Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks

Wei-Long Zheng, *Student Member, IEEE*, and Bao-Liang Lu, *Senior Member, IEEE*

***Abstract*—To investigate critical frequency bands and channels, this paper introduces deep belief networks (DBNs) to constructing EEG-based emotion recognition models for three emotions: positive, neutral and negative. We develop an EEG dataset acquired from 15 subjects. Each subject performs the experiments twice at the interval of a few days. DBNs are trained with differential entropy features extracted from multichannel EEG data. We examine the weights of the trained DBNs and investigate the critical frequency bands and channels. Four different profiles of 4, 6, 9, and 12 channels are selected. The recognition accuracies of these four profiles are relatively stable with the best accuracy of 86.65%, which is even better than that of the original 62 channels. The critical frequency bands and channels determined by using the weights of trained DBNs are consistent with the existing observations. In addition, our experiment results show that neural signatures associated with different emotions do exist and they share commonality across sessions and individuals. We compare the performance of deep models with shallow models. The average accuracies of DBN, SVM, LR, and KNN are 86.08%, 83.99%, 82.70%, and 72.60%, respectively.**

***Index Terms*—Affective computing, deep belief networks, EEG, emotion recognition.**

I. Introduction

EMOTION research is an interdisciplinary field that encompasses research in computer science, psychology, neuroscience, and cognitive science. For neuroscience, researchers aim to find out the neural circuits and brain mechanisms of emotion processing. For psychology, there exist many basic theories of emotion from different researchers and it is important to build up computational models of emotion. For computer science, we focus on developing practical applications such as estimation of task workload [1] and driving fatigue detection [2].

In multimedia context analysis, for example, there is a large sematic gap between the high-level cognition in the human brain and the low-level features in raw digit data. As the emerging big data of social media, it is difficult to tag the contents reliably, especially for affective factors, which are hard to describe across different cultures and language backgrounds. So it is necessary to build an emotion model to automatically recognize the affective tags implicitly [3]. The field of Affective Computing (AC) aspires to narrow the communicative gap between the highly emotional human and the emotionally challenged computer by developing computational systems that recognize and respond to human emotions [4]. The detection and modeling of human emotions are the primary studies of affective computing using pattern recognition and machine learning techniques. Although affective computing has achieved rapid development in recent years, there are still many open problems to be solved [5], [6].

Among various approaches to emotion recognition, the method based on electroencephalography (EEG) signals is more reliable because of its high accuracy and objective evaluation in comparison with other external appearance clues like facial expression and gesture [7]. To deeply understand the brain response under different emotional states can fundamentally advance the computational models for emotion recognition. Various psychophysiology studies have demonstrated the correlations between human emotions and EEG signals [8]–[10]. Moreover, with the quick development of wearable devices and dry electrode techniques [11]–[14], it is now possible to implement EEG-based emotion recognition from laboratories to real-world applications, such as driving fatigue detection and mental state monitoring [15]–[19].

However, EEG signals have low signal-to-noise ratio (SNR) and are often mixed with much noise when collected. The more challenge problem is that, unlike image or speech signals, EEG signals are temporal asymmetry and nonstationary [20]. So to analyze EEG signals is a hard task. Traditional manual feature extraction and feature selection for EEG are crucial to affective modeling and require specified domain knowledge. The popular feature selection methods for EEG signal analysis are principal component analysis (PCA) and Fisher projection. In general, the cost of these traditional feature selection methods increases quadratically with respect to the number of features considered [21]. What's more, these methods cannot preserve the original

Manuscript received October 10, 2014; revised March 13, 2015; accepted April 13, 2015. Date of publication May 08, 2015; date of current version November 04, 2015. This work was supported in part by the grants from the National Natural Science Foundation of China (Grant No. 61272248), the National Basic Research Program of China (Grant 2013CB329401), the Science and Technology Commission of Shanghai Municipality (Grant 13511500200), the Open Funding Project of National Key Laboratory of Human Factors Engineering (Grant HF2012-K-01), and the European Union Seventh Framework Program (Grant 247619). *(Corresponding author: Bao-Ling Lu.)*

The authors are with the Center for Brain-Like Computing and Machine Intelligence, Department of Computer Science and Engineering, Shanghai Jiao Tong University and the Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: weilonglive@gmail.com; bllu@sjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAMD.2015.2431497

domain information such as channels and frequency bands that are very important for understanding brain response. Recent developing deep learning techniques in machine learning community allow automatic feature extraction and feature selection and can eliminate the limitation of handcrafted features [5]. Deep learning allows automatically feature selection at the same time with training classification models by bypassing the computational cost in feature selection phase.

In the past few years, researchers focused on finding the critical frequency bands and channels for EEG-based emotion recognition with different methods. Li and Lu [22] proposed a frequency band searching method to choose an optimal band for emotion recognition and their results showed that the gamma band (roughly 30-100 Hz) is suitable for EEG-based emotion classification with emotional still images as stimuli. It is also interesting that what would be good positions to place electrodes for emotion recognition when using only few electrodes. Bos [23] chose the following montage: Fpz/right mastoid for arousal recognition, $F3/F4$ for valence recognition, and left mastoid as ground. Her results indicated that $F3$ and $F4$ are the most suitable electrode positions to detect emotional valence. Combining the existing results, Valenzi [24] obtained a pool of eight electrodes: $AF3$, $AF4$, $F3$, $F4$, $F7$, $F8$, $T7$, and $T8$ and achieved an average classification rate of 87.5% with these eight electrodes. However, how to select the critical channels and frequency bands and how to evaluate selected pools of electrodes have not been fully investigated yet.

Since 2006, deep learning has emerged in machine community [25] and has generated a great impact in signal and information processing. Many deep architecture models are proposed such as deep auto-encoder [26], convolution neural network [27], [28] and deep belief network [29]. Deep architecture models achieve successful results and outperform shallow models (e.g. MLP, SVMs, CRFs) in many challenge tasks, especially in speech and image domains [29]–[31]. Recently deep learning methods are also successfully applied to physiological signal processing such as EEG, electromyogram (EMG), electrocardiogram (ECG), and skin resistance (SC), and achieve comparable results in comparison with other conventional methods [5], [32]–[34].

In this paper, we focus on investigating critical frequency bands and critical channels for efficient EEG-based emotion recognition. Here, we introduce deep learning methodologies to deal with these two problems. First, to shed light on the relationship between emotional states and change of EEG signals, we devise a protocol that subjects are asked to elicit their own emotions when watching three types of emotional movies (positive, neutral and negative). After that, we extract efficient features called differential entropy [35], [36] from multichannel EEG data, and then we train deep belief networks with differential entropy features as inputs. By analyzing the weight distributions learned from the trained deep belief networks, we choose different setups for frequency bands and channels and compare the performance of different feature subsets. We also compare the deep learning methods with feature-based shallow models like kNN, logistic regression and SVM, in order to explore the advantages of deep learning and the feasibility of applying unsupervised feature learning to EEG-based emotion recognition.

The main contributions of this paper can be described as the following aspects. First, considering the feature learning and feature selection properties of deep neural networks, we introduce deep learning methodologies to emotion recognition based on multichannel EEG data. By analyzing the weight distributions learned from the trained deep belief networks, we investigate different electrode set reductions and define the optimal electrode placement which outperforms original full channels with less computational cost and more feasibility in real world applications. And we show the superior performance of deep models over shallow models like kNN, logistic regression and SVM. The experiment results also indicate that the differential entropy features extracted from EEG data possess accurate and stable information for emotion recognition. We find that neural signatures associated with positive, neutral and negative emotions in channels and frequency bands do exist.

The layout of the paper is as follows. In Section II, we give a brief overview of related research on emotion recognition using EEG, as well as the use of deep learning methodologies for physiological signals. A systematic description of signal analysis methods and classification procedure for feature extraction and construction of deep belief networks is given in Section III. Section IV gives the motivation and rationale for our emotion experimental setting. A detailed description of all the materials and protocol we used is presented. In Section V, the detailed parameters for different classifiers are given and we systematically compare the performance of deep belief networks with other shallow models. Then we investigate different electrode set reductions and neural signatures associated with different emotions according to the weight distributions obtained from the trained deep neural networks. In Section VI, we discuss the problems in emotion recognition studies. Finally, in Section VII, we present conclusions.

II. Related Work

With the fast development of wearable devices and dry electrode techniques [11]–[14], it enables us to record and analyze the brain activity in natural settings. This development is leading to a new trend that integrates brain-computer interfaces (BCIs) with emotional factors. Emotional brain-computer interfaces are closed-loop affective computing systems, which build interactive environments [37]. Fig. 1 shows the emotional brain-computer interface cycle. Emotional brain-computer interfaces consist of the following six main phases. First, users are exposed to designed or real-world stimuli according to the protocol. The brain activities are recorded as EEG simultaneously. Then the raw data will be preprocessed to remove noise and artifacts. Some relevant features will be extracted and a classifier will be trained based on the extracted features. After identifying user current emotional states, a feedback can be implemented to respond to the users.

