Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

A personality-guided affective brain–computer interface for implementation of emotional intelligence in machines*

Shaojie LI§, Wei LI§, Zejian XING, Wenjie YUAN, Xiangyu WEI, Xiaowei ZHANG†‡, Bin HU†‡

School of Information Science and Engineering, Lanzhou University, Lanzhou 730099, China

†E-mail: zhangxw@lzu.edu.cn; bh@lzu.edu.cn

Received Oct. 14, 2021; Revision accepted July 28, 2022; Crosschecked Aug. 2, 2022

Abstract: Affective brain–computer interfaces have become an increasingly important topic to achieve emotional intelligence in human–machine collaboration. However, due to the complexity of electroencephalogram (EEG) signals and the individual differences in emotional response, it is still a great challenge to design a reliable and effective model. Considering the influence of personality traits on emotional response, it would be helpful to integrate personality information and EEG signals for emotion recognition. This study proposes a personality-guided attention neural network that can use personality information to learn effective EEG representations for emotion recognition. Specifically, we first use a convolutional neural network to extract rich temporal and regional representations of EEG signals, and a special convolution kernel is designed to learn inter- and intra-regional correlations simultaneously. Second, inspired by the fact that electrodes within distinct brain scalp regions play different roles in emotion recognition, a personality-guided regional-attention mechanism is proposed to further explore the contributions of electrodes within a region and between regions. Finally, attention-based long short-term memory is designed to explore the temporal dynamics of EEG signals. Experiments on the AMIGOS dataset, which is a dataset for multimodal research for affect, personality traits, and mood on individuals and groups, show that the proposed method can significantly improve the performance of subject-independent emotion recognition and outperform state-of-the-art methods.

Key words: Electroencephalogram (EEG); Emotion recognition; Attention mechanism; Personality traits
https://doi.org/10.1631/FITEE.2100489 **CLC number:** TP39

1 Introduction

Emotional intelligence (EI) is an important part of human intelligence, and involves the ability to monitor one's own and others' feelings and emotions, discriminate among them, and use this information to guide one's thinking and actions (Salovey and Mayer, 1990). EI plays an important role in making artifact machines have human intelligence, and based on the way human–machine interaction largely imitates human–human interaction (Picard et al., 2001). Emotion is one of the most important cognitive activities in human beings, and adopting this cognitive ability into intelligent machines can enrich and facilitate the collaboration between human and machine. In this context, the ability to identify emotion is one of the hallmarks of EI (Maaoui and Pruski, 2010). Therefore, the research on identification of human

§ These authors contributed equally to this work
‡ Corresponding authors

* Project supported by the National Key R&D Program of China (No. 2019YFA0706200), the National Natural Science Foundation of China (Nos. 62072219 and 61632014), and the National Basic Research Program (973) of China (No. 2014CB744600)

ORCID: Shaojie LI, https://orcid.org/0000-0003-2432-0482; Wei LI, https://orcid.org/0000-0002-5388-129X; Xiaowei ZHANG, https://orcid.org/0000-0001-8562-416X; Bin HU, https://orcid.org/0000-0003-3514-5413

emotional states is one of the hot spots in the field of human–machine collaboration. Emotion recognition has been widely used in various areas, such as health care (Alhussein, 2016; Shen J et al., 2022) and music recommendation systems (Ayata et al., 2018).

The emergence and development of brain–computer interfaces (BCIs) provide an effective method for computer to perceive mental states of human. BCIs consist of the technology that converts signals generated by brain activity into control signals for external devices without the participation of peripheral nerves or muscles (Wolpaw et al., 2002). With the advent of BCI research, the idea of affective BCIs which decodes emotional experience from brain signals arose (Mühl et al., 2014). Affective BCI is a technology that is able to detect, influence, and stimulate affective states, and can be used to modify human–computer interaction (Steinert and Friedrich, 2020). Electroencephalogram (EEG) is collected using several electrodes located on the surface of human head, reflecting the potential neural activity directly (Ding et al., 2020). Compared with other behavioral signals, such as facial expressions and speech, EEG signals reflect changes in the central nervous system during human emotional expressions (Alarcão and Fonseca, 2019), and provide reliable information for emotions compared to visual cues and audio cues (Shu and Wang, 2017). In addition, the use of EEG is noninvasive, fast, and inexpensive, and it is neither painful nor uncomfortable making EEG a preferred method for studying brain's responses to emotional stimuli (Niemic, 2002).

Traditionally, most emotion recognition methods extract hand-crafted features from EEG signals and apply supervised machine learning methods to classify different emotional states. For example, Mohammadi et al. (2017) extracted entropy and energy features from five frequency bands and then used K-nearest neighbor (KNN) and support vector machine (SVM) to detect emotional states. Lan et al. (2016) combined higher-order crossings (HOC) and statistical features and used SVM to classify emotions. Bhardwaj et al. (2015) extracted power spectral density (PSD) and energy features and then fed them into SVM and linear discriminant analysis (LDA) classifiers. Duan et al. (2013) extracted differential entropy (DE) features from five frequency bands and proved that DE features perform better than traditional frequency-domain features. However, hand-crafted feature engineering is usually complicated and time-consuming. Moreover, such methods rely heavily on domain knowledge (Zhang DL et al., 2018), thus lacking the ability to mine useful latent hand-crafted features.

Recently, with the great success of deep learning in different fields (He et al., 2016; Vaswani et al., 2017), many researchers have introduced deep learning methods into EEG-based emotion recognition. Yang et al. (2018) constructed two-dimensional (2D) representations of EEG signals and applied a parallel convolutional recurrent neural network (CRNN) for emotion recognition. Li X et al. (2016) encapsulated EEG signals into grid-like frames and fed them into a hybrid deep learning model that combines the convolutional neural network (CNN) and RNN. Alhagry et al. (2017) proposed a deep learning method to recognize emotion from raw EEG signals, which used long short-term memory (LSTM) to learn features from EEG signals and used the dense layer to classify these features. Wen et al. (2017) rearranged the original channels of an EEG using Pearson correlation coefficient, and the rearranged EEGs were fed into the CNN.

Although numerous EEG-based emotion recognition methods have made tremendous progress in recent years, it is still a complex problem because of diverse open challenges. First, emotion is a highly subjective psychophysiological process, which is often associated with psychological and demographic factors, such as personality, gender, and age. Kehoe et al. (2012) demonstrated several relationships between personality and emotional processing. Fossum and Barrett (2000) found that personality characteristics of neuroticism and extraversion are associated with the experience of negative and positive emotions, respectively. Thus, it would be better to incorporate these factors into affective BCIs. Second, neurological studies have shown that the activities in different brain regions are obviously different during emotion expression, and that these activities are related mainly to a variety of brain regions (Etkin et al., 2011; Kragel and LaBar, 2016). In addition, many studies have found differences in electrodes' activities between people with different personalities (Zhao GZ et al., 2018b). Extraversion was predicted to be associated with great activities of electrodes within the relative left frontal region and increased activities of electrodes within the right posterior region, and introversion was associated with great activities of electrodes within the relative right frontal region and decreased activities of electrodes in the right posterior region (Schmidtke and Heller, 2004). Consequently, exploring the differences in electrode activities effectively would be helpful to improve the performance of emotion recognition.

Motivated by neurological findings and remarkable success achieved by recent deep neural networks (DNNs), we propose a CRNN with personality-guided attention mechanism (PA-CRNN) for EEG-based emotion recognition. By integrating personality factors and EEG signals, the proposed method can not only incorporate personality factors to learn high-level-region representations, but also promote the learning of rich and discriminative representations for emotion recognition. To be specific, PA-CRNN combines attention-based CNN and RNN to explore the complex relations of electrodes within distinct brain scalp regions in the CNN and extract discriminative representations in the RNN. A novel personality-guided attention mechanism is proposed to use personality information to guide attention mechanisms in CNN and RNN. The proposed method is evaluated on the AMIGOS dataset which is publicly available, and empirical evaluations demonstrate that PA-CRNN outperforms state-of-the-art methods.

The main contributions of this paper are summarized as follows: (1) We explore other possible factors that cause significant changes in EEG signals during emotion recognition, especially personality, and show their effectiveness in further strengthening EEG feature extraction. (2) We propose a personality-guided attention mechanism to exploit regional and temporal information of EEG signals, to extract discriminative and informative representations during emotion recognition. (3) By conducting visual analysis of attention weights, we demonstrate that the proposed model has admirable comprehensibility.

