



# Nuclear Norm Regularized Deep Neural Network for EEG-Based Emotion Recognition

Shuang Liang<sup>1,2</sup>, Mingbo Yin<sup>3</sup>, Yecheng Huang<sup>3</sup>, Xiubin Dai<sup>1</sup> and Qiong Wang<sup>4\*</sup>

<sup>1</sup> Smart Health Big Data Analysis and Location Services Engineering Lab of Jiangsu Province, Nanjing University of Posts and Telecommunications, Nanjing, China, <sup>2</sup> State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China, <sup>3</sup> School of Computer Science and Technology, Nanjing Tech University, Nanjing, China, <sup>4</sup> Guangdong Provincial Key Laboratory of Computer Vision and Virtual Reality Technology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

## OPEN ACCESS

### Edited by:

Yizhang Jiang,  
Jiangnan University, China

### Reviewed by:

Yuanpeng Zhang,  
Nantong University, China  
Xiaoqing Gu,  
Changzhou University, China

### \*Correspondence:

Qiong Wang  
wangqiong@siat.ac.cn

### Specialty section:

This article was submitted to  
Emotion Science,  
a section of the journal  
Frontiers in Psychology

Received: 20 April 2022

Accepted: 24 May 2022

Published: 29 June 2022

### Citation:

Liang S, Yin M, Huang Y, Dai X and  
Wang Q (2022) Nuclear Norm  
Regularized Deep Neural Network for  
EEG-Based Emotion Recognition.  
*Front. Psychol.* 13:924793.  
doi: 10.3389/fpsyg.2022.924793

Electroencephalography (EEG) based emotion recognition enables machines to perceive users' affective states, which has attracted increasing attention. However, most of the current emotion recognition methods neglect the structural information among different brain regions, which can lead to the incorrect learning of high-level EEG feature representation. To mitigate possible performance degradation, we propose a novel nuclear norm regularized deep neural network framework (NRDNN) that can capture the structural information among different brain regions in EEG decoding. The proposed NRDNN first utilizes deep neural networks to learn high-level feature representations of multiple brain regions, respectively. Then, a set of weights indicating the contributions of each brain region can be automatically learned using a region-attention layer. Subsequently, the weighted feature representations of multiple brain regions are stacked into a feature matrix, and the nuclear norm regularization is adopted to learn the structural information within the feature matrix. The proposed NRDNN method can learn the high-level representations of EEG signals within multiple brain regions, and the contributions of them can be automatically adjusted by assigning a set of weights. Besides, the structural information among multiple brain regions can be captured in the learning procedure. Finally, the proposed NRDNN can perform in an efficient end-to-end manner. We conducted extensive experiments on publicly available emotion EEG dataset to evaluate the effectiveness of the proposed NRDNN. Experimental results demonstrated that the proposed NRDNN can achieve state-of-the-art performance by leveraging the structural information.

**Keywords:** electroencephalography (EEG), emotion recognition, affective brain-computer interface (aBCI), structural information, nuclear norm regularization

## INTRODUCTION

Affective brain computer interface (aBCI) can establish an effective communication pathway between brain and devices (Mühl et al., 2014). Emotion recognition enables aBCI to accurately perceive the affective states of brains, which has attracted increasing attention (Fragopanagos and Taylor, 2005). Naturally, there exist several patterns of emotion expression, such as voice signals (Ang et al., 2002), facial expressions (Xiaohua et al., 2017), body gestures (Yan et al., 2014),

electromyogram (EMG) signals (Cheng and Liu, 2008), electrocardiogram (ECG) signals (Agrafioti et al., 2012), and electroencephalogram (EEG) signal (Zheng, 2017). Among the above techniques, EEG is the most extensively adopted to record brain electrical activities caused by emotional fluctuations because of its portability and non-invasive way (Liu et al., 2017; Zheng, 2017).

As shown in **Figure 1**, a classical EEG-based aBCI system can be divided into four parts (Li J. et al., 2019), namely, signal acquisition, preprocess, feature extraction, and emotion recognition. For the acquisition of EEG signal, the electrical signal of brain activity can be efficiently obtained by the non-invasive electrodes along the scalp. Then, some kinds of filters, e.g., Butterworth and Chebyshev (Bustamante et al., 2015), are adopted to preprocess the original EEG signals to clean the noise. Subsequently, the affective EEG data can be transformed into a suitable feature representation by using the domain-specific feature extractors. In general, the extracted features are mainly represented as follows: (1) time feature; (2) frequency feature; (3) time-frequency feature; and (4) spatial feature. For example, Hjorth Features (Petranonakis and Hadjileontiadis, 2010) and independent component analysis (ICA) (Iacoviello et al., 2015) are the widely used time domain feature extractors. Wavelet transform (WT) (Mazumder, 2019) and wavelet packet decomposition (WPD) (Ting et al., 2008) are commonly adopted as the affective EEG feature extractors in time-frequency domain. Besides, fast Fourier transform (FFT) (Murugappan and Murugappan, 2013) and autoregressive (AR) model (Atyabi et al., 2016) are two widely used frequency domain affective EEG feature extractors. Common spatial pattern (CSP) is the typical spatial feature extractors for EEG data (Ramoser et al., 2000). The main principle of CSP is to learn an optimal spatial projection, which can maximize the variance of two classes by simultaneous diagonalization of their covariance matrices. Besides the methods mentioned above, many other feature extractors have also been developed for affective EEG feature extraction, such as differential entropy (DE) method (Duan et al., 2013), which is widely used for extracting EEG features in time-spatial domain. After feature extraction, many classifiers can be exploited to emotion recognition, such as support vector machine (SVM) (Zhang et al., 2018), k-nearest neighbor (KNN) (Tang et al., 2019), and Linear Discriminant Analysis (LDA) (Zhang et al., 2013). The recognition results are finally feedback to the users.

Although the abovementioned methods have shown their efficacy in EEG-based aBCI system, these methods belong to shallow learning methods that cannot exploit deep EEG feature representations with the powerful deep learning framework. Therefore, many deep learning methods have been developed for EEG-based emotion recognition. For example, Zheng and Lu (2015) proposed to utilize deep belief network to construct emotion recognition model. Pandey and Seeja (2019) developed a multilayer perceptron-based neural network for EEG-based emotion recognition. In Song et al. (2018), proposed to use graph convolution neural network extract EEG features. Besides, Li Y. et al. (2019) developed a spatial-temporal deep neural network model for emotion recognition. In addition to the above methods, many other deep models have been

exploited for decoding EEG, some of them greatly advanced the performance in feature representation and classification. Readers can refer to the in-depth systematic review in Alarcao and Fonseca (2017) for details. Although existing deep learning methods exhibit powerful feature learning capability in dealing with EEG-based emotion data, they have not considered the different contributions of individual brain regions to EEG feature representation and pattern classification.

Recently, neuroscience researches have shown that human emotions are correlated to multiple cerebral cortex regions, such as orbitofrontal cortex and ventromedial prefrontal cortex (Lotfi et al., 2014). Hence, the EEG signals associated with different brain regions might provide different contribution to emotion recognition (Lindquist et al., 2012). In view of this, Li Y. et al. (2019) assigned a set of weights to EEG signals within different brain regions to strengthen or weaken their contributions to EEG decoding. Besides, Park and Chung (2019) selected some good local regions by interquartile range and then adopted local CSP to extract their features. Despite promising progress, most of the current methods do not take into account the reliability of structural information among different brain regions, which can lead to the incorrect learning of high-level EEG feature representation.