One of the goals of affective neuroscience is to examine whether patterns of brain activities for specific emotions exist, and whether these patterns are to some extent common across individuals. Various studies have examined the neural correlates of emotions. Davidson *et al.* [38], [39] showed that frontal EEG asymmetry is related to approach and withdrawal emotions, with approach tendencies reflected in left frontal activity

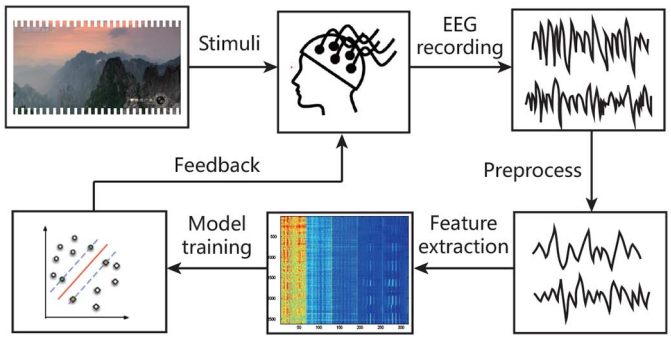


Fig. 1. Emotional brain-computer interface cycle.

and withdrawal tendencies reflected in relative right-frontal activity. Sammler *et al.* [8] investigated the EEG correlates of the processing of pleasant and unpleasant music. They found that pleasant music is associated with an increase of frontal midline theta power. Knyazev *et al.* [9] proposed gender differences in implicit and explicit processing of emotional facial expressions with the event-related theta synchronization. Mathersul *et al.* [10] investigated the relationships among nonclinical depression/anxiety and lateralized frontal/parietotemporal activity on the basis of both negative mood and alpha EEG. Their findings supported predictions for frontal but not posterior regions. Wang *et al.* [40] indicated that for positive and negative emotions, the subject-independent features are mainly on right occipital lobe and parietal lobe in alpha band, the parietal lobe and temporal lobe in beta band, and left frontal lobe and right temporal lobe in gamma band. Martini *et al.* [41] found that an increase in P300 and late positive potential and an increase in gamma activity during viewing unpleasant pictures as compared to neutral ones. They suggested that the full elaboration of unpleasant stimuli requires a tight interhemispheric communication between temporal and frontal regions, which is realized by means of phase synchronization at about 40 Hz. However, most of the existing experiments on passive BCI use a very controlled approach with time locked stimuli using ERP analysis, especially in psychology. This ideal experimental setting limits the range of real-world conditions and hard to be generalized to natural settings in a real environment.

Various studies in affective computing community try to build computational models to estimate emotional states using machine learning techniques. Lin *et al.* [42] applied machine learning algorithms to categorize EEG signals according to subject self-reported emotional states during music listening. They obtained an average classification accuracy of 82.29% for four emotions (joy, anger, sadness, and pleasure) across 26 subjects. Soleymani *et al.* [3] proposed a user-independent emotion recognition method using EEG, pupillary response and gaze distance, which achieved the best classification accuracies of 68.5% for three labels of valence and 76.4% for three labels of arousal using a modality fusion across 24 participants. Hadjidimitriou *et al.* [43] employed three time-frequency distributions (spectrogram, Hilbert-Huang spectrum, and Zhao-Atlas-Marks transform) as features to classify ratings of liking and familiarity. They also investigated the time course of music-induced affect responses and the role of familiarity. Li and Lu [22] proposed a frequency band searching method to choose an optimal band, into which the recorded EEG signal is filtered. They used common spatial patterns (CSP) and linear-SVM to classify two emotions (happiness and sadness). Their experimental results indicated that the gamma band (roughly 30-100 Hz) is suitable for EEG-based emotion classification. Wang *et al.* [40] systematically compared three kinds of EEG features (power spectrum feature, wavelet feature and nonlinear dynamical feature) for emotion classification. They proposed an approach to track the trajectory of emotion changes with manifold learning.

Recently, deep learning methods are applied to processing physiological signals such as EEG, EMG, ECG, and SC. Martinez *et al.* [5] trained an efficient deep convolution neural network to classify four cognitive states (relaxation, anxiety, excitement and fun) using skin conductance and blood volume pulse signals. They indicated that the proposed deep learning approach can outperform traditional feature extraction and selection methods and yield a more accurate affective model. Martin *et al.* [44] applied deep belief nets and hidden Markov model to detect sleep stage using multimodal clinical sleep datasets. Their results of using raw data with a deep model were comparable to handmade feature approach. To address two challenges of small sample problem and irrelevant channels, Li *et al.* [34] proposed a DBN based model for affective state recognition from EEG signals and compared it with five baselines with improvement of 11.5% to 24.4%. Zheng *et al.* [33] trained a deep belief network with differential entropy features extracted from multichannel EEG as input and achieved the best classification accuracy of 87.62% for two emotional categories in comparison with the state-of-the-art methods. In our previous work [32], we proposed a deep belief network based method to select the critical channels and frequency bands for three emotions (positive, neutral and negative). The experimental results showed that the selected channels and frequency bands could achieve comparable accuracies in comparison with that of the total features. In this paper, we extend our previous work to multichannel EEG processing and further investigate the weight distributions of trained deep neural networks, which reflects crucial neural signatures for emotion recognition.

The problem of electrode set reduction is commonly studied to reduce the computational complexity and ignore the irrelative noise. The optimal electrodes placement is usually defined according to some statistical factors like correlation coefficient, F-score and accuracy rate. Some studies shared the same pool of electrodes for restrict of commercial EEG device like Emotiv[1]. In [42], Lin *et al.* identified 30 subject-independent features that were most relevant to emotional processing across subjects according to F-score criterion and explored the feasibility of using fewer electrodes to characterize the EEG dynamics during music listening. The identified features were primarily derived from electrodes placed near the frontal and the parietal lobes. Valenzi *et al.* [24] selected a set of eight electrodes: $AF3$, $AF4$, $F3$, $F4$, $F7$, $F8$, $T7$, and $T8$, and achieved a promising result of 87.5% for four emotions. A similar study is proposed by Li *et al.* [34], which applied a DBN based model for

[1]http://emotiv.com/

affective states recognition from EEG signals to deal with two problems: small number of samples and noisy channels. They proposed a DBN-based channels selection method. Their interesting observation is that data in irrelevant channels randomly update the parameters in the DBN model, and data in critical channels update the parameters in the DBN model according to the related patterns. However, they did not explore the performance of these critical channels. In this paper, we proposed a novel electrode selection method through the weight distributions obtained from the trained deep neural networks instead of statistical parameters and show its superior performance over original full pool of electrodes.

Although various approaches have been proposed for EEG-based emotion recognition, most of the experimental results cannot be compared directly for different setups of experiments. There is still a lack of publicly available emotional EEG datasets. To the best of our knowledge, the popular publicly available emotional EEG datasets are MAHNOB HCI [3] and DEAP [45]. The first one includes EEG, physiological signals, eye gaze, audio, and facial expressions of 30 people when watching 20 emotional videos. The subjects self-reported their felt emotions using arousal, valence, dominance, and predictability as well as emotional keywords. The DEAP dataset includes the EEG and peripheral physiological signals of 32 participants when watching 40 one-minute music videos. It also contains participants' rate of each video in terms of the levels of arousal, valence, like/dislike, dominance, and familiarity. For reproducing the results in this paper and enhancing the cooperation in related research fields, the dataset used in this study is freely available to the academic community[2].

III. Methods

A. Preprocessing

According to the response of the subjects, only the experiment epochs when the target emotions were elicited were chosen for further analysis. The raw EEG data was downsampled to 200 Hz sampling rate. The EEG signals were visually checked and the recordings seriously contaminated by EMG and EOG were removed manually. EOG was also recorded in the experiments, and later used to identify blink artifacts from the recorded EEG data. In order to filter the noise and remove the artifacts, the EEG data was processed with a bandpass filter between 0.3 to 50 Hz. After performing the preprocessing, we extracted the EEG segments corresponding to the duration of each movie. Each channel of the EEG data was divided into the same-length epochs of 1s without overlapping. There were about 3300 clean epochs for one experiment. Features were further computed on each epoch of the EEG data. All signal processing was performed in the Matlab software.

B. Feature Extraction

An efficient feature called differential entropy (DE) [35], [36] extends the idea of Shannon entropy and is used to measure the complexity of a continuous random variable [46]. Since EEG data has the higher low frequency energy over high frequency energy, DE has the balance ability of discriminating EEG pattern between low and high frequency energy, which was first introduced to EEG-based emotion recognition by Duan *et al.*[36].

The original calculation formula of differential entropy is defined as

$$h(X) = -\int_X f(x)log(f(x))dx. \tag{1}$$

If a random variable obeys the Gaussian distribution $N(\mu, \sigma^2)$, the differential entropy can simply be calculated by the following formulation:

$$h(X) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(x-\mu)^2}{2\sigma^2} \log \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(x-\mu)^2}{2\sigma^2} dx = \frac{1}{2} \log 2\pi e\sigma^2. \tag{2}$$

It has been proven that, for a fixed length EEG segment, differential entropy is equivalent to the logarithm energy spectrum in a certain frequency band [35]. So differential entropy can be calculated in five frequency bands (delta: 1-3 Hz, theta: 4-7 Hz, alpha: 8-13 Hz, beta: 14-30 Hz, gamma: 31-50 Hz) with time complexity $O(KN \log N)$, where K is the number of electrodes, and N is the size of samples.

For a specified EEG sequence, we used a 256-point Short-Time Fourier Transform with a nonoverlapped Hanning window of 1s to extract five frequency bands of EEG signals. Then we calculated differential entropy for each frequency band. Since each frequency band signal has 62 channels, we extracted differential entropy features with 310 dimensions for a sample.

