

EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks

Tengfei Song^{ID}, Wenming Zheng^{ID}, *Member, IEEE*, Peng Song^{ID}, *Member, IEEE*, and Zhen Cui^{ID}

Abstract—In this paper, a multichannel EEG emotion recognition method based on a novel dynamical graph convolutional neural networks (DGCNN) is proposed. The basic idea of the proposed EEG emotion recognition method is to use a graph to model the multichannel EEG features and then perform EEG emotion classification based on this model. Different from the traditional graph convolutional neural networks (GCNN) methods, the proposed DGCNN method can dynamically learn the intrinsic relationship between different electroencephalogram (EEG) channels, represented by an adjacency matrix, via training a neural network so as to benefit for more discriminative EEG feature extraction. Then, the learned adjacency matrix is used to learn more discriminative features for improving the EEG emotion recognition. We conduct extensive experiments on the SJTU emotion EEG dataset (SEED) and DREAMER dataset. The experimental results demonstrate that the proposed method achieves better recognition performance than the state-of-the-art methods, in which the average recognition accuracy of 90.4 percent is achieved for subject dependent experiment while 79.95 percent for subject independent cross-validation one on the SEED database, and the average accuracies of 86.23, 84.54 and 85.02 percent are respectively obtained for valence, arousal and dominance classifications on the DREAMER database.

Index Terms—EEG emotion recognition, adjacency matrix, graph convolutional neural networks (GCNN), dynamical convolutional neural networks (DGCNN)

1 INTRODUCTION

EMOTION recognition plays an important role in the human-machine interaction [1], which enables machine to perceive the emotional states of human beings so as to make machine more ‘sympathetic’ in the human-machine interaction. Basically, emotion recognition methods can be divided into two categories. The first one is based on non-physiological signals, such as facial expression images [2], [3], [4], [5], [6], body gesture [7], and voice signal [8]. The second one is based on physiological signal, such as electroencephalogram (EEG) [9], electromyogram (EMG) [10], and electrocardiogram (ECG) [11]. Among the various types of physiological signals, EEG signal is one of the most commonly used ones, which is directly captured from the brain cortex and hence it would be advantageous to reflect the mental states of human beings. With the rapid development of dry EEG electrode techniques and the EEG

signal processing methods, EEG emotion recognition has received more and more attentions in recent years [12], [13], [14], [15].

Basically, there are two major ways to describe human’s emotions [16], i.e., the discrete basic emotion description approach and the dimension approach. For the discrete basic emotion description approach, the emotions are classified into a set of discrete status, e.g., the six basic emotions (i.e., joy, sadness, surprise, fear, anger, and disgust) [17]. Different from the discrete emotion description approach, the dimension approach describes emotions in continuous form, in which the emotions are characterized by three dimensions (valence, arousal and dominance) [18], [19] or simply two dimensions (valence and arousal), in which the valence dimension mainly characterizes how positive or negative the emotions are, whereas the arousal dimension aims to characterize the degree of how excited or apathetic the emotions are [16].

The research of applying EEG signal to the emotion recognition can be traced back to work of Musha et al. in [20]. During the past decades, many machine learning and signal processing methods are proposed to deal with the EEG emotion recognition [21], [22]. A typical EEG emotion recognition method usually consists of two major parts, i.e., discriminative EEG feature extraction and emotion classification. Basically, the EEG features used for emotion recognition can be generally divided into two kinds, i.e., time-domain feature type and frequency-domain feature type. The time domain features, e.g., Hjorth feature [23], fractal dimension feature [24] and higher order crossing feature [25], mainly capture the temporal information of EEG signals. Different from the time-domain feature, the frequency-domain feature aims to

- T. Song is with the Key Laboratory of Child Development and Learning Science, Ministry of Education, and School of Information Science and Engineering, Southeast University, Nanjing 210096, China. E-mail: songtf@seu.edu.cn.
- W. Zheng is with the Key Laboratory of Child Development and Learning Science, Ministry of Education, and School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China. E-mail: wenming_zheng@seu.edu.cn.
- P. Song is with the School of Computer and Control Engineering, Yantai University, Yantai 264005, China. E-mail: pengsongseu@gmail.com.
- Z. Cui is with the School of Computer Science, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, P.R. China. E-mail: zhen.cui@njjust.edu.cn.

Manuscript received 10 Oct. 2017; revised 28 Jan. 2018; accepted 2 Mar. 2018.

Date of publication 21 Mar. 2018; date of current version 4 Sept. 2020.

(Corresponding author: Wenming Zheng.)

Recommended for acceptance by B. Hu.

Digital Object Identifier no. 10.1109/TAFFC.2018.2817622

capture the EEG emotion information from the frequency point of view. One of the most commonly used frequency-domain feature extraction methods is to decompose the EEG signal into several 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) [21], [26], [27], [28], and then extract EEG features from each frequency band, respectively. The commonly used EEG features include the differential entropy (DE) feature [29], [30], the power spectral density (PSD) feature [31], the differential asymmetry (DASM) feature [24], the rational asymmetry (RASM) feature [32] and the differential caudality (DCAU) feature [19].

To deal with EEG emotion classification problem, there are many methods appeared in the literatures [33], among which the method of using deep neural networks (DNN) [19] had been demonstrated to be one of the most successful one. Convolutional neural networks (CNN) is one of the most famous DNN approaches and had been widely used to cope with various classification problems, such as image classification [34], [35], [36], [37], object detection [38], tracking [39] and segmentation [40]. Although CNN model had been demonstrated to be very powerful in dealing with the classification problems, it is notable that previous applications of CNN focus more on the local feature learning from image, video and speech, in which the data points of the signal are continuously changed. For another feature learning problems, such as the feature learning from transportation network and brain network, the traditional CNN method may not be well suitable because the signals are discrete and discontinuous in the spatial domain. In this case, graph based description methods [41], [42] would provide a more appropriate way.

Graph neural network (GNN) [43] aims to build the neural networks under the graph theory to cope with the data in graph domain. Graph convolutional neural network (GCNN) [44] is an extension of the traditional CNN method by combining CNN with spectral theory [45]. Compared with classical CNN method, GCNN would be more advantageous in dealing with the discriminative feature extraction of signals in the discrete spatial domain [46]. More importantly, the GCNN method provides an effective way to describe the intrinsic relationship between different nodes of the graph, which would provide a potential way to explore the relationships among the multiple EEG channels during the EEG emotion recognition.

Motivated by the success of the GCNN model, in this paper we will investigate the multichannel EEG emotion recognition problem using graph representation approach, in which each EEG channel corresponds to a vertex node whereas the connection between two different vertex nodes corresponds to an edge of the graph. Although GCNN can be used to describe the connections among different nodes according to their spatial positions, we should predetermine the connections among the various EEG channels before applying it to building the emotion recognition model. On the other hand, it is notable that the spatial position connections among the EEG channels are different from the functional connections among them. In other words, a closer spatial relationship may not guarantee a closer functional relationship, whereas the functional relationship would be useful for the discriminative EEG feature extraction in

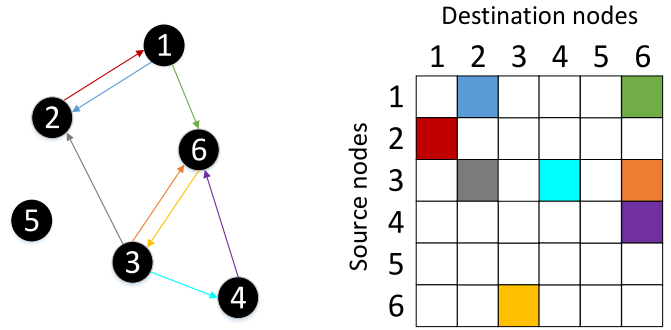


Fig. 1. Example of a directed graph and the corresponding adjacency matrix, where the left part is the connections of six nodes and right part is the adjacency matrix.

emotion recognition. Consequently, it is not reasonable to predetermine the connections of the graph nodes according to their spatial positions.