2 Related works

2.1 Functional differences in brain scalp regions

Neuroscience research has shown that different brain regions have specific functionalities while coordinating together to accomplish various tasks, and that the specific contributions of different brain regions to classification may change due to the learning of different tasks (Zhang YH et al., 2017). In this context, Chen et al. (2019) divided EEG electrodes into several regions and proposed a new form of connectivity matrix based on the division to adapt the concept of the convolution operation of the CNN for attention deficit hyperactivity disorder (ADHD) studies. Zhang XW et al. (2022) integrated eye movement information into EEGs, divided electrodes into groups, and then used group-sparse canonical correlation analysis (GSCCA) to investigate group structure information. Recent studies have found that humans' emotional response is closely related to some regions of the cerebral cortex (Lotfi and Akbarzadeh-T, 2014). Due to the different contributions of electrodes within distinct brain scalp regions, exploiting the spatial and topographical information of EEG signals to improve classification performance is an attractive and promising solution. Cui et al. (2020) used one-dimensional (1D) convolution layers to learn time–frequency representations and used 2D convolution layers to capture regional information among physically adjacent channels; meanwhile, their study took the asymmetry property of emotion responses into account using the asymmetric feature extractor. However, there are still no physiological explanations for 2D convolution. Li Y et al. (2022) adopted a bidirectional LSTM (BiLSTM) network to capture intrinsic spatial relationships of EEG electrodes; then, a region–attention layer was introduced to learn a set of weights to strengthen or weaken the contributions of electrodes in one brain region, but they needed to extract a set of hand-crafted features first. Different from these methods, we devise a novel attention-based network to automatically and simultaneously learn the inter- and intra-regional relationships of EEG electrodes.

2.2 Influences of different factors on emotion

Emotional experiences can be influenced by people's psychological and demographic factors, such as personality, age, and gender (Orgeta, 2009; Chaplin, 2015; Chevalier et al., 2015). By taking age and gender into account, Rukavina et al. (2016) demonstrated that their approach can lead to an improvement of emotion recognition accuracy. Personality can be described by the "Big-Five" factor model according to five traits, namely, extraversion, neuroticism, agreeableness, conscientiousness, and openness (Perugini and Di Blas, 2002). Relationships between personality and emotion have been studied extensively. Zhao GZ et al. (2018b) recognized an individual's personality traits by analyzing brain activity when he/she watched a series of emotional film clips. Using functional magnetic resonance imaging (fMRI) in a group of women, Kehoe et al. (2012) investigated the relationships among extraversion, neuroticism, and emotional perception. Larsen and Ketelaar (1991) proved that extraverts and neurotics are differentially sensitive to stimuli that generate positive and negative affect, respectively. Accordingly, research on emotion can take advantage of individual differences in sensitivities to situational cues and predispositions to emotional states (Revelle and Scherer, 2009). However, how the information from these factors is being fit as a feature for the classifiers and how the classifiers are being designed are important to capture the correlations between them (Martínez-Tejada et al., 2020).

Fiterau et al. (2017) proposed a method named ShortFuse, which incorporated demographic factors into the model via weights that were shared across the temporal domain. Zhang XW et al. (2020) integrated gender and age into a 1D CNN via an attention mechanism, which could prompt the 1D CNN to explore the complex correlations between EEG signals and demographic factors for the detection of depression, and the results showed that their method is superior to the unitary 1D CNN without gender or age. Zhao SC et al. (2018) proposed to recognize personalized emotions by jointly modeling personality and physiological signals in a hypergraph learning framework.

3 Methodology

Consider an EEG trial $\boldsymbol{X} \in \mathbb{R}^{E\times T}$, where E is the number of EEG electrodes and T is the number of time points. First, the sliding window method is applied to split the EEG trial into several temporal slices. Suppose that an EEG trial $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n]$ contains n temporal slices, where $\boldsymbol{x}_i \in \mathbb{R}^{E\times L}$ and L is the length of a temporal slice. The training set $\varPhi = \left\{\boldsymbol{X}^{(i)}\right\}_{i=1}^{N}$ contains N EEG trials; each EEG trial $\boldsymbol{X}^{(i)}$ corresponds to a label y_i, where $y_i \in \{1, 2, \cdots, C\}$ and C is the number of class labels. Our goal is to learn discriminative EEG features and use these features to classify different emotional states. The overview of the proposed approach is given in Fig. 1, and details of each technical component are described in the subsections.

3.1 Brain scalp regions

First, the divisions of the brain scalp and the groups of EEG electrodes are introduced for presenting the proposed approach. Generally, an EEG records electrical activities of cerebral cortical neurons near the scalp through multiple electrodes. The

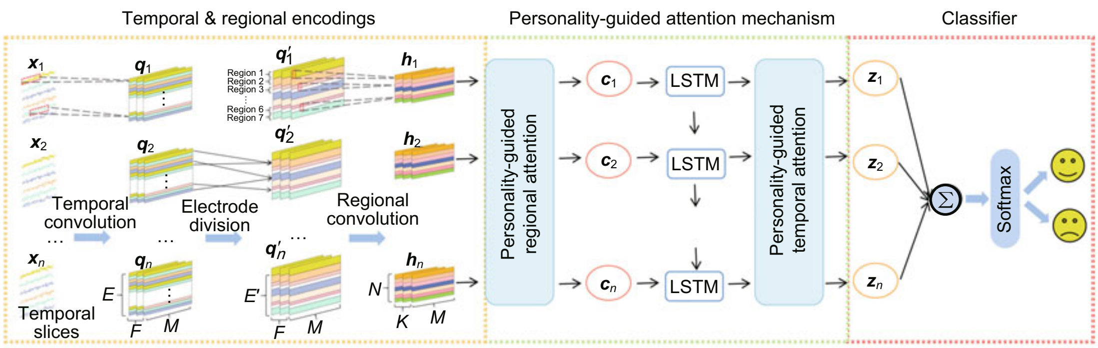

Fig. 1 Flow diagram of our model

(1) Temporal and regional encodings (yellow) first use temporal convolution to extract temporal representation $\boldsymbol{q}_i$ from each EEG slice $\boldsymbol{x}_i$. Then, electrode division is applied to organize electrodes into categories according to their locations. Next, regional convolution is used to extract regional representation $\boldsymbol{h}_i$, which can simultaneously explore inter- and intra-regional correlations of electrodes. (2) Personality-guided attention mechanism (green) includes regional attention and temporal attention, which can explore attentive regional and temporal dynamics by integrating the personality factor. (3) Finally, feature vectors will be fed into one fully connected layer for emotion recognition (red). EEG: electroencephalogram; LSTM: long short-term memory. References to color refer to the online version of this figure

electrodes are usually embedded in electrode caps and placed in accordance with the International 10–20 System. According to the EEG electrode division (Zhang YH et al., 2017; Chen et al., 2019; Zhang T et al., 2020; Ding et al., 2021; Zhang GH et al., 2021; Barkana et al., 2022; Li Y et al., 2022), we group EEG electrodes into seven regions to explore the relationships of electrodes within distinct brain scalp regions. Specifically, the brain scalp is divided into seven topographical regions of interest (ROI): left frontal, left temporal, left occipital, right occipital, right temporal, right frontal, and parietal regions. The standard EEG electrode placement and partition of the scalp are shown in Fig. 2.

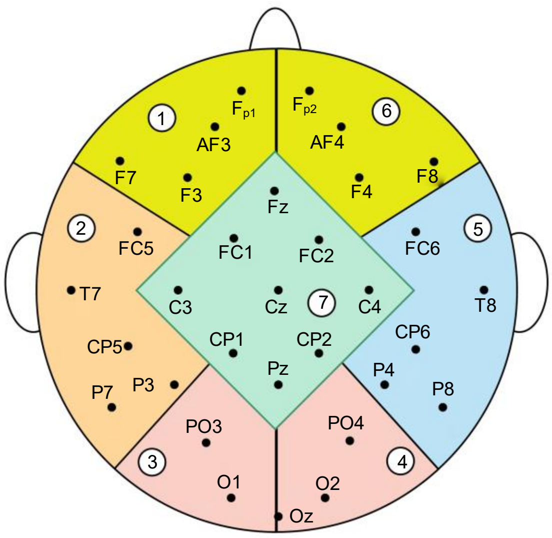

Fig. 2 An illustration of the divisions of the 32 EEG electrodes based on the International 10–20 System
The numbers 1, 2, 3, 4, 5, 6, and 7 represent the left frontal, left temporal, left occipital, right occipital, right temporal, right frontal, and parietal regions, respectively

3.2 Temporal and regional encodings

Due to the powerful learning ability of CNN and the characteristics of EEG signals, multiscale 1D CNN is used to extract potential representations of each EEG electrode only in the time dimension. The multiscale kernel size depends on the EEG signal sampling rate. From the frequency perspective, setting the length of the temporal kernel at half the sampling rate allows for capturing frequency information at 2 Hz and above (Lawhern et al., 2018). In addition, many studies have shown that emotional states are more related to high frequency of EEG signals (Alarcão and Fonseca, 2019). Therefore, the multiscale kernel size is set to different ratios of the EEG sampling rate to learn more diverse high-frequency representations. The ratios in this work are 0.5, 0.25, and 0.125. From the time perspective, a small-size kernel can learn short-term temporal information, and a big-size kernel can extract long-term temporal information. Formally, given an EEG trial $\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]$, each temporal slice $\boldsymbol{x}_i$ will be operated by consecutive convolution with multiscale kernels, and the exponential linear unit (ELU) is used as the activation function, which can result in superior performance for several existing CNN models of EEG signals (Schirrmeister et al., 2017; Lawhern et al., 2018; Waytowich et al., 2018; Farahat et al., 2019; Ma et al., 2021). Then, average pooling is applied to reduce the feature dimension and prevent overfitting. Let $\boldsymbol{q}_i$ denote the output of the convolution operation of the i^{th} slice. Then, this whole process can be expressed as follows:

$$\boldsymbol{q}_i=\text{Avgpool}(\text{ELU}(\text{Conv}(\boldsymbol{x}_i))), \tag{1}$$

where $\boldsymbol{q}_i\in\mathbb{R}^{F\times E\times M}$, F is the number of temporal filters, E is the number of electrodes, and M is the output length after convolution. Then, the generated temporal representation $\boldsymbol{Q}=[\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_n]$ is used as the input to the regional convolution layer.