As the previous studies suggested [38], [47], the asymmetrical brain activity (lateralization in left-right direction and caudality in frontal-posterior direction) seems to be effective in the emotion processing. So we also computed differential asymmetry (DASM) and rational asymmetry (RASM) features [36] as the differences and ratios between the DE features of 27 pairs of hemispheric asymmetry electrodes ($Fp1$, $F7$, $F3$, $FT7$, $FC3$, $T7$, $P7$, $C3$, $TP7$, $CP3$, $P3$, $O1$, $AF3$, $F5$, $F7$, $FC5$, $FC1$, $C5$, $C1$, $CP5$, $CP1$, $P5$, $P1$, $PO7$, $PO5$, $PO3$, and $CB1$ of the left hemisphere, and $Fp2$, $F8$, $F4$, $FT8$, $FC4$, $T8$, $P8$, $C4$, $TP8$, $CP4$, $P4$, $O2$, $AF4$, $F6$, $F8$, $FC6$, $FC2$, $C6$, $C2$, $CP6$, $CP2$, $P6$, $P2$, $PO8$, $PO6$, $PO4$, and $CB2$ of the right hemisphere). DASM and RASM are, respectively, defined as

$$DASM = DE(X_{\text{left}}) - DE(X_{\text{right}}) \tag{3}$$

and

$$RASM = DE(X_{\text{left}})/DE(X_{\text{right}}) \tag{4}$$

where X_{left} and X_{right} represent the pairs of electrodes on the left and right hemisphere. We define DCAU features as the differences between DE features of 23 pairs of frontal-posterior electrodes ($FT7$- $TP7$, $FC5$- $CP5$, $FC3$- $CP3$, $FC1$- $CP1$, FCZ- CPZ, $FC2$- $CP2$, $FC4$- $CP4$, $FC6$- $CP6$, $FT8$- $TP8$, $F7$- $P7$, $F5$- $P5$, $F3$- $P3$, $F1$- $P1$, FZ- PZ, $F2$- $P2$,

[2]http://bcmi.sjtu.edu.cn/seed/index.html

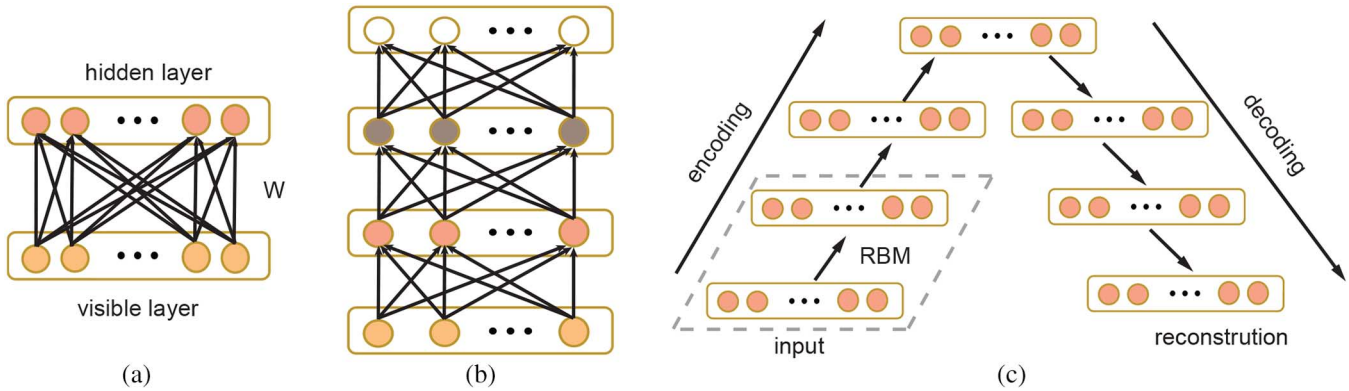


Fig. 2. (a) A RBM contains the hidden layer neurons connected to the visible layer neurons with weights W. (b) A DBN using supervised fine-tuning of all layers with backpropagation. (c) The graphical depiction of unrolled DBN using unsupervised fine-tuning of all layers with backpropagation.

$F4$- $P4$, $F6$- $P6$, $F8$- $P8$, $FP1$- $O1$, $FP2$- $O2$, FPZ- OZ, $AF3$- $CB1$, and $AF4$- $CB2$). DCAU is defined as

$$DCAU = DE(X_{\text{frontal}})/DE(X_{\text{posterior}}) \tag{5}$$

where X_{frontal} and $X_{\text{posterior}}$ represent the pairs of frontal-posterior electrodes.

For comparison, we also extracted conventional power spectral density (PSD) as baseline. The dimensions of PSD, DE, DASM, RASM, and DCAU features are 310, 310, 135, 135, and 115, respectively. We applied the linear dynamic system (LDS) approach to further filter out irrelative components and take temporal dynamics of emotional states into account [48].

C. Classification with Deep Belief Networks

Deep Belief Network is a probabilistic generative model with deep architecture, which characterizes the input data distribution using hidden variables [25], [29]. Each layer of the DBN consists of a restricted Boltzmann machine (RBM) with visible units and hidden units, as shown in Fig. 2(a). There are no visible-visible connections and no hidden-hidden connections. The visible and hidden units have a bias vector, c and b, respectively.

A DBN is constructed by stacking a predefined number of RBMs on top of each other, where the output from a lower-level RBM is the input to a higher-level RBM, as shown in Fig. 2(b). An efficient greedy layer-wise algorithm is used to pre-train each layer of networks.

In an RBM, the joint distribution $P(v, h; \theta)$ over the visible units v and hidden units h, given the model parameters θ, is defined in terms of an energy function $E(v, h; \theta)$ as

$$P(v, h; \theta) = \frac{\exp(-\mathrm{E}(\mathrm{v}, \mathrm{h}; \theta))}{Z} \tag{6}$$

where $Z = \sum_v \sum_h \exp(-E(v, h; \theta))$ is a normalization factor, and the marginal probability that the model assigns to a visible vector v is

$$P(v; \theta) = \frac{\sum_h \exp(-E(v, h; \theta))}{Z}. \tag{7}$$

For a Gaussian (visible)-Bernoulli (hidden) RBM, the energy function is defined as

$$E(v, h; \theta) = -\sum_{i=1}^{I}\sum_{j=1}^{J} w_{ij} v_i h_j - \frac{1}{2}\sum_{i=1}^{I}(v_i - b_i)^2 - \sum_{j=1}^{J} a_j h_j \tag{8}$$

where w_{ij} is the symmetric interaction term between visible unit v_i and hidden unit h_j, b_i and a_j are the bias term, and I and J are the numbers of visible and hidden units. The conditional probabilities can be efficiently calculated as

$$P(h_j = 1 | v; \theta) = \sigma\left(\sum_{i=1}^{I} w_{ij} v_i + a_j\right) \tag{9}$$

$$P(v_i = 1 | h; \theta) = N\left(\sum_{j=1}^{J} w_{ij} h_j + b_i, 1\right) \tag{10}$$

where $\sigma(x) = 1/(1 + \exp(x))$, and v_i takes real values and follows a Gaussian distribution with mean $\sum_{j=1}^{J} w_{ij} h_j + b_i$ and variance one.

Taking the gradient of the log likelihood $\log p(v; \theta)$, we can derive the update rule for adjusting RBM weights as

$$\Delta w_{ij} = E_{\text{data}}(v_i h_j) - E_{\text{model}}(v_i h_j) \tag{11}$$

where $E_{\text{data}}(v_i h_j)$ is the expectation observed in the training set and $E_{\text{model}}(v_i h_j)$ is the same expectation under the distribution defined by the model. But $E_{\text{model}}(v_i h_j)$ is intractable to compute so the contrastive divergence approximation to the gradient is used, where $E_{\text{model}}(v_i h_j)$ is replaced by running the Gibbs sampler initialized at the data for one full step. Sometimes momentum in weight update is used for preventing getting stuck in local minima and regularization prevents the weights from getting too large [49].

In this paper, training is performed in three steps: 1) unsupervised pretraining of each layer; 2) unsupervised fine-tuning of all layers with backpropagation; and 3) supervised fine-tuning of all layers with backpropagation. For unsupervised fine-tuning, n RBMs are unrolled to form a $2n-1$ directed encoder and decoder network that can be fine-tuning with backpropagation [25], [49]. Fig. 2(c) shows the graphical depiction of unrolled DBN. The goal of training this deep autoencoder is to learn the weights and biases between each

TABLE I
DETAILS OF FILM CLIPS USED IN OUR EMOTION EXPERIMENT

No.	Emotion label	Film clips sources
1	negative	Tangshan Earthquake
2	negative	1942
3	positive	Lost in Thailand
4	positive	Flirting Scholar
5	positive	Just Another Pandora's Box
6	neutral	World Heritage In China

layer such that the reconstruction and the input are as close to each other as possible. For supervised fine-tuning, a label layer is added to the top of pretrained DBN and the weights are updated through error backpropagation.

IV. EXPERIMENTS

A. Stimuli

It is important to design efficient and reliable emotion elicitation stimuli for emotion experiments. Nowadays, there are various kinds of stimuli used in emotion research like image, music, metal imagery, and films. Compared to other stimuli, emotional films have several advantages. The existing studies have already evaluated the reliability and efficiency of film clips to elicitation [50], [51]. Emotional films contain both scene and audio, which can expose subjects to more real-life scenarios and elicit strong subjective and physiological changes. So in our experiment, we chose some emotional movie clips to help subjects elicit their own emotions. There are totally fifteen clips in one experiment and each of them lasts for about 4 min. There are three categories of emotions (positive, neutral, and negative) evaluated in this paper and each emotion has five corresponding emotional clips. All the movie clips were carefully chosen as stimuli to help elicit subjects' right emotions from a preliminary study. Since all of the subjects are native Chinese, we selected the emotional clips from Chinese films. The details of the film clips used in this study are listed in Table I.