To alleviate the limitations of the GCNN method, in this paper we propose a novel dynamical graph convolutional neural networks (DGCNN) model for learning discriminative EEG features as well as the intrinsic relationship, e.g., the functional relationship, among the various EEG channels. Specifically, to learn the relationships among the various EEG channels, we propose a novel method to construct the connections among the various vertex nodes of the graph by learning an adjacency matrix. However, different from the traditional GCNN method that predetermines the adjacency matrix before the model training, the proposed DGCNN method learns the adjacency matrix in a dynamic way, i.e., the entries of the adjacency matrix are adaptively updated with the changes of graph model parameters during the model training. Consequently, in contrast to the GCNN method, the adjacency matrix learned by the DGCNN would be more useful because it captures the intrinsic connections of the EEG channels and hence it would be able to improve the discriminant abilities of the networks.

The remainder of this paper is organized as follows: In Section 2, we briefly review the preliminaries of graph theory. In Section 3, we propose the DGCNN model and the EEG emotion recognition method based on this model. Extensive experiments are conducted in Section 4. Finally, we conclude the paper in Section 5.

2 GRAPH PRELIMINARY

In this section, we introduce some preliminary knowledge about the graph representation and the spectral graph filtering, which are the basis for our DGCNN method.

2.1 Graph Representation

A directed and connected graph can be defined as $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{W}\}$, in which \mathcal{V} represents the set of nodes with the number of $|\mathcal{V}| = N$ and \mathcal{E} denotes the set of edges connecting these nodes. Let $\mathbf{W} \in \mathbb{R}^{N \times N}$ denote an adjacency matrix describing the connections between any two nodes in \mathcal{V} , in which the entry of \mathbf{W} in the i th row and j th column, denoted by w_{ij} , measures the importance of the connection between the i th node and the j th one. Fig. 1 illustrates an example of a graph containing six vertex nodes and the

edges connecting the nodes of the graph, as well as the adjacency matrix associated with the graph, where the different color arrows in the left-hand side of the figure denote the edges connecting the source nodes to destination nodes, whereas the right-hand side of the figure is the illustration of the corresponding adjacency matrix.

The commonly used methods to determine the entries w_{ij} of the adjacency matrix \mathbf{W} include the distance function method [47] and K-nearest neighbor (KNN) rule method [48]. A typical distance function would be the Gaussian kernel function, which can be expressed as

$$w_{ij} = \begin{cases} \exp(-\frac{[\text{dist}(i,j)]^2}{2\theta^2}), & \text{if } \text{dist}(i,j) \leq \tau, \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where τ and θ are two fixed parameters and $\text{dist}(i,j)$ denotes the distance between the i th node and the j th one.

2.2 Spectral Graph Filtering

The spectral graph theory has been successfully used for building expander graphs [49], spectral clustering [50], graph visualization [51] and other applications [52]. Spectral graph filtering, also called graph convolution, is a popular signal processing method for graph data operation, in which Graph Fourier Transform (GFT) [47] is a typical example.

Let \mathbf{L} denote the Laplacian matrix of the graph \mathcal{G} . Then, \mathbf{L} can be expressed as

$$\mathbf{L} = \mathbf{D} - \mathbf{W} \in \mathbb{R}^{N \times N}, \quad (2)$$

where $\mathbf{D} \in \mathbb{R}^{N \times N}$ is a diagonal matrix and the i th diagonal element can be calculated by $D_{ii} = \sum_j w_{ij}$.

For a given spatial signal $\mathbf{x} \in \mathbb{R}^N$, its GFT is expressed as follows:

$$\hat{\mathbf{x}} = \mathbf{U}^T \mathbf{x}, \quad (3)$$

where $\hat{\mathbf{x}}$ denotes the transformed signal in the frequency domain, \mathbf{U} is an orthonormal matrix that can be obtained via the singular value decomposition (SVD) of the graph Laplacian matrix \mathbf{L} [45]

$$\mathbf{L} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T, \quad (4)$$

in which the columns of $\mathbf{U} = [\mathbf{u}_0, \dots, \mathbf{u}_{N-1}] \in \mathbb{R}^{N \times N}$ constitute the Fourier basis, and $\mathbf{\Lambda} = \text{diag}([\lambda_0, \dots, \lambda_{N-1}])$ is a diagonal matrix.

From (3), the inverse of GFT can be expressed as the following form:

$$\mathbf{x} = \mathbf{U} \mathbf{U}^T \mathbf{x} = \mathbf{U} \hat{\mathbf{x}}. \quad (5)$$

Then, the convolution of two signals \mathbf{x} and \mathbf{y} on the graph, denoted by $\mathbf{x} *_G \mathbf{y}$, can be expressed as [44]

$$\mathbf{x} *_G \mathbf{y} = \mathbf{U}((\mathbf{U}^T \mathbf{x}) \odot (\mathbf{U}^T \mathbf{y})), \quad (6)$$

where \odot denotes the element-wise Hadamard product.

Now let $g(\cdot)$ be a filtering function and then a signal \mathbf{x} filtered by $g(\mathbf{L})$ can be expressed as

$$\mathbf{y} = g(\mathbf{L})\mathbf{x} = g(\mathbf{U} \mathbf{\Lambda} \mathbf{U}^T)\mathbf{x} = \mathbf{U} g(\mathbf{\Lambda}) \mathbf{U}^T \mathbf{x}, \quad (7)$$

where $g(\mathbf{\Lambda})$ is expressed as

$$g(\mathbf{\Lambda}) = \begin{bmatrix} g(\lambda_0) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g(\lambda_{N-1}) \end{bmatrix}. \quad (8)$$

It is notable that the filtering operation of (7) is equivalent to the graph convolution of the signal \mathbf{x} with the vector of $\mathbf{U} \text{diag}(g(\mathbf{\Lambda}))$ due to the following formulation:

$$\begin{aligned} \mathbf{y} &= g(\mathbf{L})\mathbf{x} = \mathbf{U} g(\mathbf{\Lambda}) \mathbf{U}^T \mathbf{x} \\ &= \mathbf{U}(\text{diag}(g(\mathbf{\Lambda})) \odot (\mathbf{U}^T \mathbf{x})) \\ &= \mathbf{U}((\mathbf{U}^T (\mathbf{U} \text{diag}(g(\mathbf{\Lambda})))) \odot (\mathbf{U}^T \mathbf{x})) \\ &= \mathbf{x} *_G (\mathbf{U} \text{diag}(g(\mathbf{\Lambda}))). \end{aligned} \quad (9)$$

3 DGCNN FOR EEG EMOTION RECOGNITION

In this section, we first propose the DGCNN model and then apply it to the EEG emotion recognition problem, in which the adjacency matrix \mathbf{W} that characterizes the relationships of the various vertex nodes is dynamically learned instead of being predetermined [44].

