PAPER

Investigating EEG-based functional connectivity patterns for multimodal emotion recognition

To cite this article: Xun Wu *et al* 2022 *J. Neural Eng.* **19** 016012

View the article online for updates and enhancements.

Journal of Neural Engineering

PAPER

Investigating EEG-based functional connectivity patterns for multimodal emotion recognition

Xun Wu[1], Wei-Long Zheng[1,7] (iD), Ziyi Li[1] (iD) and Bao-Liang Lu[1,2,3,4,5,6,*] (iD)

1   Center for Brain-Like Computing and Machine Intelligence, Department of Computer Science and Engineering, Shanghai Jiao Tong University, 800 Dongchuan Rd., Shanghai 200240, People's Republic of China
2   Clinical Neuroscience Center, RuiJin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin 2nd Rd., Shanghai, 200020, People's Republic of China
3   RuiJin-Mihoyo Laboratory, RuiJin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin 2nd Rd., Shanghai 200020, People's Republic of China
4   The Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering, Shanghai Jiao Tong University, 800 Dongchuan Rd., Shanghai 200240, People's Republic of China
5   Brain Science and Technology Research Center, Shanghai Jiao Tong University, 800 Dongchuan Rd., Shanghai 200240, People's Republic of China
6   Qing Yuan Research Institute, Shanghai Jiao Tong University, 800 Dongchuan Rd., Shanghai 200240, People's Republic of China
7   Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, MA 02319, United States of America
*   Author to whom any correspondence should be addressed.

E-mail: bllu@sjtu.edu.cn

Abstract

Objective. Previous studies on emotion recognition from electroencephalography (EEG) mainly rely on single-channel-based feature extraction methods, which ignore the functional connectivity between brain regions. Hence, in this paper, we propose a novel emotion-relevant critical subnetwork selection algorithm and investigate three EEG functional connectivity network features: strength, clustering coefficient, and eigenvector centrality. *Approach.* After constructing the brain networks by the correlations between pairs of EEG signals, we calculated critical subnetworks through the average of brain network matrices with the same emotion label to eliminate the weak associations. Then, three network features were conveyed to a multimodal emotion recognition model using deep canonical correlation analysis along with eye movement features. The discrimination ability of the EEG connectivity features in emotion recognition is evaluated on three public datasets: SEED, SEED-V, and DEAP. *Main results.* The experimental results reveal that the strength feature outperforms the state-of-the-art features based on single-channel analysis. The classification accuracies of multimodal emotion recognition are $95.08 \pm 6.42\%$ on the SEED dataset, $84.51 \pm 5.11\%$ on the SEED-V dataset, and $85.34 \pm 2.90\%$ and $86.61 \pm 3.76\%$ for arousal and valence on the DEAP dataset, respectively, which all achieved the best performance. In addition, the brain networks constructed with 18 channels achieve comparable performance with that of the 62-channel network and enable easier setups in real scenarios. *Significance.* The EEG functional connectivity networks combined with emotion-relevant critical subnetworks selection algorithm we proposed is a successful exploration to excavate the information between channels.

1. Introduction

Emotion plays a crucial role in many aspects of our daily lives, such as social communication and decision-making. According to the Gartner hype cycle in 2019 [1], emotion artificial intelligence (AI) is one of the 21 emerging technologies that will significantly impact our society over the next 5 to 10 years. Emotion AI, also known as artificial emotional intelligence or affective computing [2], aims at enabling machines to offer the capabilities to recognize, understand, and process emotions. Compared with the rich studies on the motor brain-computer interface (BCI), the recently emerging affective BCI

(aBCI) [3] faces distinct challenges since the brain functional connectivity networks involving emotions are not well investigated [4]. The aBCI technology aims to advance the human-computer interaction systems with the assistance of various devices to detect the affective states from neurophysiological signals. Therefore, the major challenge facing emotion AI and aBCI in the primary stage lies in emotion recognition [5].

In recent years, extensive endeavors have been devoted to emotion recognition. Among a variety of emotion recognition approaches, the modalities used to detect affective states primarily comprise two categories: the external behavioral signals, including facial expression [6], speech [7], body language, etc and the internal physiological signals [8], containing electroencephalography (EEG) [9], electrocardiography (ECG) [10], respiration, galvanic skin response, etc. These two categories have their own prominent properties. The external behavioral signals outperform in terms of convenience of data collection, while the physiological signals are believed to be more objective and reliable in conveying emotions. As a result, multimodal emotion recognition has become the major trend, since it may leverage the complementary representation properties of different modalities. Nevertheless, most existing studies have focused on the fusion of visual and audio signals [11, 12], while few studies have combined the behaviors with physiological signals [13, 14].

Among the physiological modalities, EEG has exhibited outstanding performance in emotion recognition and is promising in elucidating the basic neural mechanisms underlying emotion [15, 16]. Moreover, the fusion of EEG and eye tracking data has been shown efficient in multimodal emotion recognition, with increasing interests among research communities [17, 18]. In this paper, we adopt the EEG signals, along with eye movement data or peripheral physiological signals, to classify different emotions.

Most existing studies on EEG-based emotion recognition have relied on single-channel analysis [15–18], where EEG features are independently extracted within each EEG channel in different brain regions. In contrast, studies on cognitive science and neuroimaging have demonstrated that emotion is a complex behavioral and physiological reaction that involves circuits in multiple cerebral regions [19]. In addition, studies in neuroscience and neuropsychiatry have revealed that patients with cognitive defect psychophysiological diseases such as autism, schizophrenia and major depressive disorder present decreased brain functional connectivity by both functional magnetic resonance imaging (fMRI) and EEG [20]. Furthermore, studies on neuroimaging based on fMRI have indicated that brain functional connectivity may offer the potential of representing the fingerprints in profiling individuals [21], as well as

the ability of individuals to sustain attention [22]. These results have provided evidence for the connection between cognition and brain functional connectivity. However, few studies have explored the emotion-relevant brain functional connectivity patterns. The study of emotion recognition from the perspective of the brain functional connectivity network remains to be further investigated and may eventually lead to the understanding of the underlying neurological mechanisms behind how emotions are processed in the brain.

In this paper, we aim to investigate the emotion-relevant brain functional connectivity patterns. We propose a novel emotion-relevant critical subnetwork selection algorithm to improve the recognition performance and evaluate the performance of the EEG connectivity feature for multimodal emotion recognition with respect to three public datasets: SEED [17], SEED-V [23], and DEAP [24]. Figure 1 depicts the overall framework of our proposed multimodal emotion recognition method. The main contributions of our work lie in the following aspects:
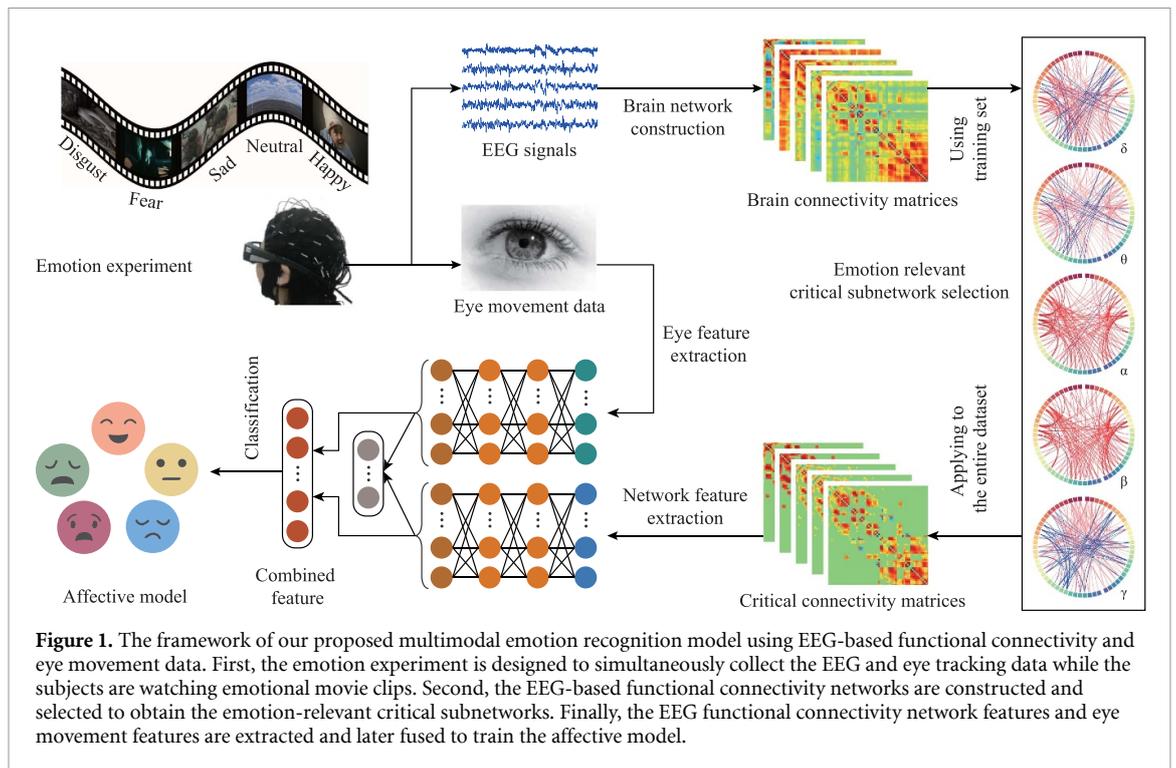
(a) We propose a novel emotion-relevant critical subnetwork selection algorithm to remove the weak and emotion-irrelevant associations in brain networks and investigate three EEG connectivity features: strength, clustering coefficient, and eigenvector centrality.

(b) We demonstrate the outstanding performance of the EEG connectivity feature and its complementary representation properties with eye movement data in multimodal emotion recognition.

(c) We reveal the emotion associated brain functional connectivity patterns and the potential of applying the brain networks based on fewer EEG electrodes to build aBCI systems in real scenarios.

The remainder of this paper is organized as follows. Section 2 introduces the related literature regrading multimodal emotion recognition and brain functional connectivity analysis. Section 3 describes the emotion experimental design. Section 4 presents the proposed multimodal emotion recognition framework based on the brain functional connectivity. Section 5 analyzes the experimental results and describes the future work. Finally, a brief conclusion will be presented in section 6.

2. Related work

2.1. Emotion recognition

Various modalities have been exploited to detect affective states in the past few decades. With the advent of computer vision and speech recognition,

Figure 1. The framework of our proposed multimodal emotion recognition model using EEG-based functional connectivity and eye movement data. First, the emotion experiment is designed to simultaneously collect the EEG and eye tracking data while the subjects are watching emotional movie clips. Second, the EEG-based functional connectivity networks are constructed and selected to obtain the emotion-relevant critical subnetworks. Finally, the EEG functional connectivity network features and eye movement features are extracted and later fused to train the affective model.

research on emotion recognition using facial expression and speech has gained prevalence [6, 7]. Hasani and Mahoor [25] proposed an enhanced neural network architecture that consists of a 3D version of the Inception-ResNet network followed by a long short-term memory (LSTM) unit for emotion recognition from facial expressions in videos. They employed four databases in classifying different emotions, including anger, fear, disgust, sadness, neutrality, contempt, happiness, and surprise. Trigeorgis *et al* [26] presented an end-to-end speech emotion recognition framework created by combining the convolutional neural network (CNN) and LSTM models. Schirmer and Adolphs [27] studied different modalities with respect to emotion perception, including facial expression, voice, and touch. The authors suggested that the mechanisms of these modalities have their own specializations and that together, they could lead to holistic emotion judgment.

Apart from the external behavioral modalities, the internal physiological signals have also attracted the attention of numerous researchers due to their objectivity and reliability. Zhang *et al* [28] conducted respiration-based emotion recognition using the sparse autoencoder and logistic regression model. Nardelli *et al* [29] performed valence-arousal emotion recognition based on the heart rate variability derived from the ECG signals. Atkinson and Campos [30] improved the EEG-based emotion recognition by combining the mutual information-based feature selection methods with kernel classifiers. Liu *et al* [31] constructed a real-time movie-induced emotion

recognition system to continuously detect the discrete emotional states in the valence-arousal dimension. Zheng *et al* [15] systematically evaluated the performance of different feature extraction, feature smoothing, feature selection and classification models for EEG-based emotion recognition. Their results indicated that stable neural patterns do exist within and across sessions. Among these modalities, EEG has been proven to be promising for emotion recognition and demonstrates competence for revealing the neurological mechanisms behind emotion processing.