B. Subjects

Fifteen subjects (7 males and 8 females; MEAN: 23.27, STD: 2.37) with self-reported normal or corrected-to-normal vision and normal hearing participated in the experiments. All participants were right-handed and were students from Shanghai Jiao Tong University. We selected the subjects using the Eysenck Personality Questionnaire (EPQ). The EPQ is a questionnaire to assess the personality traits of a person devised by Eysenck *et al.*[52]. They initially conceptualized personality as three biologically based independent dimensions of temperament measured on a continuum: Extraversion/Introversion, Neuroticism/Stability and Psychoticism/Socialisation. It seems that not every subject can elicit specific emotions immediately, even with the stimuli. The subjects who are extraverted and have stable moods tend to elicit the right emotions throughout the emotion experiments. So from the feedback of the EPQ questionnaires, we selected these subjects to participate in the emotion experiments.

Fig. 3. The experiment scene.

In advance, the subjects were informed about the procedure. The subjects were instructed to sit comfortably, watch the forthcoming movie clips attentively, and refrain as much as possible from overt movements. Fig. 3 shows the experiment scene. The subjects got paid for their participation in the experiments. Each subject participated in the experiment twice at an interval of one week or longer.

C. Protocol

We performed the experiments in a quiet environment in the morning or early in the afternoon. EEG was recorded using an ESI NeuroScan System at a sampling rate of 1000 Hz from 62-channel electrode cap according to the international 10-20 system. The layout of EEG electrodes on the cap is shown in Fig. 4. To remove eye-movement artifacts, we recorded the electrooculogram. The frontal face videos were also recorded from the camera mounted in front of the subjects. There are totally fifteen sessions in one experiment. There is a 5s hint before each clip, 45s for self-assessment and 15s for rest after each clip in one session. For self-assessment, the questions are following Philippot [53]: 1) what they had actually felt in response to viewing the film clip; 2) have they watched this movie before; 3) have they understood the film clip. Fig. 5 shows the detailed protocol.

V. EXPERIMENT RESULTS

A. Neural Patterns

After extracting differential entropy features from five frequency bands (Delta, Theta, Alpha, Beta, and Gamma), we further investigate the neural patterns associated with different emotions. The DE feature map of one experiment is shown in Fig. 6. We find that there exist specific neural patterns in high frequency bands for positive, neutral and negative emotions through time-frequency analysis. For positive emotion, it shows that energy of beta and gamma frequency bands increases whereas neutral and negative emotions have lower energy of beta and gamma frequency bands. While the neural patterns of neutral and negative emotions have similar patterns in beta and gamma bands, neutral emotions have higher energy of alpha oscillations. These findings provide fundamental evidences for

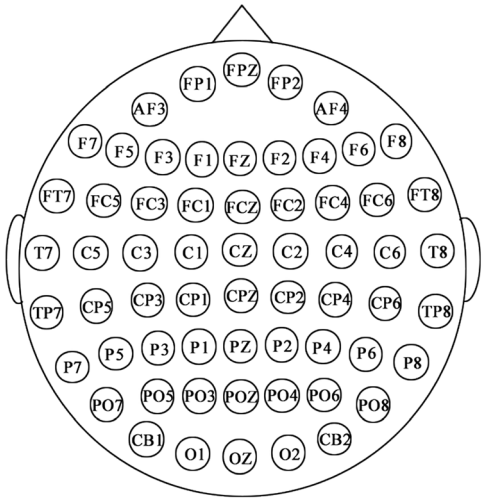


Fig. 4. The EEG cap layout for 62 electrodes.

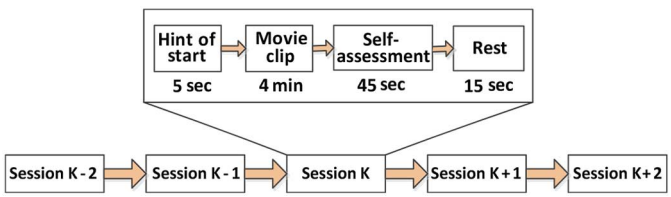


Fig. 5. Protocol of the EEG experiment.

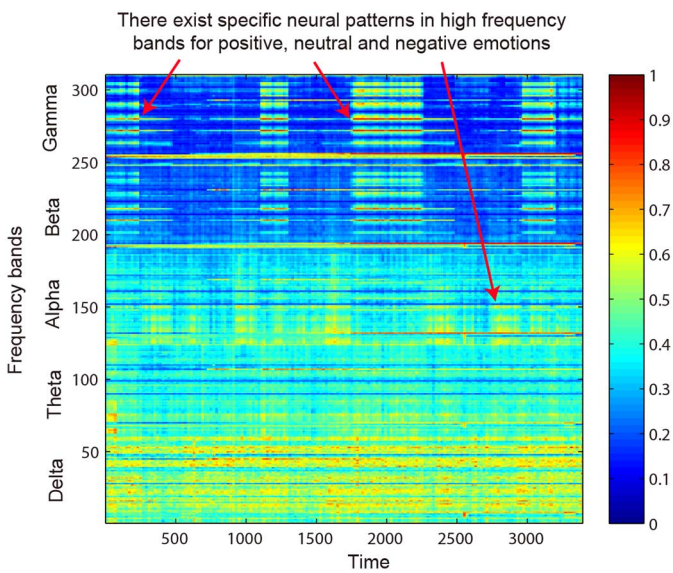


Fig. 6. The DE feature map in one experiment, where the time frames are on the horizontal axis, and the DE features are on the vertical axis.

understanding the mechanism of emotion processing in the brain.

The observed frequencies have been divided into specific groups, as specific frequency ranges are more prominent in certain states of mind. Previous neuroscience studies [54], [55] have shown that EEG alpha bands reflect attentional processing and beta bands reflect emotional and cognitive processing in the brain. Li and Lu [22] also showed that gamma bands of EEG are suitable for emotion classification with emotional images as stimuli. Our findings are consistent with the existing results. When participants watch neutral stimuli, they tend to be more relaxed and less attentional, which evoke alpha responses. And when processing positive emotion, the energy of beta and gamma response enhance.

TABLE II
THE DETAILS OF PARAMETERS USED IN DIFFERENT CLASSIFIERS

Classifiers	Parameter Details
kNN	k=5
LR $L2$	$L2$-regularized, tune the regularization in [1.5:10] with a step of 0.5
SVM	Linear kernel, search space $2^{[-10:10]}$ with a step of one for C
DBN	Structure with 2 hidden layers: search optimal numbers of neurons at the first and the second hidden layers in the ranges of [200:500] and [150:500], respectively, with step of 50.
	Mini-batch size: 201
	Unsupervised and supervised Learning rate: 0.5, 0.6
	Momentum parameter: 0.1
	Activation function: sigmoid function

B. Classifier Training

In this paper, we systematically compare the classification performance of four classifiers, K nearest neighbor (kNN), logistic regression (LR), support vector machine (SVM), and deep belief networks (DBNs) for EEG-based emotion recognition. These classifiers use the DE features aforementioned as inputs. In the emotion experiments, we collect the EEG data from fifteen subjects and each subject has done the experiments twice at intervals of about one week. There are totally 30 experiments evaluated here. The training data and the test data are from different sessions of the same experiment. The training data contains nine sessions of data while the test data contains other six sessions of data from the same experiment.

Table II shows the details of parameters used in different classifiers. For kNN, we use $k = 5$ for baseline in comparison with other classifiers. For LR, we employ $L2$-regularized LR and we tune the regularization parameter in [1.5:10] with a step of 0.5. We also use SVM to classify the emotional states for each EEG segment. The basic idea of SVM is to project input data onto a higher dimensional feature space via a kernel transfer function, which is easier to be separated than that in the original feature space. We use LIBSVM software [56] to implement the SVM classifier and employ linear kernel. We search the parameter space $2^{[-10:10]}$ with a step of one for C to find the optimal value.

For deep neural networks, we construct a DBN with two hidden layers. We search the optimal numbers of neurons in the first and the second hidden layers with step of 50 in the ranges

TABLE III
THE MEAN ACCURACIES AND STANDARD DEVIATIONS (%) OF SVM AND DNN FOR DIFFERENT KINDS OF FEATURES

Feature	Classifier	Delta	Theta	Alpha	Beta	Gamma	Total
PSD	SVM	58.03/15.39	57.26/15.09	59.04/15.75	**73.34/15.20**	71.24/16.38	59.60/15.93
	DNN	60.05/16.66	55.03/13.88	52.79/15.38	60.68/21.31	**63.42/19.66**	61.90/16.65
DE	SVM	60.50/14.14	60.95/10.20	66.64/14.41	80.76/11.56	79.56/11.38	**83.99/09.72**
	DNN	64.32/12.45	60.77/10.42	64.01/15.97	78.92/12.48	79.19/14.58	**86.08/08.34**
DASM	SVM	48.87/10.49	53.02/12.76	59.81/14.67	**75.03/15.72**	73.59/16.57	72.81/16.57
	DNN	48.79/09.62	51.59/13.98	54.03/17.05	69.51/15.22	70.06/18.14	**72.73/15.93**
RASM	SVM	47.75/10.59	51.40/12.53	60.71/14.57	74.59/16.18	74.61/15.57	**74.74/14.79**
	DNN	48.05/10.37	50.62/14.02	56.15/15.28	70.31/15.62	68.22/18.09	**71.30/16.16**
DCAU	SVM	55.92/14.62	57.16/10.77	61.37/15.97	75.17/15.58	76.44/15.41	**77.38/11.98**
	DNN	54.58/12.81	56.94/12.54	57.62/13.58	70.70/16.33	72.27/16.12	**77.20/14.24**

of [200:500] and [150:500], respectively. We set the unsupervised learning rate and supervised learning rate as 0.5 and 0.6, respectively, in the experiment. We also use momentum in the weight update to prevent getting stuck in local minima. Before putting the DE features into DBN, the values of these features are scaled between 0 and 1 by subtracting the mean, divided by the standard deviation and finally adding 0.5. We implement DBN with the DBNToolbox Matlab code [44] in this study.