3.1 DGCNN Model for EEG Emotion Recognition

Let \mathbf{W}^* denote the optimal adjacency matrix to be learned. Then, the graph convolution of the signal \mathbf{x} with the vector of $\mathbf{U}^* \text{diag}(g(\mathbf{\Lambda}^*))$ defined by the spatial filtering $g(\mathbf{L}^*)$ can be expressed as

$$\mathbf{y} = g(\mathbf{L}^*)\mathbf{x} = \mathbf{U}^* g(\mathbf{\Lambda}^*) \mathbf{U}^{*T} \mathbf{x}, \quad (10)$$

where the \mathbf{L}^* can be calculated from \mathbf{W}^* based on (2), and $\mathbf{\Lambda}^* = \text{diag}([\lambda_0^*, \dots, \lambda_{N-1}^*])$ is a diagonal matrix.

Since it is difficult to directly calculate the expression of $g(\mathbf{\Lambda}^*)$, we simplify this calculation by using the K order Chebyshev polynomials [44] to replace the polynomial expansion of $g(\mathbf{\Lambda}^*)$, such that the calculation becomes much easier and faster. Specifically, let λ_{max}^* denote the largest element among the diagonal entries of $\mathbf{\Lambda}^*$ and $\tilde{\mathbf{\Lambda}}^*$ denote the normalized $\mathbf{\Lambda}^*$, i.e., $\tilde{\mathbf{\Lambda}}^* = 2\mathbf{\Lambda}^*/\lambda_{max}^* - \mathbf{I}_N$, such that the diagonal elements of $\tilde{\mathbf{\Lambda}}^*$ lie in the interval of $[-1, 1]$, where \mathbf{I}_N is the $N \times N$ identity matrix.

Under the K order Chebyshev polynomials framework, we obtain that $g(\mathbf{\Lambda}^*)$ can be approximated by

$$g(\mathbf{\Lambda}^*) = \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\mathbf{\Lambda}}^*), \quad (11)$$

where θ_k is the coefficient of Chebyshev polynomials, and $T_k(x)$ can be recursively calculated according to the following recursive expressions:

$$\begin{cases} T_0(x) = 1, T_1(x) = x, \\ T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), \quad k \geq 2. \end{cases} \quad (12)$$

According to (11), we obtain that the graph convolution operation defined in (10) can be rewritten by

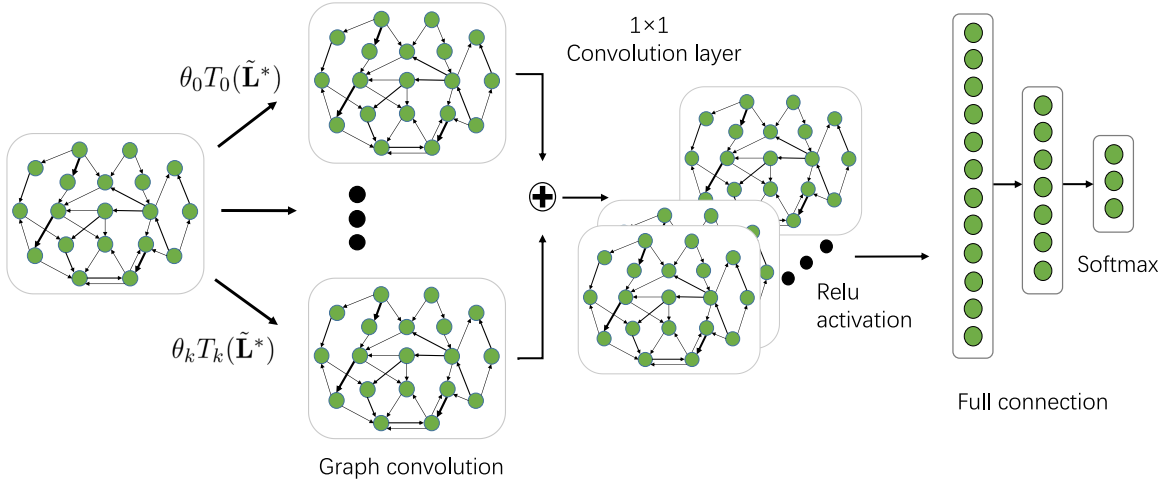


Fig. 2. The framework of the DGCNN model for EEG emotion recognition, which consists of the dynamical graph convolution operation via the learned graph connections, convolution layer with 1×1 kernel, Relu activation and the full connection. The inputs of the model are the EEG features extracted from multiple frequency bands, e.g., five frequency bands (δ band, θ band, α band, β band, and γ band), in which each EEG channel is represented as a node of the graph. The outputs are the predicted labels through softmax.

$$\begin{aligned}
 \mathbf{y} &= \mathbf{U}^* g(\mathbf{\Lambda}^*) \mathbf{U}^{*T} \mathbf{x} \\
 &= \sum_{k=0}^{K-1} \mathbf{U}^* \begin{bmatrix} \theta_k T_k(\tilde{\lambda}_0^*) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_k T_k(\tilde{\lambda}_{N-1}^*) \end{bmatrix} \mathbf{U}^{*T} \mathbf{x} \\
 &= \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\mathbf{L}}^*) \mathbf{x},
 \end{aligned} \quad (13)$$

where $\tilde{\mathbf{L}}^* = 2\mathbf{L}^* / \lambda_{max}^* - \mathbf{I}_N$.

The expression of (13) means that calculating the graph convolution of \mathbf{x} can be expressed as the combination of the convolutional results of \mathbf{x} with each of the Chebyshev polynomial components. Based on the expression of (13), we propose the DGCNN model for EEG emotion recognition. The framework of the proposed DGCNN model is illustrated in Fig. 2, which consists of four major layers, i.e., the dynamical graph convolutional layer, the 1×1 convolutional layer, the Relu activation layer, and the full connection layer.

Specifically, the input of the DGCNN model corresponds to the EEG features extracted from multiple frequency bands, e.g., five frequency bands (δ band, θ band, α band, β band, and γ band), in which each EEG channel is represented as a node in the DGCNN model. Following the graph filtering operation is a 1×1 convolution layer, which aims to learn the discriminative features among the various frequency domains. Moreover, to realize the non-linear mapping capability of the network, the Relu activation function [53] is adopted to ensure that the outputs of the graph filtering layer are non-negative. Finally, the outputs of activation function are further sent to a multi-layer full connection network and a softmax function is also used to predict the desired class label information of the input EEG features.

3.2 Algorithm for DGCNN

To optimize the optimal network parameters, we adopt the back propagation (BP) method to iteratively update the network parameters until the optimal or suboptimal solutions

are achieved. For this purpose, we define a loss function based on cross entropy cost, which is expressed as the following form:

$$Loss = cross_entropy(\mathbf{1}, \mathbf{I}^p) + \alpha \|\Theta\|, \quad (14)$$

where $\mathbf{1}$ and \mathbf{I}^p denote the actual label vector of training data and the predicted one, respectively, Θ denotes all the model parameters and α is the trade-off regularization weight. The cross entropy function $cross_entropy(\mathbf{1}, \mathbf{I}^p)$ aims at measuring the dissimilarity between the actual emotional labels and the desired ones while the regularization $\alpha \|\Theta\|$ aims to prevent over-fitting of the model parameters learning.

