where the p -th group contains nodes from number n_{ts} to number n_{te} and \mathbf{Z} denotes the nodes after graph coarsening. Actually, this operation is to calculate the average values of neighbor nodes, which can be regarded as the specific region. The detailed location division for EEG channels clustering is displayed in Fig. 3.

Sparse Graphic Representation

During the training process, we employ the cross entropy loss function to measure the dissimilarity between predicted labels and real labels. Specifically, we investigate whether sparse graphic connections between EEG channels are helpful to recognize various emotion states. Therefore, we define the loss function as the following form:

$$loss = cross_entropy(\mathbf{l}, \mathbf{l}^p) + \sum_{j=1}^d \alpha_j \|\mathbf{A}_j^*\|_1, \quad (8)$$

in which \mathbf{l} and \mathbf{l}^p represent the real label vector and the predicted one, respectively, α is the trade-off parameter and $\|\cdot\|_1$ denotes the l_1 -norm. We introduce l_1 -norm to constrain the adjacency matrices $[\mathbf{A}_1^*, \dots, \mathbf{A}_d^*]$, which presents a type of sparse graphic representation. The sparsity of graph connection is controlled by α .

Experiment

Datasets and Settings

SEED The SJTU Emotion EEG Database (SEED) recorded 15 subjects' EEG data (7 males and 8 females), when they were watching 15 Chinese clip videos. Each video lasts for about 4 minutes. These videos are applied to elicit three types of emotions, i.e., neutrality, positivity and negativity. For each subject, this recording process is repeated in 3 different periods corresponding to 3 sessions and each session contains 15 trials of EEG signals. All EEG signals are divided into 1-second samples for classification. To compare with previous literatures, we obey same subject-dependent and subject-independent protocols strictly (Zheng and Lu 2015; 2016). For subject-dependent protocol, we use the first 9 trials of EEG data as training data and remaining 6 ones as testing data, and then the mean accuracy of 15 subjects is evaluated. For subject-independent protocol, we employ the leave-one-subject-out cross validation strategy.

MPED The Multi-Modal Physiological Emotion Database (MPED) contains four types of physiological signals of 23 subjects (10 males and 13 females), which are recorded when they are watching 28 film clips with seven types of emotions, i.e., joy, funny, anger, fear, disgust, sadness and neutrality. Therefore, there are 28 trials of EEG data for each subject. Here, we only use EEG data for emotion recognition. Similarly, all EEG signals are divided into 1-second samples. 21 trials of EEG data are served as training data and the rest 7 trials of EEG data are served as testing data. We follow the protocols in Song et al. strictly (Song et al. 2019). *Protocol one*: Eight types of combinations in form of positive-neutral-negative are conducted and the average accuracy of these combination are calculated for comparison. *Protocol two*: Seven emotions are divided into three categories, i.e., positive, neutral and negative emotions, for classification. Joy and funny are classified into positive emotion, while angry, sadness, fear and disgust are classified into negative emotion. *Protocol three*: Seven types of emotions are presented for classification.

Implementation Details

For the input of our model, we employ the extracted EEG features, i.e., differential entropy (DE) for SEED dataset and short-time Fourier transform (STFT) spectrum for MPED dataset, which are consistent with former studies.

For the proposed graph convolution part, the number of EEG channels (n) is set to 62, the number of frequency bands (d) is set to 5, the order of graph convolution (K) is set to 8 and the transformed dimension (d') is set to 32. For graph coarsening, original 62 nodes are clustered into 17 nodes. The dimensions of hidden state and memory cell in LSTM are both set to 64. In our loss function, the trade-off parameters, i.e., $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and α_5 are set to $10^{-4}, 10^{-5}, 10^{-5}, 10^{-5}$ and 10^{-5} , respectively. The learning rate is set to 0.001. In our system, whole model is implemented by TensorFlow.

Experiment Results

Verification of IAG structure To validate the efficiency of our IAG model, we conduct extensive experiments using different structures on two published EEG discrete emotion datasets.

As shown in Table 1, we present the results using our IAG model without instance-adaptive branch (IA) and multi-level and multi-graph convolution (MMG), respectively. Without IA structure, we can see that the classification accuracies for EEG emotion recognition decrease a lot, especially in subject-independent protocol on SEED and protocol three on MPED. Without MMG structure, the model also achieves low classification accuracies. Specifically, we can see that both IA and MMG are important to improve the classification accuracies for EEG emotion recognition. IA structure provides a more flexible way to represent the intrinsic connections among different EEG channels. Besides, MMG represents EEG features from different frequency bands with multiple graphs and explores a high-level relationships, which is helpful to extract discriminative features for EEG emotion recognition.

