IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 11 July 2022, accepted 20 July 2022, date of publication 25 July 2022, date of current version 29 July 2022.
Digital Object Identifier 10.1109/ACCESS.2022.3193768

RESEARCH ARTICLE

Deep Neural Network for Emotion Recognition Based on Meta-Transfer Learning

HENGYAO TANG, GUOSONG JIANG, AND QINGDONG WANG
Computer School, Huanggang Normal University, Huanggang, Hubei 438000, China
Corresponding author: Qingdong Wang (qingdongwang@hgnu.edu.cn)

This work was supported in part by the Natural Science Foundation of Hubei Province, China, under Grant 2019CFC868.

ABSTRACT In recent years, many EEG-based emotion recognition methods have been proposed, which can achieve good performance on single-subject data. However, when the models are applied to cross-subject scenarios, due to the existence of subject differences, these models are often difficult to accurately identify the emotions of new subjects, which is not conducive to the practical application of the models. Many transfer learning methods have been applied to cross-subject EEG emotion recognition tasks to reduce the effect of subject differences. Most of them need to be trained with source data of many subjects and calibrated with more data of target subjects to obtain better classification performance on target subjects. However, this process relies on a large amount of training data to guarantee the final effect. This paper proposed a meta-transfer learning model for emotion recognition. The model can reduce the amount of data required by the meta-learning optimization algorithm. Even if only a small amount of data is used for training, it can achieve good performance, thereby reducing the cost of EEG acquisition and labeling, and it is also conducive to the model for new subjects. Finally, this paper conducts cross-subject emotion recognition experiments based on two public datasets SEED and SEED-IV. The experimental results show that the performance of the proposed meta-transfer learning method is better than the baseline method, and can rapid adaptation to unknown subjects while reducing training data.

INDEX TERMS EEG signal, emotion recognition, transfer learning, meta-learning.

I. INTRODUCTION

As a basic physiological activity of our human body, emotion has a great influence on speech, behavior, and other aspects. In simple terms, positive emotions are often beneficial to physical and mental health and improve the efficiency of study and work. while negative emotions will bring the opposite effect, which is not good for the physical and mental health of the human body and is not conducive to study and work [1], [2].

Studying emotion recognition methods, on the one hand, for each individual can make people better understand their own and others' psychological state, better manage their emotions and improve their work and study efficiency. The study of emotion belongs to the field of cognitive science and social psychology, so the method of studying emotion recognition can enrich the theory of emotion psychology and has a wide range of applications in reality. Firstly, accurate emotion recognition helps individuals integrate into the social environment. By analyzing the emotions of others and ourselves, we can calmly cope with and adapt to the environment. Secondly, emotion recognition is very necessary for psychological counseling, clients are often under great pressure or psychological barriers. Only by accurately identifying the emotional state of the client can they maintain a good counseling relationship. Furthermore, emotion recognition can provide strong theoretical support for diseases such as emotional disorders. Finally, emotion plays a role in polygraph detection technology. It plays an important role in combining emotion recognition theory with polygraph technology, and the results have a high reference value. In general, emotion recognition has profound research significance and broad application prospects [3]–[6].

Emotion recognition is possible using images, text, sounds, etc. Although image, text, and sound data are easier to collect and can achieve good results, human facial expressions and

The associate editor coordinating the review of this manuscript and approving it for publication was Hong-Mei Zhang.

voices are easily camouflaged, which may not necessarily reflect real and objective emotions. Emotion recognition based on physiological signals can avoid such situations. Physiological signals are objective responses to human physiological situations and cannot be easily camouflaged. Although emotion recognition based on physiological signals has made great progress in recent years, emotions varied widely between subjects. To address subject variability, some researchers use transfer learning methods to improve performance on new subjects. But most of these methods still require a lot of data dependencies.

To solve the above problems, this paper proposed a deep learning model based on meta-transfer learning, which can automatically discriminate and classify the physiological signals generated by the human body in different types of emotional states. To improve the performance of the model in cross-subject scenarios, this paper used the meta-learning method Reptile [7] to improve the generalization ability of the model. Through meta-learning, on the one hand, the model can adapt to new subjects faster with a small amount of data, which is very valuable in practical application scenarios. On the other hand, it can improve the performance of the model and avoid the problem of the accuracy of the model being too low after transfer. Our contributions can be summarized as follows: the proposed meta-transfer learning method can use a small number of samples for training. In addition, the proposed method can reduce the subject difference problem. This paper conducts cross-subject emotion recognition experiments based on two public datasets SEED and SEED-IV. The experimental results show that the performance of the proposed meta-transfer learning method is better than the baseline method, and can rapid adaptation to unknown subjects while reducing training data.