When applying the BP method to dynamically learn the optimal adjacency matrix \mathbf{W}^* of the DGCNN model, we have to calculate the partial derivative of the loss function with respect to \mathbf{W}^* , which is formulated as

$$\frac{\partial Loss}{\partial \mathbf{W}^*} = \begin{pmatrix} \frac{\partial Loss}{\partial w_{11}^*} & \cdots & \frac{\partial Loss}{\partial w_{1N}^*} \\ \vdots & \ddots & \vdots \\ \frac{\partial Loss}{\partial w_{N1}^*} & \cdots & \frac{\partial Loss}{\partial w_{NN}^*} \end{pmatrix}, \quad (15)$$

where w_{ij}^* denotes the i th row and j -th column element of \mathbf{W}^* . According to the chain rule, the calculation of $\frac{\partial Loss}{\partial w_{ij}^*}$ can be expressed as

$$\frac{\partial Loss}{\partial w_{ij}^*} = \frac{\partial cross_entropy(\mathbf{1}, \mathbf{I}^p)}{\partial \tilde{\mathbf{L}}^*} \cdot \frac{\partial \tilde{\mathbf{L}}^*}{\partial w_{ij}^*} + \alpha \frac{\partial \|\Theta\|}{\partial w_{ij}^*}. \quad (16)$$

After calculating the partial derivative of $\frac{\partial Loss}{\partial \mathbf{W}^*}$, we can use the following rule to update the optimal adjacency matrix \mathbf{W}^* :

$$\mathbf{W}^* = (1 - \rho) \mathbf{W}^* + \rho \frac{\partial Loss}{\partial \mathbf{W}^*},$$

where ρ denotes the learning rate of the network.

Algorithm 1 summarizes the detailed procedures of training the DGCNN model in EEG emotion recognition.

Algorithm 1. Procedures of Training Optimal DGCNN Model for EEG Emotion Recognition

Require: Multichannel EEG features associated with multiple frequency bands, the class labels corresponding to the EEG features, the number of Chebyshev polynomial order K , the learning rate ρ ;

Ensure: The desired adjacency matrix \mathbf{W}^* and model parameters of DGCNN;

- 1: Initialize the adjacency matrix \mathbf{W}^* and other model parameters;
- 2: **repeat**
- 3: Regularizing the elements of the matrix \mathbf{W}^* using Relu operation such that the elements are non-negative;
- 4: Calculating the Laplacian matrix \mathbf{L}^* ;
- 5: Calculating the normalized Laplacian matrix $\tilde{\mathbf{L}}^*$;
- 6: Calculating the Chebyshev polynomial items $T_k(\tilde{\mathbf{L}}^*)$ ($k = 0, 1, \dots, K-1$);
- 7: Calculating $\sum_{k=0}^{K-1} \theta_k T_k(\tilde{\mathbf{L}}^*) \mathbf{x}$;
- 8: Calculating the 1×1 convolution results and regularizing the results using the Relu operation;
- 9: Calculating the results of the full connection layer;
- 10: Calculating the loss function using (14);
- 11: Updating the adjacency matrix

$$\mathbf{W}^* = (1 - \rho)\mathbf{W}^* + \rho \frac{\partial Loss}{\partial \mathbf{W}^*}$$

and other model parameters;

- 12: **until** the iterations satisfy the predefined algorithm convergence condition.

4 EXPERIMENTS

In this section, we conduct extensive experiments on two emotional EEG databases that are commonly used in EEG emotion recognition to evaluate the effectiveness of the proposed DGCNN method. The first one is the SJTU Emotion EEG Database (SEED) [19] while the second one is the DREAMER [54].

4.1 Emotional EEG Databases

The SEED database contains EEG data of 15 subjects (7 males and 8 females), which are collected via 62 EEG electrodes from the subjects when they are watching fifteen Chinese film clips with three types of emotions, i.e., negative, positive and neutral. To avoid making subjects fatigue, the whole experiment will not last for a long time and the duration of each film clip is about 4 minutes. The stimulus materials can be understood without explanation. Consequently, all the EEG signals will be categorized into one of three kinds of emotion states (positive, neutral and negative). The data collection lasted for 3 different periods corresponding to 3 sessions, and each session corresponds to 15 trials of EEG data such that there are totally 45 trials of EEG data for each subject. In addition, an additional subjective self-assessment for each subject is also carried out after the subjects watching the film clips in order to guarantee that the collected EEG data share the same emotion states as the film clips presented to the subjects.

The DREAMER database consists of EEG data via 14 EEG electrodes of 23 subjects (14 males and 9 females). To build this database, 18 film clips are used for eliciting 9 different

TABLE 1
Number of the Five Types of EEG Features Extracted from Each Frequency Band on SEED Database

| Feature type | Number of EEG features per sample | | | | |
|--------------|-----------------------------------|----------|----------|---------|----------|
| | δ | θ | α | β | γ |
| PSD | 62 | 62 | 62 | 62 | 62 |
| DE | 62 | 62 | 62 | 62 | 62 |
| DASM | 27 | 27 | 27 | 27 | 27 |
| RASM | 27 | 27 | 27 | 27 | 27 |
| DCAU | 23 | 23 | 23 | 23 | 23 |

emotions, i.e., amusement, excitement, happiness, calmness, anger, disgust, fear, sadness and surprise. Each film clip lasts for a period time between 65 to 393s, which is thought to be sufficient for eliciting single emotions. The data collection begins with a neutral film clip watching to help the subjects return to the neutral emotion state in each new trial of data collection and also to serve as the baseline signals. After watching a film clip, the self-assessment manikins (SAM) were used to acquire subjective assessments of valence, arousal and dominance. Finally, all the EEG signals are assigned with three binary states (low/high valence, low/high arousal and low/high dominance).

4.2 EEG Emotion Recognition Experiments on SEED Database

In this part, we conduct two kinds of experiments to evaluate the EEG emotion recognition performance using the proposed DGCNN method. The first kind of experiment is subject-dependent whereas the second one is subject-independent.

4.2.1 Subject-Dependent Experiments on SEED Database

In subject-dependent experiments, we adopt the same experimental protocol as that of [19] to evaluate the proposed DGCNN method. Specifically, for all the 15 trials of EEG data associated with one session of one subject, the first 9 trials of EEG data are used to serve as the training set and remaining 6 ones as the testing set. Then, the recognition accuracy corresponding to each period is obtained for each subject. Finally, the average classification accuracy and standard deviation over two sessions of all the 15 subjects are calculated.

We investigate five kinds of features to evaluate the proposed EEG emotion recognition method, i.e., the differential entropy feature, the power spectral density feature, the differential asymmetry feature, the rational asymmetry feature, and the differential causality feature. The features are respectively extracted in each of the frequency bands (δ band, θ band, α band, β band, and γ band). Moreover, to extract the EEG feature, each trial of EEG signal flow is partitioned into a set of blocks, where each block contains 1s of EEG signals. In this case, we can extract five kinds of features from each block. The number of EEG features extracted from each frequency band corresponding to the various feature types are summarized in Table 1.

Based on the EEG features shown in Table 1, we have trained the DGCNN model shown in Fig. 2 for the EEG emotion recognition. In this case, each vertex node of the graph in the DGCNN model is associated with five EEG