To investigate the efficiency of spatial graph pooling, we conduct extensive experiments using three types of structures, i.e., IAG without pooling, IAG with spatial pooling and IAG with k-means pooling, which are shown in Table 1. For our spatial graph pooling, we divide the EEG channels into 17 spatial regions, which are shown in Fig. 3. For k-means graph pooling, certain centers (FPz, F5, Fz, F6, FC5, FC6, C5, Cz, C6, CP5, CPz, CP6, P5, P6, PO5, POz, PO6) are given in the initial stage. Specifically, we can see that spatial graph pooling is more advantageous to abstract useful information than k-means graph pooling. The k-means graph pooling is based on the distance of features between different nodes, which may destroy the dynamic information of EEG signals and induce the low classification results.

Additionally, we also explore the performance of sparse graphic representation by constraining five adjacency matrices, i.e., $[\mathbf{A}_1^*, \dots, \mathbf{A}_d^*]$. The trade-off parameters, i.e., $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, and α_5 , are set to $10^{-4}, 10^{-5}, 10^{-5}, 10^{-5}$ and 10^{-5} , respectively. Most existing works indicate that low frequency band contains lower discriminative information for EEG emotion recognition. So we give a large

Table 1: The mean accuracies (ACC) and standard deviations (STD) on SEED dataset and MPED dataset.

Method	SEED		MPED		
	Subject-dependent ACC / STD (%)	Subject-independent ACC / STD (%)	Protocol one ACC / STD (%)	Protocol two* ACC / F1 (%)	Protocol three ACC / STD (%)
IAG (w/o IA, pooling)	90.25 / 08.46	81.61 / 08.57	72.31 / 12.51	69.72 / 64.64	36.31 / 09.84
IAG (w/o MMG, pooling)	91.42 / 07.34	82.10 / 09.31	72.82 / 12.00	69.88 / 63.83	37.20 / 10.05
IAG (w/o pooling)	93.87 / 05.95	85.01 / 08.24	73.11 / 12.34	72.33 / 67.18	40.09 / 10.23
IAG (k-means pooling)	86.39 / 09.45	74.36 / 11.11	67.34 / 12.35	68.22 / 51.79	35.69 / 09.37
IAG (spatial pooling)	94.89 / 06.16	85.24 / 06.86	73.95 / 11.34	72.76 / 66.76	39.10 / 09.65
IAG + sparse	95.44 / 05.48	86.30 / 06.91	74.77 / 10.75	73.58 / 68.41	40.38 / 08.75

*Protocol two in MPED is designed for EEG emotion recognition with unbalanced data such that ACC and F1 score are suggested performance metrics. ‘w/o’ denotes ‘without’.

weight to make it sparser to suppress the useless information. From the result shown in Table 1, sparse graphic representation is effective to improve the classification accuracies for EEG emotion recognitions. The sparse graphic representation captures useful graph connections and ignores redundant relationships so as to extract more discriminative features.

Table 2: The mean accuracies (ACC) and standard deviations (STD) on SEED database for subject-dependent EEG emotion recognition experiment.

Method	ACC/STD(%)
SVM (Suykens and Vandewalle 1999)	83.99 / 09.72
CCA (Thompson 2005)	77.63 / 13.21
DBN (Zheng and Lu 2015)	86.08 / 08.34
GCNN (Defferrard 2016)	87.40 / 09.20
DANN (Ganin et al. 2016)	91.36 / 08.30
GRSLR (Li et al. 2018c)	87.39 / 08.64
DGCNN (Song et al. 2018)	90.40 / 08.49
BiDANN (Li et al. 2018b)	92.38 / 07.04
R2G-STNN (Li et al. 2019)	93.38 / 05.96
IAG	95.44 / 05.48

Comparison with the state-of-the-art method To further validate the proposed IAG, we compared our model with the start-of-the-art methods on SEED and MPED, respectively, including linear support vector machine (SVM), canonical correlation analysis (CCA), deep believe network (DBN), graph convolutional neural network (GCNN), domain adversarial neural network (DANN), graph regularization sparse linear regression (GRSLR), dynamical graph convolutional neural network (DGCNN), kernel principal component analysis (KPCA), transfer component analysis (TCA), transductive parameter transfer (TPT), bi-hemispheres domain adversarial neural network (BiDANN), BiDANN-S, spatial-temporal recurrent neural network (STRNN), a hierarchical spatial-temporal neural network (R2G-STNN) and attention long-short time memory networks (A-LSTM).