II. RELATED WORK

EEG is an important time series data [8], [9], which can be used in the fields of emotion recognition, motor imagery [10], sleep staging [11]–[14] and so on. In recent years, EEG signals have been widely used in emotion recognition. In the early days, manually extracted features and traditional machine learning classification methods were used for emotion recognition. For example, Bahari *et al.* [15] employed a recursive graph-based nonlinear k-nearest neighbor classifier to identify different emotions. Wang *et al.* [16] adopted a support vector machine classifier based on frequency domain features to classify different types of emotions. However, traditional machine learning techniques are greatly limited in both feature design and feature selection and require a lot of expert knowledge.

To overcome the above limitations, deep learning methods [17] are proposed and applied, deep learning usually uses a computational model composed of multiple computational layers to learn deep representations of data. Driven by the achievements of deep learning in computer vision, speech recognition, and natural language processing, many researchers have also applied deep learning to EEG emotion recognition. Researchers are usually based on deep learning models that extract features from EEG signals for emotion recognition. Commonly used frequency domain features include differential entropy feature [18], power spectral density feature [19], DCAM feature [20], RASM feature [21], DCAU feature [22] and so on. Al-Nafjan *et al.* [23] adopted a deep neural network based on PSD features to identify emotions. Zheng *et al.* [24] found that DE features extracted from EEG signals are an accurate and stable classification feature, proposed a deep belief network, and used DE features to identify emotions. Yang *et al.* [25] proposed a hierarchical network that uses five frequency bands of DE features to identify different emotions. For the extraction of temporal features, some researchers use deep learning models to automatically obtain dynamic information based on raw EEG signals. For example, Fourati *et al.* [26] proposed an echo-state network that uses recurrent layers to project raw EEG signals into a high-dimensional state space. Alhagry *et al.* [27] used EEG raw signals as input and achieved good emotion recognition results through a two-layer long short-term memory network. Ma *et al.* [28] proposed a multimodal residual model whose temporal weights are shared among the multimodalities. In terms of spatial feature extraction, Li *et al.* [29] proposed a hierarchical convolutional neural network based on EEG topography to obtain spatial information between different channels. To capture the interactions between different brain regions, Jung *et al.* [30] converted EEG signals into image-based representations and obtained satisfactory results. Song *et al.* [31] proposed a novel bihemispheric dissimilarity model to learn asymmetric disparities between the two hemispheres for EEG emotion recognition. Zhang *et al.* [32] proposed a graph convolutional wide network using graph structure to study deeper information.

In emotion recognition, the same model often has a large performance degradation when it is transferred between different subjects. To solve the problem of the performance degradation of EEG emotion recognition across subjects, many researchers have made attempts. Chai *et al.* [33] proposed an Adaptive Subspace Feature Matching (ASFM) method, which mainly uses Principal Component Analysis (PCA) for projection transformation to reduce the data distribution difference between the source and target domains. Lin *et al.* [34] proposed a signal filtering strategy based on the robust PCA (RPCA) and verified its performance on a cross-day binary emotion recognition task. Chai *et al.* [35] proposed a Subspace Alignment Autoencoder (SAAE), which utilizes stacked autoencoders to transform the differential entropy features of two domains into a domain-invariant subspace, and uses the kernel main component analysis, graph regularization, and maximum mean difference to reduce feature distribution differences between two domains. Yin and Zhang [36] proposed an adaptive Stacking Denoising Autoencoder (SDAE), which adaptively updates the weights of shallow neurons in the model using enhanced test samples and their pseudo-labels during the testing phase.