TABLE 2
Comparisons of the Average Accuracies and Standard Deviations (%) of Subject Dependent EEG-Based Emotion Recognition Experiments on SEED Database Among the Various Methods

| Feature | Method | δ band | θ band | α band | β band | γ band | all ($\delta, \theta, \alpha, \beta, \gamma$) |
|---------|----------|----------------------|---------------------|----------------------|----------------------|----------------------|---|
| DE | DBN [19] | 64.32 / 12.45 | 60.77 / 10.42 | 64.01 / 15.97 | 78.92 / 12.48 | 79.19 / 14.58 | 86.08 / 8.34 |
| | SVM [19] | 60.50 / 14.14 | 60.95 / 10.20 | 66.64 / 14.41 | 80.76 / 11.56 | 79.56 / 11.38 | 83.99 / 9.72 |
| | GCNN | 72.75 / 10.85 | 74.40 / 8.23 | 73.46 / 12.17 | 83.24 / 9.93 | 83.36 / 9.43 | 87.40 / 9.20 |
| | DGCNN | 74.25 / 11.42 | 71.52 / 5.99 | 74.43 / 12.16 | 83.65 / 10.17 | 85.73 / 10.64 | 90.40 / 8.49 |
| PSD | DBN [19] | 60.05 / 16.66 | 55.03 / 13.88 | 52.79 / 15.38 | 60.68 / 21.31 | 63.42 / 19.66 | 61.90 / 16.65 |
| | SVM [19] | 58.03 / 15.39 | 57.26 / 15.09 | 59.04 / 15.75 | 73.34 / 15.20 | 71.24 / 16.38 | 59.60 / 15.93 |
| | GCNN | 69.89 / 13.83 | 70.92 / 9.18 | 73.18 / 12.74 | 76.21 / 10.76 | 76.15 / 10.09 | 81.31 / 11.26 |
| | DGCNN | 71.23 / 11.42 | 71.20 / 8.99 | 73.45 / 12.25 | 77.45 / 10.81 | 76.60 / 11.83 | 81.73 / 9.94 |
| DASM | DBN [19] | 48.79 / 9.62 | 51.59 / 13.98 | 54.03 / 17.05 | 69.51 / 15.22 | 70.06 / 18.14 | 72.73 / 15.93 |
| | SVM [19] | 48.87 / 10.49 | 53.02 / 12.76 | 59.81 / 14.67 | 75.03 / 15.72 | 73.59 / 16.57 | 72.81 / 16.57 |
| | GCNN | 57.07 / 6.75 | 54.80 / 9.09 | 62.97 / 13.43 | 74.97 / 13.40 | 73.28 / 13.67 | 76.00 / 13.32 |
| | DGCNN | 55.93 / 9.14 | 56.12 / 7.86 | 64.27 / 12.72 | 73.61 / 14.35 | 73.50 / 16.6 | 78.45 / 11.84 |
| RASM | DBN [19] | 48.05 / 10.37 | 50.62 / 14.02 | 56.15 / 15.28 | 70.31 / 15.62 | 68.22 / 18.09 | 71.30 / 16.16 |
| | SVM [19] | 47.75 / 10.59 | 51.40 / 12.53 | 60.71 / 14.57 | 74.59 / 16.18 | 74.61 / 15.57 | 74.74 / 14.79 |
| | GCNN | 59.70 / 5.65 | 55.91 / 8.82 | 59.97 / 14.27 | 79.45 / 13.32 | 79.73 / 13.22 | 84.06 / 12.86 |
| | DGCNN | 57.79 / 6.90 | 55.79 / 8.10 | 61.58 / 12.63 | 75.79 / 13.07 | 82.32 / 11.54 | 85.00 / 12.47 |
| DCAU | DBN [19] | 54.58 / 12.81 | 56.94 / 12.54 | 57.62 / 13.58 | 70.70 / 16.33 | 72.27 / 16.12 | 77.20 / 14.24 |
| | SVM [19] | 55.92 / 14.62 | 57.16 / 10.77 | 61.37 / 15.97 | 75.17 / 15.58 | 76.44 / 15.41 | 77.38 / 11.98 |
| | GCNN | 62.60 / 12.88 | 65.05 / 8.35 | 66.41 / 11.06 | 77.28 / 11.55 | 78.68 / 13.00 | 79.02 / 11.27 |
| | DGCNN | 63.18 / 13.48 | 62.55 / 7.96 | 67.71 / 10.74 | 78.68 / 10.81 | 80.05 / 13.03 | 81.91 / 10.06 |

features corresponding to the five frequency bands. It is notable that the number of vertex nodes in the graph would be different for different feature types. Specifically, for both PSD and DE features, the number of vertex nodes in the graph is 62 whereas it would be 27 nodes for both DASM and RASM because there are only 27 features for these two feature types. Similarly, the number of vertex nodes in the graph would be 23 for DCAU feature type.

Table 2 summarizes experimental results in terms of the average EEG emotion recognition accuracy and the standard deviation of the DGCNN method under the five different EEG feature types (PSD, DE, DASM, RASM and DCAU) and the five different frequency bands, where the results annotated by “all” mean that the EEG features associated with all of the five frequency bands are combined together in the experiments such that each vertex node of the graph is associated with five EEG features. For the comparison purpose, we also include the experimental results of [19] with deep belief networks (DBN) [55] and support vector machine (SVM) [56] in Table 2. Moreover, we also conduct the same experiments using the GCNN method to serve as a baseline method to evaluate the performance of DGCNN. Here it should be noted that, for the GCNN model, the elements of the adjacency matrix are predetermined according to the spatial relationship of the EEG channels (Fig. 3 illustrates the spatial relationship of the 62 EEG channels used to construct the adjacency matrix for the GCNN model, i.e., there would be a direct connection between two EEG channels if they are closely related.).

From Table 2, we can observe the following major points:

- Among the five kinds of EEG features, the DE feature was demonstrated to be better than most of the other ones in terms of the average recognition accuracy. For each emotion recognition method, the best average recognition accuracy was achieved when all

the five frequency bands were used together. Especially, the best average recognition accuracy of DE feature was as high as 90.4 percent (with standard deviation of 8.49 percent) when all the five frequency bands were used.

- Among the four EEG emotion recognition methods, both DGCNN and GCNN achieve better average recognition accuracies than the other two methods (SVM and DBN). Compared with GCNN, however, DGCNN demonstrates to be more powerful in classifying the EEG emotion classes in most cases when the same EEG feature is used. This is most likely due to the optimization for determining the entries of the adjacency matrix in the GCNN model, which makes

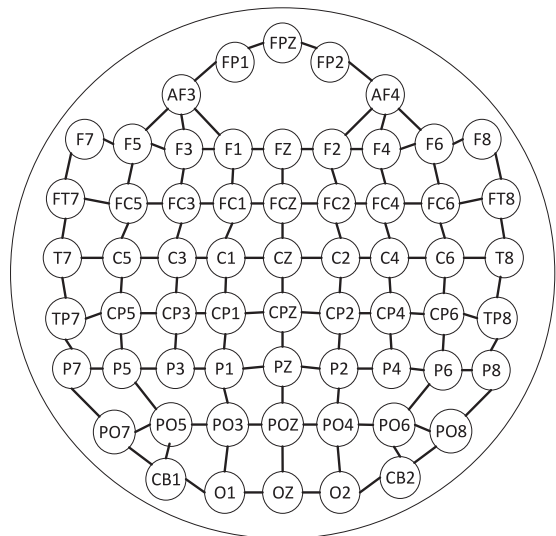


Fig. 3. Illustration of the connections among the 62 EEG channels, which is used for constructing the adjacency matrix of GCNN.

TABLE 3
The Average Accuracies and Standard Deviations (%) of Subject Independent LOSO Cross Validation EEG-Based Emotion Recognition Experiments on SEED Database Using DGCNN

| Method | Feature | δ band | θ band | α band | β band | γ band | all ($\delta, \theta, \alpha, \beta, \gamma$) |
|--------|---------|---------------------|---------------------|----------------------|---------------------|----------------------|---|
| DGCNN | DE | 49.79 / 10.94 | 46.36 / 12.06 | 48.29 / 12.28 | 56.15 / 14.01 | 54.87 / 17.53 | 79.95 / 9.02 |
| | PSD | 50.36 / 9.66 | 48.85 / 9.44 | 43.39 / 8.22 | 56.39 / 10.9 | 51.81 / 10.88 | 64.27 / 13.80 |
| | DASM | 44.31 / 6.67 | 43.79 / 6.31 | 44.64 / 8.94 | 47.53 / 11.95 | 47.13 / 10.89 | 52.50 / 11.92 |
| | RASM | 43.42 / 4.07 | 42.69 / 7.11 | 42.77 / 7.43 | 49.52 / 10.22 | 46.37 / 12.1 | 58.46 / 10.08 |
| | DCAU | 48.52 / 10.23 | 52.44 / 9.26 | 44.46 / 8.39 | 52.55 / 13.31 | 50.97 / 13.66 | 65.19 / 10.49 |

it more accurate to characterize the relationship between the various EEG channels.