In EEG-based emotion recognition, numerous EEG features have been exploited to enhance the performance of aBCI systems. The conventional EEG features could be categorized into temporal domain, frequency domain, and time-frequency domain [32]. In the temporal domain, the most commonly used EEG features mainly include the fractal dimension and higher order crossings [33]. Due to the non-stationary essence of the EEG signals and the fact that raw EEG signals are usually contaminated with artifacts and noises, the frequency domain features such as power spectral density (PSD) [34], higher order spectra [33], and differential entropy (DE) [35] and the time-frequency domain features such as wavelet features [32] and Hilbert-Huang spectra [33, 36] have demonstrated outstanding performance in the EEG-based emotion recognition systems. However, these conventional EEG feature extraction methods are based on single-channel analysis, which neglects the EEG-based functional connectivity networks in association with different emotions.

2.2. Brain functional connectivity

Brain connectivity has long been studied in the fields of neuroscience and neuroimaging to explore the essential nature of the cerebrum. According to the attributes of connections, brain connectivity could be classified into three modes: structural connectivity, functional connectivity, and effective connectivity [37]. These modes separately correspond to the biophysical connections between neurons or neural elements, the statistical relations between anatomically unconnected cerebral regions, and the directional causal effects from one neural element to another.

Recently, increasing evidence has indicated that a link does exist between brain functional connectivity and multiple psychophysiological diseases with cognitive deficiency. Murias *et al* [38] found that robust patterns of EEG connectivity are apparent in autism spectrum disorders in the resting state. Yin *et al* [39] concluded that the EEG-based functional connectivity in schizophrenia patients tends to be slower and less efficient. Ho *et al* [40] indicated that adolescent depression typically relates to the inflexibly elevated default mode network connections based on fMRI. Whitton *et al* [41] suggested that elevations in high frequency EEG-based functional connectivity may represent a neural pattern for the recurrent illness course of major depressive disorder. Xu *et al* [42] compared two coherence-based EEG functional connectivity network features, statistical properties and common spatial pattern (SPN), between psychogenic nonepileptic seizures (PNES) and epilepsy patients based on fisher discriminate analysis (FDA) and support vector machine (SVM), through which they demonstrate the feasibility of SPN in distinguishing PNES from epilepsy.

In the past years, a growing number of preliminary research efforts on EEG-based emotion recognition have attempted to employ the brain functional connectivity networks. Dasdemir *et al* [43] directly used the connectivity metric of phase locking value (PLV) as the EEG feature in distinguishing the positive and negative emotions. Lee and Hsieh [44] tested three different connectivity metrics, correlation, coherence, and phase synchronization index, in classifying the positive, neutral, and negative emotions. Li *et al* [45] also studied these three emotions by combing the PLV-based network features with 6 power-based features utilized in [15]. F-score was applied to select features with strong discriminative ability. Al-Shargie *et al* [46] utilized similar strategy. They concatenated PLV-based network features and PSD features calculated from the autoregressive model parameters to distinguish the positive, neutral, and negative emotions.

Some studies applied deep learning methods to the connectivity feature extraction or the brain network construction processes. Moon *et al* [47] utilized CNN to model the connectivity matrices constructed by three different connectivity metrics: correlation, PLV, and phase lag index. Chao and colleagues proposed a deep learning framework with a multiband feature matrix and a capsule network to capture the frequency domain and spatial characteristics in multichannel EEG [48]. Hua *et al* [49] proposed a new method to construct brain connectivity network. They conveyed the eight-channel EEG signals to the stacked autoencoder and calculated the Euclidean distance using the values from hidden layers to form a network.

However, these studies mentioned above ignore the optimization of network features for emotion recognition task. Emotion-related information is not fully utilized neither in the network construction nor the feature extraction processes. Whether there truly exist specific connectivity patterns for different affective states remains to be elucidated. In our previous study on EEG-based emotion recognition [50], we identified the brain functional connectivity patterns of the three emotions (sad, happy and neutral) and extracted the topological features from the brain networks to recognize these emotions. In this paper, we extend this preliminary work to the three-class (sad, happy, and neutral), five-class (disgust, fear, sad, happy, and neutral), and valence-arousal dimension multimodal emotion recognition tasks.

2.3. Eye movement analysis

Studies in neuroscience and biological psychology have indicated the relation between emotion and eye movement data, especially pupil diameter and dilation response. Widmann *et al* [51] indicated that emotional arousal by novel sounds is reflected in the pupil dilation response and the P3 event-related potentials. Oliva and Anikin [52] suggested that the pupil dilation response reveals the perception of emotion valence and confidence in the decision-making process. Moreover, Black *et al* [53] showed that the eye tracking and EEG data in autism spectrum disorders are atypical during the processes of attention to and cognition of facial emotions.

In addition, eye movement data could be obtained through eye tracking glasses which are wearable, portable and noninvasive. Therefore, eye movement data, as a behavioral reaction to emotions, have been widely utilized to assist with EEG-based emotion recognition in aBCI systems. López-Gil *et al* [18] improved EEG-based emotion recognition by combining eye tracking and synchronized biometrics to detect the valence-arousal basic emotions and a complex emotion of empathy. Zheng *et al* [54] evaluated the complementary characteristics of EEG and eye movement data in classifying positive, neutral and negative emotions by fusing the DE and pupil diameter features. Lu *et al* [17] extended this preliminary work and systematically examined sixteen different eye movement features. Furthermore, their work has been

extended to the five emotions by Li *et al* [23] and Zhao *et al* [55], and the discrimination ability and stability over time of EEG and eye tracking data were also revealed. However, these research approaches were all based on single-channel analysis for the EEG signals: whether there exist complementary representation properties of EEG connectivity features and eye movement data remains to be further analyzed.

2.4. Multimodal frameworks

As a complex psychological state, emotion is reflected in both physical behaviors and physiological activities [19, 56]. The collection of external behavioral data is more convenient than that of internal physiological signals, since the procedure could be accomplished without involving any invasive devices. Despite the inconvenience of data collection, the physiological signals are believed to be more objective and reliable because they are directly controlled by the peripheral or central nervous systems and forging signals such as EEG consciously seems unfeasible.

With different modalities exhibiting distinct properties, modern emotion recognition approaches have the tendency of combining multiple modalities to enhance the performance of aBCI systems. Perez-Gaspar *et al* [57] extended the evolutionary computation of artificial neural networks and hidden Markov models in classifying four emotions (angry, sad, happy, and neutral) by combining the speech with facial expressions. Tzirakis *et al* [58] also fused the auditory and visual modalities using an end-to-end valence-arousal emotion recognition model. They applied the CNN and ResNet models to extract features from speech and visual signals, respectively, which were then concatenated and fed into the LSTM model to accomplish the end-to-end training manner. Ranganathan *et al* [59] conducted a 23-class discrete emotion recognition task based on four different deep belief networks by combining a variety of modalities, including face, gesture, voice and physiological signals. Huang *et al* [60] studied the fusion of EEG and facial expression data using two decision-level fusion strategies, the sum and production rules, in detecting the four basic emotions (fear, sad, happy, and neutral).

In recent years, many researchers have suggested that the combination of EEG and eye tracking data is a promising approach for recognizing emotions in aBCI systems. López-Gil *et al* [18] combined EEG with eye tracking and biometric signals in a synchronized manner to classify emotions using multiple machine learning methods. Liu *et al* [61] applied the bimodal deep autoencoder (BDAE) neural network to fuse EEG and eye movement features by extracting the shared representations of these two modalities and detected the three basic valences for emotion (positive, neutral, and negative). Tang *et al* [62] conducted the same task using bimodal deep denoising autoencoder and bimodal-LSTM models. Zheng *et al* [63]

presented EmotionMeter for detecting the four emotions (fear, sad, happy, and neutral). Qiu *et al* [64] adopted the deep canonical correlation analysis (DCCA) model as a multimodal feature fusion network for classifying the three-class, four-class, and valence-arousal emotions. Their results suggested that DCCA outperforms BDAE and bimodal-LSTM models in multimodal emotion recognition. In this paper, we apply the DCCA model to address the multimodal emotion recognition task.

3. Emotion experiment design

3.1. Stimuli

The emotion experiments were designed to simultaneously record the EEG and eye movement signals of the five prototypical emotions (disgust, fear, sad, happy, and neutral). Many existing works have indicated the efficiency and reliability of movie clips in eliciting the subjects' emotions due to the blending of audio and visual information [17, 63]. Therefore, the movie clips were selected as the type of stimuli to better induce the subjects' affective states.

During the preliminary experiment, a stimuli pool containing emotional movie clips corresponding with the five emotions was prepared and then assessed by 20 volunteers using rating scores ranging from 0 to 5. The higher scores represented the more successful elicitation of the subjects' emotions. Eventually, 9 movie clips for each of the five emotions were selected from the stimuli pool, all of which received a mean score of 3 or higher. The durations of these clips range from 2 to 4 min.

3.2. Subjects

Sixteen subjects (six males and ten females) aged between 19 and 24 (mean: 21.62, std.: 2.48) with normal hearing and self-reported normal or corrected-to-normal vision were recruited for our emotion experiments. All subjects were the students from Shanghai Jiao Tong University and selected using the Eysenck Personality Questionnaire (EPQ), which could measure the personality of an individual in three independent dimensions: Extroversion/Introversion, Neuroticism/Stability, and Psychoticism/Socialization [65]. Those with extroverted characteristics and stable mood are more readily induced to experience the intended emotions throughout the experiment in comparison with those of other personalities. Hence, subjects that are more appropriate for the emotion experiments were selected according to the EPQ feedback.

3.3. Protocol

The emotion experiments were conducted in a laboratory environment. During the emotion experiment, the subject was required to view the emotional movie clips and relaxed as much as possible to induce
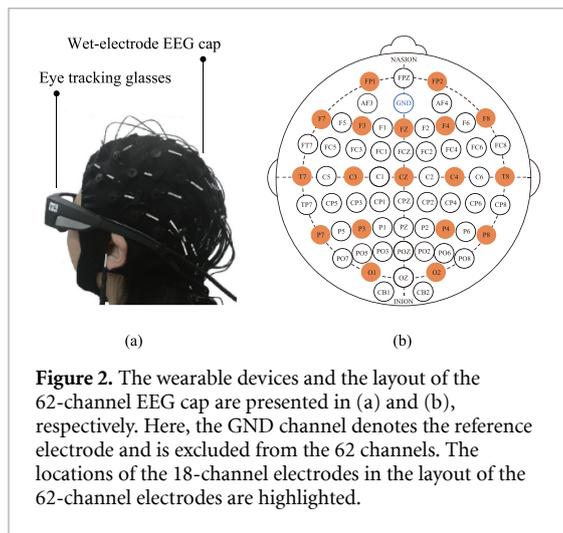
Figure 2. The wearable devices and the layout of the 62-channel EEG cap are presented in (a) and (b), respectively. Here, the GND channel denotes the reference electrode and is excluded from the 62 channels. The locations of the 18-channel electrodes in the layout of the 62-channel electrodes are highlighted.

Figure 3. The protocol of the designed emotion experiment.

their emotions. Meanwhile, their EEG and eye movement signals were simultaneously collected by the 62-channel wet-electrode cap and the SMI eye tracking glasses, respectively. The EEG data were recorded with the ESI NeuroScan System at a sampling rate of 1000 Hz. The layout of the 62-channel EEG cap is based on the higher-resolution international 10–20 system. Figure 2 presents these wearable devices and the layout of the 62-channel EEG cap. The reference channel is set between CPZ and CZ. All of the electrode impedances were kept below 5000Ω. The vertical electrooculogram (VEO) signal was acquired simultaneously by the NeuroScan System to facilitate artifact elimination.