From the results of EEG emotion recognition summarized in Table 2, Table 3 and Table 4, we have the following observation.

- The proposed IAG is superior to the recent graph-based

Table 3: The mean accuracies (ACC) and standard deviations (STD) on SEED dataset for subject-independent EEG emotion recognition experiment.

Method	ACC/STD(%)
SVM (Suykens and Vandewalle 1999)	56.73 / 16.29
KPCA (Schölkopf and Müller 1998)	61.28 / 14.62
TCA (Pan et al. 2011)	63.64 / 14.88
TPT (Sangineto et al. 2014)	76.31 / 15.89
DANN (Ganin et al. 2016)	75.08 / 11.18
DGCNN (Song et al. 2018)	79.95 / 09.02
BiDANN (Li et al. 2018b)	83.28 / 09.60
BiDANN-S (Li et al. 2018d)	84.14 / 06.87
R2G-STNN (Li et al. 2019)	84.16 / 07.63
IAG	86.30 / 06.91

methods, domain adversarial methods and LSTM-based methods. From the results shown in Table 2, our IAG achieves better classification result, which is 5.04% and 8.04% higher than graph-based methods, i.e., DGCNN and GCNN, respectively. Although they are all graph-based methods, our self-adaptive structure is more effective to characterize the intrinsic relationships between different EEG channels. Also, IAG has an improvement of 3.06% in contrast to BiDANN, which combines domain adversarial structure with LSTM. In both Table 3 and Table 4, our IAG achieves better performance, which is superior to these LSTM-based methods, i.e., STRNN, LSTM and A-LSTM.

- Our IAG improves the current state-of-the-art results on both SEED and MPED. On SEED dataset, our IAG achieves 95.44% in subject-dependent protocol and 86.30% in subject-independent protocol. On MPED dataset, the proposed IAG has improved the best performance to 74.77%, 73.58% and 40.38% in protocol one, protocol two and protocol three, respectively.
- Different datasets have different performances. We can see that there are different performances on EEG emotion recognition using different datasets or in different classification protocols. For SEED dataset, there are three emotional categories, which are organized by subject-dependent and subject-independent protocols. In subject-

Table 4: Experiment results(%) in protocol one and protocol three (average accuracies and standard deviations), as well as in protocol two (average accuracies and F1 scores) on MPED database.

Method	Protocol one (ACC/STD)	Protocol two (ACC/F1)	Protocol three (ACC/STD)
SVM (Suykens and Vandewalle 1999)	59.86 / 16.29	57.06 / 24.43	31.14 / 08.06
DBN (Zheng and Lu 2015)	65.83 / 13.20	65.98 / 59.19	29.26 / 09.19
LSTM (Sak, Senior, and Beaufays 2014)	72.09 / 14.94	71.92 / 65.12	38.55 / 08.43
STRNN (Zhang et al. 2017)	65.38 / 13.20	66.84 / 60.57	35.64 / 09.57
DGCNN (Song et al. 2018)	71.13 / 15.77	68.02 / 61.11	36.92 / 12.78
A-LSTM (Song et al. 2019)	72.93 / 13.19	71.57 / 67.74	38.74 / 07.75
IAG	74.77 / 10.75	73.58 / 68.41	40.38 / 08.75

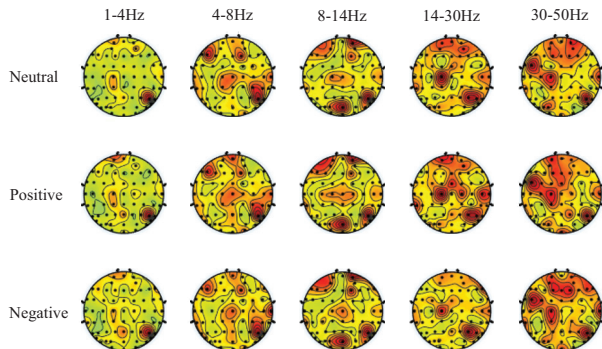


Figure 4: The visualization of degree centrality of EEG electrodes in our IAG model on SEED dataset.

dependent protocol, most methods achieve higher accuracies. In contrast, the classification accuracies in subject-independent protocol will be lower, due to the individual difference. For MPED dataset, there are seven emotion categories. The unbalanced distribution of training data and more categories are the key points to induce lower classification accuracies.