- Among the five EEG frequency bands, both β and γ frequency bands achieve better recognition results than the other three ones in most cases, which indicates that the higher frequency bands may be more closely related with the emotion activities whereas the lower frequency bands are less related with the emotion activities.

4.2.2 Subject-Independent Experiments on SEED

In the subject-independent experiments, we adopt the leave-one-subject-out (LOSO) cross-validation strategy to evaluate the EEG emotion recognition performance of the proposed DGCNN method. Specifically, in LOSO cross-validation experimental protocol, the EEG data of 14 subjects are used for training the model and the EEG data of the rest one subject is used as testing data. The experiments are repeated such that the EEG data of each subject are used once as the testing data. The average classification accuracies and standard deviations corresponding to the five kinds of EEG features are respectively calculated.

Table 3 summarizes the experimental results in terms of the average EEG emotion recognition accuracy and the standard deviation of DGCNN under the different kinds of EEG features and the different frequency bands. From Table 3, we obtain the following major points:

- For each kind of feature, the recognition accuracies associated with higher frequency bands are better than the ones of lower frequency bands.
- For each kind of feature, the best recognition accuracy is obtained when combining the features extracted from the five frequency bands together for emotion recognition. Especially, when the five frequency bands are used and DE feature is adopted, the best recognition accuracy (= 79.95%) is achieved.

Considering that the subject-independent emotion recognition task can be seen as a cross-domain emotion recognition problem, we adopt several popular cross-domain recognition methods, including transfer component analysis (TCA) [57], KPCA [58], transductive SVM (T-SVM) [59] and transductive parameter transfer (TPT) [60], to serve as the baseline methods. Since the use of the DE features combined with the five frequency bands ($\delta, \theta, \alpha, \beta$, and γ) had been demonstrated to be the most effective features in the EEG emotion recognition, we adopt these features to compare the EEG emotion recognition performance among the five methods. Fig. 4 summarizes the experimental results. From the results of Fig. 4, we can see that the proposed DGCNN

method achieves the highest recognition accuracy (= 79.95%) among the five methods. In addition, we can also see that the standard deviation (= 9.02%) of the proposed DGCNN is much lower than that of TCA, KPCA, T-SVM, and TPT, which indicates that DGCNN is much more stable compared with the other four methods.

4.3 EEG Emotion Recognition Experiments on DREAMER Database

In this part, we conducted experiments on the DREAMER database to evaluate the EEG emotion recognition performance using the proposed DGCNN method. Before the experiments, we adopted the same feature extraction method as that of [54] to extract a set of PSD features. During the feature extraction, we first cropped the EEG signals corresponding to the last 60 seconds of each film clip and then decomposed the signals into θ (4-8 Hz), α (8-13 Hz), and β (13-20 Hz) frequency bands. For each frequency band, the 60s EEG signals are further segmented into a set of 59 blocks by sliding a window with the size of 256 EEG points, in which there are half overlap between two subsequent blocks. Finally, the PSD features were calculated from the EEG signal of each block. As a result, we obtain 14 features associated with the 14 EEG channels in total from each block and we concatenate them into a 14-dimensional feature vector to represent a EEG data sample. In this way, we totally obtain 59 EEG data samples associated with each frequency band from each session, which are used for our EEG emotion recognition evaluation of the DGCNN method.

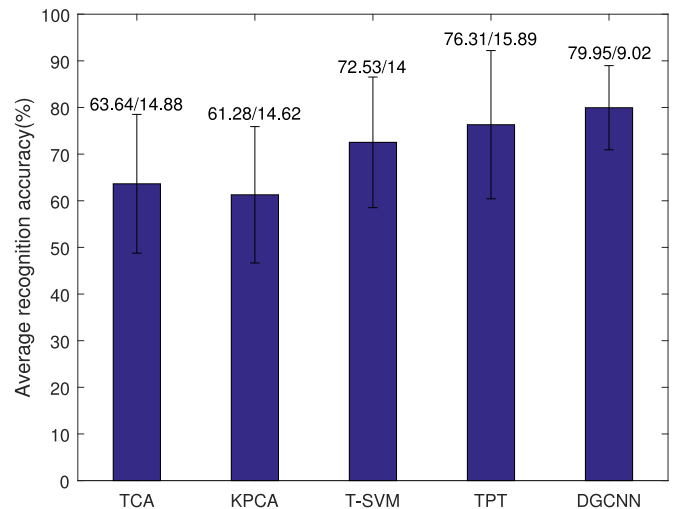


Fig. 4. Comparisons of the EEG emotion recognition accuracies and standard deviations among TCA, KPCA, T-SVM, TPT, and DGCNN on SEED database.

TABLE 4
Number of the PSD EEG Features of Each EEG Sample
Associated with Each Frequency Band on DREAMER Database

| Feature type | Number of EEG features per sample | | |
|--------------|-----------------------------------|----------|---------|
| | θ | α | β |
| PSD | 14 | 14 | 14 |

Table 4 shows the information about the EEG features we extracted from the DREAMER database.

To evaluate the EEG emotion recognition performance of the proposed DGCNN method, we adopt the subject-dependent leave-one-session-out cross-validation strategy to carry out the experiments. Specifically, for all the 18 sessions (corresponding to the 18 film clips) of experiments for each subject, we choose the EEG data belonging to one session as the testing data and use the ones belonging to the other 17 sessions as the training data. The emotion classification accuracy for the testing data is calculated based on the training data. This procedure is repeated for 18 trials such that the EEG data samples of each session have been used once as the testing data. Then, the overall classification accuracy of this subject is achieved by averaging the recognition accuracies of all the 18 trials. Finally, we use the average emotion classification accuracy obtained by averaging all the 23 subjects to evaluate the emotion recognition performance of the DGCNN method.

Table 5 shows the experimental results of the DGCNN method with respect to different emotion dimensions (i.e., valence, arousal and dominance). For comparison purpose, three state-of-the-art methods, i.e., SVM, graph regularized sparse linear discriminant analysis (GraphSLDA) [61] and group sparse canonical correlation analysis (GSCCA) [9], are also used to conduct the same experiments and the experimental results are included in Table 5.

From Table 5, we can obtain the following major points:

- The proposed DGCNN method achieves much higher classification accuracies than the other three state-of-the-art methods, in which the classification accuracies could be as high as 86.23 percent for valence classification, 84.54 percent for arousal classification, and 85.02 percent for dominance classification, respectively.
- The proposed DGCNN achieves more stable results than SVM, GraphSLDA, and GSCCA in terms of the standard deviations, in which the standard deviations of DGCNN are 12.29 percent for valence, 10.18 percent for arousal and 10.25 percent for dominance, respectively.