In this paper, there were 15 trials in total in each experiment, where each of the five emotions corresponds to 3 movie clips. Moreover, each subject was required to perform three sessions of the experiment on different days with an interval longer than three days. To better elicit the subjects' emotions, there was no repetition of movie clips within or across the three sessions. Thus, the aforementioned 9 movie clips for each emotion were randomly divided into three groups and later constructed the three sessions. As studies in cognitive science have indicated that emotion varies in a fluent and smooth manner, the order of play of these movie clips in one experiment was elaborately designed according to the following criteria: (a) avoiding sudden changes in emotion, such as clips of the emotion of disgust followed by clips of happiness; (b) utilizing movie clips of neutral emotion as a cushion between two opposite emotions.

Figure 3 illustrates the protocol of the designed emotion experiment. During each trial of the experiment, the movie clip was guided by 15 s of a brief introduction about the content and the emotion to be elicited and ended with 15 or 30 s of self-assessment and relaxation for the subjects to mitigate their emotions. Particularly, the resting time was 30 seconds after disgust or fear emotions, and 15 s after the other three emotions. In general, the duration of
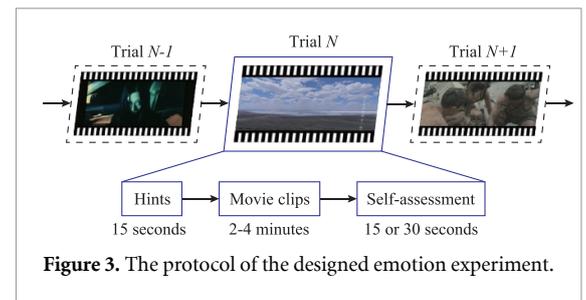
one experiment was approximately 55 min. The self-assessment required subjects to rate their arousal level evoked by the stimuli from 0 to 5. Here, 0 indicates no corresponding emotion is evoked while 5 indicates the strongest corresponding emotion is evoked; 16 subjects rated 45 (15×3) different trails. There are in total 720 (16×45) rating sheets of which only 58 rating sheets have scores less than 3, demonstrating the effectiveness of our emotional stimuli.

This multimodal emotion dataset is named SEED-V, of which the V means there are five emotion included in SEED-V. SEED-V was firstly introduced in our previous conference paper [23]. We provide a more systematic introduction about it in this paper. SEED-V is an expansion of our public emotion EEG dataset SEED[8]. The SEED dataset [17], which was proposed in 2015 contains 62-channel EEG signals and eye movement data corresponding to three emotions (sad, happy, and neutral) from nine subjects five males and four females, mean: 23.27, std.: 2.37), with each subject performing the experiments three times. Thus, there are 27 experiments in total and each experiment contains 15 trials, with 5 movie clips for each of the three emotions.

3.4. Ethic statement

The emotion experiments have been approved by the Scientific & Technical Ethics Committee of the Bio-X Institute at Shanghai Jiao Tong University. All of the subjects participating in our experiments were informed of the experimental procedures and signed the informed consent document before the emotion experiments.

4. Methodology

4.1. Preprocessing

The SEED-V dataset follows the same preprocessing process as SEED. The raw EEG signals collected during the emotion experiments are usually of high resolution and contaminated by surrounding artifacts, which hampers both the processing and the analysis of the emotion-relevant brain neural activities. We applied the Curry 7 software to remove the irrelevant artifacts. First, we adopted the Constant option in the Baseline block of Curry 7 to do the baseline

---

[8] http://bcmi.sjtu.edu.cn/home/seed/.

correction. Second, in the Filter Parameter block, a bandpass filter between 1 and 50 Hz and a notch filter of 50 Hz was applied. Finally, the artifacts caused by eye movements were eliminated based on VEO signal in the Artifact Reduction block. Curry 7 can automatically locate blink artifacts in each channel according to VEO signal and remove them using Principal Component Analysis (PCA). We needed to visually inspect the time period of each located blink artifact and chose the best location parameters for each subject.

Then, we downsampled the EEG signals to 200 Hz to expedite the processing procedures. For further exploration of the frequency-specific brain functional connectivity patterns, the EEG data were filtered with the five bandpass filters corresponding to the five frequency bands (δ: 1–4 Hz, θ: 4–8 Hz, α: 8–14 Hz, β: 14–31 Hz, and γ: 31–50 Hz).

It has been proven that pupil diameter is subject to the ambient luminance in addition to the emotion stimuli materials [66]. Fortunately, according to our observation and analysis, the pupil diameter exhibits its consistency in response to the same emotional stimuli material across different subjects. Thus, principal component analysis was adopted to eliminate the luminance reflex of the pupil and to preserve the emotion-relevant components [17, 54, 63].

4.2. Brain functional connectivity network

The EEG-based brain functional connectivity networks consist of vertices and edges, which could be represented by the EEG electrodes and the associations between pairs of EEG signals from two different channels, respectively [67].

4.2.1. Vertex selection

Although the wet-electrode EEG cap was adopted in our emotion experiments within the laboratory environment due to its reliability, the dry-electrode device with fewer EEG channels offers great convenience and portability in developing aBCI systems under actual scenario conditions. Thus, the question of whether the brain functional connectivity networks comprised of fewer EEG channels could exhibit considerable performance in emotion recognition remains unexplored.

Tong and colleagues [68] conducted two experimental paradigms: a laboratory experiment which utilize a 62-channel wet-electrode cap based on ESI NeuroScan System and a real-scenario experiment which utilize an 18-channel dry-electrode cap based on the DSI-24 System. They demonstrated that the DSI-24 wearable sensing EEG headset is quite portable and appropriate for real scenarios. Inspired by this work, we wonder if the proposed model using 18 channels can achieve comparable results compared with 62 channels based on ESI NeuroScan System.

The results may facilitate the reduction in the number of electrodes in real-scenario emotion recognition research. However, it should be mentioned that the influence of electrode sensitivities on signal quality needs to be taken into account when applying 18 channels on other wearable EEG devices.

The layout of DSI-24 headset electrodes conform to the normal international 10–20 system. Fortunately, these 18 electrodes can be perfectly mapped into the same locations in the layout of the higher-resolution international 10–20 system. Figure 2(b) displays the locations of the 18-channel dry-electrode reflected in the layout of the 62-channel wet-electrode device.

In this paper, since the raw EEG signals were acquired with the 62-channel wet-electrode device, we constructed the brain connectivity network with 62 vertices in total. Furthermore, we selected these 18 electrodes and compared the performances of EEG connectivity features extracted from the brain networks constructed with two categories of vertices: 18-channel and 62-channel.

4.2.2. Edge measurement

To measure the associations between pairs of EEG signals recorded from different channels, we compared two connectivity metrics in this paper: Pearson's correlation coefficient and spectral coherence. The former can measure the linear relation between two EEG signals x and y, which is defined as:

$$\rho_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y}, \qquad (1)$$

where $\text{cov}(x,y)$ denotes the covariance between x and y, and σ_x and σ_y are the respective standard deviations.

Distinguished from the correlation that measures the connectivity between two EEG channels in the temporal domain, coherence can measure the connectivity between two signals x and y at frequency f in the frequency domain, which could be written as:

$$C_{x,y}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)}, \qquad (2)$$

where $P_{xy}(f)$ is the cross power spectral density between x and y, and $P_{xx}(f)$ and $P_{yy}(f)$ are the respective power spectral densities.

4.2.3. Network construction

In some previous work, the complementary characteristics of the state-of-the-art DE feature and eye movement data in emotion recognition have been revealed, where features were extracted with a 4-second nonoverlapping time window in SEED and SEED-V dataset, while DEAP dataset used a 2-second nonoverlapping time window [17, 55, 63, 69, 70]. Considering further fusion of our proposed EEG
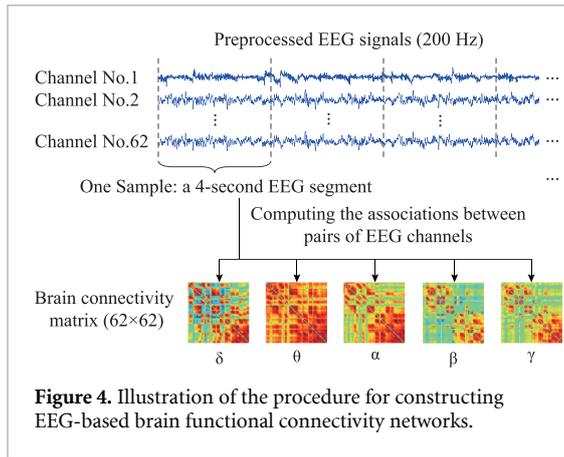
Figure 4. Illustration of the procedure for constructing EEG-based brain functional connectivity networks.

connectivity feature with eye movement data and comparison with the DE feature, we constructed the brain functional connectivity network using the same time window. Taking SEED and SEED-V dataset as example, the detailed procedure of brain network construction is depicted in figure 4.

First, the preprocessed EEG signals were segmented with a 4-second nonoverlapping time window. As a result, each sample is represented by a 4-second EEG segment of 62 channels. Since the five bandpass filters were employed during the preprocessing, each sample actually contains five 4-second EEG segments corresponding to the five frequency bands.

Second, the associations between pairs of EEG channels were computed using the connectivity metric (correlation or coherence) in each frequency band for each sample. Therefore, each brain network corresponds to a 62×62 symmetric connectivity matrix, where elements denote the association weights between pairs of EEG channels. Since the value of self-correlation is always equal to 1, elements on the main diagonal of connectivity matrices are usually set to zero and will not be used in later analysis [37].

Finally, five brain connectivity networks were acquired for each sample, corresponding to the five frequency bands. For the brain networks constructed with 18 channels, there is no need to repeat the above procedures. We could directly select the corresponding elements in the 62×62 connectivity matrix to construct the 18×18 connectivity matrix.

4.3. Emotion-relevant critical subnetwork selection

Although the raw EEG signals were preprocessed to remove the noises, there still remain certain minor artifacts that may not be eliminated during the preprocessing phase. Unfortunately, these artifacts may further lead to the weak associations in the brain networks, which eventually results in obscureness in profiling the brain network topology. By convention, this problem is resolved by directly discarding these weak associations according to an absolute or a proportional threshold after sorting the association

weights [37, 67]. However, this method fails to take the targeted task into consideration, thus offering no guarantee that the preserved stronger associations are truly task relevant. Therefore, we have proposed an emotion-relevant critical subnetwork selection approach to address this issue.

The goal is to explore the universal emotion-relevant brain functional connectivity patterns among different subjects. Therefore, we utilized samples in training sets of all subjects to select the emotion-relevant critical subnetworks. Nevertheless, it should be mentioned that the affective models trained in this paper are subject-dependent. A subject-independent adaptation model still needs further research to fully explore the EEG-based functional connectivity patterns across subjects [71–75]. For analysis of the brain connectivity patterns in different frequency bands, we selected a total of five critical subnetworks corresponding to the five frequency bands. Assuming that L is the set of emotion labels, there are thus $|L|$ categories of emotions. The emotion-relevant critical subnetwork selection approach for one specific frequency band is summarized into the following three phases:

(a) Averaging phase: all brain networks in training sets are averaged over all samples and all subjects for each emotion: thus, $|L|$ averaged brain networks corresponding to $|L|$ emotions could be obtained.
(b) Thresholding phase: for each averaged brain network, the same proportional threshold is applied to solely preserve the strongest associations. Thus, we attain critical edges for each emotion.
(c) Merging phase: along with the original vertices, the critical edges in all $|L|$ averaged brain networks are merged together to construct the emotion-relevant critical subnetwork.

Here, the proportional threshold represents the proportion of the preserved connections relative to all connections in the brain network. Particularly, the threshold value was tuned as a hyperparameter. Specifically, we utilized SVM to select the best threshold values for 18-channel model and 62-channel model, respectively. The selection range is [0.0, 1.0] with step size of 0.01. Details of threshold tuning is described in section 5.1.

This procedure is also described in algorithm 1. Suppose that the connectivity matrices X along with corresponding labels Y in the training sets for one frequency band are defined as:

$$X = \{x^i\}_{i=1}^M \ (x^i \in \mathbb{R}^{N \times N})$$
$$Y = \{y^j\}_{i=1}^M \ (y^j \in L), \tag{3}$$

where M and N represent the number of samples and the number of vertices, respectively.