Recently, certain methods have been developed to capture the structural information within the EEG feature, such as support matrix machine (SMM) (Luo et al., 2015), Robust SMM (Zheng et al., 2018), and deep stacked SMM (Hang et al., 2020). In Luo et al. (2015), a spectral elastic net regularization was combined with the hinge loss to formulate a matrix classifier, named SMM, which uses nuclear norm to exploit the structural information within EEG feature matrices. Based on SMM, robust SMM (Zheng et al., 2018) was proposed to eliminate outliers within EEG signals and construct a matrix classifier using the recovered clean data. Besides, Hang et al. (2020) adopted SMM as the basic building block to construct a deep stacked SMM, which inherits the characteristic of SMM that can learn the structural information of data as well as the powerful capability of deep representation learning. Although these methods have achieved promise EEG classification performance, they take pre-extracted EEG features as input, which heavily relies on the expertise.

In this study, a novel nuclear norm regularized deep neural network framework (NRDNN) is proposed to capture the structural information among different brain regions in affective EEG decoding. To learn high-level feature representations of multiple brain regions, the proposed NRDNN utilizes different deep neural networks for decoding EEG signals within multiple regions. In view of different brain regions may have different functions for the EEG emotion recognition, NRDNN introduces a region-attention layer to automatically learn weights of different brain regions to strengthen or weaken their corresponding contributions. To leverage the structural information among different brain regions, the weighted feature representations of multiple brain regions are stacked into a feature matrix, the nuclear norm regularization with the hinge loss is used for EEG-based emotion recognition. Besides, NRDNN can be efficiently optimized through standard end-to-end manner. To validate the effectiveness of the proposed method, we conducted extensive experiments on publicly

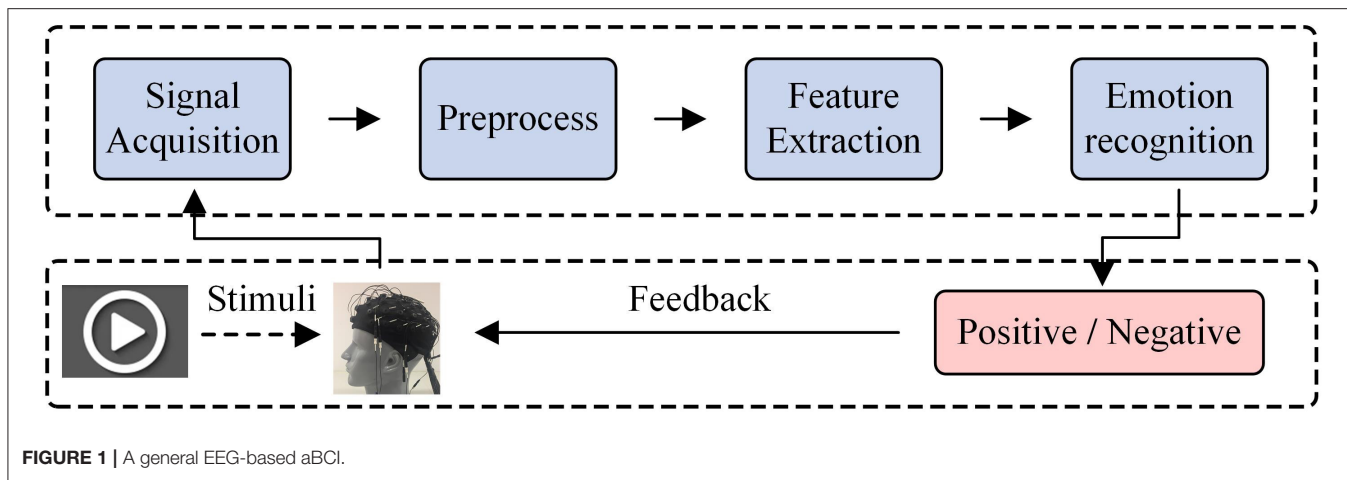


FIGURE 1 | A general EEG-based aBCI.

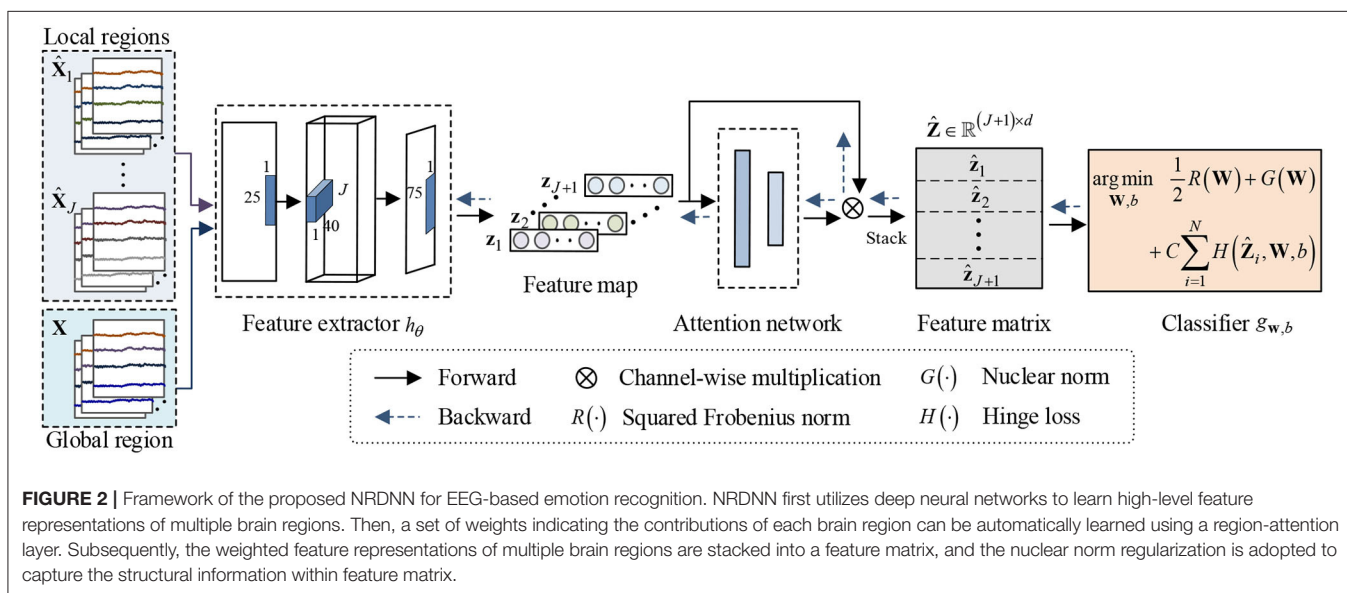


FIGURE 2 | Framework of the proposed NRDNN for EEG-based emotion recognition. NRDNN first utilizes deep neural networks to learn high-level feature representations of multiple brain regions. Then, a set of weights indicating the contributions of each brain region can be automatically learned using a region-attention layer. Subsequently, the weighted feature representations of multiple brain regions are stacked into a feature matrix, and the nuclear norm regularization is adopted to capture the structural information within feature matrix.

available affective EEG dataset. Experimental results demonstrate that our NRDNN outperforms other comparison methods.

The remainder of this study is organized as follows: the “Nuclear norm regularized deep neural network” section illustrates the proposed NRDNN model and its learning algorithm in detail. In the “Experiments” section, extensive experiments and result analyses are presented. Finally, the conclusions of the study can be found in the “Conclusion” section.

## NUCLEAR NORM REGULARIZED DEEP NEURAL NETWORK

Neuroscience studies have shown that relevant information exists among different brain regions (Clark, 1994; Vecchio et al., 2013; Kurmukov et al., 2016). To a certain extent, the structural information among brain regions can reflect this relevant information. Since the nuclear norm of the matrix is the convex

approximation of its rank, it can directly capture its structural information between columns or rows. Hence, this study develops an end-to-end nuclear norm regularized deep learning framework by using the structural information in EEG decoding. The flowchart of proposed framework is schematized in Figure 2. As shown in Figure 2, NRDNN first utilizes deep neural networks to learn high-level feature representations of multiple brain regions, respectively. Then, a set of weights indicating the contributions of each brain region can be automatically learned using a region-attention layer. Subsequently, the weighted feature representations of multiple brain regions are stacked into a feature matrix, and the nuclear norm regularization is adopted to capture the structural information within the feature matrix.