TABLE 5
The Average Classification Accuracies and Standard Deviations (%) of Subject-Dependent EEG Emotion Recognition on DREAMER Database Using PSD Feature

| Method | Valence | Arousal | Dominance |
|-----------|----------------------|----------------------|----------------------|
| SVM | 60.14 / 33.34 | 68.84 / 24.94 | 75.84 / 20.76 |
| GraphSLDA | 57.70 / 13.89 | 68.12 / 17.53 | 73.90 / 15.85 |
| GSCCA | 56.65 / 21.50 | 70.30 / 18.66 | 77.31 / 15.44 |
| DGCNN | 86.23 / 12.29 | 84.54 / 10.18 | 85.02 / 10.25 |

5 CONCLUSIONS AND DISCUSSIONS

In this paper, we have proposed a novel DGCNN model for EEG emotion recognition and conducted experiments on SEED and DREAMER EEG emotion databases, respectively, to evaluate the effectiveness of the proposed method. Both subject-dependent experiments and subject-independent cross validation experiments on SEED database had been conducted and the experimental results indicated that the DGCNN method achieves better recognition performance than the state-of-the-art methods such as the SVM, DBN, KPCA, TCA, T-SVM and TPT methods. Especially, when the DE features of five frequency bands are combined together, the average recognition accuracy of DGCNN can be as high as 90.40 percent for the case of subject dependent experiments and can be as high as 79.95 percent for the case of subject independent cross validation experiments. On the DREAMER database, the average accuracies of valence, arousal, and dominance using the proposed DGCNN are 86.23, 84.54 and 85.02 percent respectively, which are higher than SVM, GraphSLDA and GSCCA. The better recognition performance of DGCNN is most likely due to the following major points:

- The use of nonlinear neural network of DGCNN makes it much more powerful to deal with the nonlinear discriminative feature learning;
- The graph representation of DGCNN provides a useful way to characterize the intrinsic relationships among the various EEG channels, which is advantageous for extracting the most discriminative features for the emotion recognition task;
- DGCNN adaptively learns the intrinsic relationships of EEG channels through optimization of the adjacency matrix \mathbf{W} . In contrast to DGCNN, the GCNN model determines the values of \mathbf{W} prior to the model learning stage. Consequently, using DGCNN model would be more accurate than GCNN to characterize the relationships of EEG channels.

Additionally, it is notable that the diagonal elements of the adjacency matrix indicate the contributions of the EEG channels to the EEG emotion recognition. Hence, the adjacency matrix would provide a potential way to find out what are the most contributive EEG channels in the EEG emotion recognition, which would be advantageous to further improve the EEG emotion recognition performance. We leave this interesting topic as our future work.

Another interesting issue that could be investigated in future is about the data scale generalization. Although the proposed DGCNN method had demonstrated to be a good method in dealing with the EEG emotion recognition, it is also notable that the data scales of the EEG databases used in the experiments are still relatively small, which would not be enough for learning more powerful deep neural network models and hence may limit the further performance improvement in this research. Consequently, an EEG database with much larger scales is desired in solving this problem and this is also a major task of our future work.

ACKNOWLEDGMENTS

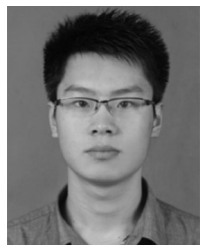
This work was supported in part by the National Basic Research Program of China under Grant 2015CB351704, in

part by the National Natural Science Foundation of China under Grant 61572009, Grant 61703360 and Grant 61772276, in part by the Key Research and Development Program of Jiangsu Province, China under Grant BE2016616.

REFERENCES

- [1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Process. Mag.*, vol. 18, no. 1, pp. 32–80, Jan. 2001.
- [2] K. Anderson and P. W. McOwan, "A real-time automated system for the recognition of human facial expressions," *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)*, vol. 36, no. 1, pp. 96–105, Feb. 2006.
- [3] X. Huang, G. Zhao, X. Hong, W. Zheng, and M. Pietikainen, "Spontaneous facial micro-expression analysis using spatiotemporal completed local quantized patterns," *Neurocomputing*, vol. 175, pp. 564–578, 2016.
- [4] X. Huang, S.-J. Wang, G. Zhao, and M. Pietikainen, "Facial micro-expression recognition using spatiotemporal local binary pattern with integral projection," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2015, pp. 1–9.
- [5] H. Xiaohua, S.-J. Wang, X. Liu, G. Zhao, X. Feng, and M. Pietikainen, "Discriminative spatiotemporal local binary pattern with revisited integral projection for spontaneous facial micro-expression recognition," *IEEE Trans. Affect. Comput.*, 2017, doi: [10.1109/TAFFC.2017.2713359](https://doi.org/10.1109/TAFFC.2017.2713359).
- [6] Y.-J. Liu, J.-K. Zhang, W.-J. Yan, S.-J. Wang, G. Zhao, and X. Fu, "A main directional mean optical flow feature for spontaneous micro-expression recognition," *IEEE Trans. Affect. Comput.*, vol. 7, no. 4, pp. 299–310, Oct.–Dec. 2016.
- [7] J. Yan, W. Zheng, M. Xin, and J. Yan, "Integrating facial expression and body gesture in videos for emotion recognition," *IEICE Trans. Inf. Syst.*, vol. E97-D, no. 3, pp. 610–613, 2014.
- [8] J. Ang, R. Dhillon, A. Krupski, E. Shriberg, and A. Stolcke, "Prosody-based automatic detection of annoyance and frustration in human-computer dialog," in *Proc. Int. Conf. Spoken Language Process.*, 2002, pp. 2037–2040.
- [9] W. Zheng, "Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis," *IEEE Trans. Cogn. Develop. Syst.*, vol. 9, no. 3, pp. 281–290, Sep. 2017.
- [10] B. Cheng and G. Liu, "Emotion recognition from surface EMG signal using wavelet transform and neural network," in *Proc. 2nd Int. Conf. Bioinf. Biomed. Eng.*, 2008, pp. 1363–1366.
- [11] F. Agraftioti, D. Hatzinakos, and A. K. Anderson, "ECG pattern analysis for emotion detection," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 102–115, Jan.–Mar. 2012.
- [12] Y. M. Chi, Y.-T. Wang, Y. Wang, C. Maier, T.-P. Jung, and G. Cauwenberghs, "Dry and noncontact EEG sensors for mobile brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 2, pp. 228–235, Mar. 2012.
- [13] Y.-J. Huang, C.-Y. Wu, A. M.-K. Wong, and B.-S. Lin, "Novel active comb-shaped dry electrode for EEG measurement in hairy site," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 256–263, Jan. 2015.
- [14] X. Li, B. Hu, S. Sun, and H. Cai, "EEG-based mild depressive detection using feature selection methods and classifiers," *Comput. Methods Programs Biomed.*, vol. 136, pp. 151–161, 2016.
- [15] Y.-J. Liu, M. Yu, G. Zhao, J. Song, Y. Ge, and Y. Shi, "Real-time movie-induced discrete emotion recognition from EEG signals," *IEEE Trans. Affect. Comput.*, 2017, doi: [10.1109/TAFFC.2017.2660485](https://doi.org/10.1109/TAFFC.2017.2660485).
- [16] W. Zheng, H. Tang, and T. S. Huang, "Emotion recognition from non-frontal facial images," in *Emotion Recognition: A Pattern Analysis Approach*, A. Konar and A. Chakraborty, Eds., 1st ed. Hoboken, NJ, USA: Wiley, 2014, pp. 183–213.
- [17] E. L. Van den Broek, "Ubiquitous emotion-aware computing," *Pers. Ubiquitous Comput.*, vol. 17, no. 1, pp. 53–67, 2013.
- [18] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, no. 4, pp. 261–292, 1996.
- [19] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Trans. Auton. Mental Develop.*, vol. 7, no. 3, pp. 162–175, Sep. 2015.
- [20] T. Musha, Y. Terasaki, H. A. Haque, and G. A. Ivamitsky, "Feature extraction from EEGs associated with emotions," *Artif. Life Robot.*, vol. 1, no. 1, pp. 15–19, 1997.
- [21] L. Aftanas, N. Reva, A. Varlamov, S. Pavlov, and V. Makhnev, "Analysis of evoked EEG synchronization and desynchronization in conditions of emotional activation in humans: Temporal and topographic characteristics," *Neuroscience Behavioral Physiology*, vol. 34, no. 8, pp. 859–867, 2004.
- [22] S. Saney and J. A. Chambers, *EEG Signal Processing*. Hoboken, NJ, USA: Wiley, 2013.
- [23] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalography Clinical Neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.
- [24] Y. Liu and O. Sourina, "Real-time fractal-based valence level recognition from EEG," in *Proc. Trans. Comput. Sci. XVIII*, 2013, pp. 101–120.
- [25] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from EEG using higher order crossings," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 186–197, Mar. 2010.
- [26] R. J. Davidson, "What does the prefrontal cortex 'do' in affect: perspectives on frontal EEG asymmetry research," *Biol. Psychology*, vol. 67, no. 1, pp. 219–234, 2004.
- [27] M. Li and B.-L. Lu, "Emotion classification based on gamma-band EEG," in *Proc. 31st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2009, pp. 1223–1226.
- [28] 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.*, 2011, pp. 667–670.
- [29] L.-C. Shi, Y.-Y. Jiao, and B.-L. Lu, "Differential entropy feature for EEG-based vigilance estimation," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2013, pp. 6627–6630.
- [30] W.-L. Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, "EEG-based emotion classification using deep belief networks," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2014, pp. 1–6.
- [31] C. A. Frantzidis, C. Bratsas, C. L. Papadelis, E. Konstantinidis, C. Pappas, and P. D. Bamidis, "Toward emotion aware computing: An integrated approach using multichannel neurophysiological recordings and affective visual stimuli," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 589–597, May 2010.
- [32] Y.-P. Lin, C.-H. Wang, T.-P. Jung, T.-L. Wu, S.-K. Jeng, J.-R. Duann, and J.-H. Chen, "EEG-based emotion recognition in music listening," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1798–1806, Jul. 2010.
- [33] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using EEG signals: A survey," *IEEE Trans. Affect. Comput.*, 2017, doi: [10.1109/TAFFC.2017.2714671](https://doi.org/10.1109/TAFFC.2017.2714671).
- [34] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2015, pp. 91–99.
- [35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
- [36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [37] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 248–255.
- [38] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
- [39] M. Danelljan, A. Robinson, F. S. Khan, and M. Felsberg, "Beyond correlation filters: Learning continuous convolution operators for visual tracking," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 472–488.
- [40] Y. Li, H. Qi, J. Dai, X. Ji, and Y. Wei, "Fully convolutional instance-aware semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2359–2367.
- [41] C. Godsil and G. F. Royle, *Algebraic Graph Theory*. Berlin, Germany: Springer, 2013.
- [42] D. A. Spielman, "Spectral graph theory and its applications," in *Proc. 48th Annu. IEEE Symp. Found. Comput. Sci.*, 2007, pp. 29–38.
- [43] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Jan. 2009.
- [44] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 3844–3852.
- [45] F. R. Chung, *Spectral Graph Theory*. Rhode, Island: American Mathematical Soc., 1997.