Algorithm 1. The proposed emotion-relevant critical subnetwork selection algorithm in a frequency band.

Input: Connectivity matrices X, labels Y, label set L, threshold t, and vertices V

Output: The emotion-relevant critical subnetwork G^*

1: **for** each $c \in L$ **do**

2:      Average matrices with the same emotion label c:
$$x_c = mean_{y^i==c}(x^i)$$

3:      Sort the upper triangular elements of matrix x_c based on the absolute value of the association weights:
$$x_c^* = sort(abs(triu(x_c)))$$

4:      Derive the critical edges corresponding to the strongest associations using threshold t:
$$E_c = index(x_c^*(1 : t * N * (N-1)/2))$$

5: **end for**

6: Merge the critical edges of all emotions:
$$E^* = union_{c \in L}(E_c)$$

7: Construct the emotion-relevant critical subnetwork:
$$G^* = (V, E^*)$$

8: **return** G^*

4.4. Feature extraction

4.4.1. EEG functional connectivity network features

In this paper, we extracted EEG features from the perspective of brain functional connectivity networks. The essence of the emotion-relevant critical subnetwork primarily consists in the network topology. According to the five critical subnetworks in the five corresponding frequency bands, we could derive the critical connectivity matrices for each sample in the entire dataset. Precisely, if one edge belongs to the critical subnetwork, the corresponding association weight in the matrix will remain unmodified; otherwise, it will be set to zero, thus simulating the process of discarding this edge from the brain network. The critical connectivity matrices were subsequently fed into the Brain Connectivity Toolbox [37] to extract the three topological features: strength, clustering coefficient, and eigenvector centrality.

Assume that the selected emotion-relevant brain functional connectivity network of one sample is regarded as an indirect graph $G = (V, E^*)$, where V and E^* represent the sets of vertices and critical edges, respectively. There are in total N vertices in the brain network. Suppose that the corresponding symmetric connectivity matrix of G is $A = (a_{ij})$, where $i, j = 1, 2, \ldots, N$, a_{ij} denotes the association between two vertices v_i and v_j, and $a_{ij} = a_{ji}$. According to [37], we could provide rigorous definitions for the three EEG functional connectivity network features as below.

The strength feature F_S is a basic measurement of the network topology, which could be written as:

$$F_S = \left[s_{i+}\big|_{i=1}^N, s_{i-}\big|_{i=1}^N, \sum_{i=1}^N s_{i+}, \sum_{i=1}^N s_{i-} \right], \quad (4)$$

where s_{i+} and s_{i-} represent the sum of the positive and negative associations connected to vertex v_i, respectively, and are computed as:

$$s_{i+} = \sum_{\substack{a_{ij}>0 \\ j=1,2,\ldots,N}} a_{ij}, \quad (5)$$

$$s_{i-} = \sum_{\substack{a_{ij}<0 \\ j=1,2,\ldots,N}} a_{ij}. \quad (6)$$

The clustering coefficient feature F_C is a measurement of the brain functional segregation that primarily quantifies the clusters within the brain network, which is defined as:

$$F_C = \left[c_{i+}\big|_{i=1}^N, c_{i-}\big|_{i=1}^N, \sum_{i=1}^N c_{i+}, \sum_{i=1}^N c_{i-} \right], \quad (7)$$

where c_{i+} and c_{i-} represent the clustering coefficient vector for the positive and negative associations of vertex v_i, respectively. The clustering coefficient is equivalent to the fraction of triangles around a vertex and is calculated as:

$$c_{i+} = \frac{2t_{i+}}{k_i(k_i - 1)}, \quad (8)$$

$$c_{i-} = \frac{2t_{i-}}{k_i(k_i - 1)}, \quad (9)$$

where k_i denotes the total numbers of neighbors for vertex v_i, and t_{k+} and t_{k-} are the positive and negative weighted geometric means of triangles around v_i, respectively. The triangles around a vertex are represented as:

$$t_{i+} = \sum_{\substack{j,h \in M_i \\ a_{ij}, a_{ih}, a_{jh} > 0}} (a_{ij} a_{ih} a_{jh})^{1/3}, \quad (10)$$

$$t_{i-} = \sum_{\substack{j,h \in M_i \\ a_{ij}, a_{ih}, a_{jh} < 0}} (a_{ij} a_{ih} a_{jh})^{1/3}, \quad (11)$$

where $M_i = \{v_j | e_{ij} \in E^*\}$ is the neighborhood of vertex v_i.

The eigenvector centrality feature F_E evaluates the significance of an individual vertex in interacting with other vertices and thus serving a crucial role in network resilience. A vertex is more central and influential if the vertices it connects have high centrality. The eigenvector centrality score of vertex v_i could be defined as:

$$F_E(i) = \frac{1}{\lambda} \sum_{j=1}^N a_{ij} F_E(k). \quad (12)$$

The eigenvector centrality of vertex v_i is proportional to the average of centrality values of its neighbors. This equation could be easily transformed to the eigenvector equation using the vector notations:

$$\mathbf{A} F_E = \lambda F_E, \quad (13)$$

where we define λ is the maximal eigenvalue of $\mathbf{A}$ [76].

In general, the dimensions of the strength, clustering coefficient, and eigenvector centrality features in each frequency band are $2N+2$, $2N+2$, and N, respectively.

4.4.2. Eye movement features

First, eye movement parameters were calculated using the BeGaze[9] analysis software of the SMI eye tracking glasses, including pupil diameter, fixation duration, blink duration, saccade, and event statistics. Subsequently, the statistics of these eye movement parameters were derived, thus obtaining the 33-dimensional eye movement feature. The detailed description of the extracted eye movement feature could be found in the previous work [17, 55].

4.5. Classification

As aforementioned, the variation of emotion is fluent and smooth, which should be reflected in the attributes of the extracted features. In addition, the extracted EEG features are typically of high dimensionality and may contain unrelated and redundant information, which increases the unnecessary computation and time costs. Hence, a feature smoothing method, linear dynamical system (LDS) [77] was applied to tackle this issue before feeding features into the feature fusion model DCCA.

4.5.1. Deep canonical correlation analysis model

The DCCA model can fuse features from multimodal data by learning the most relevant features and forming the shared representations [78]. Figure 5 presents the architecture of the DCCA model, which comprises three parts: the stacked nonlinear layers (L2 and L3), CCA calculation, and feature fusion layer.

Assume that the transformed features for two modalities X_1 and X_2 are separately denoted by $H_1 = f_1(X_1; \theta_1)$ and $H_2 = f_2(X_2; \theta_2)$, where f_1 and f_2 are the respective nonlinear transformations, and θ_1 and θ_2 are the corresponding parameters. Thus, the optimization function is written as:

$$(\theta_1^*, \theta_2^*) = \underset{(\theta_1, \theta_2)}{\arg\max}\, \mathrm{corr}\big(f_1(X_1; \theta_1), f_2(X_2; \theta_2)\big). \quad (14)$$

Suppose that the centered data matrices are $\bar{H}_1$ and $\bar{H}_2$, and r_1 and r_2 are the respective regularization parameters: hence, the correlation of the transformed features could be calculated as:

$$\mathrm{corr}(H_1, H_2) = ||T||_{\mathrm{tr}} = \mathrm{tr}(T'T)^{1/2}, \quad (15)$$

where

$$T = \hat{\Sigma}_{11}^{-1/2} \hat{\Sigma}_{12} \hat{\Sigma}_{22}^{-1/2},$$
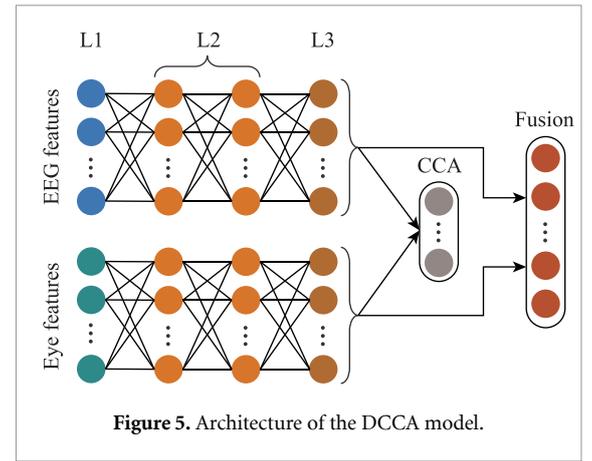$$\hat{\Sigma}_{11} = \frac{1}{m-1}\bar{H}_1\bar{H}_1' + r_1 I,$$

Figure 5. Architecture of the DCCA model.

$$\hat{\Sigma}_{22} = \frac{1}{m-1}\bar{H}_2\bar{H}_2' + r_2 I,$$
$$\hat{\Sigma}_{12} = \frac{1}{m-1}\bar{H}_1\bar{H}_2'. \quad (16)$$

In particular, the gradient of $\mathrm{corr}(H_1, H_2)$ could be computed using singular value decomposition. The parameter updating is accomplished by using the negative value of correlation as the loss function. Thus, minimizing loss is equivalent to maximizing correlation. The feature fusion layer is defined as the weighted average of the two transformed features [64]. Finally, the fused multimodal feature is fed into the SVM to train the affective model.

In this paper, the cross validation and grid search methods were adopted to tune the hyperparameters in different experiments. The search scope was defined as follows: Supposing that the numbers of nodes in L1, L2, and L3 layers of DCCA are n_1, n_2, and n_3, respectively, these three hyperparameters are searched in the space where $n_1 \geqslant n_2 \geqslant n_3$ and $n_1, n_2, n_3 \in \{32, 64, 128, 256\}$. So n_1, n_2, n_3 have 20 combinations in total. The learning rate is tuned from set $\{10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}\}$. These hyperparameters were optimized according to experiments.

4.5.2. Experiment setups

In this paper, we evaluated the proposed approaches on three public datasets: SEED [17], SEED-V [23], and DEAP [24]. For model evaluation, we used different cross validation methods on different databases with consideration of consistency with previous studies and the amount of emotion data available. For the SEED dataset, the three-class (sad, happy, and neutral) emotion classification task was conducted. The training and test sets were the first 9 trials and the last 6 trials, respectively, which was the same as in [17, 61, 62, 69]. For the SEED-V dataset, the five-class (disgust, fear, sad, happy, and neutral) emotion classification task was performed with a three-fold cross validation strategy, which followed the same setups as in [55, 69]. SEED and SEED-V followed the same experiment protocol for data collection and several sessions with the same participant were done

on different days in both datasets, which was one of the advantages over DEAP to study the stable neural patterns over time. The preprocessing of EEG and eye movement signals was the same in SEED and SEED-V.

The DEAP dataset contains 32-channel EEG signals and 8-channel peripheral physiological signals from 32 subjects (16 males, 16 females) aged between 19 and 37 (mean age 26.9) with labels in the valence-arousal dimension. Each subject watched 40 one-minute music videos. The EEG signals were preprocessed with a bandpass filter between 4 and 45 Hz. For the DEAP dataset, we built the brain networks using solely 32 channels in the four frequency bands (without the δ band) with a 2-second nonoverlapping time window. The peripheral physiological feature was 48-dimensional. In addition, two binary (arousal-level, valence-level) classification tasks were conducted with a ten-fold cross validation strategy. We divided the continuous arousal and valence dimensions into two categorical levels (low/high) with the threshold of 5. The DEAP database includes continuous levels of arousal, valence, like/dislike, dominance, and familiarity. However, we focused on the arousal and valence dimensions in this study, which are the most common components in emotion models. The setups for the DEAP dataset are in accordance with [61, 62, 69].

5. Experimental results and discussion

5.1. Threshold tuning for subnetwork construction
As mentioned in section 4.3, the proportional threshold is applied to preserve the strongest associations in the averaged brain networks. It is a crucial value for subnetwork construction. We utilized SVM to select the best threshold values. The hyperparameter penalty coefficient C of SVM is tuned from 2^{-10} to 2^{10} with the step size of the index setting to 1. Taking SEED-V as example, the best thresholds differ according to the categories of edge measurements and vertices, so the threshold tuning process should be applied in the following four cases: 18-channel network with correlation edges, 64-channel network with correlation edges, 18-channel network with coherence edges and 64-channel network with coherence edges. To reduce time overhead, the threshold tuning includes two steps:

(a) Rough tuning: The threshold range is set as [0.0, 1.0] with step size of 0.1. The best threshold a found in the rough tuning process will be utilized to construct the threshold search space in fine tuning.