C. Classification Performance

The mean accuracies (standard deviations) of DBN and SVM with the DE features from different frequency bands in thirty experiments of fifteen subjects are shown in Table III. It should be noted that 'Total' in Table III represents the direct concatenation of five frequency bands of EEG data in this paper. First, we compare the performance of the DE feature on different frequency bands (Delta, Theta, Alpha, Beta, and Gamma). As we can see from Table III, Gamma and Beta frequency bands perform better than other frequency bands. These results confirm that beta and gamma oscillation of brain activity are more related with emotion processing than other frequency oscillations, which is consistent with our above findings in time-frequency analysis.

We also compare the performance of different features. From the results, we can see that the DE features from total frequency bands achieve the best classification accuracy of 86.08% and lowest standard deviation of 8.34% for DBN. For SVM, we can make a similar conclusion that the DE features from the total frequency bands perform the best. These results show the superior performance of the DE features in comparison with other kinds of features. While the asymmetric features (DASM, RASM and DCAU) have much fewer dimensions than the PSD and DE features, they can achieve comparable accuracies, which prove that the asymmetrical brain activity (lateralization in left-right direction and caudality in frontal-posterior direction) is meaningful in emotion processing.

One of the essential questions for EEG-based emotion recognition is whether it is reliable and robust to recognize emotion in different time for each subject. In order to find a solution to this problem, each subject was asked to participate in the experiment twice at intervals of 1 wk or longer. And we evaluate our models with different EEG data acquired at different time slots. From the results, we come to the conclusion that our models can achieve similar prediction accuracies for each subject's twice experiments, despite manifest differences between people's psychology and slight difference of conductance for different experiments. These results also show the potential strength of the proposed method to identify emotion in different time.

Using the DE features from five frequency bands as inputs, the means and standard deviations of accuracies of kNN, LR, SVM, and DBN are 72.60%/13.16%, 82.70%/10.38%, 83.9%/9.72%, 86.08%/8.34%, respectively. The best accuracy of all-frequency-band features is achieved with DBN, followed by SVM, LR, and last kNN. The results show that the DBN models outperform over other models with higher mean accuracy and lower standard deviations. The DBN model achieves 2.09% higher accuracy and 1.38% lower standard deviation than SVM. While the accuracies vary between different subjects, DBN outperforms other conventional methods for most subjects according to the results of total frequency bands. There are many factors that may affect the classification accuracies between the subjects, including subjects' education background, sociability and their true evoked emotional state when participating in the experiments.

The confusion matrix of different classifiers on one experiment for one subject is shown in Fig. 7, which shows the details of strength and weakness of different classifiers. Each row of the confusion matrix represents the target class and each column represents the predicted class that a classifier outputs. The element (i, j) is the percentage of samples in class i that was classified as class j. From the results in Fig. 7, we can see that in general, positive emotion can be recognized with high accuracies, while negative emotion is most difficult to recognize. For kNN, LR and SVM, they confuse negative emotion with neutral and positive emotion, and cannot classify negative emotion very well. However, DBN can significantly improve the classification accuracies for negative emotion. SVM performs

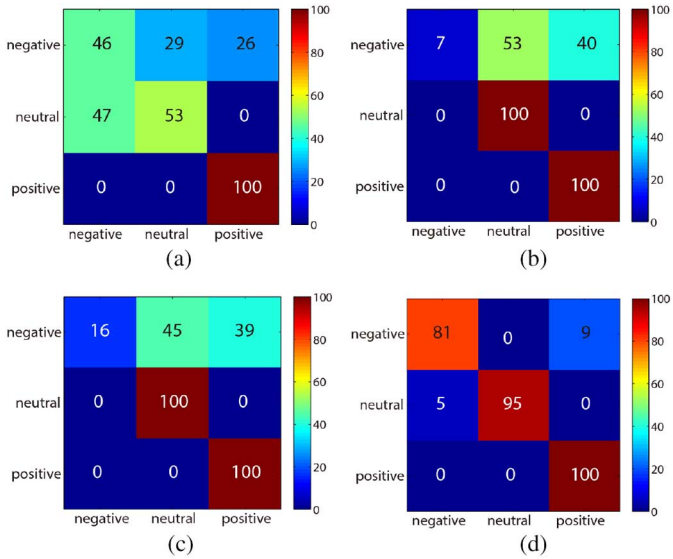


Fig. 7. The confusion matrix of different classifiers on one experiment for one subject. Here the number inside the figures denotes the recognition accuracy in percentage. (a) KNN. (b) LR. (c) SVM. (d) DBN.

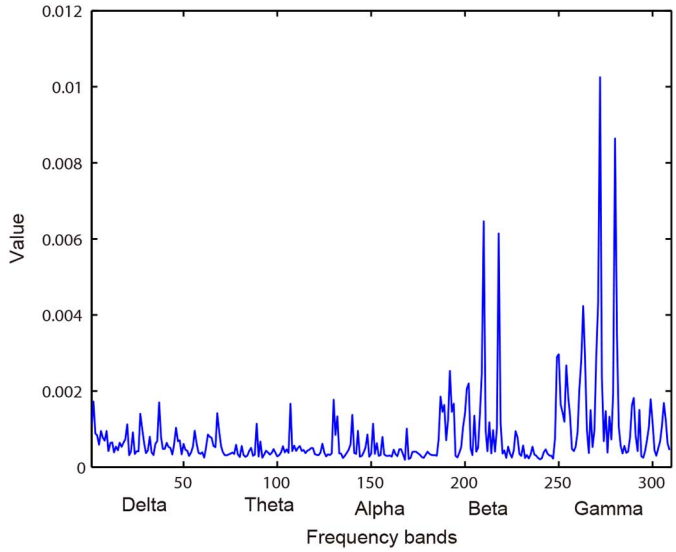


Fig. 8. The mean absolute weight distribution of the trained DBNs learned with the features of direct concatenation of five frequency bands of EEG data.

slightly better than LR and can predict more negative emotion samples accurately. These results show that the deep learning method using DBN has an ability to perform feature selection task to filter out the unrelated features and achieves a better classification accuracy. Feature extraction and feature selection are crucial in the process of emotion modeling. The efficiency of DBN can combine feature extraction and feature selection when doing unsupervised and supervised learning. We will further analyze the powerful representations learned from deep belief networks and how it can select the critical channels and critical frequency bands through weight distributions learned from the deep models in the next session.

The aforementioned experimental results show that DBN methods obtain higher accuracy and lower standard deviation than SVM, LR, and kNN. The reliability of classification performance achieved suggests that such neural signatures associated with positive, neutral and negative emotions do exist. The classification accuracies indicate the possibility of a neural architecture for emotions, and provide modest support for a biologically basic view.

D. Electrode Reduction

In the earlier discussions, we propose the critical frequency bands for emotion recognition through time frequency analysis. Another problem is how to determine critical brain areas associated with emotion recognition. According to our previous work [57], electrode set reduction can not only reduce the computational complexity, but also filter out irrelative noise. Since some EEG channels are irrelevant to emotion recognition [57], these irrelevant channels need more computational cost, introduce noise to emotion recognition, and degrade the performance of trained models. Various studies focus on this problem and try to find the optimal electrodes placement in different tasks. The optimal electrode placement is usually defined according to some statistical factors like correlation coefficient, F-score and accuracy rate in the literature [24], [40], [42]. Some studies share the same pool of electrodes for restrict of commercial EEG device [19], [58].

In this study, we first collect signals of multichannel EEG as many as 62 channels. Then we find the critical channels and frequency bands through analyzing the weight distributions of the trained deep belief networks. Li *et al.* pointed that the EEG data from irrelevant channels are irrelevant to emotion recognition tasks, and the weights of these channels tend to be distributed randomly [34]. According to the rules of knowledge representation, if a particular feature is important, there should be a larger number of neurons involved in representing it in the network [59]. Following this knowledge representing rule in neural network, we assume that the weights of critical channels tend to be updated to certain high values, which can represent how important they are for emotion recognition models. Here, we choose four different setups of electrodes placements and compare their performance with that of full 62 electrodes.

The efficiency of DBN can combine feature extraction and feature selection when doing unsupervised and supervised learning. Fig. 8 shows the mean absolute weight distribution of the trained DBNs in the first layers, where the features are direct concatenation of five frequency bands of EEG data. From Fig. 8, we can see that the high peaks are mostly located at beta and gamma bands. Since the larger weights of corresponding dimensions of inputs contribute more to the output of the neurons in neural networks, this phenomena indicates that the feature components of beta and gamma bands contain more important discriminative information for the tasks learned by the neural networks. In other words, the critical frequency bands for emotion recognition are beta and gamma bands. This observation is consistent with our previous finding [22], [32], [33].