- [46] F. P. Such, S. Sah, M. A. Dominguez, S. Pillai, C. Zhang, A. Michael, N. D. Cahill, and R. Ptucha, "Robust spatial filtering with graph convolutional neural networks," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 6, pp. 884–896, Sep. 2017.
- [47] 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.
- [48] L. J. Grady and J. R. Polimeni, *Discrete Calculus: Applied Analysis on Graphs for Computational Science*. Berlin, Germany: Springer, 2010.
- [49] S. Hoory, N. Linial, and A. Wigderson, "Expander graphs and their applications," *Bulletin Amer. Math. Soc.*, vol. 43, no. 4, pp. 439–561, 2006.
- [50] U. Von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, 2007.
- [51] U. Naumann and O. Schenk, *Combinatorial Scientific Computing*. Boca Raton, FL, USA: CRC Press, 2012.
- [52] D. M. Cvetkovic, "Applications of graph spectra: An introduction to the literature," *Appl. Graph Spectra*, vol. 13, no. 21, pp. 7–31, 2009.
- [53] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learning*, 2010, pp. 807–814.
- [54] S. Katsigiannis and N. Ramzan, "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 1, pp. 98–107, Jan. 2018.
- [55] G. E. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, no. 5, 2009, Art. no. 5947.
- [56] C. Cortes and V. Vapnik, "Support vector machine," *Mach. Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [57] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 199–210, Feb. 2011.
- [58] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a Kernel eigenvalue problem," *Neural Comput.*, vol. 10, no. 5, pp. 1299–1319, 1998.
- [59] R. Collobert, F. Sinz, J. Weston, and L. Bottou, "Large scale transductive SVMs," *J. Mach. Learning Res.*, vol. 7, no. Aug, pp. 1687–1712, 2006.
- [60] E. Sangineto, G. Zen, E. Ricci, and N. Sebe, "We are not all equal: Personalizing models for facial expression analysis with transductive parameter transfer," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 357–366.
- [61] Y. Li, W. Zheng, Z. Cui, and X. Zhou, "A novel graph regularized sparse linear discriminant analysis model for EEG emotion recognition," in *Proc. Int. Conf. Neural Inf. Process.*, 2016, pp. 175–182.



Tengfei Song received the BS degree in communication engineering from Hohai University, Jiangsu, China, in 2016. He is currently working towards the PhD degree in the Department of Information and Communication Engineering, Southeast University, China. His research interests include computer vision, machine learning, and pattern recognition.



Wenming Zheng (M'08) received the BS degree in computer science from Fuzhou University, Fuzhou, China, in 1997, the MS degree in computer science from Huaqiao University, Quanzhou, China, in 2001, and the PhD degree in signal processing from Southeast University, Nanjing, China, in 2004. Since 2004, he has been with the Research Center for Learning Science, Southeast University. He is currently a professor with the Key Laboratory of Child Development and Learning Science, Ministry of Education, Southeast University. His research interests include affective computing, pattern recognition, machine learning, and computer vision. He is an associated editor of the *IEEE Transactions on Affective Computing*, an associated editor of the *Neurocomputing* and also a member of associate editors-in-chief of the *Visual Computer*. He is a member of the IEEE.



Peng Song received the BS degree in EE from Shandong University of Science and Technology, China, in 2006, the ME and PhD degrees in EE both from Southeast University, China, in 2009 and 2014, respectively. He is currently an associate professor with the School of Computer and Control Engineering, Yantai University, China. From 2007 to 2008, he was a research intern with Microsoft Research Asia. From 2009 to 2011, he worked as a software engineer with Motorola. His current main research interests include affective

computing, speech signal processing, and machine learning. He is a member of the IEEE.



Zhen Cui received the PhD degree in computer science from the Institute of Computing Technology (ICT), Chinese Academy of Science (CAS), Beijing, in Jun. 2014. He was a research fellow with the Department of Electrical and Computer Engineering, National University of Singapore (NUS) from Sep. 2014 to Nov. 2015. He also spent half a year as a research assistant on Nanyang Technological University (NTU) from Jun. 2012 to Dec. 2012. Now he is a professor of Nanjing University of Science and Technology, China. His research interests include computer vision, pattern recognition and machine learning, especially focusing on deep learning, manifold learning, sparse coding, face detection/alignment/recognition, object tracking, image super resolution, emotion analysis, etc.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/csdl.