(b) Fine tuning: The threshold range is set as $[a − 0.1, a + 0.1]$ with step size of 0.01. Threshold that achieves the highest accuracy rate is selected as the final threshold a_{final}.
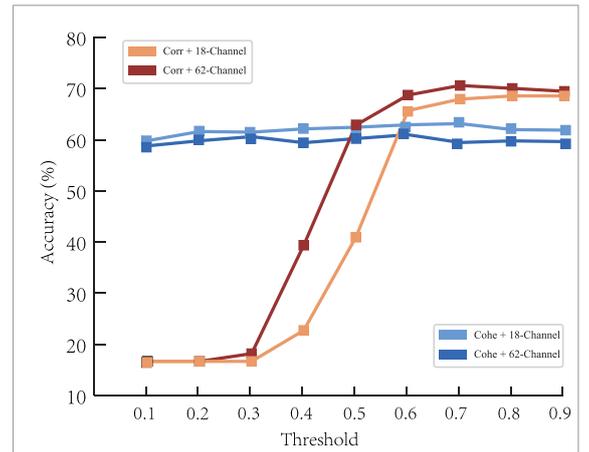

Figure 6. Threshold rough tuning results of SEED-V.
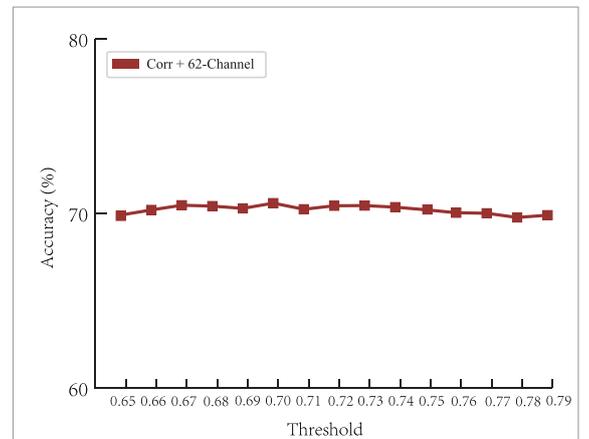

Figure 7. Threshold fine tuning results of SEED-V (64-channel network with correlation edges).

Table 1. The best thresholds in four cases with respect to SEED-V.

	18-channel	64-channel
Corr.	0.83	0.70
Cohe.	0.65	0.24

Figure 6 depicts the rough tuning results of SEED-V dataset. The best threshold ranges differ in these four cases. Taking 64-channel network with correlation edges as example, the best range is [0.6, 0.8]. Figure 7 depicts the fine tuning results under this circumstance. Finally, the best threshold is 0.7 with the accuracy rate of $70.59 \pm 8.03\%$. For other cases with respect to SEED-V, the best thresholds are shown in table 1. According to section 5.3.1, strength features with correlation as the connectivity metric achieves the best performance, so we only verify the discrimination ability of this feature on SEED and DEAP datasets, for which the best thresholds are 0.83 and 0.9, respectively.

Table 2. Performance (%) of conventional thresholding and critical subnetwork thresholding methods with respect to the SEED-V dataset using 18 channels.

	Conventional	Critical subnetwork
Mean.	55.64	**84.45**
Std.	**5.27**	6.10

5.2. Conventional thresholding vs critical subnetwork thresholding

The conventional thresholding method directly sets weak associations to 0 according to a proportional threshold after sorting the association weights [37, 67]. However, this approach is task-unrelevant. In this study, we proposed an emotion-relevant critical subnetwork selection approach. To demonstrate the superiority of our method, we compare the performance of both methods on SEED-V with 18 channels using DCCA.

The experiment setups in conventional thresholding method were exactly the same as our method. The strength feature with correlation as connectivity metric was applied. The threshold value was tuned as a hyperparameter ranging from [0.0, 1.0] with step size of 0.01. The best threshold value for conventional thresholding method is 0.9.

Table 2 shows the performance comparison. The best classification performance values (%) of conventional thresholding and critical subnetwork thresholding methods are 55.64 ± 5.27 and 84.45 ± 6.10, respectively. Our method shows significant superiority with accuracy rate improved by around 30%. This result demonstrates that compared with the conventional method, our proposed method extracts the emotion-relevant subnetworks actively to reduce the influence of noise in the functional connectivity.

5.3. Discrimination ability

To demonstrate the discrimination ability of the EEG functional connectivity network features in emotion recognition, we conduct EEG-based emotion recognition for the three datasets. The LIBLINEAR tool kit is adopted to achieve SVM classifier with L2-regularized and L1-loss, of which penalty coefficient C is the only hyperparameter. Applying the best threshold a_{final}, the tuning process includes two steps:

(a) Rough tuning: The search space of penalty coefficient is set as $2^{[-10,10]}$ with step size of 0.1. The best penalty coefficient 2^c found in the rough tuning process will be utilized to construct the search space in fine tuning.

(b) Fine tuning: The search space is set as $2^{[c-2,c+2]}$ with step size of 0.01. Penalty coefficient $2^{c_{final}}$ that achieves the highest accuracy rate is selected as the final hyperparameter value.

5.3.1. Experimental results on the SEED-V dataset

For the SEED-V dataset, we constructed the EEG-based brain functional connectivity networks using two different categories of vertices and two different edge measurements, then extracted three EEG functional connectivity network features from the brain networks.

Table 3 presents the five-class emotion recognition performance of these features. We could observe that the strength feature exhibits outstanding performance regardless of the number of vertices and connectivity metric. This may be because the strength feature could intuitively reflect the emotion associated connectivity of the entire brain regions. In general, the strength and eigenvector centrality features exhibit higher accuracy with correlation as the connectivity metric, whereas the clustering coefficient feature exhibits better performance with coherence.

Correlation and coherence measure the similarity of paired signals in two aspects. Correlation focuses on the temporal domain whereas coherence focuses on the frequency domain. According to the analyses of Guevara *et al* [79], coherence is affected by the power and phrase changes and represents the stability of paired signals in power and phase relationship. Correlation is independent on amplitudes and indicates the time coupling and waveform similarity. These two methods construct different network patterns, which makes differences in the information extraction efficiency of network features. Our results show that quantifying the clusters (i.e. clustering coefficient feature) within brain network is more effectively to refine information when the network edge is coherence between paired signals. Strength and eigenvector centrality features, on the other hand, are more efficient when correlation is applied.
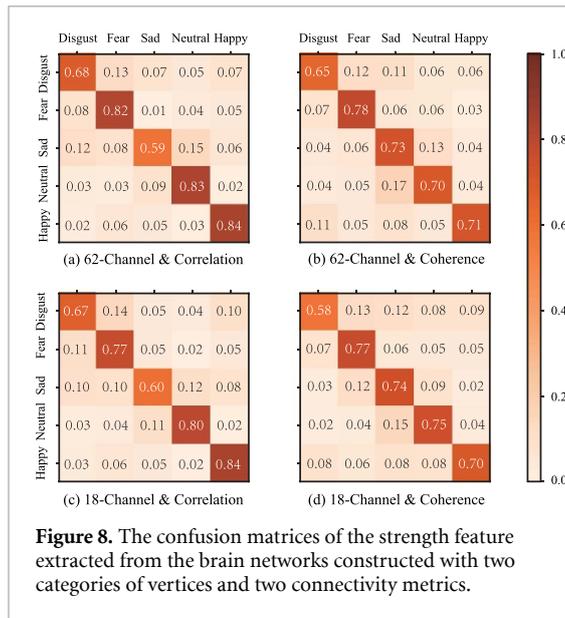
The features extracted from 18-channel-based brain networks exhibit considerable performance compared with those of 62-channel networks, which indicates that the EEG functional connectivity network features extracted from the brain networks constructed with fewer channels are promising for actual scenarios of emotion recognition applications in aBCI systems.

In the previous work [50], the authors have demonstrated that the EEG functional connectivity network features considerably outperform the PSD feature and that they are superior to those directly using the connectivity metrics as features. In this paper, the best classification accuracy of $74.05 \pm 7.09\%$ achieved by the strength feature defeats the value of $69.50 \pm 10.28\%$ attained by the single-channel-based state-of-the-art DE feature in the work of [55] for the same dataset.

To further analyze the capability of the best feature, the strength, in recognizing each of the five emotions, the confusion matrices are displayed in figure 8. It could be observed that the strength feature is superior in detecting the emotion of happiness,

Table 3. Mean accuracy (%) and standard deviation (%) of the EEG connectivity features in classifying the five emotions with respect to the SEED-V dataset.

Vertices	62-Channel				18-Channel			
Edges	Correlation		Coherence		Correlation		Coherence	
Stat. Features	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Strength	**74.05**	**7.09**	71.89	7.20	**72.63**	**8.26**	71.46	6.08
Clustering coefficient	56.35	11.14	71.18	7.82	51.94	8.63	64.14	6.10
Eigenvector centrality	68.78	9.39	60.80	7.18	67.89	11.13	65.57	7.38

Figure 8. The confusion matrices of the strength feature extracted from the brain networks constructed with two categories of vertices and two connectivity metrics.

followed by the emotions of neutrality and fear with correlation as connectivity metric. In contrast, from the perspective of coherence as connectivity metric, the strength feature exhibits the best performance in recognizing the fear emotion and exhibits similar performances with respect to the emotions of sadness, neutrality, and happiness.

Generally, the strength feature with correlation as connectivity metric could achieve better performance in classifying all emotions, except sadness, in comparison with that of coherence. Overall, the EEG feature exhibits fair performance in recognizing the emotion of disgust and could be easily confused by the sad and neutral emotions, the results for which are in accordance with previous findings [55]. In addition, the strength feature with the 18-channel approach exhibits considerable performance compared with that of the 62-channel approach. Particularly, the strength feature with the 18-channel approach could achieve the same classification accuracy of 84% with that of the 62-channel approach with respect to recognizing the emotion of happiness.

5.3.2. Experimental results on SEED and DEAP datasets

On SEED and DEAP datasets, we utilize the best EEG functional connectivity network feature, the strength, to further verify its discrimination ability in classifying emotions.

On the SEED dataset, the three-class emotion recognition accuracy achieved by the strength feature is $80.17 \pm 7.12\%$, which is higher than the value of $78.51 \pm 14.32\%$ [17] attained by the DE feature. On the DEAP dataset, the performances of the strength feature for two binary classification tasks (arousal-level and valence-level) are $73.42 \pm 4.67\%$ and $76.10 \pm 4.49\%$, respectively. These results considerably outperform those of 62.0% and 57.6% [24] achieved by the PSD feature, as well as those of 68.28% and 66.73% [80] attained using the capsule network.

These results demonstrate the discrimination ability of the EEG functional connectivity network features in classifying three-class emotions (sad, happy, and neutral), five-class emotions (disgust, fear, sad, happy, and neutral), and valence-arousal dimension. Additionally, the strength feature outperforms the most commonly used PSD feature and the state-of-the-art DE feature.

5.4. Complementary representation properties

In this section, the DCCA model is adopted to combine the EEG signals with other modalities for multimodal emotion recognition with respect to the three datasets. Here, the best EEG functional connectivity network feature, strength with correlation as connectivity metric, is utilized for the evaluation.

5.4.1. Experimental results on the SEED-V dataset

The combination of EEG and eye movement data for the SEED-V dataset was implemented using two different feature fusion strategies: feature-level fusion (FLF) and DCCA. The FLF is a direct concatenation of two feature vectors from EEG and eye movements within the same time windows according to the time stampers provided by the recording systems. The concatenated feature vectors were used as the inputs of classifiers for training.

The experimental results are displayed in table 4. The best classification performance values (%) based on the EEG connectivity feature, eye movement data, FLF, and DCCA approaches are 74.05 ± 7.09, 65.21 ± 7.60, 78.03 ± 6.07, and 84.51 ± 5.11, respectively. These results indicate that the combination of

Table 4. Performance (%) of two single modalities and two multimodal fusion strategies in classifying the five emotions with respect to the SEED-V dataset.