Inspired by the previous psychophysiological study (Ding et al., 2022), 1D convolution is used to automatically capture the correlations of electrodes between different scalp regions and ultimately generate regional representations. To this end, the order of EEG electrodes should be carefully arranged according to their different locations in brain scalp regions, and electrodes in the same region should be arranged together. In this study, all electrodes are rearranged into seven regions, as shown in Fig. 2. Then, 1D convolution is performed on the arranged temporal features in the electrode dimension to learn regional information; the kernel size and convolution step are set to $(r,1)$, where r is the maximum number of electrodes in all brain scalp regions. Because the number of electrodes in each region is different, we add electrodes with zero values for some regions to ensure the same size. This setting of 1D convolution can simultaneously learn inter- and intra-regional correlations by sharing the convolution kernel. Let $\boldsymbol{Q}'=[\boldsymbol{q}'_1,\boldsymbol{q}'_2,\cdots,\boldsymbol{q}'_n]$ be the regional convolution layer input, where $\boldsymbol{q}'_i\in\mathbb{R}^{F\times E'\times M}$ and E' is the number of electrodes after rearranging and adding values. After specific designed convolution, the regional representations of each slice are as follows:

$$\boldsymbol{h}_i = \mathrm{Avgpool}(\mathrm{ELU}(\mathrm{Conv}(\boldsymbol{q}_i'))), \tag{2}$$

where $\boldsymbol{h}_i \in \mathbb{R}^{K\times N\times M}$, K is the number of regional filters, and N and M are the height and width of the features, respectively. In addition, N is the same size as the number of brain scalp regions. The final encoded temporal-regional representation $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_n]$ is fed to the personality-guided regional-attention layer.

3.3 Personality-guided regional-attention encoding

In the emotion recognition process, not all electrodes contribute equally to the classification task. To further explore the correlations between activities of regional electrodes and people with different personalities, the multi-dimensional self-attention mechanism (Shen T et al., 2018) is integrated into the model to learn attention weights for all regions. The attention weights denote the degrees of contribution of different regions in the emotion recognition process. Different from traditional self-attention mechanisms, which compute only a single scalar score for each input, multi-dimensional self-attention computes a feature-wise score vector for each input, so it can select the features that have more specific meaning.

As shown in Fig. 3, we first permute each slice representation $\boldsymbol{h}_i$ to $\boldsymbol{h}_i' \in \mathbb{R}^{N\times(K\times M)}$. Thus, we can obtain every regional representation of each slice $\boldsymbol{r}_{ik} \in \mathbb{R}^{(K\times M)\times 1}$ $(k = 1, 2, \cdots, N)$, which denotes the k^{th} regional representation of the i^{th} temporal slice and is a row of $\boldsymbol{h}_i'$. Then, we concatenate $\boldsymbol{r}_{ik}$ and the personality vector and obtain $\boldsymbol{r}_{ik}' \in \mathbb{R}^{(K\times M+5)\times 1}$, which can let personality information guide the attention process and generate more informative regional representations. The personality-guided regional-attention mechanism can be formalized as follows:

$$\boldsymbol{r}_{ik}' = [\boldsymbol{r}_{ik}; \boldsymbol{p}], \tag{3}$$

$$\boldsymbol{\mu}_{ik} = \tanh(\boldsymbol{W}_{\mathrm{r}}^1 \boldsymbol{r}_{ik}' + \boldsymbol{b}_{\mathrm{r}}^1), \tag{4}$$

$$\boldsymbol{\alpha}_{ik} = \frac{\exp(\boldsymbol{W}_{\mathrm{r}}^2 \boldsymbol{\mu}_{ik})}{\sum_{k=1}^{N} \exp(\boldsymbol{W}_{\mathrm{r}}^2 \boldsymbol{\mu}_{ik})}, \tag{5}$$

$$\boldsymbol{c}_i = \sum_{k=1}^{N} \boldsymbol{\alpha}_{ik} \otimes \boldsymbol{r}_{ik}, \tag{6}$$

where $\boldsymbol{p}$ represents the personality vector, "$[\cdot]$" stands for the concatenation operation, and "$\otimes$" stands for the element-wise product. Here, the personality-guided attention mechanism can be regarded as a two-layer neural network. The hidden representation of $\boldsymbol{r}_{ik}'$ is computed through one fully connected layer to obtain $\boldsymbol{\mu}_{ik}$. Then, we compute the weight of the k^{th} regional representation and obtain a normalized weight $\boldsymbol{\alpha}_{ik}$ via softmax activation. Finally, we calculate the sum of all regional representation vectors weighted by their attention weights to obtain $\boldsymbol{c}_i$, which is the uniform representation of all N scalp regions. $\boldsymbol{W}_{\mathrm{r}}^1$, $\boldsymbol{b}_{\mathrm{r}}^1$, and $\boldsymbol{W}_{\mathrm{r}}^2$ are the regional-attention parameters that are randomly initialized and jointly learned during the training process. The final encoded uniform representation $\boldsymbol{C} = [\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_n]$ is used as the input to the personality-guided temporal-attention layer.

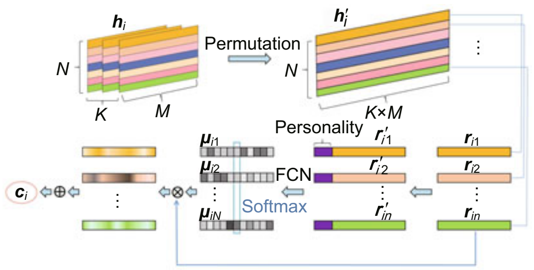

Fig. 3 An illustration of the personality-guided regional-attention mechanism

"$\otimes$" and "$\oplus$" denote the element-wise product and element-wise summation, respectively

3.4 Personality-guided temporal-attention encoding

The personality-guided temporal-attention layer consists of a two-layer LSTM and a personality-guided self-attention mechanism. LSTM network has been successfully used to capture the long-term temporal dynamics of sequences (Rao et al., 2021; Xu et al., 2021). We construct a two-layer LSTM to discover the temporal dynamics of different encoded EEG temporal slices. The output of the LSTM network is denoted as $\boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_n]$, where $\boldsymbol{z}_i \in \mathbb{R}^l$ represents the LSTM's hidden state of the i^{th} temporal slice and l is the dimension of the hidden state for each LSTM unit.

Because people with different personalities have differences in emotional expressions (Fossum and

Barrett, 2000; Vuoskoski and Eerola, 2011; Furnes et al., 2019), personality-guided self-attention is adopted to extract discriminative temporal representations, which integrate personality factors to compute weights for each EEG temporal slice to explore the intrinsic importance of each slice. We concatenate the hidden state outputs and the personality vector to compute the attention weights. The personality-guided self-attention for LSTM can be formalized as follows:

$$\boldsymbol{z}_i' = [\boldsymbol{z}_i; \boldsymbol{p}], \tag{7}$$

$$\boldsymbol{\beta}_i = \tanh(\boldsymbol{W}_{\mathrm{t}}^1 \boldsymbol{z}_i' + \boldsymbol{b}_{\mathrm{t}}^1), \tag{8}$$

$$\boldsymbol{\gamma}_i = \frac{\exp(\boldsymbol{W}_{\mathrm{t}}^2 \boldsymbol{\beta}_i)}{\sum_{i=1}^{n} \exp(\boldsymbol{W}_{\mathrm{t}}^2 \boldsymbol{\beta}_i)}, \tag{9}$$

$$\boldsymbol{v} = \sum_{i=1}^{n} \boldsymbol{\gamma}_i \otimes \boldsymbol{z}_i, \tag{10}$$

where $\boldsymbol{\beta}_i$ is a normalized weight via a softmax activation, and $\boldsymbol{v}$ is the uniform representation of the whole slices that is computed by the sum of all hidden states weighted by their attention weights. $\boldsymbol{W}_{\mathrm{t}}^1$, $\boldsymbol{b}_{\mathrm{t}}^1$, and $\boldsymbol{W}_{\mathrm{t}}^2$ are the temporal-attention parameters that are randomly initialized and jointly learned during the training process.

The ultimate representation $\boldsymbol{v}$ of EEG signals will be classified using a fully connected hidden layer with softmax activation, and the cross-entropy error loss over all labeled samples is evaluated as

$$\mathcal{L} = -\sum_{i} \hat{Y}_i \log P_i, \tag{11}$$

where $\hat{Y}_i$ is the ground-truth label and P_i is the predicted result.