### Brain Region EEG Feature Learning

Given a trial of raw affective EEG signal, we aim to identify the emotion states by decoding this signal. Suppose the EEG signal has  $m$  channels, each of which has a  $t$  time point. Thus, the EEG signal can be represented as a two-dimensional matrix

$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m] \in \mathbb{R}^{m \times t}$ . Here,  $\mathbf{x}_i = [x_1^i, x_2^i, \dots, x_t^i] \in \mathbb{R}^t$  denotes the  $i$ -th channel of the EEG signal and  $x_j^i, j = 1, 2, \dots, t$  denotes the value of the  $i$ -th channel at the time point  $t$ .

According to the principle of brain regions (Vecchio et al., 2013), we divide the entire EEG signal  $\mathbf{X}$  into  $J$  parts located in different brain regions. Without loss of generality, EEG signal located in the  $j$ -th brain region can be represented as follows:

$$\hat{\mathbf{X}}_j = [\mathbf{x}_1^j, \mathbf{x}_2^j, \dots, \mathbf{x}_{m_j}^j] \in \mathbb{R}^{m_j \times d}, j = 1, 2, \dots, J. \quad (1)$$

Here, we use  $\mathbf{x}_k^j$ , which represents the EEG signal of the  $k$ -th channel located in the  $j$ -th brain region.  $m_j$  denotes the number of channels, which is located in the  $j$ -th brain region. Besides, we have  $m_1 + m_2 + \dots + m_J = m$ .

In general, a deep classification model  $f$  can be decomposed as  $f = g \circ h$ , in which  $h_\theta: \mathbf{X} \rightarrow \mathbb{R}^d$  parameterized by the network weight  $\theta$  that maps the input EEG signal  $\mathbf{X}$  to the high-level feature representation space. Besides,  $g_{w,b}: \mathbf{Z} \rightarrow [0, 1]^K$  parameterized by the weight  $\mathbf{W}$  and bias  $b$  that maps the feature representation to the final output.

Currently, many deep neural networks can be used for EEG feature extraction, such as Shallow ConvNet (SConvNet) and Deep ConvNet (DConvNet) developed in Schirrneister et al. (2017), and EEGNet developed in Lawhern et al. (2018). However, these widely used neural networks focus on motor imagery EEG classification. Hence, we slightly modify SConvNet to form the backbone network for affective EEG decoding (called AConvNet for simplicity), and the detailed network architecture of AConvNet is given in Table 1. To learn deep feature representation of the  $j$ -th brain region, feature extractor  $h_{\theta_j}: \hat{\mathbf{X}}_j \rightarrow \mathbf{z}_j$  parameterized by the network weight  $\theta_j$  that maps the EEG signal located in the  $j$ -th brain region to feature  $\mathbf{z}_j$ . Here,  $\mathbf{z}_j, j = 1, 2, \dots, J$  denotes the deep EEG feature of the  $j$ -th brain region. In addition, we also apply AConvNet to learn the high-level EEG feature representation of the global brain region, which can be represented as  $h_\theta: \mathbf{X} \rightarrow \mathbf{z}_{J+1}$ .

## Discriminative Feature Identification

As pointed earlier, human emotions are correlated to multiple cerebral cortex regions, such as orbitofrontal cortex and ventromedial prefrontal cortex (Lotfi et al., 2014). Hence, the EEG signals acquired from different brain regions would contribute differently to emotion recognition (Lindquist et al., 2012). To identify the contribution of different brain regions, we first reshape all the local and global EEG feature representations into a feature map, which can be represented as follows:

$$\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_J, \mathbf{z}_{J+1}] \in \mathbb{R}^{1 \times d \times (J+1)}. \quad (2)$$

Thus, identification of the contribution of different brain regions equals to assign a set of appropriate weights to  $J+1$  channels. To achieve this goal, we use squeeze-and-excitation (SE) block (Hu et al., 2018) to adaptively emphasize informative channels and suppress the less useful ones, as shown in Figure 3.

To abstract the information of different brain regions, the global average pooling is used to produce channel-wise statistics  $\mathbf{s} \in \mathbb{R}^{J+1}$ , which can be obtained by

$$s_j = \frac{1}{d} \sum_{i=1}^d z_j^i, j = 1, 2, \dots, J+1, \quad (3)$$

where  $z_j^i$  and  $s_j$  denote the  $i$ -th and the  $j$ -th element of  $\mathbf{z}_j$  and  $\mathbf{s}$ , respectively. To capture the channel-wise dependencies, the following gating operator and an activation is utilized:

$$\mathbf{u} = \sigma(\mathbf{W}_2 \cdot \delta(\mathbf{W}_1 \cdot \mathbf{s})), \quad (4)$$

where  $\delta(\cdot)$  and  $\sigma(\cdot)$  denote the activation functions ReLU (Nair and Hinton, 2010) and sigmoid, respectively. Besides,  $\mathbf{W}_1 \in \mathbb{R}^{\frac{L+1}{r} \times (J+1)}$  and  $\mathbf{W}_2 \in \mathbb{R}^{(J+1) \times \frac{L+1}{r}}$ . Here,  $\mathbf{u} \in \mathbb{R}^{J+1}$  represents the weights of multiple channels, which can reflect the contributions of brain regions.

Finally, the weighted feature representation of brain region can be represented as the channel-wise multiplication between the scale  $u_j$  and the feature map  $\mathbf{z}_j$ :

$$\hat{\mathbf{z}}_j = u_j \cdot \mathbf{z}_j, j = 1, 2, \dots, J, J+1, \quad (5)$$

where  $u_j$  denotes the  $j$ -th element of  $\mathbf{u}$ .

## Learning Structural Information

To capture the structural information among multiple brain regions, we first stack the weighted feature representation of multiple brain regions into a feature matrix:

$$\hat{\mathbf{Z}} = [\hat{\mathbf{z}}_1, \hat{\mathbf{z}}_2, \dots, \hat{\mathbf{z}}_J, \hat{\mathbf{z}}_{J+1}] \in \mathbb{R}^{(J+1) \times d}. \quad (6)$$

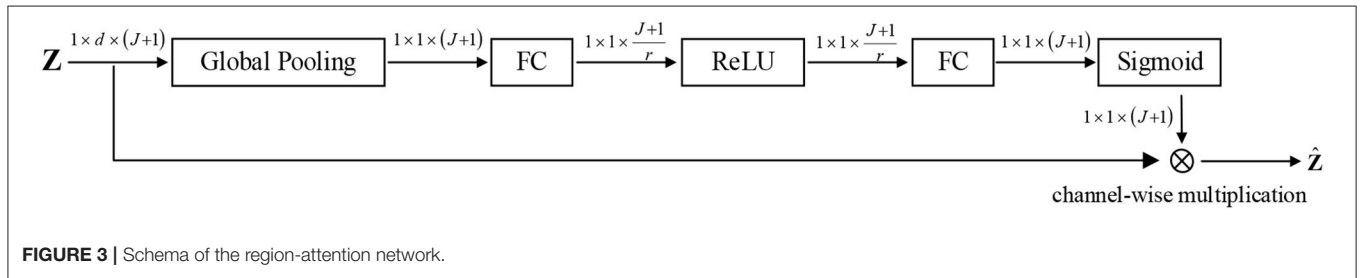
Then, we focus on construct a matrix classifier, i.e.,  $g_{w,b}: \hat{\mathbf{Z}} \rightarrow y$ , which can exploit the structural information to help the emotion recognition. Hence, the classifier  $g$  can be formulated as follows:

$$\arg \min_{\mathbf{W}, b} \frac{1}{2} R(\mathbf{W}) + G(\mathbf{W}) + C \sum_{i=1}^N H(\hat{\mathbf{z}}_i, \mathbf{W}, b), \quad (7)$$

where  $\mathbf{W} \in \mathbb{R}^{(J+1) \times d}$  and  $b$  represent the regression matrix and bias, respectively.  $C > 0$  is the trade-off parameter.  $R(\mathbf{W}) = \text{tr}(\mathbf{W}^T \mathbf{W}) = \|\mathbf{W}\|_F^2$  is the squared Frobenius norm of regression matrix  $\mathbf{W}$ , which can be used to control the complexity of model and avoid the overfitting problem.  $G(\mathbf{W}) = \tau \|\mathbf{W}\|_*$  denotes the nuclear norm of  $\mathbf{W}$ , where  $\tau > 0$  is the penalty parameter. As the convex approximation of the rank of regression matrix  $\mathbf{W}$ , nuclear norm can grasp the structural information within the feature matrix  $\hat{\mathbf{Z}}_i, i = 1, 2, \dots, N$ . Besides, we adopt the widely used hinge loss as the loss function because of its ability in sparseness and robustness modeling.

**TABLE 1** | Network architecture of AConvNet.

Modules	Layers	Operation	Parameters	Size
Input				$m \times d$
Feature extractor	Reshape			$1 \times m \times d$
	Convolution	Conv2D	$1 \times 25, 40$	$40 \times m \times (d-24)$
	Convolution	Conv2D	$m \times 1, 40$	$40 \times 1 \times (d-24)$
	Normalization	BatchNorm	/	$40 \times 1 \times (d-24)$
	Activation	Square	/	$40 \times 1 \times (d-24)$
	Pooling	AveragePool	$1 \times 75, 15$	$40 \times 1 \times [(d-99)//r]$
	Activation	log	/	$40 \times 1 \times [(d-99)//r]$
	Flatten		/	$40 [(d-99)15 + 1]$
Classification	Fully connected	Dense	$40 [(d-99)15 + 1] \times 300$	300
	Fully connected	Dense	$300 \times K$	$K$



**FIGURE 3** | Schema of the region-attention network.

**TABLE 2** | The EEG electrodes associated with each brain region.

Brain region	Electrode name
Frontal	Fp1,Fp2,AF3,AF4, F7,F3,Fz,F4,F8
Temporal	T7,T8
Central	FC5,FC1,FC2,FC6,C3,Cz,C4
Parietal	CP1,CP2,CP5,CP6,P7,P3,Pz,P4,P8,PO3,PO4
Occipital	O1,Oz,O2

$$\nabla_{\mathbf{W}} = \begin{cases} \frac{1}{2} \cdot \frac{\partial R}{\partial \mathbf{W}} + \frac{\partial G}{\partial \mathbf{W}}, H(\cdot) \leq 0 \\ \frac{1}{2} \cdot \frac{\partial R}{\partial \mathbf{W}} + \frac{\partial G}{\partial \mathbf{W}} + C \sum_{i=1}^N \frac{\partial H}{\partial \mathbf{W}}, H(\cdot) \leq 0 \end{cases} \quad (10)$$

$$\nabla_b = \begin{cases} 0, H(\cdot) \leq 0 \\ C \sum_{i=1}^N \frac{\partial H}{\partial b}, H(\cdot) \leq 0 \end{cases} \quad (11)$$

where the gradient of nuclear norm  $\partial G/\partial \mathbf{W}$  could be calculated according to Papadopoulos and Lourakis (2000).

## EXPERIMENTS

In this section, the proposed NRDNN is evaluated on the publicly available affective EEG datasets [i.e., DEAP dataset (Koelstra et al., 2011)]. The affective EEG datasets are first preprocessed. Then, we introduce the comparison methods and their parameter settings. The experimental results are subsequently presented and discussed in detail. Finally, we conclude this study.

### Affective EEG Data Preparation

The DEAP dataset contains multiple channel physiological signals for analyzing human emotional states. It is composed of 32-channel EEG signals recorded from 32 subjects. The sampling rate is set to 512 Hz. All subjects are required to watch 40 one-min long music video so that their various emotions are stimulated accordingly. Therefore, there are 40 trials per subject, each of which corresponds to affective EEG

$$H(\hat{\mathbf{Z}}_i, \mathbf{W}, b) = \max \left( 0, 1 - y_n \cdot \left\{ \text{tr}(\mathbf{W}^T \hat{\mathbf{Z}}_i) + b \right\} \right). \quad (8)$$

Finally, the prediction of test emotion data  $\tilde{\mathbf{Z}}$  using classifier  $g$  can be represented as follows:

$$g(\tilde{\mathbf{Z}}) = \text{tr}(\mathbf{W}^T \tilde{\mathbf{Z}}) + b. \quad (9)$$

### Optimization

To optimize the parameter of deep classification model  $f$ , we use the stochastic gradient descent (SGD) method to optimize the objective function in Equation (7), so that the end-to-end training of both feature extractor  $h$  and classifier  $g$  can be carried out *via* standard backpropagation. The partial derivatives of the objective function with respect to the regression matrix  $\mathbf{W}$  and bias  $b$  can be computed efficiently as follows:

**TABLE 3** | Classification performances (ACC) of our NRDNN against the comparison methods.

Subjects	Comparison methods							
	SVM	SMM	EEGNet	AConvNet	FConvNet	DLSVM	DNN	NRDNN
S01	0.4750	0.5250	0.6250	0.6250	0.7250	0.7280	<b>0.7500</b>	<b>0.7500</b>
S02	0.5500	0.5750	0.6000	0.6250	0.6500	0.7000	0.7000	<b>0.7250</b>
S03	0.5750	0.6000	0.5500	0.6000	0.6250	0.6375	<b>0.6500</b>	<b>0.6500</b>
S04	0.5500	0.4750	0.6750	0.6750	0.7000	0.7050	0.7250	<b>0.7500</b>
S05	0.5500	0.7000	0.6500	0.7000	0.7250	0.7250	<b>0.7750</b>	<b>0.7750</b>
S06	0.5500	0.7500	0.7000	0.7250	0.7000	0.7500	<b>0.7750</b>	<b>0.7750</b>
S07	0.5750	0.7000	0.6250	0.7000	0.7250	0.7250	<b>0.7500</b>	<b>0.7500</b>
S08	0.5250	0.6250	0.6000	0.6250	0.6250	0.6000	0.6500	<b>0.7000</b>
S09	<b>0.7000</b>	0.5250	0.6000	0.5750	0.6500	0.6750	0.6500	<b>0.7000</b>
S10	0.5250	0.6250	0.6250	0.6250	0.6500	0.6750	0.7000	<b>0.7250</b>
S11	0.4000	0.6500	0.6500	0.6750	0.6500	0.6750	0.7500	<b>0.7750</b>
S12	0.4750	0.5500	0.6250	0.7000	0.7000	0.7000	0.7000	<b>0.7250</b>
S13	0.5500	0.5500	0.6500	0.6500	0.7000	0.7000	0.7250	<b>0.7500</b>
S14	0.5750	0.6000	0.5750	0.5750	0.6000	0.6250	0.6500	<b>0.7000</b>
S15	0.5000	0.5250	0.6250	0.6500	0.6750	0.6750	<b>0.7500</b>	<b>0.7500</b>
S16	0.5250	0.4000	0.6000	0.6000	0.6250	0.6250	<b>0.6750</b>	<b>0.6750</b>
Avg.	0.5375	0.5859	0.6234	0.6453	0.6703	0.6825	0.7109	<b>0.7297</b>

The best classification results are boldfaced.