Vertices	62-Channel		18-Channel	
Stat.	Mean	Std.	Mean	Std.
EEG	74.05	7.09	72.63	8.26
EYE	65.21	7.60	65.21	7.60
FLF	78.03	6.07	78.02	7.30
DCCA	**84.51**	**5.11**	**84.45**	**6.10**

Table 5. Classification performance (%) of different works in multimodal emotion recognition on the SEED-V dataset.

Works	Method	Mean	Std.
Zhao et al [55]	FLF	73.65	8.90
	BDAE	79.70	4.76
Liu et al [69]	Max	73.17	9.27
	Fuzzy	73.24	8.72
	DCCA	83.08	7.11
Our method	FLF	78.03	6.07
(62-channel)	DCCA	**84.51**	**5.11**
Our method	FLF	78.02	7.30
(18-channel)	DCCA	**84.45**	**6.10**

Figure 9. Confusion matrices of five-class emotion recognition using: (a) eye movement data, (b) EEG signals, (c) FLF, and (d) DCCA models.

Figure 10. Confusion graph of the EEG functional connectivity network feature and eye movement data in classifying the five emotions: disgust, fear, sadness, happiness, and neutrality.

the EEG connectivity feature and eye movement data could enhance the performance of five-class emotion recognition. Moreover, the DCCA model may find the shared space to be more related to emotion. In addition, the fusion based on the 18-channel EEG connectivity feature and eye movement data also achieves considerable classification performance.

Table 5 presents the performance of our proposed EEG feature compared with the single-channel-based state-of-the-art DE feature [55, 69] for the multimodal emotion recognition task with respect to the SEED-V dataset. All these methods applied fusion methods. BDAE is a feature fusion approach which combines EEG and Eye movement features by two deep autoencoder models [55]. Max fusion and fuzzy fusion are decision-level fusion approaches, of which the aim is to combine the classification results of many small classifiers [69]. These results reveal that our proposed EEG connectivity feature outperforms the DE feature in combination with eye movement data based on DCCA when classifying the five emotions, whether using the 62-channel or 18-channel-based functional connectivity networks.

To investigate the capabilities of EEG connectivity feature and eye movement data in detecting each specific emotion, the confusion matrices are displayed in figure 9. Here, the EEG feature is the strength feature with 62 channels and the correlation connectivity metric. It could be observed that both EEG and eye movement data exhibit potential in classifying the emotions of fear, happiness, and neutrality. In particular, the EEG connectivity feature dominates the recognition of the happiness emotion, while eye movement data excel at detecting the fear emotion.
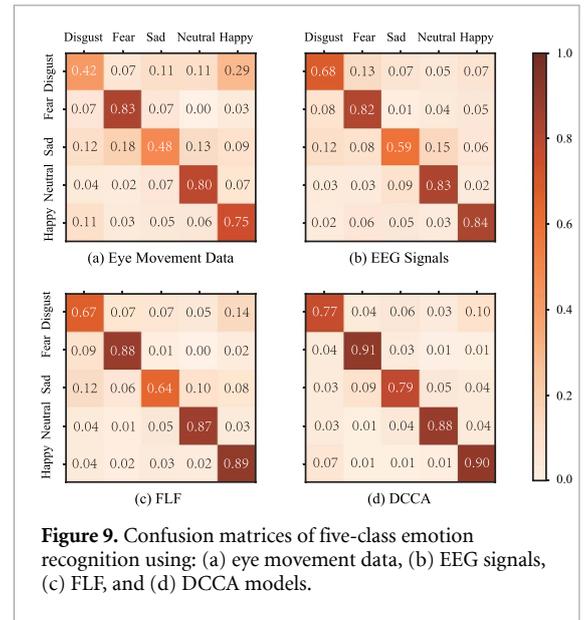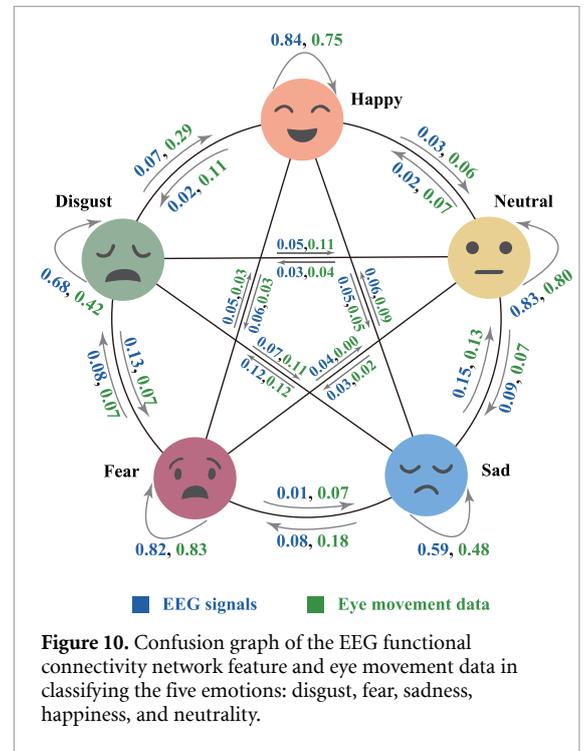
The confusion graph of these two modalities is also presented in figure 10.

In comparison with the single modality affective model, the last two confusion matrices in figure 9 indicate that the multimodal fusion strategies could indeed improve the classification performance for all of the five emotions. These results demonstrate the complementary representation properties of the EEG connectivity feature and eye movement data in classifying the five emotions.

5.4.2. Experimental results on SEED and DEAP datasets

The classification performances of our work and several existing works with respect to the SEED dataset

Table 6. Classification performance (%) of different works in multimodal emotion recognition on the SEED dataset. All methods were evaluated with the same dataset split for training and test sets, i.e. 9/6 trial split.

Works	Method	Mean	Std.
Lu *et al* [17]	FLF	83.70	—
	Fuzzy	87.59	—
Song *et al* [81]	DGCNN	90.40	8.49
Liu *et al* [61]	BDAE	91.01	8.91
Tang *et al* [62]	Bimodal-LSTM	93.97	7.03
Liu *et al* [69]	DCCA	94.58	**6.16**
Our method	DCCA	**95.08**	6.42

Table 7. Classification performance (%) of different works in multimodal emotion recognition on the DEAP dataset. All reported performance were evaluated using 10-fold cross validation. All methods used the same threshold of 5 for valence and arousal except Xing *et al* [82], where the ratings below 4.5 were labeled as low and those above 5.5 were labeled as high.

Works	Method	Arousal		Valence	
		Mean	Std.	Mean	Std.
Xing *et al* [82]	SAE-LSTM	74.38	—	81.10	—
Liu *et al* [61]	BDAE	80.50	3.39	85.20	4.47
Tang *et al* [62]	Bimodal-LSTM	83.23	2.61	83.82	5.01
Yin *et al* [83]	MESAE	84.18	—	83.04	—
Liu *et al* [69]	DCCA	84.33	**2.25**	85.62	**3.48**
Our method	DCCA	**85.34**	2.90	**86.61**	3.76

are displayed in table 6. The existing works are based on the DE feature and eye movements. We could observe that the best performance of $95.08 \pm 6.42\%$ is achieved by our work, which combines the strength feature with eye movement data to detect three emotions (happiness, neutrality, and sadness). These results further verify that the combination of EEG and eye movements could enhance the classification performance.

Table 7 presents the classification performances of our work and several existing works with respect to the DEAP dataset. The existing works are based on the combination of peripheral physiological features with the PSD [82, 83] or DE [61, 62, 69] features. The highest classification accuracy of the two binary classification tasks, $85.34 \pm 2.90\%$ for the arousal-level and $86.61 \pm 3.76\%$ for the valence-level, are both obtained by our work. These results reveal that the strength feature is also superior to the PSD and DE features in fusion with peripheral physiological signals.

5.5. Critical frequency bands
In this section, we evaluate the critical frequency band of the EEG functional connectivity network feature on the SEED-V dataset. Figure 11 presents the classification performance of different frequency bands using the strength feature with correlation as the connectivity metric. The result demonstrates that 18-channel and 62-channel networks with correlation edges show similar characteristics, i.e. the β and γ frequency bands are superior in classifying the five emotions in comparison with other bands, which is in accordance with the results attained by the DE feature [34, 50]. The topology similarity of these two
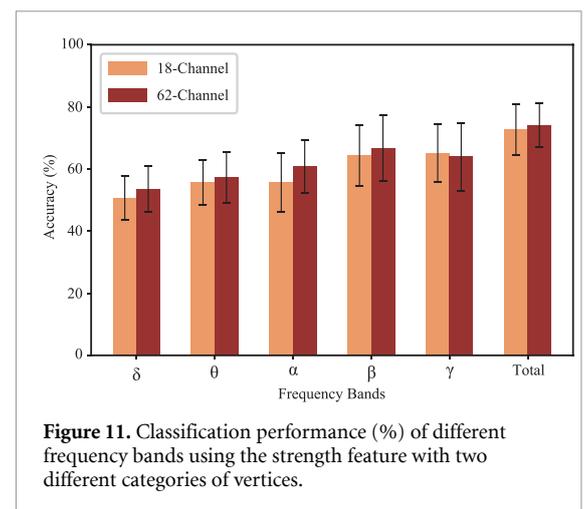

Figure 11. Classification performance (%) of different frequency bands using the strength feature with two different categories of vertices.

networks may be indicated. Additionally, the frequency bands with the 18-channel approach achieve comparable performance with that of the 62-channel approach, which implies the possibility of applying 18 electrodes to detect emotions in real scenario applications. It should be noted that the performance of the 18-channel approach is obtained by selecting 18 out of 62 channels in the ESI NeuroScan System instead of direct recordings from a wearable device. The electrode sensitivity could influence the signal quality and might degrade the task performance.