4 Experiments

4.1 Dataset

To verify the effectiveness of the proposed method, extensive experiments were conducted on the AMIGOS dataset (Miranda-Correa et al., 2021). This dataset contained only 14 EEG channels, including AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. Considering that there were no electrodes in the parietal region, we grouped these 14 electrodes into six regions, namely, left frontal, left temporal, left occipital, right occipital, right temporal, and right frontal. During the experiments, the signals were recorded from 40 subjects when they were watching 20 videos, including 16 short videos that were individually played and 4 long videos in groups. The participants rated the self-assessment of arousal, valence, liking, and dominance on a scale from 1 to 9 after watching a video. The durations of the videos were 51–150 s (short) and 841–1410 s (long). Moreover, 5 s baseline data were collected before each affect video. The Big-Five personality traits were measured with a Big-Five marker scale questionnaire (Perugini and Di Blas, 2002), in which for each personality trait, 10 descriptive adjectives were rated with a seven-point scale and a mean was calculated. This study primarily used arousal and valence dimensions during the experiments, because these dimensions can effectively represent various aspects of emotion. In this study, we divided the data into two classes (high/low valence and arousal) according to the mean value of each dimension and obtained two binary classification tasks. Details of the AMIGOS dataset are shown in Table 1.

Table 1 Description of the AMIGOS dataset

Parameter	Description
Number of subjects	40
Numbers of videos	16 (short) and 4 (long)
Video duration	51–150 s (short) and 841–1410 s (long)
Self-assessment	Arousal and valence
Rating scale	1–9
Recorded EEG signals	14 channels, 128 Hz
Big-Five personality traits	Extraversion, neuroticism agreeableness, openness, and conscientiousness
Personality trait scale	1–7

4.2 Data preprocessing

In the experiments, we used 16 short videos of the preprocessed EEG recordings from the AMIGOS dataset. Data preprocessing consisted of downsampling data to 128 Hz, averaging EEG data to the common reference, and applying a bandpass frequency filter between 4.0 and 45.0 Hz. Then, we used the baseline removal method proposed by Yang et al. (2018) to improve recognition accuracy. After baseline removal, we partitioned the preprocessed signals into 20 s nonoverlapping segments for data augmentation. This yielded a final sample of 2127 trials from 36 participants.

4.3 Implementation details

This work used leave-one-subject-out cross-validation to evaluate the performances of several methods for subject-independent emotion recognition, where one subject was used as the testing set and the remaining subjects were used as the training set. The number of temporal slices of an EEG trial was set to 20, and the length of each slice was 128. The sizes of the temporal convolution kernels were 1×64, 1×32, and 1×16. The regional convolution kernel size was set to 3×1, for which the maximum number of electrodes in all regions was 3, according to the location of the 14 electrodes in the AMIGOS dataset. The number of filters was set to 64, and the pooling size was 1×4 with a stride of 4. In addition, the number of hidden states in the LSTM was set to 128. The dropout rate in the fully connected prediction layer was set to 0.5. Batch normalization was adopted to reduce internal covariate shift. All models were implemented with PyTorch framework and optimized using Adam with the learning rate set to 0.001.

4.4 Results

In this subsection, we conducted a comprehensive comparison with existing EEG-based emotion recognition methods in the literature. Meantime, we carried out ablation studies to verify the necessity of attention mechanisms in the proposed approach and compared our method with several state-of-the-art strategies that incorporating the personality factor differently. We computed the accuracy, precision, and F1 score to evaluate the results of all the methods. In addition, the receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) were computed for better comparison.

4.4.1 Performance comparison for different methodologies and ablation studies

First, the proposed approach was compared with several traditional classifiers and deep learning methods. The traditional classifiers include SVM with Gaussian kernel and decision tree (DT). According to Duan et al. (2013), we extracted DE features from the raw EEG signals with respect to theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz), and gamma (31–50 Hz) frequency subbands because emotional states are more related to the high frequency of EEG signals (Alarcão and Fonseca, 2019), and then concatenated them as the final feature vector of size 56 (4×14). Furthermore, we searched the parameter space [10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}] to determine the optimal parameter C for SVM. In DT, we used the Gini index to measure the split quality and set the parameters to default values. The deep learning methods included CRAM (Zhang DL et al., 2019), EEGNet (Lawhern et al., 2018), and DeepConvNet (Schirrmeister et al., 2017). CRAM used a CNN to extract spatial and temporal encodings from EEG signals, and used a recurrent attention mechanism to discover the attentive temporal dynamics (Zhang DL et al., 2019). EEGNet used depthwise and separable convolutions to construct an EEG-specific model that encapsulated well-known EEG feature extraction concepts (Lawhern et al., 2018). DeepConvNet used temporal and spatial filters for end-to-end EEG analysis (Schirrmeister et al., 2017). In addition, we carried out ablation studies to verify the necessity of attention mechanisms in the proposed approach. We designed four models to demonstrate the effectiveness of the personality-guided regional- and temporal-attention mechanisms. They were CNN-RNN (which removed both personality-guided regional- and temporal-attention factors), PACNN-RNN (which removed personality-guided temporal attention), CNN-PARNN (which removed personality-guided regional attention), and A-CRNN (which did not use personality information).

As shown in Table 2 and Fig. 4, the proposed PA-CRNN had obvious advantages over existing EEG-based methods in terms of both arousal and valence dimensions. First, the proposed PA-CRNN evidently outperformed recent deep learning methods (EEGNet, CRAM, and DeepConvNet) and traditional methods relying on hand-crafted features. Specifically, in the classification task on arousal dimension, the proposed PA-CRNN achieved an improvement of 11.1%–16.2% in average accuracy compared with the other methods. Similarly, PA-CRNN achieved an improvement of 6.8%–13% in average precision and 7%–14.6% in average F1 score. In the classification task on valence dimension, the proposed PA-CRNN improved the average recognition accuracy by 10.2%–15.3%, average precision by 5.4%–18.8%, and average F1 score by 7.0%–15.7%. In addition, the results from the paired-sample t-test showed that our method was significantly better than the other EEG-based methods.

Second, we carried out ablation studies to verify the necessity of each attention mechanism in the proposed approach. As shown in Fig. 5, the proposed personality-guided attention mechanism is able to improve the accuracy remarkably by up to 10% on arousal and 8.9% on valence. This observation is consistent with previous study results showing that integrating personality factors could contribute to emotion recognition performance (Zhao SC et al., 2018). PACNN-RNN achieved an improvement of 4.3% and 1.8% respectively, in the accuracy in comparison with the baseline framework CNN-RNN on the arousal and valence dimensions, because personality-guided regional attention further explores the regional dynamics and extracts more discriminative regional representations. In addition, CNN-PARNN improved the accuracy by approximately 2% on two dimensions compared to CNN-RNN, because the personality-guided temporal attention extracts attentive information according to the importance of each time step. A-CRNN improved the accuracy by approximately 5% on two dimensions compared with the baseline framework CNN-RNN. Without the guidance of personality information, A-CRNN still outperformed PACNN-RNN and CNN-PARNN with the help of regional and temporal attention modes. Finally, after introducing the personality traits, the prior information was used to guide the regional- and temporal-attention mechanisms, which further improved the classification performance. This demonstrates that our approach could extract more discriminative regional and temporal representations by exploiting both attention mechanisms simultaneously and this results in better performance.

Table 2 Performance comparison among different EEG modeling methodologies

Model		Accuracy (%)		Precision (%)		F1 score (%)	
		Arousal	Valence	Arousal	Valence	Arousal	Valence
SVM	Mean	50.8**	55.2**	53.7**	44.6**	53.0**	41.8**
	Std	6.5	6.5	13.3	10.8	9.7	10.3
DT	Mean	50.0**	52.1**	53.7**	41.8**	51.0**	41.2**
	Std	7.5	6.3	13.2	13.0	8.5	10.5
DeepConvNet	Mean	51.9**	50.6**	54.0**	54.1*	58.4*	49.9**
	Std	11.6	6.7	20.1	18.9	19.1	18.9
EEGNet	Mean	55.1**	50.1**	57.2*	41.4**	58.6*	43.3**
	Std	12.6	10.3	15.6	17.3	16.1	16.8
CRAM	Mean	50.5**	50.2**	51.0**	40.7**	53.2**	43.1**
	Std	10.0	10.7	16.3	14.2	15.8	13.9
PA-CRNN	Mean	**66.2**	**65.4**	**64.0**	**59.5**	**65.6**	**56.9**
	Std	**8.5**	**8.8**	**8.7**	**7.7**	**16.0**	**13.4**

* for paired-sample t-test at significance level $\alpha = 0.05$ and ** for paired-sample t-test at significance level $\alpha = 0.01$. Best results are in bold