**TABLE 4** | Classification performances (F1) of our NRDNN against the comparison methods.

Subjects	Comparison methods							
	SVM	SMM	EEGNet	AConvNet	FConvNet	DLSVM	DNN	NRDNN
S01	0.4747	0.4861	0.5807	0.6248	0.7206	0.7255	0.7469	<b>0.7500</b>
S02	0.4000	0.4133	0.5657	0.6389	0.6616	0.7222	0.7020	<b>0.7234</b>
S03	0.5726	0.4984	0.5343	0.5990	0.6248	0.6375	<b>0.6491</b>	<b>0.6491</b>
S04	0.5396	0.4130	0.6484	0.6698	0.6703	0.6918	0.7163	<b>0.7475</b>
S05	0.5500	0.6581	0.6465	0.6931	0.7163	0.7206	0.7382	<b>0.7566</b>
S06	0.4872	0.5098	0.6000	0.6204	0.6429	0.6667	0.6257	<b>0.6894</b>
S07	0.5402	0.4805	0.6190	0.6238	<b>0.7025</b>	0.6925	0.6865	0.7024
S08	0.5223	0.5636	0.6000	0.6132	0.6190	0.5733	0.6267	<b>0.6931</b>
S09	0.6970	0.4473	0.5908	0.5248	0.6354	0.6698	0.6419	<b>0.7000</b>
S10	0.5247	0.5807	0.6229	0.6229	0.6491	0.6647	0.6970	<b>0.7234</b>
S11	0.3600	0.5333	0.6491	0.6577	0.6419	0.6698	0.7396	<b>0.7630</b>
S12	0.4667	0.4357	0.6132	0.6992	0.7000	0.6992	0.7000	<b>0.7234</b>
S13	0.4643	0.5396	0.6011	0.6491	0.6703	0.6970	0.7163	<b>0.7494</b>
S14	0.5248	0.5442	0.5616	0.5726	0.5908	0.5943	0.6465	<b>0.6800</b>
S15	0.4885	0.3866	0.6229	0.6491	0.6577	0.6647	<b>0.7500</b>	0.7494
S16	0.5175	0.3750	0.5833	0.5833	0.3846	0.4398	<b>0.6647</b>	<b>0.6647</b>
Avg.	0.5081	0.4916	0.6025	0.6276	0.6430	0.6581	0.6905	<b>0.7166</b>

The best classification results are boldfaced.

data stimulated by one music video. After each trial, all subjects are required to perform self-assessments on five dimensions, i.e., valence (from sad to joyful), arousal (from calm to excited), dominance (from submissive to dominant), liking (related to the preference of participants), and familiarity (related to the prior experience of participants). In addition to the rating range of familiarity, which is distributed from 1 (weakest) to

5 (strongest), the remaining dimensions range from 1 to 9. Referring to Yang et al. (2018), we adopt 5 as the threshold of the valence dimension to divide EEG trials into two categories, i.e., if the valence rating is greater (smaller) than 5, it is positive (negative). In this study, we downsampled 32-channel affective EEG signals to 128 Hz. We then bandpass filtered EEG signals between 4 and 45-Hz frequency band. Without loss of

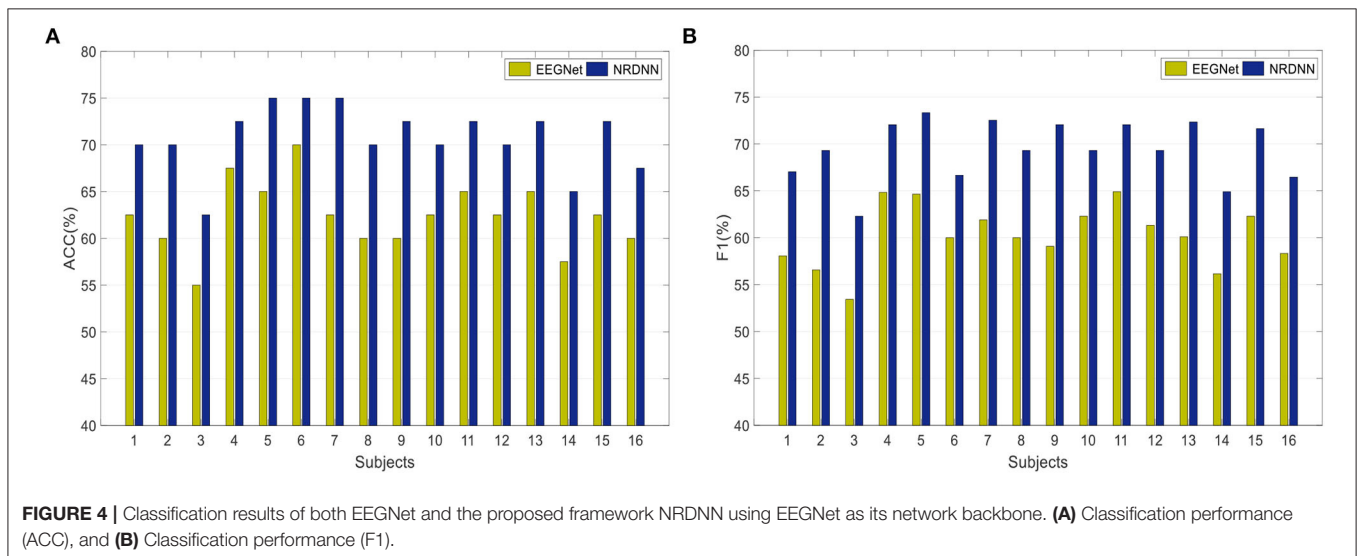
**TABLE 5** | Classification performances (AUC) of our NRDNN against the comparison methods.

Subjects	Comparison methods							
	SVM	SMM	EEGNet	AConvNet	FConvNet	DLSVM	DNN	NRDNN
S01	0.3885	0.3985	0.4862	0.5238	0.6241	0.6479	0.6717	<b>0.6892</b>
S02	0.2146	0.2652	0.3409	0.5126	0.5000	0.5379	0.5732	<b>0.6010</b>
S03	0.4520	0.3409	0.3972	0.4596	0.4975	0.4950	0.5177	<b>0.5202</b>
S04	0.4167	0.2839	0.5208	0.5651	0.5339	0.5776	0.6042	<b>0.6198</b>
S05	0.4714	0.5417	0.5859	0.6120	0.6380	0.6484	0.6250	<b>0.6667</b>
S06	0.3400	0.2667	0.4233	0.4533	<b>0.5533</b>	0.5300	0.5400	0.5467
S07	0.4524	0.3512	0.6012	0.4970	<b>0.6399</b>	<b>0.6399</b>	0.6250	0.6190
S08	0.4116	0.3965	0.4722	0.4722	0.4874	0.4091	0.4899	<b>0.5732</b>
S09	0.5600	0.3425	0.4725	0.4275	0.4950	0.5025	0.5500	<b>0.5925</b>
S10	0.3500	0.3775	0.4650	0.4675	0.5000	0.5225	0.5800	<b>0.6125</b>
S11	0.1563	0.2630	0.4661	0.4479	0.4583	0.4974	0.5677	<b>0.6042</b>
S12	0.3058	0.2281	0.4236	0.5639	0.5564	0.5539	0.5664	<b>0.5990</b>
S13	0.2864	0.4271	0.4348	0.5703	0.5217	0.5703	0.5985	<b>0.6496</b>
S14	0.3500	0.3600	0.3900	0.4225	0.4375	0.4175	0.5300	<b>0.5375</b>
S15	0.2725	0.1350	0.4200	0.4600	0.4550	0.4700	<b>0.6000</b>	0.5800
S16	0.4747	0.4027	0.5120	0.5387	0.3947	0.4400	<b>0.6613</b>	0.6480
Avg.	0.3689	0.3363	0.4632	0.4996	0.5183	0.5287	0.5813	<b>0.6037</b>

The best classification results are boldfaced.