III. PRELIMINARIES

Definition 1: Emotion recognition EEG samples. To construct emotion recognition data, the subjects were induced to use video clips to induce emotions, and at the same time, they were recorded using a multi-lead EEG recording device to obtain the original EEG signals. After preprocessing, the raw EEG signals are segmented in units of every second, that is, a series of samples are formed. Each sample $S = [S_1, S_2, \cdots, S_N] \in \mathbb{R}^{N\times T}$, where S_i represents the EEG sequence of the i-th EEG channel, N represents the number of EEG channels, and T represents the number of sampling points in one second. Meanwhile, for each sample, its label Y is given according to the type of video that evokes emotion, i.e. the emotional state of the subject. In the SEED dataset used in this paper, the emotional state includes three emotions: positive, neutral, and negative, while in the SEED-IV dataset, it includes four emotions: positive, neutral, sad, and fear.

Definition 2: Emotion recognition classification task. The research goal of this paper is to find the mapping relationship between the EEG signal sequence and the emotional state of the subjects. The emotion recognition task is: given a multi-channel EEG sequence sample $S = [s_1, s_2, \ldots, s_n]$, to determine the emotional state Y of the subject when the sample is collected. For each EEG sample X, they are the input of the model in this paper. The subject's emotional state $Y$is the one-hot encoding of multi-category emotions, which is the target of model fitting.

IV. EMOTION RECOGNITION METHODS BASED ON META-TRANSFER LEARNING

Emotions vary greatly among different subjects. To address the differences between subjects, some researchers use transfer learning methods to improve performance on new subjects, but most of these methods still require a lot of data dependencies. This paper proposed a method using a meta-learning strategy for cross-subject emotion recognition, which enables the model to quickly adapt to new subjects through meta-learning strategies while reducing data dependencies in cross-subject contexts by sampling meta-tasks. The proposed meta-learning-based emotion recognition model includes two parts: a meta-learning framework and a feature extractor. Specifically, the meta-learning framework is a training strategy that learns the initialization parameters of a feature extraction classifier by meta-training a meta-task composed of data from different subjects, so that the classifier can adapt to unknown subjects quickly and cost-effectively. The feature extractor is a deep neural network that achieves high-performance emotion recognition classification by extracting EEG features. Specifically, the input EEG signals are extracted separately through a 3D DenseNet network, and finally, emotion classification is performed through a fully connected layer.

A. META-TRANSFER LEARNING

The core of meta-learning is to learn, which is a learning method that is more in line with our human learning style. There are many learning objectives of meta-learning. The hyperparameters of the model, the structure of the model, and the initialization weights of the model can all be used as the goals of meta-learning. The meta-learning method of learning the initialization weight of the model has attracted the attention and application of researchers in recent years. Compared with other meta-learning methods, this type of method is usually easy to deploy and can be applied to various basic neural networks and quickly adapt to mission scenarios. In this paper, a meta-learning method is used to learn the initialization weight of the model. Through meta-learning, on the one hand, the model can achieve better performance with fewer samples and reduce the cost of model training. On the other hand, meta-learning can improve the generalization ability of the model, and the speed of the model adapting to new subjects will also be accelerated.

In recent years, Finn *et al.* [37] proposed a Model-Agnostic Meta-Learning (MAML), which is optimized based on the model algorithm. Afterward, it can have a stronger generalization ability and stronger ability to adapt to unknown data. Firstly, MAML meta-learning needs to define a batch of meta-tasks, which are a set of tasks with similar goals, but they may have different data distributions due to some factors such as individual differences and spatio-temporal differences. For each meta-task, assuming that task T follows a distribution $p(T)$, the model with parameters θ is represented by the function $f(x,\theta)$. For a task T_i, its task loss can be expressed as $\nabla_\theta \mathrm{L_T}(f_\theta)$. For a multi-class task, that is, the cross-entropy loss function, when the gradient descent method is used to update, the model parameter θ is updated according to the following equation:

$$\theta_i' = \theta - \alpha\nabla_\theta \mathrm{L}_{\mathrm{T}_i}(f_\theta), \tag{1}$$

where α is the learning rate of the inner layer optimization of meta-training. Similar to other machine learning methods, by minimizing the loss on task T_i, the performance of the model on task T_i can be optimized. The difference is that the unit of training in MAML is a meta-task, rather than a single sample, the goal of the meta-task is to reduce the total loss of the model on different tasks:

$$\min_\theta \sum_{\mathrm{T}_i\sim p(\mathrm{T})} \mathrm{L}_{\mathrm{T}_i}\left(f_{\theta_i}\right) = \sum_{\mathrm{T}_i\sim p(\mathrm{T})} \mathrm{L}_{\mathrm{T}_i}\left(f_{\theta-\alpha\nabla_\theta \mathrm{L}_{\mathrm{I}_j}(f_\theta)}\right). \tag{2}$$