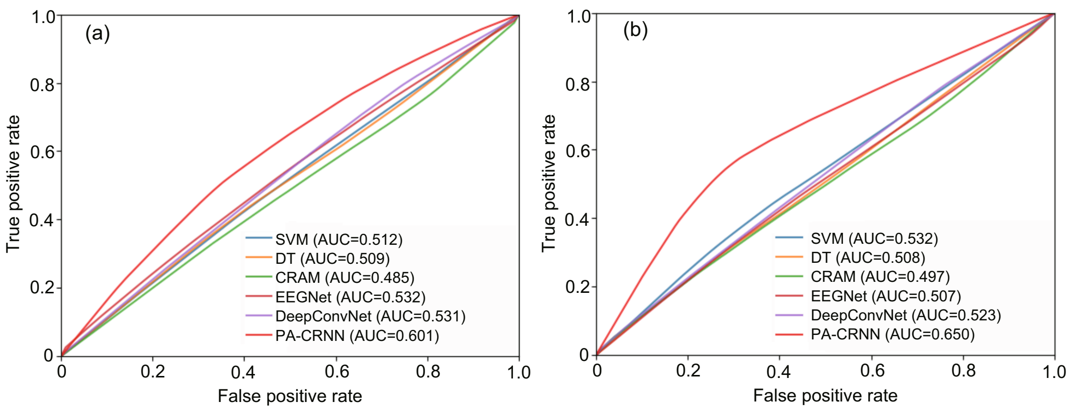

Fig. 4 Comparison of the ROC curves among different EEG modeling methodologies: (a) arousal; (b) valence

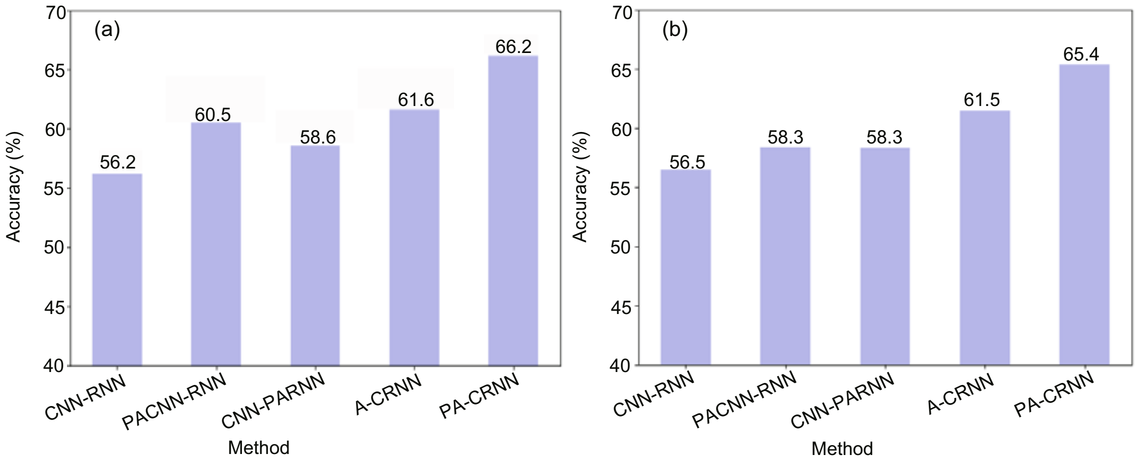

Fig. 5 Average classification accuracy of our model after ablation of different modules: (a) arousal; (b) valence

4.4.2 Performance comparison among different incorporating strategies of the personality factor

In addition, our approach was compared with several state-of-the-art strategies for incorporating the personality factor. For the traditional classifiers, we concatenated DE features and personality vector as inputs, and used SVM and DT as classifiers. For the deep learning methods, we compared our incorporating strategy with three state-of-the-art methods. Fiterau et al. (2017) proposed ShortFuse, which integrated demographic information via convolutions parameterized by the demographics, where the weights were shared across the temporal domain of convolutions. van Leeuwen et al. (2019) proposed a method that concatenated the age factor with EEG representations generated by a CNN and fed them into a fully connected linear layer for classification. Zhang XW et al. (2020) proposed a CNN that integrated gender and age into a 1D CNN via an attention mechanism, which could prompt the 1D CNN to explore complex correlations between EEG signals and demographic factors.

As shown in Table 3 and Fig. 6, the proposed method achieved the best performance compared with classical methods in terms of both arousal and valence dimensions. On the arousal dimension, the proposed PA-CRNN improved the average recognition accuracy by about 15% in comparison with the other methods, and improved the average precision and F1 score by 9.4%–13.1% and 9.4%–15.4%, respectively. On the valence dimension, the proposed PA-CRNN improved the recognition accuracy by 10.0%–14.8%, precision by 14.3%–17.5%, and F1 score by 8.9%–15.0%. According to the results of paired-sample t-test, our method obtained statistically significant improvements over the other methods. Moreover, when compared with the methods that do not consider personality traits, state-of-the-art methods with personality traits did not foster classification performance improvement, or there was only insignificant improvement. By contrast, the proposed method consistently achieved stable and satisfactory performance on both classification tasks. The results demonstrate the effectiveness of PA-CRNN, which integrates personality information and EEG signals via the attention mechanism and achieves the best performance among all compared methods.

4.4.3 Visual analysis of regional attention weights

To present the role of personality-guided attention mechanism intuitively, we depicted the attention weight distribution over various brain scalp regions for each Big-Five personality dimension, which can be used to evaluate the potential dynamic interplay between personality traits and brain scalp regions and their contributions to emotion recognition. For each Big-Five personality dimension, we first selected subjects whose scores were higher than average value on the corresponding dimension. Then, the regional-attention weight vectors of different trials belonging to the selected subjects on each dimension were averaged to obtain the mean weight vectors of different personality types for visualization. We plotted the

Table 3 Performance comparison among different incorporating strategies for the personality factor

Model		Accuracy (%)		Precision (%)		F1 score (%)	
		Arousal	Valence	Arousal	Valence	Arousal	Valence
SVM-P	Mean	50.3**	55.4**	53.6**	44.8**	51.9**	41.9**
	Std	7.5	6.1	14.0	9.9	10.3	10.0
DT-P	Mean	51.0**	51.8**	54.5**	42.0**	50.2**	42.0**
	Std	7.0	6.0	14.2	10.9	10.6	8.7
Fiterau et al. (2017)'s	Mean	50.4**	52.7**	50.9**	45.2**	50.5**	44.8**
	Std	11.6	10.6	20.5	12.7	19.4	15.6
van Leeuwen et al. (2019)'s	Mean	51.5**	50.6**	54.6**	43.3**	56.2**	48.0**
	Std	10.8	9.2	15.2	11.0	14.4	16.8
Zhang XW et al. (2020)'s	Mean	50.8**	52.3**	53.8**	43.7**	52.7**	47.2**
	Std	12.7	9.7	15.3	11.4	14.4	11.2
PA-CRNN	Mean	**66.2**	**65.4**	**64.0**	**59.5**	**65.6**	**56.9**
	Std	**8.5**	**8.8**	**8.7**	**7.7**	**16.0**	**13.4**

SVM-P: support vector machine with personality factor; DT-P: decision tree with personality factor. ** for paired-sample t-test at significance level $\alpha = 0.01$. Best results are in bold

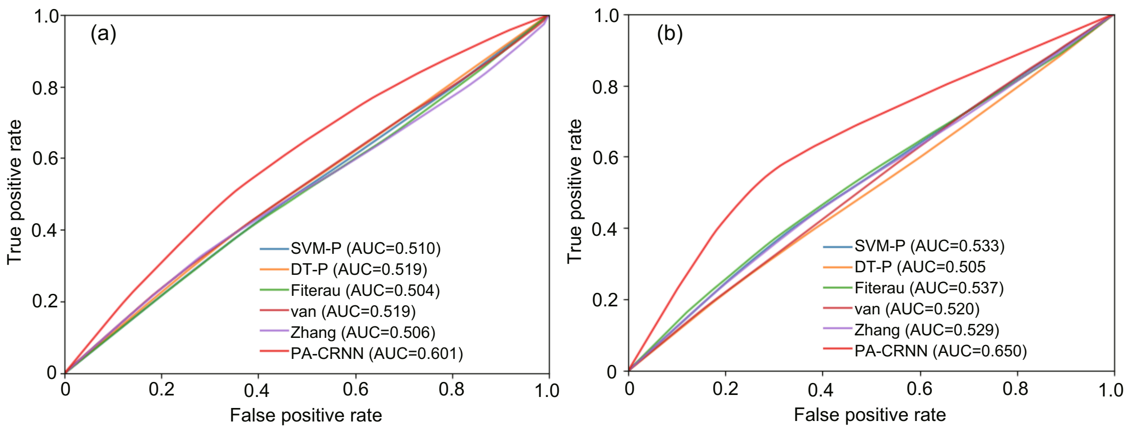

Fig. 6 Comparison of the ROC curves among different incorporating strategies for the personality factor: (a) arousal; (b) valence (Fiterau: Fiterau et al. (2017); van: van Leeuwen et al. (2019); Zhang: Zhang XW et al. (2020))