The visualization of degree centrality in the graph for EEG emotion recognition

The degree centrality is a validated index measuring connectivity of a node with the other nodes, which has been widely used to evaluate the importance of the nodes in the graph (Zhang, Cheng, and Qu 2007). To further evaluate the importance of nodes sub-serving the emotion recognition of our IAG, in this section, we visualize the degree centrality of each scalp EEG electrode based on graphic connections. In the graphic model, the adjacency matrix \mathbf{A} characterizes the connections between nodes. In an adjacency matrix, values of the i -th row and the i -th column represent the graphic weights connected with the i -th nodes. The degree centrality C_i of the i -th EEG electrode can be calculated by

$$C_i = \sum_{n=1}^{62} \mathbf{A}_{i,n} + \sum_{m=1}^{62} \mathbf{A}_{m,i} - 2\mathbf{A}_{i,i}, (i = 1, \dots, 62). \quad (9)$$

For SEED dataset, Fig. 4 shows the degree centrality for different EEG electrodes based on our IAG model. Due to

the multi-graph module in IAG, we can see that there are different graphic connection patterns for different frequency bands. Particularly, more EEG electrodes in high frequency bands are connected with each other, also with stronger connectivity than that in low frequency bands. These results indicate that the signals in high frequency bands are more informative for emotion recognition. Notably, this conclusion is highly consistent with former studies (Hadjidimitriou and Hadjileontiadis 2012). Across emotion states, however, we observed similar scalp distribution of degree centrality. That is to say, although our model can adaptively change the graphic connections, our IAG model is effective to capture the EEG electrodes of most importance, which are commonly shared for emotion recognition across different emotion states.

Conclusion

In this paper, we proposed a graph-based method, called IAG, to tackle individual differences and model the relationship among different EEG regions for EEG emotion recognition. The graph connections in our IAG is self-adaptive along with input EEG data so as to characterize the individual differences. With the generated graphs, the multi-level and multi-graph convolutional operation and graph coarsening are applied to extract more discriminative features for classification. In addition, sparse graphic representation is helpful for EEG emotion recognition to some extent. The experiment results have validated the efficiency of our IAG, which is superior to other deep learning methods and achieves the state-of-the-art performance.

Acknowledgments

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1305200, in part by the National Natural Science Foundation of China under Grant 61921004, Grant 61572009, Grant 61902064, and Grant 81971282, in part by the Fundamental Research Funds for the Central Universities under Grant 2242018K3DN01 and Grant 2242019K40047 and in part by the Scientific Research Foundation of Graduate School of Southeast University under Grant YBPY1941.