5.6. Brain functional connectivity patterns
In this section, we investigate the brain functional connectivity patterns based on the SEED-V dataset. The emotion-relevant critical subnetworks are selected through three phases: averaging, thresholding, and merging. The numbers of subnetworks attained
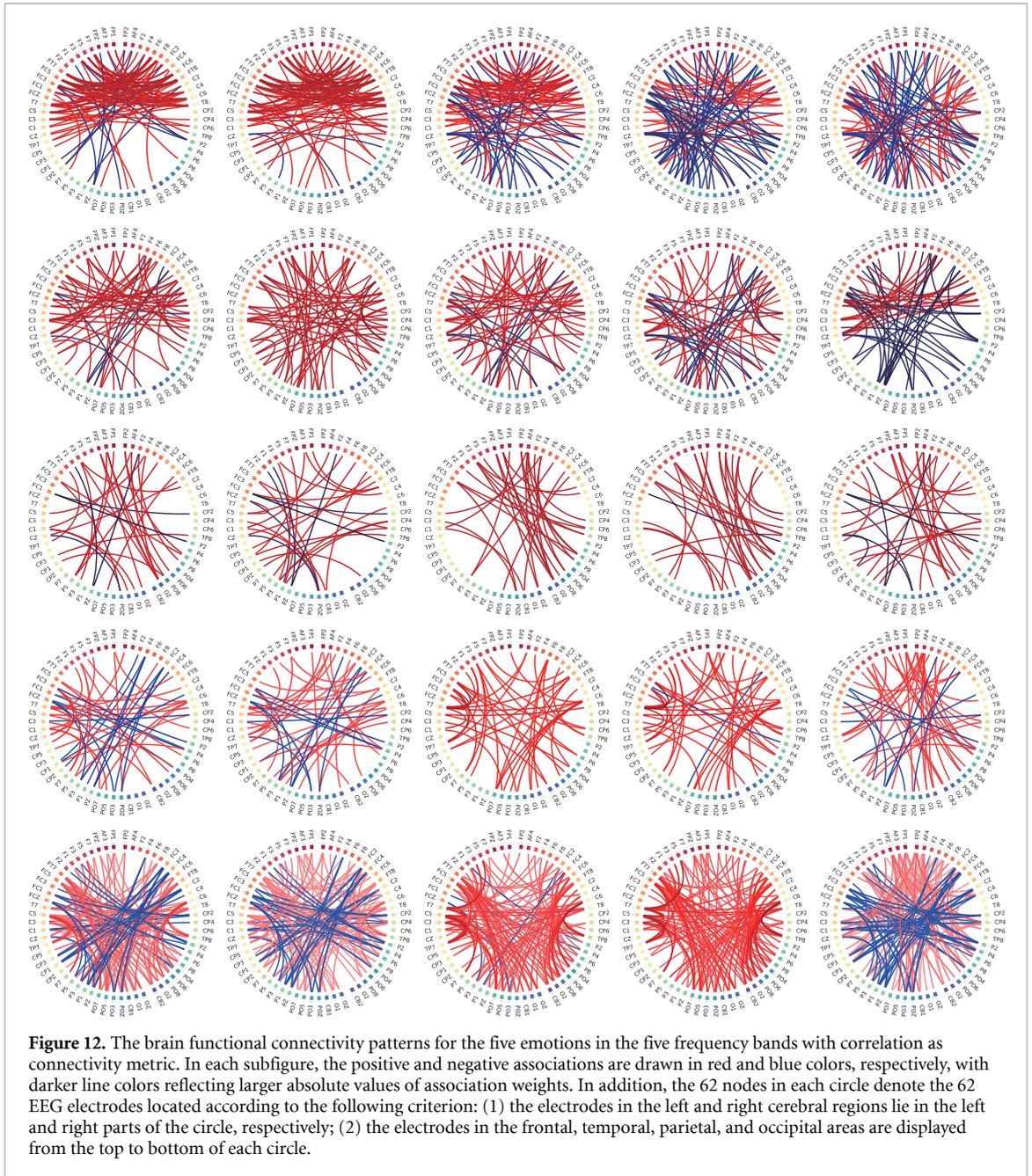
Figure 12. The brain functional connectivity patterns for the five emotions in the five frequency bands with correlation as connectivity metric. In each subfigure, the positive and negative associations are drawn in red and blue colors, respectively, with darker line colors reflecting larger absolute values of association weights. In addition, the 62 nodes in each circle denote the 62 EEG electrodes located according to the following criterion: (1) the electrodes in the left and right cerebral regions lie in the left and right parts of the circle, respectively; (2) the electrodes in the frontal, temporal, parietal, and occipital areas are displayed from the top to bottom of each circle.

after each phase are 25, 25, and 5, respectively. Specifically, 25 corresponds to the five emotions in the five frequency bands, while 5 refers to the five frequency bands, since the critical connections of the five emotions in each frequency band are merged together during the third phase.

In this paper, the 25 subnetworks attained after the thresholding phase are adopted to analyze the frequency-specific brain functional connectivity patterns in association with the five emotions. To analyze both the positive and negative connections, we visualize the brain functional connectivity networks based on the connectivity metric of correlation.

Considering that the subnetworks are selected using samples from training sets of all participants and that a three-fold cross validation strategy is

utilized, the emotion-relevant critical subnetworks are calculated three times. The results demonstrate that stable connectivity patterns are exhibited across these three calculations.

Figure 12 presents the 25 critical subnetworks associated with the five emotions in the five frequency bands averaged over three folds. For better analysis of the distinct connectivity patterns for each emotion in each frequency band, we display all of the critical connections except for the intersections among the five emotions.

It could be observed that the positive correlation connectivity is much higher in the frontal lobes in the δ band for the negative affective states, including the emotions of disgust, fear, and sadness. In particular, for the disgust emotion, stronger positive

connectivity is exhibited in the γ bands within both the left and right brain regions, and stronger negative connectivity is observed between the left and right brain regions. However, the fear emotion in the γ band is dominated by the stronger negative connectivity, and there are much weaker positive connections compared with those of the disgust emotion. In addition, much more positive connectivity in the θ band is exhibited for the fear emotion.

The fact that the functional connectivity patterns are quite similar for the sad and neutral emotions could account for the confusion between the sad and neutral emotions, which is consistent with previous findings [23, 34]. Nevertheless, the connectivity in the δ band tends to be positive within the frontal areas and negative in the left brain regions for the sadness emotion, while negative in larger brain areas for the neutral emotion.

In terms of the happiness emotion, the entire cerebral areas are much more active in the δ band with both positive and negative correlation connectivity. Moreover, in the θ band, negative connectivity is revealed between the frontal and parietal lobes, with positive connectivity between the frontal and temporal lobes. In the γ band, the functional connectivity patterns for the emotions of happiness, fear, and disgust are more similar, which may be originated from the fact that amygdala voxels contribute to these three emotions [84]. Overall, these results are in accordance with findings in the literature based on fMRI that the brain regions contributing to the emotion classification are predominated in the frontal and parietal lobes [84].

5.7. Future work

In this study, we mainly focus on the subject-dependent models. However, there are some open questions that are worth exploring in the future. For example, how could the models generalize across different subjects? i.e. developing subject-independent models. Some factors such as individual differences and temporal evolution should be considered. Recently transfer learning methods including domain adaptation and domain generalization have been shown efficient to deal with these problems [71–75].

We demonstrated that our proposed model can achieve comparable performance when using 18 channels compared with that using 62 channels by directly selecting 18 out of 62 channels in the ESI NeuroScan System. However, when applying 18 channels on other wearable EEG devices, the different frequency sensitivities may influence the signal quality and thereby affect emotion recognition results. Meanwhile, whether 18-channel network can be applied to datasets using different EEG acquisition devices such as DEAP are not discussed as we only did a preliminary exploration of the reduction in the number of electrodes in this paper. More in-depth research about this field is needed.

We recruited sixteen subjects aged between 19 and 24 (mean: 21.62) for SEED-V dataset. The ratio of women to men is 0.6. Compared with other physiological-signal based emotional datasets such as DEAP (32 subjects), DREAMER (23 subjects) and so on, the number of subjects seems moderate. However, it is still insufficient compared with conventional healthy EEG datasets. Meanwhile, all these subjects in SEED-V are college student with extroverted characteristics and stable mood, which limits our research to a particular personality and age profile. To conduct more rigorous and comprehensive research, expanding the number of subjects and increasing the diversity of subjects are essential.

6. Conclusion

In this paper, we have proposed a novel emotion-relevant critical subnetwork selection algorithm and evaluated three EEG connectivity features (strength, clustering coefficient, and eigenvector centrality) on three public datasets: SEED, SEED-V, and DEAP. The experimental results have revealed that the emotion associated brain functional connectivity patterns do exist. The strength feature is the best EEG connectivity feature and outperforms the state-of-the-art DE feature based on single-channel analysis. Furthermore, we have performed the multimodal emotion recognition using the DCCA model based on the EEG connectivity feature. The classification accuracies are $95.08 \pm 6.42\%$ on the SEED dataset, $84.51 \pm 5.11\%$ on the SEED-V dataset, and $85.34 \pm 2.90\%$ and $86.61 \pm 3.76\%$ on the DEAP dataset. These results have demonstrated the complementary representation properties between the EEG connectivity feature and eye movement data. Additionally, the results have indicated that the brain functional connectivity networks based on the 18-channel approach are promising for multimodal emotion recognition applications in aBCI systems under actual scenario situations.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://bcmi.sjtu.edu.cn/home/seed/.

Acknowledgments

ORCID iDs

Wei-Long Zheng ● https://orcid.org/0000-0002-9474-6369

Ziyi Li ● https://orcid.org/0000-0002-8944-741X

Bao-Liang Lu ● https://orcid.org/0000-0001-8359-0058

References

[1] Panetta K 2019 5 trends appear on the Gartner hype cycle for emerging technologies (available at: www.gartner.com/smarterwithgartner/5-trends-appear-on-the-gartner-hype-cycle-for-emerging-technologies-2019)

[2] Poria S, Cambria E, Bajpai R and Hussain A 2017 A review of affective computing: from unimodal analysis to multimodal fusion *Inf. Fusion* **37** 98–125

[3] Mühl C, Allison B, Nijholt A and Chanel G 2014 A survey of affective brain computer interfaces: principles, state-of-the-art and challenges *Brain-Comput. Interfaces* **1** 66–84

[4] Shanechi M M 2019 Brain-machine interfaces from motor to mood *Nat. Neurosci.* **22** 1554–64

[5] Thanapattheerakul T, Mao K, Amoranto J and Chan J H 2018 Emotion in a century: a review of emotion recognition *Proc. 10th Int. Conf. on Advances in Information Technology* (ACM) pp 1–8

[6] Byoung K 2018 A brief review of facial emotion recognition based on visual information *Sensors* **18** 401

[7] Schuller Born W 2018 Speech emotion recognition: two decades in a nutshell, benchmarks and ongoing trends *Commun. ACM* **61** 90–9

[8] Shu L, Xie J, Yang M, Ziyi Li, Zhenqi Li, Liao D, Xiangmin X and Yang X 2018 A review of emotion recognition using physiological signals *Sensors* **18** 2074

[9] Alarcao S M and Fonseca M J 2017 Emotions recognition using EEG signals: a survey *IEEE Trans. Affect. Comput.* **10** 374–93

[10] Cheng X, Wang Y, Dai S, Zhao P and Liu Q 2019 Heart sound signals can be used for emotion recognition *Sci. Rep.* **9** 6486

[11] Poria S, Chaturvedi I, Cambria E and Hussain A 2016 Convolutional MKL based multimodal emotion recognition and sentiment analysis *2016 IEEE 16th Int. Conf. on Data Mining (ICDM)* (IEEE) pp 439–48

[12] Zhang Z, Ringeval F, Dong B, Coutinho E, Marchi E and Schüller Born 2016 Enhanced semi-supervised learning for multimodal emotion recognition *2016 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE) pp 5185–9

[13] Povolny F, Matejka P, Hradis M, Popková A, Otrusina Lır, Smrz P, Wood I, Robin C and Lamel L 2016 Multimodal emotion recognition for AVEC 2016 challenge *Proc. 6th Int. Workshop on Audio/Visual Emotion Challenge* (ACM) pp 75–82

[14] Soleymani M, Pantic M and Pun T 2011 Multimodal emotion recognition in response to videos *IEEE Trans. Affect. Comput.* **3** 211–23

[15] Zheng W-L, Zhu J-Y and Lu B-L 2019 Identifying stable patterns over time for emotion recognition from EEG *IEEE Trans. Affect. Comput.* **10** 417–29

[16] García-Martínez B, Martinez-Rodrigo A, Alcaraz R and Fernández-Caballero A 2019 A review on nonlinear methods using electroencephalographic recordings for emotion recognition *IEEE Trans. Affect. Comput.* **12** 801–20

[17] Lu Y-F, Zheng W-L, Li B-B and Lu B-L 2015 Combining eye movements and EEG to enhance emotion recognition *Int. Conf. on Artificial Intelligence* vol 15 (Citeseer) pp 1170–6

[18] López-Gil J-M, Virgili-Gomá J, Gil R, Guilera T, Batalla I, Soler-González J and García R 2016 Method for improving EEG based emotion recognition by combining it with synchronized biometric and eye tracking technologies in a non-invasive and low cost way *Front. Comput. Neurosci.* **10** 85

[19] Mauss I B and Robinson M D 2009 Measures of emotion: a review *Cognit. Emotion* **23** 209–37

[20] Zhan Y *et al* 2014 Deficient neuron-microglia signaling results in impaired functional brain connectivity and social behavior *Nat. Neurosci.* **17** 400

[21] Finn E S, Shen X, Scheinost D, Rosenberg M D, Huang J, Chun M M, Papademetris X and Constable R T 2015 Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity *Nat. Neurosci.* **18** 1664

[22] Smith S 2016 Linking cognition to brain connectivity *Nat. Neurosci.* **19** 7

[23] Li T-H, Liu W, Zheng W-L and Lu B-L 2019 Classification of five emotions from EEG and eye movement signals: discrimination ability and stability over time *2019 9th Int. IEEE/ Conf. on Neural Engineering (NER)* (IEEE) pp 607–10