**TABLE 6** | Statistical significance comparisons of ACC and F1 of NRDNN and other comparison methods.

Metrics	NRDNN vs. SVM	NRDNN vs. SMM	NRDNN vs. EEGNet	NRDNN vs. AConvNet	NRDNN vs. FConvNet	NRDNN vs. DLSVM	NRDNN vs. DNN
ACC	1.71E-07	3.57E-06	8.18E-13	1.09E-08	5.63E-07	5.75E-06	1.50E-04
F1	2.84E-07	1.06E-08	1.65E-11	2.305E-08	3.11 E-05	8.00E-05	2.20E-05
AUC	4.59E-07	1.93E-09	5.75E-08	3.97E-08	1.16E-05	1.31E-05	4.14E-04



generality, we only take the first half subjects to evaluate the effectiveness of the proposed NRDNN, in order to reduce the training time.

The EEG signals in DEAP database were recorded with 32 electrodes following the international 10/20 system. According to the spatial locations of electrodes, we grouped the 32 electrodes

into 5 brain regions. **Table 2** summarizes the EEG electrodes located in each brain region in detail (Li Y. et al., 2019).

## Experimental Setup

### Baseline Methods

In the experiments, the proposed method was compared with the following comparison methods by using the aforementioned affective EEG classification tasks: (1) Support vector machine (SVM) (Zhang et al., 2018), (2) Support matrix machine (SMM) (Luo et al., 2015), (3) EEGNet (Lawhern et al., 2018), (4) Shallow ConvNet for affective EEG (AConvNet) (Schirrmeyer et al., 2017), (5) Fusion ConvNet (FConvNet) (Liang et al., 2020), (6) Deep learning with SVM (DLSVM) (Tang, 2013), (7) NRDNN without region-attention layer (DNN), and (8) Our NRDNN.

### Implementation Details

As the format of the input data of SMM should be matrices, the principal component analysis (PCA) (Placidi et al., 2016) was adopted to reduce the dimension of EEG data to  $32 \times 16$  matrix features. Then, the obtained two-dimensional EEG features were reshaped into vectors, which were used as the input for SVM. Besides, AConvNet was used as the network backbone for FConvNet. Referring to Liang et al. (2020), we took the high-level feature representations of multiple brain regions as multiple views, which were then classified using cross-entropy loss. For DLSVM, we also used AConvNet as its network backbone. The obtained high-level feature representations of multiple brain regions were concatenated and then classified using SVM in an end-to-end manner. The trade-off parameter  $C$  of SVM, SMM, DLSVM, and our NRDNN was decided through searching from the set  $\{1e-2, 1e-1, 1e0, 1e1, 1e2\}$ . The parameter  $\tau$  of SMM and our NRDNN was decided through searching from the set  $\{1e-3, 2e-3, 5e-3, 1e-2, 2e-2, 5e-2, 1e-1, 2e-1, 5e-1, 1e0\}$ . For all comparison methods, the optimal parameters  $C$  and  $\tau$  were chosen by using the 5-fold cross-validation method on the training dataset. For deep learning methods EEGNet, AConvNet, FConvNet, DLSVM, and NRDNN, the batch size and epoch were set to 40 and 1,000, respectively. The learning rate was dynamically changed during optimization using the formula as follows (Pei et al., 2018):  $\eta_p = \eta_0 / (1 + \alpha p)^\beta$ , in which  $p$  linearly changes from 0 to 1,  $\eta_0 = 1e-3$ ,  $\alpha=10$ , and  $\beta = 0.75$ . Besides, the parameter  $r$  is set to 2 in the region-attention network.

Following the evaluation protocol developed by Lan et al. (2018), we used the leave-one-subject-out cross-validation method to evaluate the affective EEG classification performance on each subject. The following metrics (Chen et al., 2020) on the test dataset were adopted, i.e., Accuracy (ACC), F1score (F1), and the area under the receiver operating characteristics curve (AUC). Herein,  $ACC = (TP+TN)/(TP+FN+FP+TN)$  and  $F1=2 \times PPV \times SEN/(PPV+SEN)$ , in which the positive predictive value ( $PPV = TP/(TP+FP)$ ) and sensitivity ( $SEN = TP/(TP+FN)$ ). Generally, the higher are the metric values, the better affective is the EEG classification performance.

## Experimental Results Analysis

The classification accuracy (ACC), F1 score (F1), and AUC of all comparison methods on 16 subjects are presented in **Tables 3–5**. The best classification results are boldfaced. From these classification results, we can obtain the following observations. In terms of ACC, matrix learning method SMM can obtain better classification performance than vector-based classifier SVM. This is because SMM can exploit the correlation within EEG feature matrices to improve the classification performance. In addition, deep learning methods, such as EEGNet, AConvNet, FConvNet, DLSVM, and NRCNN, can yield better classification results than shallow methods, such as SVM and SMM, in almost all cases. Compared with shallow methods, deep neural networks can automatically learn high-level feature representations from the raw data, resulting in better EEG decoding performance. It is notable that our NRDNN can obtain the best classification performance than other deep learning methods. The promising results are mainly attributed to the fact that NRDNN can not only learn high-level feature representations of EEG signals located in multiple brain regions but also capture the structural information among different brain regions. The experimental results verify the fact that the structural information among different brain regions is conducive to boost the decoding performance of affective EEG signals.

As summarized in **Tables 3–5**, we can observe that the proposed method obtains the highest average classification results. Specifically, the proposed NRDNN achieves promising average results of 72.97, 71.66, and 60.37% in terms of ACC, F1, and AUC. Compared with the baseline SVM, the absolute average of ACC, F1, and AUC increases by 19.22, 20.85, and 23.48%, respectively. NRDNN outperforms the results of SMM by an average of 14.38, 22.5, and 26.74%, which validates the high-level feature learning capability of our NRDNN. Compared with DLSVM that does not leverage the structural information among multiple brain regions, the average classification results of NRDNN are increased by 4.72, 5.85, and 7.5% in terms of ACC, F1, and AUC, respectively. Besides, NRDNN outperforms EEGNet by 10.63, 11.41, and 14.5%, and yields 8.44, 8.9, and 10.41% higher average classification results than AConvNet. NRDNN is superior to FConvNet by 5.94, 7.36, and 8.54% in terms of ACC, F1, and AUC, respectively. Furthermore, NRDNN outperforms the results of DNN by an average of 1.88, 2.61, and 2.24%, which validates the contributions of different brain regions that can be automatically identified by our NRDNN. These experimental results demonstrate the effectiveness of the proposed NRDNN.

## DISCUSSION

To evaluate the statistical significance of the experimental results, we further perform pairwise two-tailed  $t$ -test (Zheng et al., 2018) to verify whether there exist significant differences with a confidence level of 95% between the proposed NRDNN and the comparison methods. The statistical significance comparisons of ACC and F1 of NRDNN and other comparison methods are given in **Table 6**. The  $p$ -value less than 0.05 expresses that



significant differences exist between the proposed NRDNN and the comparison methods. We highlighted the  $p$ -values that are less than 0.05 in boldface. As summarized in **Table 6**, we can see that the null hypothesis can be rejected with 95% confidence level in each case. The statistical results verify that the proposed NRDNN significantly outperformed the comparison methods. This further indicated the capability of the NRDNN to exploit the structural information among multiple brain regions, as well as the powerful high-level affective EEG feature learning capability. The above experimental results illustrate that the proposed NRDNN is suitable for the classification of affective EEG data.