To clearly explore the critical channels selected by the trained DBNs, we further project the mean weight distribution to the brain scalp. Fig. 9 depicts the weight distribution of different brain regions in five frequency bands. These results show that

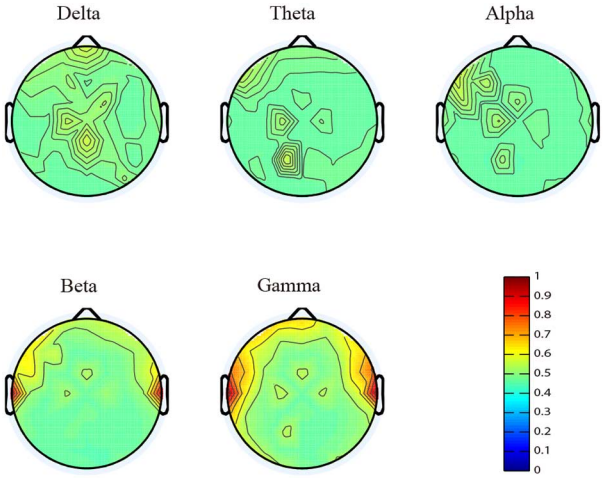


Fig. 9. The weight distribution of different brain regions in five frequency bands.

the neural signatures and patterns associated with positive, neutral and negative emotions do exist. The lateral temporal and prefrontal brain areas activate more than other brain areas in beta and gamma frequency bands.

There is often an interference of facial muscular activities in the EEG signals. Muscle artifacts can affect the patterns of EEG signals. Soleymani *et al.* [60] thought that the correlation between the EEG features and continuous valence was caused by a combination of the effect from the facial expression and brain activities in their study. However, we think that the topographs in Fig. 9 are not due to muscle artifact, but rather brain activity with the following reasons: 1) the significant EMG activities often happens in higher frequency bands (up to 350 Hz), while the raw EEG signals are preprocessed with a bandpass filter between 0.3 to 50 Hz and the recordings seriously contaminated by EMG are removed manually in our study; 2) the subjects are not asked to show their facial expressions explicitly, but rather stay still throughout the experiments; 3) the findings of these neural patterns are consistent with previous emotion studies with EEG [22], [42], [54], [57], [61]. Therefore, we think that the neural patterns shown in Fig. 9 come from the brain activities.

Next, we examine whether the activation patterns underlying positive, neutral and negative emotions could be reduced to a small pool of channels and the performance could be enhanced significantly. We design four different profiles of electrode placements according to the features of high peaks in the weight distribution and asymmetric properties in emotion processing. Fig. 10 shows the four different profiles evaluated in this paper: (a) four channels: $FT7$, $FT8$, $T7$ and $T8$; (b) six channels: $FT7$, $FT8$, $T7$, $T8$, $TP7$ and $TP8$; (c) nine channels: $FP1$, FPZ, $FP2$, $FT7$, $FT8$, $T7$, $T8$, $TP7$ and $TP8$; (d) 12 channels: $FT7$, $FT8$, $T7$, $T8$, $C5$, $C6$, $TP7$, $TP8$, $CP5$, $CP6$, $P7$ and $P8$. The electrodes of profiles (a), (b), and (d) are located in the lateral temporal areas and profile (c) adds three extra prefrontal electrodes.

We extract the PSD, DE, DASM, RASM, and DCAU features of these four profiles and compare their performance with that of full 62 channels. Since the selected pools of electrode sets are

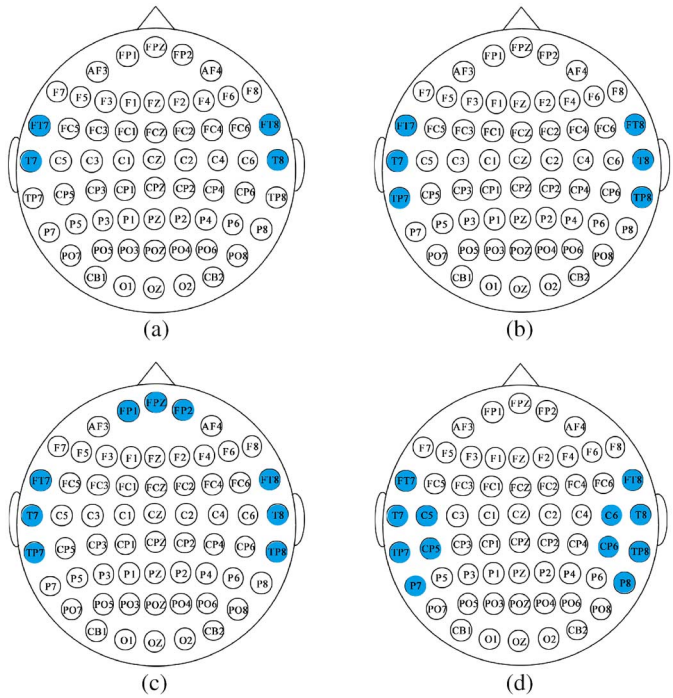


Fig. 10. Four different profiles of selected electrode placements according to the features of high peaks in the weight distribution and asymmetric properties in emotion processing: (a) 4 channels: $FT7$, $FT8$, $T7$ and $T8$; (b) 6 channels: $FT7$, $FT8$, $T7$, $T8$, $TP7$ and $TP8$; (c) 9 channels: $FP1$, FPZ, $FP2$, $FT7$, $FT8$, $T7$, $T8$, $TP7$ and $TP8$; (d) 12 channels: $FT7$, $FT8$, $T7$, $T8$, $C5$, $C6$, $TP7$, $TP8$, $CP5$, $CP6$, $P7$ and $P8$.

reduced to comparably low dimensions as input and these critical channels are selected by deep neural networks after training, it is better to evaluate the performance of these critical channels for emotion recognition models with SVM, which has no explicit feature selection properties. Table IV shows the mean accuracies and standard deviations (%) of SVM for different profiles of electrode sets. For the 4 channels profile, we can see that it can achieve comparably high and stable accuracies of 82.88%/10.92% with the DE features of total frequency bands. With only these four electrodes, our model can achieve the best mean accuracy of 82.88%, which is just slight lower than the accuracy of 83.99% for the full 62 electrodes. What's more, these four electrodes are located at the lateral temporal area, which are easy to mount in real world scenarios. These results suggest the possibility of developing a wearable EEG device for implementing emotion recognition systems for real-world applications.

The best mean accuracies and standard deviations of the 4 channels, the 6 channels, the 9 channels and the 12 channels are 82.88%/10.92%, 85.03%/9.63%, 84.02%/10.34%, 86.65%/8.62%, respectively, while the best mean accuracy and standard deviation of the full 62 channels are 83.99%/9.72%. For all profiles, the DE features attain the best performance among the existing EEG features. These results confirm the conclusion that the DE features are more suitable for EEG-based emotion recognition. Compared to the six channels, the nine channels profile adds extra three frontal electrodes FP1, FPZ, and FP2, which attains slight about 1 percentage lower than six channels

TABLE IV
THE MEAN ACCURACIES AND STANDARD DEVIATIONS (%) OF SVM FOR DIFFERENT PROFILES OF ELECTRODES SETS. (A) 4 CHANNELS; (B) 6 CHANNELS; (C) 9 CHANNELS; (D) 12 CHANNELS

(a)

Feature	Delta	Theta	Alpha	Beta	Gamma	Total
PSD	51.38/14.22	48.39/9.04	57.97/15.86	64.28/14.31	66.60/15.90	74.09/15.73
DE	47.84/11.47	48.52/9.19	57.62/16.63	69.89/15.28	69.20/14.87	**82.88/10.92**
DASM	42.24/5.97	40.55/7.15	45.72/9.15	46.35/12.41	47.41/12.48	70.00/14.84
RASM	41.73/5.89	40.14/6.98	45.54/9.03	45.94/12.13	47.86/12.62	69.21/14.76

(b)

Feature	Delta	Theta	Alpha	Beta	Gamma	Total
PSD	54.99/14.48	56.27/10.59	63.12/14.77	72.59/15.07	74.36/15.37	70.53/14.69
DE	49.99/12.77	57.55/12.56	66.02/12.57	75.82/13.94	75.28/14.28	**85.03/9.63**
DASM	42.50/9.86	44.11/9.43	53.84/12.62	60.26/12.79	61.23/14.04	74.31/11.90
RASM	41.24/9.37	44.17/9.62	53.82/9.44	61.19/14.21	62.58/14.31	73.81/12.34
DCAU	37.56/7.72	41.14/7.22	43.56/9.73	45.41/9.99	45.65/11.08	71.82/14.28

(c)

Feature	Delta	Theta	Alpha	Beta	Gamma	Total
PSD	59.95/15.43	59.03/11.51	70.11/13.28	78.81/11.87	79.03/13.97	69.69/14.43
DE	55.68/14.75	60.65/12.64	71.28/14.20	80.19/10.21	81.33/11.27	**84.02/10.34**
DASM	44.43/10.94	47.25/10.53	58.34/10.29	68.97/15.62	66.38/15.83	75.56/10.23
RASM	43.44/11.64	45.94/10.37	58.17/10.90	65.89/16.23	65.83/14.64	74.80/10.81
DCAU	37.56/7.72	41.14/7.22	43.56/9.73	45.41/9.99	45.65/11.08	71.82/14.28

(d)

Feature	Delta	Theta	Alpha	Beta	Gamma	Total
PSD	57.55/14.68	62.73/13.80	65.87/15.81	75.80/12.72	75.68/12.79	62.92/15.64
DE	54.70/13.45	62.13/13.69	68.18/14.90	77.60/13.58	77.86/14.35	**86.65/8.62**
DASM	46.94/10.42	46.85/9.92	59.45/12.37	69.04/15.19	70.61/14.42	75.86/14.06
RASM	45.88/12.21	45.97 /10.28	58.97/11.14	68.80/15.29	70.58/13.66	75.70/13.74
DCAU	37.56/7.72	41.14/7.22	43.56/9.73	45.41/9.99	45.65/11.08	71.82/14.28

profile. However, it can attain higher accuracies for some individuals and highest accuracies in beta and gamma bands in comparison with other electrodes reduction. These results indicate that the discriminative information of the frontal electrodes is mostly from the beta and gamma oscillations and the patterns of these three frontal electrodes from total frequency bands may not be stable for training models. The profiles of the six channels, the nine channels, and the 12 channels with SVMs achieve better performance than the 62 channels. Moreover, the 12 channels profile with SVM attains the highest accuracies and lowest standard deviations (86.65%/8.62%), even better than the original full 62 channels with SVM (83.99%/9.72%) and deep belief networks (86.08%/8.34%). From these results, we can see that reducing the electrodes by selecting the critical channels can not only save computational cost, but also significantly improve the performance and robustness of emotion recognition models, which are very meaningful for developing wearable devices for brain-computer interfaces with adapting to human emotions in real world applications.