arithmetic mean of each regional-attention weight vector on the corresponding brain scalp region for intuitive interpretations, in which the areas with deep red color mean significant contributions of the corresponding scalp regions. As shown in Fig. 7, people with different personality traits present different regional-attention weights. On the whole, subjects who reported high levels of extraversion, agreeableness, conscientiousness, and openness show more sharp contrast in attention weights for the arousal label. For the valence label, a sharp contrast in attention weights appears in people with higher levels of extraversion, agreeableness, neuroticism, and openness. This finding is broadly in line with the results of the analysis about the contributions of EEG data, age, sex, and personality traits to emotion recognition processes (Martínez-Tejada et al., 2020). For all the personality traits, the larger weights are concentrated mainly in the left and right frontal regions. These results are consistent with previous study results showing that human emotional expression is closely related to the activities of the frontal cerebral cortex and its spatially corresponding electrodes (Lindquist and Barrett, 2012; Özerdem and Polat, 2017; Barkana et al., 2022; Topic et al., 2022). In addition, individuals high in extraversion have larger regional weights in the left and right temporal regions. Higher agreeableness is related with brain activity in the occipital regions simultaneously. Higher conscientiousness varies with brain activity in both temporal and parietal areas, while higher neuroticism is associated with activities over the left temporal region and left parietal area. For the openness trait, regional weights of the right frontal and temporal areas are greater than those of the other regions, especially for the arousal label. The findings are consistent with several existing study results demonstrating that each personality trait is strongly associated with some characteristic brain scalp regions (Kehoe et al., 2012; Zhao GZ et al., 2018b; Klados et al., 2020; Li WY et al., 2020). Furthermore, it is noteworthy that the attention weights are significantly asymmetrical between the left and right hemispheres of brain, which coincides with the conclusions in many previous studies that hemispheric asymmetry of EEG can be used as a sensitive feature in affective computing and emotion recognition (Zhao GZ et al., 2018a, 2018b, 2018c; Klados et al., 2020; Martínez-Tejada et al., 2020; Zhang GH et al., 2021). In comparison with some methods that treat all electrodes in the brain scalp regions equally, the proposed attention mechanism can capture the intrinsic spatial relationships of EEG electrodes within a region and between regions, on the one hand, and explore the complex relationships between personality traits and scalp regions and their contributions to emotion recognition task, on the other hand. The proposed method achieves better performance in dealing with EEG signals and improves neurophysiological and neurocognitive interpretations further.

5 Discussion

The extensive experiments on AMIGOS dataset show that the proposed method can effectively integrate personality traits and EEG signals, and this model outperforms all the baseline methods. Before proceeding, it is noteworthy that many previous studies used the k-fold cross-validation strategy to evaluate the performance of their emotion recognition methods. In practice, it is more valuable to design a robust subject-independent emotion recognition model. Therefore, we used the leave-one-subject-out cross-validation method to evaluate the generalization performance across different subjects in this study. Due to the impact of individual differences, it is still difficult to build a robust subject-independent model and obtain good results using

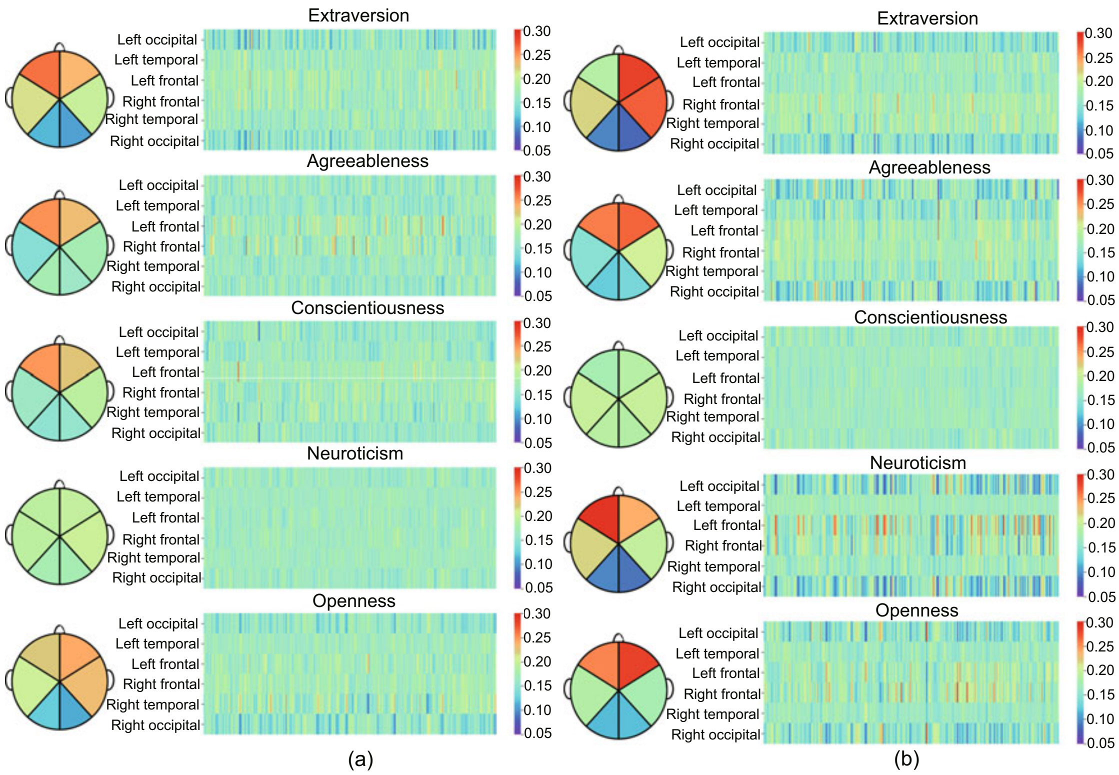

Fig. 7 Visualization of the regional-attention weights for each Big-Five personality dimension on the arousal (a) and valence (b) labels

For each matrix, the rows represent different brain scalp regions and the columns correspond to different dimensions of the regional-attention weight vectors. References to color refer to the online version of this figure the leave-one-subject-out cross-validation strategy (Zhang XW et al., 2021). Inspired by the neurophysiological and neurocognitive findings, we integrated personality traits and EEG signals using a multi-dimensional self-attention mechanism to reduce the influence of individual differences and learned rich high-level representations simultaneously. As shown in Table 2, the proposed method shows significant advantages compared with classical methods on both valence and arousal dimensions. The results justify the important role of personality traits and the effectiveness of our approach. Such observations reveal the potential to extract rich EEG features for performance improvement through incorporation of psychological and demographic factors into emotion recognition. Nevertheless, as shown in Table 3, adding personality traits does not improve classification performance in all cases. Interestingly, the results are consistent with recent study results (Martínez-Tejada et al., 2020). It may be owing to the manner in which the information from the personality questionnaires is being fit as a feature for the classifiers and how the classifiers are designed for emotion recognition (Martínez-Tejada et al., 2020). In this study, personality traits are used to guide the regional attention in CNN and temporal attention in LSTM to promote the learning of rich high-level representations. The results demonstrate the effectiveness of the proposed method in integrating personality traits and EEG signals.

In addition, visual analysis of the attention weight distribution demonstrates that emotion perception is closely related to the activities of the frontal cerebral cortex and its spatially corresponding electrodes. Furthermore, stable personality traits that can influence the precision of an individual's emotion processing can be strongly associated with some characteristic brain scalp regions. These findings can be supported by many existing neurophysiological and neurocognitive studies proving that different personalities could be connected with individual differences in brain function (Vuoskoski and Eerola, 2011; Kehoe et al., 2012; Zhao GZ et al., 2018b; Klados et al., 2020; Li WY et al., 2020).

Moreover, it can be found that the attention weights between the left and right hemispheres of the brain are not symmetrical. For instance, the weights of the right hemisphere are obviously greater than those of the left for higher agreeableness on valence, but for higher neuroticism, the opposite is true. Recently, numerous studies have taken the asymmetry of information into account in affective computing and emotion recognition (Zhao GZ et al., 2018a, 2018b, 2018c; Klados et al., 2020; Martínez-Tejada et al., 2020; Zhang GH et al., 2021). For example, Li Y et al. (2021) proposed a bihemisphere domain adversarial neural network (BiDANN) model to extract asymmetric brain features. In summary, the proposed method could integrate personality traits into an EEG deep learning framework in a suitable way to achieve personalized emotion recognition models. By integrating other demographic information into the proposed knowledge-guided attention mechanism, this framework could be verified on other available benchmark datasets such as DEAP (Koelstra et al., 2012) and SEED (Zheng and Lu, 2015), which are datasets for emotion analysis.

6 Conclusions

In this study, we proposed a novel approach that considers both regional and temporal information to learn effective EEG representations for emotion recognition. To facilitate the learning of regional and temporal representations, a novel personality-guided attention mechanism was proposed to combine EEG signals and personality traits for emotion recognition. Experimental results showed that the proposed method effectively captures the relationships between EEG signals and personality traits. The approach significantly outperformed classical EEG-based emotion recognition methods, and it was superior to several state-of-the-art strategies for incorporating the personality factor. This study facilitates affective BCIs and takes a step toward the full implementation of EI in machines.