To obtain a better insight into the classification result of our nuclear norm regularized deep neural network framework, we further investigated the effects of different network backbones on the classification performance. **Figure 4A** presents the ACCs of both EEGNet and the proposed framework NRDNN using EEGNet as its network backbone. **Figure 4B** gives the F1s. It can be found that NRDNN yields better results than the baseline EEGNet in all cases. In terms of ACC, NRDNN is superior to EEGNet by 7.5, 10, 7.5, 5, 10, 5, 12.5, 10, 12.5, 7.5, 7.5, 7.5, 7.5, 10, and 7.5% on 16 subjects, respectively. Compared to EEGNet that does not leverage the deep features of multiple brain regions and their structural information, the classification F1s of NRDNN are increased by 8.96, 12.74, 8.86, 7.22, 8.68, 6.67, 10.63, 9.31, 12.98, 7.02, 7.15, 7.99, 12.23, 8.75, 9.34, and 8.14%, respectively. The average ACC and F1 of NRDNN are 70.78 and 69.41%. The absolute values are increased by 8.44 and 9.16% compared with EEGNet.

Overall, the proposed NRDNN improves the affective EEG classification performance using different network backbones. The abovementioned results validate that NRDNN can effectively learn deep features of multiple brain regions and their corresponding structural information using the nuclear norm regularization. To summarize, NRDNN integrates the powerful deep feature learning capability and the structural information learning ability of matrix classifier. The experimental results demonstrate that the proposed NRDNN framework could achieve better classification performance than the comparison methods.

## CONCLUSION

In this study, we first presented a deep neural network, named AConvNet, for affective EEG decoding. Based on AConvNet, we further proposed a novel nuclear norm regularized deep neural network framework called NRDNN. The proposed NRDNN

can effectively learn high-level feature representations of EEG signals located in multiple brain regions using AConvNet, as well as discriminate the contributions of multiple brain regions using a set of automatically learned weights. Besides, NRDNN can exploit the structural information among multiple brain regions using the introduced nuclear norm regularization. The proposed NRDNN can be carried out in an efficient end-to-end fashion. Extensive experimental results on publicly available emotion dataset demonstrate the superiority of our NRDNN.

Despite the promising classification performance of NRDNN, there is still room for further improvement. For example, the development of more advance attention mechanism is conducive to the identification of the contribution of different brain regions. Besides, extending the proposed method to multi-class classification is another interesting direction. Furthermore, more powerful discriminative high-level features with both spatial and temporal information of EEG signals can further improve the performance of EEG-based emotion recognition. We will address these issues in the future studies.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: <http://www.eecs.qmul.ac.uk/mmv/datasets/deap/download.html>.

## AUTHOR CONTRIBUTIONS

SL is responsible for study design and manuscript writing. MY and YH are responsible for data processing and data analysis. XD is responsible for manuscript editing. QW is responsible for experimental design. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported in part by the National Natural Science Foundation of China under Grants (61902197). Research on development and application of intelligent vehicle testing simulator based on augmented reality was funded by the Guangdong Provincial Basic and Applied Basic Research Fund-Regional Joint Fund (2020B1515130004), the Open Project of State Key Laboratory for Novel Software Technology at Nanjing University (KFKT2020B11), and the NUPTSF (NY219034).

## REFERENCES

- Agrafioti, F., Hatzinakos, D., and Anderson, A. K. (2012). ECG pattern analysis for emotion detection. *IEEE Trans. Affect. Comput.* 3, 102–115. doi: 10.1109/T-AFFC.2011.28
- Alarcao, S., and Fonseca, M. (2017). Emotions recognition using EEG signals: a survey. *IEEE Transac. Affect Comput.* 10, 374–393. doi: 10.1109/TAFFC.2017.2714671

- Ang, J., Dhillon, R., Krupski, A., Shriberg, E., and Stolcke, A. (2002). "Prosody-based automatic detection of annoyance and frustration in humancomputer dialog," in *INTERSPEECH* (Denver, CO), 2037–2040.
- Atyabi, A., Shic, F., and Naples, A. (2016). Mixture of autoregressive modeling orders and its implication on single trial EEG classification. *Expert Syst. With Applic.* 65, 164–180. doi: 10.1016/j.eswa.2016.08.044