It should be noted that although the 12 channels profile with SVM attains the higher mean accuracies (86.65%) than the original full 62 channels with SVM (83.99%), the rest 50 channels are not 'uninformative' for the emotion recognition task. In this study, we aim to select the minimum pools of electrode sets with comparable performance from DBNs. The neighboring electrodes of the critical electrodes contain redundant discriminative information for emotion recognition, which will be removed from the optimum electrode sets. Moreover, due to the structural and functional differences of the brain across subjects, it may contain different optimum electrode sets for different subjects. Some electrodes contribute a lot for the performance of some subjects, but not for another group of subjects. Here, we aim to explore the critical channels across subjects with the mean weight values learned from DBNs.

VI. DISCUSSION

Despite significant progress of affective computing achieved in recent years, the topic of emotion recognition is still very challenging, due to the fuzzy boundaries of emotion. This paper introduces deep learning to the construction of reliable models of emotion built on brain activity. One of the challenges for affective computing is how to reliably label and evaluate the true evoked emotion. Since reliable labeled data is expensive, it is necessary and important to learn features from unlabeled data, especially for EEG data. Given that DBNs can also learn models in an unsupervised way, the large amounts of unlabeled

EEG data may also be conducive to the semi-supervised DBN training paradigm and allow it to learn more sophisticated models than other traditional supervised learners. Our experimental results show that DBNs can obtain higher classification performance and lower standard deviation in comparison with other shallow models, including kNN, LR and SVM. These findings demonstrate the potential of deep learning for affective modeling, as both manual feature extraction and automatic feature selection could be ultimately bypassed.

The experiment results indicate that beta and gamma bands of EEG data are more related to emotion recognition, which is consistent with the observations in literature. That is, higher frequency brain activities reflect emotional and cognitive processes [54]. We further select critical channels through the weight values learned from DBNs and propose the minimum pools of electrode sets for emotion recognition. Our new approach is different from the existing work. In our studies, we propose a novel critical channels and frequency bands selection method through the weight distributions learned by deep belief networks. Moreover, we examine the performance of different profiles of selected critical channels and propose optimal electrode placements for three categories of emotions. These selected critical channels can achieve relatively stable performance across all the experiments of different subjects, even better than those with the original full 62 channels.

We use a DBN model to show that specific emotional states can be identified with brain activities. The weights learned by DBNs suggest that neural signatures associated with positive, neutral and negative emotions do exist and they share commonality across individuals. These neural signatures are reliably activated across sessions and across individuals. The reliable results inform our understanding of critical channels and frequency oscillations in emotional processes and suggest the potential to infer person's emotional reaction to stimuli on the basis of neural activation.

There are also some limitations in this study. DBN training is an important consideration when applying it to practical applications. But with optimization improvements and using advanced, computing times for training RBM and DBN will certainly decrease. The class of emotions considered here is restricted to just three, i.e., positive, neutral and negative emotions. In the future work, we will apply the proposed method to data sets with a larger category of emotions.

VII. Conclusion

We have applied the DBN models to construction of EEG-based emotion recognition models for three categories of emotions (positive, neutral and negative). The 62-channel EEG signals are recorded from 15 subjects while they are watching emotional film clips with totally 30 experiments. After training the DBN models with the DE features from multichannel EEG data, we have proposed a DBN-based method to select meaningful critical channels and frequency bands through the weight distributions of the trained DBNs and have designed different profiles of electrode sets. The experimental results show that the pools of electrode sets we selected can achieve relatively stable performance across all the experiments of different subjects. The best mean accuracies and standard deviations of the four channels, the six channels, the nine channels and the 12 channels are 82.88%/10.92%, 85.03%/9.63%, 84.02%/10.34%, 86.65%/8.62%, respectively. The profile of the 12 channels with SVM obtains the highest accuracy and lowest standard deviation (86.65%/8.62%) among different pools of electrodes, even better than those of the original full 62 channels with SVM (83.99%/9.72%) and deep belief networks (86.08%/8.34%).

The experimental results also show that the DBN models obtain higher accuracy and lower standard deviation than those of shallow models like kNN, LR and SVM approaches. The reliability of classification performance suggests that specific emotional states can be identified with brain activities. The weights learned by DBNs suggests that neural signatures associated with positive, neutral and negative emotions do exist and they share commonality across individuals.

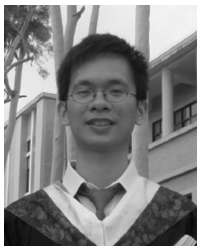
Acknowledgment

The authors would like to thank all the participants in the emotion experiments and thank the Center for Brain-Like Computing and Machine Intelligence for providing the platform for EEG experiments.

References

[1] C. A. Kothe and S. Makeig, "Estimation of task workload from EEG data: New and current tools and perspectives," in *Proc. IEEE Ann. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2011, pp. 6547–6551.
[2] L.-C. Shi and B.-L. Lu, "EEG-based vigilance estimation using extreme learning machines," *Neurocomput.*, vol. 102, pp. 135–143, 2013.
[3] M. Soleymani, M. Pantic, and T. Pun, "Multimodal emotion recognition in response to videos," *IEEE Trans. Affect. Comput.*, vol. 3, no. 2, pp. 211–223, 2012.
[4] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Trans. Affect. Comput.*, vol. 1, no. 1, pp. 18–37, 2010.
[5] H. P. Martinez, Y. Bengio, and G. N. Yannakakis, "Learning deep physiological models of affect," *IEEE Computat. Intell. Mag.*, vol. 8, no. 2, pp. 20–33, 2013.
[6] R. W. Picard, "Affective computing: Challenges," *Int. J. Human-Comp. Stud.*, vol. 59, no. 1, pp. 55–64, 2003.
[7] G. L. Ahern and G. E. Schwartz, "Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis," *Neuropsychologia*, vol. 23, no. 6, pp. 745–755, 1985.
[8] D. Sammler, M. Grigutsch, T. Fritz, and S. Koelsch, "Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music," *Psychophysiol.*, vol. 44, no. 2, pp. 293–304, 2007.
[9] G. G. Knyazev, J. Y. Slobodskoj-Plusnin, and A. V. Bocharov, "Gender differences in implicit and explicit processing of emotional facial expressions as revealed by event-related theta synchronization," *Emotion*, vol. 10, no. 5, p. 678, 2010.
[10] D. Mathersul, L. M. Williams, P. J. Hopkinson, and A. H. Kemp, "Investigating models of affect: Relationships among EEG alpha asymmetry, depression, and anxiety," *Emotion*, vol. 8, no. 4, p. 560, 2008.
[11] C. Grozea, C. D. Voinescu, and S. Fazli, "Bristle-sensors-low-cost flexible passive dry EEG electrodes for neurofeedback and bci applications," *J. Neural Eng.*, vol. 8, no. 2, p. 025008, 2011.
[12] Y. M. Chi, Y.-T. Wang, Y. Wang, C. Maier, T.-P. Jung, and G. Cauwenberghs, "Dry and noncontact EEG sensors for mobile brain–computer interfaces," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 20, no. 2, pp. 228–235, 2012.
[13] L.-F. Wang, J.-Q. Liu, B. Yang, and C.-S. Yang, "PDMS-based low cost flexible dry electrode for long-term EEG measurement," *IEEE Sens. J.*, vol. 12, no. 9, pp. 2898–2904, 2012.
[14] Y.-J. Huang, C.-Y. Wu, A.-K. Wong, and B.-S. Lin, "Novel active comb-shaped dry electrode for EEG measurement in hairy site," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 256–263, 2015.