Contributors

Shaojie LI and Wei LI designed the research and processed the data. Zejian XING and Wenjie YUAN drafted the paper. Xiangyu WEI helped organize the paper. Xiaowei ZHANG and Bin HU supervised the research, and revised and finalized the paper.

Compliance with ethics guidelines

Shaojie LI, Wei LI, Zejian XING, Wenjie YUAN, Xiangyu WEI, Xiaowei ZHANG, and Bin HU declare that they have no conflict of interest.

References

Alarcão SM, Fonseca MJ, 2019. Emotions recognition using EEG signals: a survey. *IEEE Trans Affect Comput*, 10(3):374-393. https://doi.org/10.1109/TAFFC.2017.2714671

Alhagry S, Fahmy AA, El-Khoribi RA, 2017. Emotion recognition based on EEG using LSTM recurrent neural network. *Int J Adv Comput Sci Appl*, 8(10):355-358. https://doi.org/10.14569/IJACSA.2017.081046

Alhussein M, 2016. Automatic facial emotion recognition using Weber local descriptor for e-Healthcare system. *Clust Comput*, 19(1):99-108. https://doi.org/10.1007/s10586-016-0535-3

Ayata D, Yaslan Y, Kamasak ME, 2018. Emotion based music recommendation system using wearable physiological sensors. *IEEE Trans Consum Electron*, 64(2):196-203. https://doi.org/10.1109/TCE.2018.2844736

Barkana BD, Ozkan Y, Badara JA, 2022. Analysis of working memory from EEG signals under different emotional states. *Biomed Signal Process Contr*, 71:103249. https://doi.org/10.1016/j.bspc.2021.103249

Bhardwaj A, Gupta A, Jain P, et al., 2015. Classification of human emotions from EEG signals using SVM and LDA classifiers. 2nd Int Conf on Signal Processing and Integrated Networks, p.180-185. https://doi.org/10.1109/SPIN.2015.7095376

Chaplin TM, 2015. Gender and emotion expression: a developmental contextual perspective. *Emot Rev*, 7(1):14-21. https://doi.org/10.1177/1754073914544408

Chen H, Song Y, Li XL, 2019. A deep learning framework for identifying children with ADHD using an EEG-based brain network. *Neurocomputing*, 356:83-96. https://doi.org/10.1016/j.neucom.2019.04.058

Chevalier P, Martin JC, Isableu B, et al., 2015. Impact of personality on the recognition of emotion expressed via human, virtual, and robotic embodiments. 24th IEEE Int Symp on Robot and Human Interactive Communication, p.229-234. https://doi.org/10.1109/ROMAN.2015.7333686

Cui H, Liu AP, Zhang X, et al., 2020. EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network. *Knowl Based Syst*, 205: 106243. https://doi.org/10.1016/j.knosys.2020.106243

Ding Y, Robinson N, Zeng QH, et al., 2020. TSception: a deep learning framework for emotion detection using EEG. Int Joint Conf on Neural Networks, p.1-7. https://doi.org/10.1109/IJCNN48605.2020.9206750

Ding Y, Hu X, Xia ZY, et al., 2021. Inter-brain EEG feature extraction and analysis for continuous implicit emotion tagging during video watching. *IEEE Trans Affect Comput*, 12(1):92-102. https://doi.org/10.1109/TAFFC.2018.2849758

Ding Y, Robinson N, Zhang S, et al., 2022. TSception: capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition. *IEEE Trans Affect Comput*, early access. https://doi.org/10.1109/TAFFC.2022.3169001

Duan RN, Zhu JY, Lu BL, 2013. Differential entropy feature for EEG-based emotion classification. 6th Int IEEE/EMBS Conf on Neural Engineering, p.81-84. https://doi.org/10.1109/NER.2013.6695876

Etkin A, Egner T, Kalisch R, 2011. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn Sci*, 15(2):85-93. https://doi.org/10.1016/j.tics.2010.11.004

Farahat A, Reichert C, Sweeney-Reed CM, et al., 2019. Convolutional neural networks for decoding of covert attention focus and saliency maps for EEG feature visualization. *J Neur Eng*, 16(6):066010. https://doi.org/10.1088/1741-2552/ab3bb4

Fiterau M, Bhooshan S, Fries J, et al., 2017. ShortFuse: biomedical time series representations in the presence of structured information. Proc 2nd Machine Learning for Healthcare Conf, p.59-74.

Fossum TA, Barrett LF, 2000. Distinguishing evaluation from description in the personality-emotion relationship. *Pers Soc Psychol Bull*, 26(6):669-678. https://doi.org/10.1177/0146167200268003

Furnes D, Berg H, Mitchell RM, et al., 2019. Exploring the effects of personality traits on the perception of emotions from prosody. *Front Psychol*, 10:184. https://doi.org/10.3389/fpsyg.2019.00184

He KM, Zhang XY, Ren SQ, et al., 2016. Deep residual learning for image recognition. IEEE Conf on Computer Vision and Pattern Recognition, p.770-778. https://doi.org/10.1109/CVPR.2016.90

Kehoe EG, Toomey JM, Balsters JH, et al., 2012. Personality modulates the effects of emotional arousal and valence on brain activation. *Soc Cogn Affect Neurosci*, 7(7):858-870. https://doi.org/10.1093/scan/nsr059

Klados MA, Konstantinidi P, Dacosta-Aguayo R, et al., 2020. Automatic recognition of personality profiles using EEG functional connectivity during emotional processing. *Brain Sci*, 10(5):278. https://doi.org/10.3390/brainsci10050278

Koelstra S, Muhl C, Soleymani M, et al., 2012. DEAP: a database for emotion analysis using physiological signals. *IEEE Trans Affect Comput*, 3(1):18-31. https://doi.org/10.1109/T-AFFC.2011.15

Kragel PA, LaBar KS, 2016. Decoding the nature of emotion in the brain. *Trends Cogn Sci*, 20(6):444-455. https://doi.org/10.1016/j.tics.2016.03.011

Lan ZR, Sourina O, Wang LP, et al., 2016. Real-time EEG-based emotion monitoring using stable features. *Vis Comput*, 32(3):347-358. https://doi.org/10.1007/s00371-015-1183-y

Larsen RJ, Ketelaar T, 1991. Personality and susceptibility to positive and negative emotional states. *J Pers Soc Psychol*, 61(1):132-140. https://doi.org/10.1037/0022-3514.61.1.132

Lawhern VJ, Solon AJ, Waytowich NR, et al., 2018. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *J Neur Eng*, 15(5): 056013. https://doi.org/10.1088/1741-2552/aace8c

Li WY, Hu X, Long XF, et al., 2020. EEG responses to emotional videos can quantitatively predict Big-Five personality traits. *Neurocomputing*, 415:368-381. https://doi.org/10.1016/j.neucom.2020.07.123

Li X, Song DW, Zhang P, et al., 2016. Emotion recognition from multi-channel EEG data through convolutional recurrent neural network. IEEE Int Conf on Bioinformatics and Biomedicine, p.352-359. https://doi.org/10.1109/BIBM.2016.7822545

Li Y, Zheng WM, Zong Y, et al., 2021. A bi-hemisphere domain adversarial neural network model for EEG emotion recognition. *IEEE Trans Affect Comput*, 12(2):494-504. https://doi.org/10.1109/TAFFC.2018.2885474

Li Y, Zheng WM, Wang L, et al., 2022. From regional to global brain: a novel hierarchical spatial-temporal neural network model for EEG emotion recognition. *IEEE Trans Affect Comput*, 13(2):568-578. https://doi.org/10.1109/TAFFC.2019.2922912

Lindquist KA, Barrett LF, 2012. A functional architecture of the human brain: emerging insights from the science of emotion. *Trends Cogn Sci*, 16(11):533-540. https://doi.org/10.1016/j.tics.2012.09.005

Lotfi E, Akbarzadeh-T MR, 2014. Practical emotional neural networks. *Neur Netw*, 59:61-72. https://doi.org/10.1016/j.neunet.2014.06.012

Ma WF, Gong YF, Zhou GX, et al., 2021. A channel-mixing convolutional neural network for motor imagery EEG decoding and feature visualization. *Biomed Signal Process Contr*, 70:103021. https://doi.org/10.1016/j.bspc.2021.103021

Maaoui C, Pruski A, 2010. Emotion recognition through physiological signals for human-machine communication. In: Kordic V (Ed.), Cutting Edge Robotics 2010. IntechOpen, Rijeka, Croatia. https://doi.org/10.5772/10312

Martínez-Tejada LA, Maruyama Y, Yoshimura N, et al., 2020. Analysis of personality and EEG features in emotion recognition using machine learning techniques to classify arousal and valence labels. *Mach Learn Knowl Extr*, 2(2):99-124. https://doi.org/10.3390/make2020007

Miranda-Correa JA, Abadi MK, Sebe N, et al., 2021. AMIGOS: a dataset for affect, personality and mood research on individuals and groups. *IEEE Trans Affect Comput*, 12(2):479-493. https://doi.org/10.1109/TAFFC.2018.2884461

Mohammadi Z, Frounchi J, Amiri M, 2017. Wavelet-based emotion recognition system using EEG signal. *Neur Comput Appl*, 28(8):1985-1990. https://doi.org/10.1007/s00521-015-2149-8

Mühl C, Allison B, Nijholt A, et al., 2014. A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges. *Brain Comput Interf*, 1(2):66-84. https://doi.org/10.1080/2326263X.2014.912881

Niemic CP, 2002. Studies of emotion: a theoretical and empirical review of psychophysiological studies of emotion. *J Undergrad Res*, 1:15-18.