- Bustamante, P., Lopez Celani, N., Perez, M., and Quintero M. O. (2015). "Recognition and regionalization of emotions in the arousal-valence plane," in *Proceedings of the 2015 Milano, Italy 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* p. 6042–6045. doi: 10.1109/EMBC.2015.7319769
- Chen, Y., Hang, W., Liang, S., Liu, X., Li, G., and Wang, Q. (2020). A novel transfer support matrix machine for motor imagery-based brain computer interface. *Front. Neurosci.* 14, 606949. doi: 10.3389/fnins.2020.606949
- Cheng, B., and Liu, G. (2008). "Emotion recognition from surface EMG signal using wavelet transform and neural network," in *Proceedings of the 2nd International Conference on Bioinformatics and Biomedical Engineering*, (Shanghai), p. 1363–1366. doi: 10.1109/ICBBE.2008.670
- Clark, J. M. (1994). Modularity, abstractness and the interactive brain. *Behav. Brain Sci.* 17, 67–68. doi: 10.1017/S0140525X00033380
- Duan, R., Zhu, J., and Lu, B. (2013). "Differential entropy feature for EEG-based emotion classification," in *6th International IEEE/EMBS Conference on Neural Engineering (NER)*. p. 81–84. doi: 10.1109/NER.2013.6695876
- Fragopanagos, N., and Taylor, J. G. (2005). Emotion recognition in human-computer interaction. *Neural Netw.* 18, 389–405. doi: 10.1016/j.neunet.2005.03.006
- Hang, W., Feng, W., Liang, S., Wang, Q., Liu, X., and Choi, K. S. (2020). Deep stacked support matrix machine based representation learning for motor imagery EEG classification. *Comput. Methods Progr. Biomed.* 193, 105466. doi: 10.1016/j.cmpb.2020.105466
- Hu, J., Shen, L., and Sun, G. (2018). "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (Salt Lake City, UT: IEEE), 7132–7141. doi: 10.1109/CVPR.2018.00745
- Iacoviello, D., Petracca, A., Spezialetti, M., and Placidi, G. (2015). A real-time classification algorithm for EEG-based BCI driven by self-induced emotions. *Comput. Methods Programs Biomed.* 122, 293–303. doi: 10.1016/j.cmpb.2015.08.011
- Koelstra, S., Muhl, C., and Soleymani, M. (2011). Deap: A database for emotion analysis using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15
- Kurmukov, A., Dodonova, Y., Burova, M., Mussabayeva, A., Petrov, D., and Faskowitz, J. (2016). "Topological modules of human brain networks are anatomically embedded: evidence from modularity analysis at multiple scales," in *International Conference on Network Analysis*, Cham: Springer. p. 299–308. doi: 10.1007/978-3-319-96247-4\_22
- Lan, Z., Sourina, O., Wang, L., Scherer, R., and Müller-Putz, G. R. (2018). Domain adaptation techniques for EEG-based emotion recognition: a comparative study on two public datasets. *IEEE Trans. Cogn. Develop. Syst.* 11, 85–94. doi: 10.1109/TCD.2018.2826840
- Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., and Lance, B. J. (2018). EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces. *J. Neural Eng.* 15, 056013. doi: 10.1088/1741-2552/aace8c
- Li, J., Qiu, S., Shen, Y., Liu, C. L., and He, H. (2019). Multisource transfer learning for cross-subject EEG emotion recognition. *IEEE Trans. Cybernet.* 50, 3281–3293. doi: 10.1109/TCYB.2019.2904052
- Li, Y., Zheng, W., Wang, L., Zong, Y., and Cui, Z. (2019). From regional to global brain: A novel hierarchical spatial-temporal neural network model for EEG emotion recognition. *IEEE Trans. Affect. Comput.* 13, 568–578. doi: 10.1109/TAFFC.2019.2922912
- Liang, E., Elazab, A., Liang, S., Wang, Q., and Lei, B. (2020). "Walking imagery evaluation based on multi-view features and stacked denoising auto-encoder network," in *2020 IEEE 17th International Symposium on Biomedical Imaging*. p. 1896–1899. doi: 10.1109/ISBI45749.2020.9098522
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., and Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* 35, 121–143. doi: 10.1017/S0140525X11000446
- Liu, Y. J., Yu, M., Zhao, G., Song, J., Ge, Y., and Shi, Y. (2017). Real-time movie-induced discrete emotion recognition from EEG signals. *IEEE Trans. Affect. Comput.* 9, 550–562. doi: 10.1109/TAFFC.2017.2660485
- Lotfi, E., and Akbarzadeh-T., M. R. (2014). Practical emotional neural networks. *Neural Netw.* 59, 61–72. doi: 10.1016/j.neunet.2014.06.012
- Luo, L., Xie, Y., Zhang, Z., and Li, W. (2015). "Support matrix machines," in *International Conference on Machine Learning*. (Lille), p. 938–947.
- Mazumder, I. (2019). "An analytical approach of EEG analysis for emotion recognition," in *Proceedings of the 2019 Devices for Integrated Circuit (DevIC)*. p. 256–260. doi: 10.1109/DEVIC.2019.8783331
- Mühl, C., Allison, B., Nijholt, A., and Chanel, G. (2014). A survey of affective brain computer interfaces: Principles, state-of-the-art, and challenges. *Brain Comput. Interfaces* 1, 66–84. doi: 10.1080/2326263X.2014.912881
- Murugappan, M., and Murugappan, S. (2013). "Human emotion recognition through short time Electroencephalogram (EEG) signals using Fast Fourier Transform (FFT)," in *Proceedings of the 2013 IEEE 9th International Colloquium on Signal Processing and its Applications*. (Kuala Lumpur: IEEE), p. 289–294. doi: 10.1109/CSPA.2013.6530058
- Nair, V., and Hinton, G. E. (2010). "Rectified linear units improve restricted boltzmann machines," in *International Conference on Machine Learning* (Haifa), 807–814. doi: 10.5555/3104322.3104425
- Pandey, P., and Seeja, K. (2019). "Emotional state recognition with eeg signals using subject independent approach," in *Data Science and Big Data Analytics*. Singapore: Springer. p. 117–124. doi: 10.1007/978-981-10-7641-1\_10
- Papadopoulos, T., and Lourakis, M. I. A. (2000). "Estimating the jacobian of the singular value decomposition: Theory and applications," in *European Conference on Computer Vision*. (Dublin), p. 554–570. doi: 10.1007/3-540-45054-8\_36
- Park, Y., and Chung, W. (2019). Frequency-optimized local region common spatial pattern approach for motor imagery classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* 27, 1378–1388. doi: 10.1109/TNSRE.2019.2922713
- Pei, Z., Cao, Z., Long, M., and Wang, J. (2018). "Multi-adversarial domain adaptation," in *Thirty-second AAAI Conference on Artificial Intelligence*.
- Petrantonakis, P. C., and Hadjileontiadis, L. J. (2010). Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. *IEEE Trans. Affect. Comput.* 1, 81–97. doi: 10.1109/T-AFFC.2010.7
- Placidi, G., Giamberardino, P. D., Petracca, A., Spezialetti, M., and Iacoviello, D. (2016). "Classification of Emotional Signals from the DEAP dataset," in *International Congress on Neurotechnology, Electronics and Informatics*. vol. 2, p. 15–21. doi: 10.5220/0006043400150021
- Ramoser, H., Müller-Gerking, J., and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial eeg during imagined hand movement. *IEEE Trans. Rehabil. Eng.* 8, 441–446. doi: 10.1109/86.895946
- Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., and Tangermann, M. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapp.* 38, 5391–5420. doi: 10.1002/hbm.23730
- Song, T., Zheng, W., Song, P., and Cui, Z. (2018). Eeg emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans. Affect. Comput.* 11, 532–41. doi: 10.1109/TAFFC.2018.2817622
- Tang, X., Wang, T., Du, Y., and Dai, Y. (2019). Motor imagery EEG recognition with KNN-based smooth auto-encoder. *Artif. Intell. Med.* 101, 101747. doi: 10.1016/j.artmed.2019.101747
- Tang, Y. (2013). Deep learning using support vector machines. *CoRR*, abs/0239.2.1. doi: 10.48550/arXiv.1306.0239
- Ting, W., Guo-Zheng, Y., Bang-Hua, Y., and Hong, S. (2008). EEG feature extraction based on wavelet packet decomposition for brain computer interface. *Meas. J. Int. Meas. Confed.* 41, 618–625. doi: 10.1016/j.measurement.2007.07.007
- Vecchio, F., Babiloni, C., Buffo, P., Rossini, P. M., and Bertini, M. (2013). Inter-hemispherical coupling of EEG rhythms during the perception of facial emotional expressions. *Clin. Neurophysiol.* 124, 263–272. doi: 10.1016/j.clinph.2012.03.083
- Xiaohua, H., Wang, S. J., Liu, X., Zhao, G., Feng, X., and Pietikainen, M. (2017). Discriminative spatiotemporal local binary pattern with revisited integral projection for spontaneous facial micro-expression recognition. *IEEE Trans. Affect. Comput.* 10, 32–47. doi: 10.1109/TAFFC.2017.2713359
- Yan, J., Zheng, W., Xin, M., and Yan, J. (2014). Integrating facial expression and body gesture in videos for emotion recognition. *IEICE Trans. Inf. Syst.* E97-D, 610–613. doi: 10.1587/transinf.E97.D.610
- Yang, Y., Wu, Q., Qiu, M., Wang, Y., and Chen, X. (2018). "Emotion recognition from multi-channel eeg through parallel convolutional recurrent neural network," in *International Joint Conference on Neural Networks*. (Rio de Janeiro), p. 1–7. doi: 10.1109/IJCNN.2018.8489331

- Zhang, Y., Nam, C. S., Zhou, G., Jin, J., Wang, X., and Cichocki, A. (2018). Temporally constrained sparse group spatial patterns for motor imagery BCI. *IEEE Transac. Cybernet.* 49, 3322–3332. doi: 10.1109/TCYB.2018.2841847
- Zhang, Y., Zhou, G., Zhao, Q., Jin, J., Wang, X., and Cichocki, A. (2013). Spatial-temporal discriminant analysis for ERP-based brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 21, 233–243. doi: 10.1109/TNSRE.2013.2243471
- Zheng, Q., Zhu, F., and Heng, P. A. (2018). Robust support matrix machine for single trial eeg classification. *IEEE Transac. Neural Syst. Rehabil. Eng.* 26, 551–562. doi: 10.1109/TNSRE.2018.2794534
- Zheng, W. (2017). Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis. *IEEE Transac. Cogn. Develop. Syst.* 9, 281–290. doi: 10.1109/TCDS.2016.2587290
- Zheng, W. L., and Lu, B. L. (2015). Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Transac. Auton. Mental Develop.* 7, 162–175. doi: 10.1109/TAMD.2015.2431497

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Liang, Yin, Huang, Dai and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*