[15] F. Sauvet, C. Bougard, M. Coroenne, L. Lely, P. Van Beers, M. Elbaz, M. Guillard, D. Leger, and M. Chennaoui, "In flight automatic detection of vigilance states using a single EEG channel," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 12, pp. 2840–2847, Dec. 2014.
[16] N.-H. Liu, C.-Y. Chiang, and H.-M. Hsu, "Improving driver alertness through music selection using a mobile EEG to detect brainwaves," *Sensors*, vol. 13, no. 7, pp. 8199–8221, 2013.
[17] J. B. Van Erp, F. Lotte, and M. Tangermann, "Brain-computer interfaces: Beyond medical applications," *Computer*, no. 4, pp. 26–34, 2012.
[18] B. J. Lance, S. E. Kerick, A. J. Ries, K. S. Oie, and K. McDowell, "Brain–computer interface technologies in the coming decades," in *Proc. IEEE*, 2012, vol. 100, no. Special Centennial Issue, pp. 1585–1599.
[19] P. Aspinall, P. Mavros, R. Coyne, and J. Roe, "The urban brain: Analysing outdoor physical activity with mobile EEG," *Br. J. Sports Med.*, 2013, DOI: 10.1136/bjsports-2012-091877.
[20] M. Paluš, "Nonlinearity in normal human EEG: Cycles, temporal asymmetry, nonstationarity and randomness, not chaos," *Biolog. Cybern.*, vol. 75, no. 5, pp. 389–396, 1996.
[21] M. Dash and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, vol. 1, no. 3, pp. 131–156, 1997.
[22] M. Li and B.-L. Lu, "Emotion classification based on gamma-band EEG," in *Proc. IEEE Ann. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, 2009, pp. 1223–1226.
[23] D. O. Bos, "EEG-based emotion recognition," *Influence Vis. Auditory Stimuli*, pp. 1–17, 2006.
[24] S. Valenzi, T. Islam, P. Jurica, and A. Cichocki, "Individual classification of emotions using EEG," *J. Biomed. Sci. Eng.*, vol. 7, pp. 604–620, 2014.
[25] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
[26] S. Rifai, P. Vincent, X. Muller, X. Glorot, and Y. Bengio, "Contractive auto-encoders: Explicit invariance during feature extraction," in *Proc. 28th Int. Conf. Mach. Learn. (ICML-11)*, 2011, pp. 833–840.
[27] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The Handbook Brain Theory Neural Netw.*, vol. 3361, 1995.
[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, pp. 1097–1105, 2012.
[29] G. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computat.*, vol. 18, no. 7, pp. 1527–1554, 2006.
[30] A.-R. Mohamed, D. Yu, and L. Deng, "Investigation of full-sequence training of deep belief networks for speech recognition," *Interspeech*, pp. 2846–2849, 2010.
[31] N. Jaitly and G. Hinton, "Learning a better representation of speech soundwaves using restricted Boltzmann machines," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)* , 2011, pp. 5884–5887.
[32] W.-L. Zheng, H.-T. Guo, and B.-L. Lu, "Revealing critical channels and frequency bands for EEG-based emotion recognition with deep belief network," in *Proc. IEEE 7th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, 2015, pp. 154–157.
[33] W.-L. Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, "EEG-based emotion classification using deep belief networks," in *Proc. IEEE Int. Conf. Multimed. Expo (ICME)*, Jul. 2014, pp. 1–6.
[34] K. Li, X. Li, Y. Zhang, and A. Zhang, "Affective state recognition from EEG with deep belief networks," in *Proc. IEEE Int. Conf. Bioinformat. Biomed. (BIBM)*, Dec. 2013, pp. 305–310.
[35] L.-C. Shi, Y.-Y. Jiao, and B.-L. Lu, "Differential entropy feature for EEG-based vigilance estimation," in *Proc. IEEE 35th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2013, pp. 6627–6630.
[36] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification," in *Proc. IEEE 6th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, 2013, pp. 81–84.
[37] D. Wu, C. G. Courtney, B. J. Lance, S. S. Narayanan, M. E. Dawson, K. S. Oie, and T. D. Parsons, "Optimal arousal identification and classification for affective computing using physiological signals: Virtual reality stroop task," *IEEE Trans. Affect. Comput.*, vol. 1, no. 2, pp. 109–118, 2010.
[38] R. J. Davidson and N. A. Fox, "Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants," *Science*, vol. 218, no. 4578, pp. 1235–1237, 1982.
[39] R. J. Davidson, "Anterior cerebral asymmetry and the nature of emotion," *Brain Cogn.*, vol. 20, no. 1, pp. 125–151, 1992.
[40] X.-W. Wang, D. Nie, and B.-L. Lu, "Emotional state classification from EEG data using machine learning approach," *Neurocomput.*, vol. 129, pp. 94–106, 2014.
[41] N. Martini, D. Menicucci, L. Sebastiani, R. Bedini, A. Pingitore, N. Vanello, M. Milanesi, L. Landini, and A. Gemignani, "The dynamics of EEG gamma responses to unpleasant visual stimuli: From local activity to functional connectivity," *NeuroImage*, vol. 60, no. 2, pp. 922–932, 2012.
[42] Y.-P. Lin, C.-H. Wang, T.-P. Jung, T.-L. Wu, S.-K. Jeng, J.-R. Duann, and J.-H. Chen, "EEG-based emotion recognition in music listening," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1798–1806, 2010.
[43] S. K. Hadjidimitriou and L. J. Hadjileontiadis, "EEG-based classification of music appraisal responses using time-frequency analysis and familiarity ratings," *IEEE Trans. Affect. Comput.*, vol. 4, no. 2, pp. 161–172, 2013.
[44] M. Längkvist, L. Karlsson, and A. Loutfi, "Sleep stage classification using unsupervised feature learning," *Adv. Artif. Neural Syst.*, vol. 2012, p. 5, 2012.
[45] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; using physiological signals," *IEEE Trans. Affect. Comput.* , vol. 3, no. 1, pp. 18–31, 2012.
[46] J. W. Gibbs, "Elementary principles in statistical mechanics," in *Developed With Especial Reference To The Rational Foundation Of Thermodynamics*. Cambridge, UK: Cambridge Univ. Press, 2010.
[47] Y.-P. Lin, Y.-H. Yang, and T.-P. Jung, "Fusion of electroencephalogram dynamics and musical contents for estimating emotional responses in music listening," *Front. Neurosci.*, vol. 8, no. 94, 2014.
[48] L.-C. Shi and B.-L. Lu, "Off-line and on-line vigilance estimation based on linear dynamical system and manifold learning," in *Proc. IEEE Ann. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Aug. 2010, pp. 6587–6590.
[49] D. Wulsin, J. Gupta, R. Mani, J. Blanco, and B. Litt, "Modeling electroencephalography waveforms with semi-supervised deep belief nets: Fast classification and anomaly measurement," *J. Neural Eng.*, vol. 8, no. 3, p. 036015, 2011.
[50] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cogn. Emotion*, vol. 9, no. 1, pp. 87–108, 1995.
[51] A. Schaefer, F. Nils, X. Sanchez, and P. Philippot, "Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers," *Cogn. Emotion*, vol. 24, no. 7, pp. 1153–1172, 2010.
[52] S. B. Eysenck, H. J. Eysenck, and P. Barrett, "A revised version of the psychoticism scale," *Pers. Individ. Differences*, vol. 6, no. 1, pp. 21–29, 1985.
[53] P. Philippot, "Inducing and assessing differentiated emotion-feeling states in the laboratory," *Cogn. Emotion*, vol. 7, no. 2, pp. 171–193, 1993.
[54] W. J. Ray and H. W. Cole, "EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes," *Science*, vol. 228, no. 4700, pp. 750–752, 1985.
[55] W. Klimesch, M. Doppelmayr, H. Russegger, T. Pachinger, and J. Schwaiger, "Induced alpha band power changes in the human EEG and attention," *Neurosci. Lett.*, vol. 244, no. 2, pp. 73–76, 1998.
[56] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM Trans. Intell. Syst. Technol. (TIST)*, vol. 2, no. 3, p. 27, 2011.
[57] D. Nie, X.-W. Wang, L.-C. Shi, and B.-L. Lu, "EEG-based emotion recognition during watching movies," in *Proc. IEEE 5th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, 2011, pp. 667–670.
[58] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time EEG-based human emotion recognition and visualization," in *Proc. IEEE Int. Conf. Cyberworlds (CW)*, 2010, pp. 262–269.
[59] S. S. Haykin, *Neural Networks and Learning Machines*, 3rd Ed. ed. Upper Saddle River, NJ, USA: Pearson Education, 2009.
[60] M. Soleymani, S. Asghari-Esfeden, M. Pantic, and Y. Fu, "Continuous emotion detection using EEG signals and facial expressions," in *Proc. IEEE Int. Conf. Multimed. Expo. (ICME)*, 2014, pp. 1–6.
[61] M. Balconi and C. Lucchiari, "Consciousness and arousal effects on emotional face processing as revealed by brain oscillations. a gamma band analysis," *Int. J. Psychophysiol.*, vol. 67, no. 1, pp. 41–46, 2008.



Wei-Long Zheng (S'14) received the bachelor's degree in information engineering from the Department of Electronic and Information Engineering, South China University of Technology, Guangzhou, in 2012

He is currently pursuing the Ph.D. degree in computer science with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. His research focuses on affective computing, brain-computer interface, machine learning, and pattern recognition.

Bao-Liang Lu (M'94–SM'01) received the B.S. degree in instrument and control engineering from Qingdao University of Science and Technology, Qingdao, China, in 1982, the M.S. degree in computer science and technology from Northwestern Polytechnical University, Xian, China, in 1989, and the Dr. Eng. degree in electrical engineering from Kyoto University, Kyoto, Japan, in 1994.

He was with Qingdao University of Science and Technology from 1982 to 1986. From 1994 to 1999, he was a Frontier Researcher with the Bio-Mimetic Control Research Center, Institute of Physical and Chemical Research (RIKEN), Nagoya, Japan, and a Research Scientist with the RIKEN Brain Science Institute, Wako, Japan, from 1999 to 2002. Since 2002, he has been a Full Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. He has also been an Adjunct Professor with the Laboratory for Computational Biology, Shanghai Center for Systems Biomedicine, since 2005. His current research interests include brain-like computing, neural network, machine learning, computer vision, bioinformatics, brain-computer interface, and affective computing.

Prof. Lu was the President of the Asia Pacific Neural Network Assembly (APNNA) and the General Chair of the 18th International Conference on Neural Information Processing in 2011. He is currently an Associate Editor of *Neural Networks* and a Board Member of APNNA.