[24] Koelstra S, Muhl C, Soleymani M, Lee J-S, Yazdani A, Ebrahimi T, Pun T, Nijholt A and Patras I 2011 Deap: a database for emotion analysis; using physiological signals *IEEE Trans. Affect. Comput.* **3** 18–31

[25] Hasani B and Mahoor M H 2017 Facial expression recognition using enhanced deep 3D convolutional neural networks *Proc. Conf. on Computer Vision and Pattern Recognition Workshops* pp 30–40

[26] Trigeorgis G, Ringeval F, Brueckner R, Marchi E, Nicolaou M A, Schuller Born and Zafeiriou S 2016 Adieu features? End-to-end speech emotion recognition using a deep convolutional recurrent network *2016 IEEE Conf. on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE) pp 5200–4

[27] Schirmer A and Adolphs R 2017 Emotion perception from face, voice and touch: comparisons and convergence *Trends Cognit. Sci.* **21** 216–28

[28] Zhang Q, Chen X, Zhan Q, Yang T and Xia S 2017 Respiration-based emotion recognition with deep learning *Comput. Ind.* **92** 84–90

[29] Nardelli M, Valenza G, Greco A, Lanata A and Scilingo E P 2015 Recognizing emotions induced by affective sounds through heart rate variability *IEEE Trans. Affect. Comput.* **6** 385–94

[30] Atkinson J and Campos D 2016 Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers *Expert Syst. Appl.* **47** 35–41

[31] Liu Y-J, Minjing Y, Zhao G, Song J, Yan G and Shi Y 2017 Real-time movie-induced discrete emotion recognition from EEG signals *IEEE Trans. Affect. Comput.* **9** 550–62

[32] Wang X-W, Nie D and Lu B-L 2014 Emotional state classification from EEG data using machine learning approach *Neurocomputing* **129** 94–106

[33] Jenke R, Peer A and Buss M 2014 Feature extraction and selection for emotion recognition from EEG *IEEE Trans. Affect. Comput.* **5** 327–39

[34] Zheng W-L and Lu B-L 2015 Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks *IEEE Trans. Auton. Mental Dev.* **7** 162–75

[35] Duan R-N, Zhu J-Y and Lu B-L 2013 Differential entropy feature for EEG-based emotion classification *2013 6th Int. IEEE/ Conf. on Neural Engineering (NER)* (IEEE) pp 81–84

[36] Ackermann P, Kohlschein C, Bitsch Jo Ágila, Wehrle K and Jeschke S 2016 EEG-based automatic emotion recognition: feature extraction, selection and classification methods *2016 IEEE 18th Int. Conf. on E-Health Networking, Applications and Services (Healthcom)* (IEEE) pp 1–6

[37] Rubinov M and Sporns O 2010 Complex network measures of brain connectivity: uses and interpretations *Neuroimage* **52** 1059–69

[38] Murias M, Webb S J, Greenson J and Dawson G 2007 Resting state cortical connectivity reflected in EEG coherence in individuals with autism *Biol. Psychiatry* **62** 270–3

[39] Yin Z, Jun Li, Zhang Y, Ren A, Von Meneen K M and Huang L 2017 Functional brain network analysis of schizophrenic patients with positive and negative syndrome based on mutual information of EEG time series *Biomed. Signal Process. Control* **31** 331–8

[40] Ho T C *et al* 2015 Emotion-dependent functional connectivity of the default mode network in adolescent depression *Biol. Psychiatry* **78** 635–46

[41] Whitton A E, Deccy S, Ironside M L, Kumar P, Beltzer M and Pizzagalli D A 2018 EEG source functional connectivity reveals abnormal high-frequency communication among large-scale functional networks in depression *Biol. Psychiatry Cognit. Neurosci. Neuroimaging* **3** 50

[42] Peng X, Xiong X, Xue Q, Li P, Zhang R, Wang Z, Valdes-Sosa P A, Wang Y and Yao D 2014 Differentiating between psychogenic nonepileptic seizures and epilepsy based on common spatial pattern of weighted EEG resting networks *IEEE Trans. Biomed. Eng.* **61** 1747–55

[43] Dasdemir Y, Yildirim E and Yildirim S 2017 Analysis of functional brain connections for positive–negative emotions using phase locking value *Cogn. Neurodyn.* **11** 487–500

[44] Lee Y-Y and Hsieh S 2014 Classifying different emotional states by means of EEG-based functional connectivity patterns *PLoS One* **9** e95415

[45] PeiYang Li *et al* 2019 EEG based emotion recognition by combining functional connectivity network and local activations *IEEE Trans. Biomed. Eng.* **66** 2869–81

[46] Al-Shargie F, Tariq U, Alex M, Mir H and Al-Nashash H 2019 Emotion recognition based on fusion of local cortical activations and dynamic functional networks connectivity: an EEG study *IEEE Access* **7** 143550–62

[47] Moon S-E, Jang S and Lee J-S 2018 Convolutional neural network approach for EEG-based emotion recognition using brain connectivity and its spatial information *2018 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE) pp 2556–60

[48] Chao H, Dong L, Liu Y and Baoyun L 2019 Emotion recognition from multiband EEG signals using capsnet *Sensors* **19** 2212

[49] Hua C, Wang H, Wang H, Shaowen L, Liu C and Khalid S M 2019 A novel method of building functional brain network using deep learning algorithm with application in proficiency detection *Int. J. Neural Syst.* **29** 1850015

[50] Wu X, Zheng W-L and Lu B-L 2019 Identifying functional brain connectivity patterns for EEG-based emotion recognition *2019 9th Int. IEEE/ Conf. on Neural Engineering (NER)* (IEEE) pp 235–8

[51] Widmann A, Schröger E and Wetzel N 2018 Emotion lies in the eye of the listener: emotional arousal to novel sounds is reflected in the sympathetic contribution to the pupil dilation response and the P3 *Biol. Psychol.* **133** 10–17

[52] Oliva M and Anikin A 2018 Pupil dilation reflects the time course of emotion recognition in human vocalizations *Sci. Rep.* **8** 4871

[53] Black M H, Chen N T M, Iyer K K, Lipp O V, Bölte S, Falkmer M, Tan T and Girdler S 2017 Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography *Neurosci. Biobehav. Rev.* **80** 488–515

[54] Zheng W-L, Dong B-N and Lu B-L 2014 Multimodal emotion recognition using EEG and eye tracking data *Annual Int. Conf. IEEE Engineering in Medicine and Biology Society* (IEEE) pp 5040–3

[55] Zhao L-M, Li R, Zheng W-L and Lu B-L 2019 Classification of five emotions from EEG and eye movement signals: complementary representation properties *2019 9th Int. IEEE/ Conf. on Neural Engineering (NER)* (IEEE) pp 611–4

[56] Adolphs R and Anderson D J 2018 *The Neuroscience of Emotion: A New Synthesis* (Princeton, NJ: Princeton University Press)

[57] Perez-Gaspar L-A, Caballero-Morales S-O and Trujillo-Romero F 2016 Multimodal emotion recognition with evolutionary computation for human-robot interaction *Expert Syst. Appl.* **66** 42–61

[58] Tzirakis P, Trigeorgis G, Nicolaou M A, Schuller Born W and Zafeiriou S 2017 End-to-end multimodal emotion recognition using deep neural networks *IEEE J. Sel. Top. Signal Process.* **11** 1301–9

[59] Ranganathan H, Chakraborty S and Panchanathan S 2016 Multimodal emotion recognition using deep learning architectures *2016 IEEE Conf. on Applications of Computer Vision (WACV)* (IEEE) pp 1–9

[60] Huang Y, Yang J, Liao P and Pan J 2017 Fusion of facial expressions and EEG for multimodal emotion recognition *Comput. Intell. Neurosci.* **2017** 2107451

[61] Liu W, Zheng W-L and Lu B-L 2016 Emotion recognition using multimodal deep learning *Int. Conf. on Neural Information Processing* (Springer) pp 521–9

[62] Tang H, Liu W, Zheng W-L and Lu B-L 2017 Multimodal emotion recognition using deep neural networks *Int. Conf. on Neural Information Processing* (Springer) pp 811–19

[63] Zheng W-L, Liu W, Lu Y-F, Lu B-L and Cichocki A 2018 Emotionmeter: a multimodal framework for recognizing human emotions *IEEE Trans. Cybern.* **49** 1110–22

[64] Qiu J-L, Liu W and Lu B-L 2018 Multi-view emotion recognition using deep canonical correlation analysis *Int. Conf. on Neural Information Processing* (Springer) pp 221–31

[65] Eysenck S B G, Eysenck H J and Barrett P 1985 A revised version of the psychoticism scale *Pers. Individ. Differ.* **6** 21–29

[66] Bradley M M, Miccoli L, Escrig M A and Lang P J 2008 The pupil as a measure of emotional arousal and autonomic activation *Psychophysiology* **45** 602–7

[67] Bullmore E and Sporns O 2009 Complex brain networks: graph theoretical analysis of structural and functional systems *Nat. Rev. Neurosci.* **10** 186

[68] Tong J-J, Luo Y, Ma B-Q, Zheng W-L, Lu B-L, Song X-Q and Ma S-W 2018 Sleep quality estimation with adversarial domain adaptation: from laboratory to real scenario *2018 Int. Conf. on Neural Networks (IJCNN)* (IEEE) pp 1–8

[69] Liu W, Qiu J-L, Zheng W-L and Lu B-L 2019 Multimodal emotion recognition using deep canonical correlation analysis (arXiv:1908.05349)

[70] Atyabi A and Powers D M W 2012 The impact of segmentation and replication on non-overlapping windows: an EEG study *IEEE Int. Conf. on Information Science and Technology* (IEEE) pp 668–74

[71] Pan S J and Yang Q 2009 A survey on transfer learning *IEEE Trans. Knowl. Data Eng.* **22** 1345–59

[72] Samek W, Meinecke F C and Müller K-R 2013 Transferring subspaces between subjects in brain–computer interfacing *IEEE Trans. Biomed. Eng.* **60** 2289–98

[73] Zheng W-L and Lu B-L 2016 Personalizing EEG-based affective models with transfer learning *the Twenty-Fifth Int. Conf. on Artificial Intelligence* pp 2732–8

[74] Wu D-R, Xu Y-F and Lu B-L 2020 Transfer learning for EEG-based brain-computer interfaces: a review of progress made since 2016 *IEEE Trans. Cognit. Dev. Syst.* **PP** 1–1

[75] Wan Z, Yang R, Huang M, Zeng N and Liu X 2021 A review on transfer learning in EEG signal analysis *Neurocomputing* **421** 1–14

[76] Ruhnau B 2000 Eigenvector-centrality—a node-centrality? *Soc. Netw.* **22** 357–65

[77] Duan R-N, Wang X-W and Lu B-L 2012 EEG-based emotion recognition in listening music by using support vector machine and linear dynamic system *Int. Conf. on Neural Information Processing* (Springer) pp 468–75

[78] Andrew G, Arora R, Bilmes J and Livescu K 2013 Deep canonical correlation analysis *Int. Conf. on Machine Learning* pp 1247–55

[79] Guevara M A and María C-C 1996 EEG coherence or EEG correlation? *Int. J. Psychophysiol.* **23** 145–53

[80] Chao H, Dong L, Liu Y and Baoyun L 2019 Emotion recognition from multiband EEG signals using CapsNet *Sensors* **19** 2212

[81] Song T, Zheng W, Peng S and Zhen C 2018 EEG emotion recognition using dynamical graph convolutional neural networks *IEEE Trans. Affect. Comput.* **11** 532–41

[82] Xing X, Zhenqi Li, Tianyuan X, Shu L, Hue B and Xiangmin X 2019 SAE plus LSTM: a new framework for emotion recognition from multi-channel EEG *Front. Neurorobot.* **13** 37

[83] Yin Z, Zhao M, Wang Y, Yang J and Zhang J 2017 Recognition of emotions using multimodal physiological signals and an ensemble deep learning model *Comput. Methods Programs Biomed.* **140** 93–110

[84] Saarimäki H, Gotsopoulos A, Jääskeläinen I P, Lampinen J, Vuilleumier P, Hari R, Sams M and Nummenmaa L 2015 Discrete neural signatures of basic emotions *Cereb. Cortex* **26** 2563–73