Orgeta V, 2009. Specificity of age differences in emotion regulation. *Aging Ment Health*, 13(6):818-826. https://doi.org/10.1080/13607860902989661

Özerdem MS, Polat H, 2017. Emotion recognition based on EEG features in movie clips with channel selection. *Brain Inform*, 4(4):241-252. https://doi.org/10.1007/s40708-017-0069-3

Perugini M, Di Blas L, 2002. Analyzing personality related adjectives from an eticemic perspective: the Big Five marker scales (BFMS) and the Italian AB5C taxonomy. In: de Raad B, Perugini M (Eds.), Big Five Assessment. Hogrefe & Huber Publishers, Seattle, USA, p.281-304.

Picard RW, Vyzas E, Healey J, 2001. Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Trans Patt Anal Mach Intell*, 23(10):1175-1191. https://doi.org/10.1109/34.954607

Rao HC, Wang SQ, Hu XP, et al., 2021. A self-supervised gait encoding approach with locality-awareness for 3D skeleton based person re-identification. *IEEE Trans Patt Anal Mach Intell*, early access. https://doi.org/10.1109/TPAMI.2021.3092833

Revelle WR, Scherer KR, 2009. Personality and emotion. In: Sander D, Scherer KR (Eds.), Oxford Companion to Emotion and the Affective Sciences. Oxford University Press, Oxford, UK, p.304-305.

Rukavina S, Gruss S, Hoffmann H, et al., 2016. Affective computing and the impact of gender and age. *PLoS ONE*, 11(3):e0150584. https://doi.org/10.1371/journal.pone.0150584

Salovey P, Mayer JD, 1990. Emotional intelligence. *Imaginat Cognit Pers*, 9(3):185-211. https://doi.org/10.2190/DUGG-P24E-52WK-6CDG

Schirrmeister RT, Springenberg JT, Fiederer LDJ, et al., 2017. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum Brain Mapp*, 38(11):5391-5420. https://doi.org/10.1002/hbm.23730

Schmidtke JI, Heller W, 2004. Personality, affect and EEG: predicting patterns of regional brain activity related to extraversion and neuroticism. *Pers Individ Differ*, 36(3):717-732. https://doi.org/10.1016/S0191-8869(03)00129-6

Shen J, Zhang XW, Wang G, et al., 2022. An improved empirical mode decomposition of electroencephalogram signals for depression detection. *IEEE Trans Affect Comput*, 13(1):262-271. https://doi.org/10.1109/TAFFC.2019.2934412

Shen T, Zhou TY, Long GD, et al., 2018. DiSAN: directional self-attention network for RNN/CNN-free language understanding. Proc 32nd AAAI Conf on Artificial Intelligence, p.5446-5455. https://doi.org/10.1609/aaai.v32i1.11941

Shu YY, Wang SF, 2017. Emotion recognition through integrating EEG and peripheral signals. IEEE Int Conf on Acoustics, Speech and Signal Processing, p.2871-2875. https://doi.org/10.1109/ICASSP.2017.7952681

Steinert S, Friedrich O, 2020. Wired emotions: ethical issues of affective brain-computer interfaces. *Sci Eng Ethics*, 26(1):351-367. https://doi.org/10.1007/s11948-019-00087-2

Topic A, Russo M, Stella M, et al., 2022. Emotion recognition using a reduced set of EEG channels based on holographic feature maps. *Sensors*, 22(9):3248. https://doi.org/10.3390/s22093248

van Leeuwen KG, Sun H, Tabaeizadeh M, et al., 2019. Detecting abnormal electroencephalograms using deep convolutional networks. *Clin Neurophysiol*, 130(1):77-84. https://doi.org/10.1016/j.clinph.2018.10.012

Vaswani A, Shazeer N, Parmar N, et al., 2017. Attention is all you need. Proc 31st Int Conf on Neural Information Processing Systems, p.6000-6010.

Vuoskoski JK, Eerola T, 2011. The role of mood and personality in the perception of emotions represented by music. *Cortex*, 47(9):1099-1106. https://doi.org/10.1016/j.cortex.2011.04.011

Waytowich N, Lawhern VJ, Garcia JO, et al., 2018. Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials. *J Neur Eng*, 15(6):066031. https://doi.org/10.1088/1741-2552/aae5d8

Wen ZY, Xu RF, Du JC, 2017. A novel convolutional neural networks for emotion recognition based on EEG signal. Int Conf on Security, Pattern Analysis, and Cybernetics, p.672-677. https://doi.org/10.1109/SPAC.2017.8304360

Wolpaw JR, Birbaumer N, McFarland DJ, et al., 2002. Brain–computer interfaces for communication and control. *Clin Neurophysiol*, 113(6):767-791. https://doi.org/10.1016/S1388-2457(02)00057-3

Xu SH, Rao HC, Peng H, et al., 2021. Attention-based multilevel co-occurrence graph convolutional LSTM for 3-D action recognition. *IEEE Internet Things J*, 8(21):15990-16001. https://doi.org/10.1109/JIOT.2020.3042986

Yang YL, Wu QF, Qiu M, et al., 2018. Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network. Int Joint Conf on Neural Networks, p.1-7. https://doi.org/10.1109/IJCNN.2018.8489331

Zhang DL, Yao LN, Zhang X, et al., 2018. Cascade and parallel convolutional recurrent neural networks on EEG-based intention recognition for brain computer interface. Proc 32nd AAAI Conf on Artificial Intelligence, p.1703-1710. https://doi.org/10.1609/aaai.v32i1.11496

Zhang DL, Yao LN, Chen KX, et al., 2019. A convolutional recurrent attention model for subject-independent EEG signal analysis. *IEEE Signal Process Lett*, 26(5):715-719. https://doi.org/10.1109/LSP.2019.2906824

Zhang GH, Yu MJ, Liu YJ, et al., 2021. SparseDGCNN: recognizing emotion from multichannel EEG signals. *IEEE Trans Affect Comput*, early access. https://doi.org/10.1109/TAFFC.2021.3051332

Zhang T, Cui Z, Xu CY, et al., 2020. Variational pathway reasoning for EEG emotion recognition. Proc 34th AAAI Conf on Artificial Intelligence, 2709-2716. https://doi.org/10.1609/aaai.v34i03.5657

Zhang XW, Li JL, Hou KC, et al., 2020. EEG-based depression detection using convolutional neural network with demographic attention mechanism. 42nd Annual Int Conf of the IEEE Engineering in Medicine & Biology Society, p.128-133. https://doi.org/10.1109/EMBC44109.2020.9175956

Zhang XW, Liu JY, Shen J, et al., 2021. Emotion recognition from multimodal physiological signals using a regularized deep fusion of kernel machine. *IEEE Trans Cybern*, 51(9):4386-4399. https://doi.org/10.1109/TCYB.2020.2987575

Zhang XW, Pan J, Shen J, et al., 2022. Fusing of electroencephalogram and eye movement with group sparse canonical correlation analysis for anxiety detection. *IEEE Trans Affect Comput*, 13(2):958-971. https://doi.org/10.1109/TAFFC.2020.2981440

Zhang YH, Prasad S, Kilicarslan A, et al., 2017. Multiple kernel based region importance learning for neural classification of gait states from EEG signals. *Front Neurosci*, 11:170. https://doi.org/10.3389/fnins.2017.00170

Zhao GZ, Zhang YL, Ge Y, et al., 2018a. Asymmetric hemisphere activation in tenderness: evidence from EEG signals. *Sci Rep*, 8(1):8029. https://doi.org/10.1038/s41598-018-26133-w

Zhao GZ, Ge Y, Shen BY, et al., 2018b. Emotion analysis for personality inference from EEG signals. *IEEE Trans Affect Comput*, 9(3):362-371. https://doi.org/10.1109/TAFFC.2017.2786207

Zhao GZ, Zhang YL, Ge Y, 2018c. Frontal EEG asymmetry and middle line power difference in discrete emotions. *Front Behav Neurosci*, 12:225. https://doi.org/10.3389/fnbeh.2018.00225

Zhao SC, Ding GG, Han JG, et al., 2018. Personality-aware personalized emotion recognition from physiological signals. Proc 27th Int Joint Conf on Artificial Intelligence, p.1660-1667. https://doi.org/10.24963/ijcai.2018/230

Zheng WL, Lu BL, 2015. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans Auton Ment Dev*, 7(3):162-175. https://doi.org/10.1109/TAMD.2015.2431497