References

- Bruna, J.; Zaremba, W.; Szlam, A.; and LeCun, Y. 2013. Spectral networks and locally connected networks on graphs. *arXiv preprint arXiv:1312.6203*.
- Davidson, R. J.; Abercrombie, H.; Nitschke, J. B.; and Putnam, K. 1999. Regional brain function, emotion and disorders of emotion. *Current opinion in neurobiology* 9(2):228–234.
- Defferrard, Michaël ; Bresson, X. V. P. 2016. Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in Neural Information Processing Systems*, 3844–3852.
- Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; and Lempitsky, V. 2016. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research* 17(1):2096–2030.
- Hadjidimitriou, S. K., and Hadjileontiadis, L. J. 2012. Toward an eeg-based recognition of music liking using time-frequency analysis. *IEEE Transactions on Biomedical Engineering* 59(12):3498–3510.
- Jenke, R.; Peer, A.; and Buss, M. 2014. Feature extraction and selection for emotion recognition from eeg. *IEEE Transactions on Affective Computing* 5(3):327–339.
- Kober, H.; Barrett, L. F.; Joseph, J.; Bliss-Moreau, E.; Lindquist, K.; and Wager, T. D. 2008. Functional grouping and cortical–subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *Neuroimage* 42(2):998–1031.
- LeDoux, J. E. 2000. Emotion circuits in the brain. *Annual review of neuroscience* 23(1):155–184.
- Lee, Y.-Y., and Hsieh, S. 2014. Classifying different emotional states by means of eeg-based functional connectivity patterns. *PloS one* 9(4):e95415.
- Li, R.; Wang, S.; Zhu, F.; and Huang, J. 2018a. Adaptive graph convolutional neural networks. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Li, Y.; Zheng, W.; Cui, Z.; Zhang, T.; and Zong, Y. 2018b. A novel neural network model based on cerebral hemispheric asymmetry for eeg emotion recognition. In *IJCAI*, 1561–1567.
- Li, Y.; Zheng, W.; Cui, Z.; Zong, Y.; and Ge, S. 2018c. Eeg emotion recognition based on graph regularized sparse linear regression. *Neural Processing Letters* 1–17.
- Li, Y.; Zheng, W.; Zong, Y.; Cui, Z.; Zhang, T.; and Zhou, X. 2018d. A bi-hemisphere domain adversarial neural network model for eeg emotion recognition. *IEEE Transactions on Affective Computing*.
- Li, Y.; Zheng, W.; Wang, L.; Zong, Y.; and Cui, Z. 2019. From regional to global brain: A novel hierarchical spatial-temporal neural network model for eeg emotion recognition. *IEEE Transactions on Affective Computing*.
- Liu, Y.; Sourina, O.; and Nguyen, M. K. 2011. Real-time eeg-based emotion recognition and its applications. In *Transactions on computational science XII*. Springer. 256–277.
- Luo, Z.; Liu, L.; Yin, J.; Li, Y.; and Wu, Z. 2017. Deep learning of graphs with ngram convolutional neural networks. *IEEE Transactions on Knowledge and Data Engineering* 29(10):2125–2139.
- Niepert, M.; Ahmed, M.; and Kutzkov, K. 2016. Learning convolutional neural networks for graphs. In *International conference on machine learning*, 2014–2023.
- Pan, S. J.; Tsang, I. W.; Kwok, J. T.; and Yang, Q. 2011. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks* 22(2):199–210.
- Sak, H.; Senior, A.; and Beaufays, F. 2014. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In *Fifteenth annual conference of the international speech communication association*.
- Sandryhaila, A., and Moura, J. M. 2014. Discrete signal processing on graphs: Frequency analysis. *IEEE Transactions on Signal Processing* 62(12):3042–3054.
- Sanginetto, E.; Zen, G.; Ricci, E.; and Sebe, N. 2014. We are not all equal: Personalizing models for facial expression analysis with transductive parameter transfer. In *Proceedings of the 22nd ACM international conference on Multimedia*, 357–366. ACM.
- Schölkopf, Bernhard; Smola, A., and Müller, K.-R. 1998. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation* 10(5):1299–1319.
- Song, T.; Zheng, W.; Song, P.; and Cui, Z. 2018. Eeg emotion recognition using dynamical graph convolutional neural networks. *IEEE Transactions on Affective Computing*.
- Song, T.; Zheng, W.; Lu, C.; Zong, Y.; Zhang, X.; and Cui, Z. 2019. Mped: A multi-modal physiological emotion database for discrete emotion recognition. *IEEE Access* 7:12177–12191.
- Suykens, J. A., and Vandewalle, J. 1999. Least squares support vector machine classifiers. *Neural processing letters* 9(3):293–300.
- Thompson, B. 2005. Canonical correlation analysis. *Encyclopedia of statistics in behavioral science*.
- Zhang, T.; Zheng, W.; Cui, Z.; Zong, Y.; and Li, Y. 2017. Spatial-temporal recurrent neural network for emotion recognition. *IEEE Transactions on Cybernetics* PP(99):1–9.
- Zhang, X.; Cheng, G.; and Qu, Y. 2007. Ontology summarization based on rdf sentence graph. In *Proceedings of the 16th international conference on World Wide Web*, 707–716. ACM.
- Zheng, W.-L., and Lu, B.-L. 2015. Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development* 7(3):162–175.
- Zheng, W.-L., and Lu, B.-L. 2016. Personalizing eeg-based affective models with transfer learning. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 2732–2738. AAAI Press.
- Zheng, W. 2017. Multichannel eeg-based emotion recognition via group sparse canonical correlation analysis. *IEEE Transactions on Cognitive and Developmental Systems* 9(3):281–290.