After performing meta-optimization on multiple tasks, the model parameters are updated as:

$$\theta \leftarrow \theta - \beta\nabla_\theta \sum_{\mathrm{T}_i\sim p(\mathrm{T})} \mathrm{L}_{\mathrm{T}_i}\left(f_{\theta_j}\right), \tag{3}$$

where β is the learning rate optimized for meta-training outer training. In the above MAML training process, MAML is calculated by stochastic gradient descent (SGD):

$$\begin{aligned} g_{\mathrm{MAML}} &= \frac{\partial}{\partial\phi} L_{\tau,B}\left(U_{\tau,A}(\phi)\right) \\ &= U_{\tau,A}'(\phi) L_{\tau,B}'(\not g), \quad \text{where } \overset{g}{\phi} = U_{\tau,A}(\phi). \end{aligned} \tag{4}$$

where g_{MAML} is the calculation result of SGD and ϕ is the input parameter. It can be seen from equations (4) that MAML needs to calculate the second derivative when updating, which will consume a lot of computing resources. The researchers found that the second-order derivative part is usually small and has little effect on the model update direction, so the first-order MAML (FOMAML) is proposed to simplify the gradient calculation and avoid the huge computational cost of the second-order derivative.

This paper uses an improved first-order MAML algorithm, Reptile [7], and finds another way of meta-optimization under the condition of less computational cost, that is, the parameter difference before and after the model is optimized for each task as the update direction of meta-parameter, takes $\left(\theta'-\theta\right)$ as the gradient each time the outer loop is optimized. In practical applications, for multiple meta-tasks, the mean of the model differences before and after training for each task is used as the update direction:

$$\phi \leftarrow \phi + \mathrm{o}\frac{1}{n}\sum_{i=1}^{n}\left(\phi_i-(\phi)\right) \tag{5}$$

where $\left(\theta'-\theta\right)$ is the updated gradient. In this way, the model can generalize the optimization direction of each task, thereby forming a meta-parameter θ_0 suitable for adjustment of new unknown tasks, and improving the generalization ability of the model. The researchers also demonstrated that this update method is consistent with the dominant term of the original MAML update. At the same time, Reptile uses the direction of weight update on each task as the update direction of the meta-weight, so there is no need to divide the training set and test set for each meta-task. In addition, it can be proved by equation derivation that the meta update strategy of the Reptile algorithm and the MAML algorithm is in the same line. Assuming that each task has k mini-batches, the loss of the $i \in \{1, 2, \ldots, k\}$ minibatch is denoted as L_i. When using SGD for optimization, the gradient of each minibatch can be denoted as $g_i = L_i'(\phi_i)$, the gradient of the starting point is denoted as $\bar{g}_i = L_i(\phi_1)$, and the Hessian matrix of the starting point is denoted as $\bar{H}_i = L_i'(\phi_1)$. After training on a mini-batch of data, the parameters of the model are updated to $\varphi_{i+1} = \varphi_i - \alpha g_i$. Then the gradients of MAML, FOMAML, and Reptile can be expressed as:

$$g_{\mathrm{MAML}} = \bar{g}_k - \alpha\bar{H}_k\sum_{j=1}^{k-1}\bar{g}_j - \alpha\sum_{j=1}^{k-1}\bar{H}_j\bar{g}_k + O\left(\alpha^2\right). \tag{6}$$

$$g_{\mathrm{FOMAML}} = g_k = \bar{g}_k - \alpha\bar{H}_k\sum_{j=1}^{k-1}\bar{g}_j + O\left(\alpha^2\right). \tag{7}$$

$$\begin{aligned} g_{\mathrm{Reptile}} &= -\left(\phi_{k+1}-\phi_1\right)/\alpha = \sum_{i=1}^{k} g_i = \sum_{i=1}^{k}\bar{g}_i \\ &\quad -\alpha\sum_{i=1}^{k}\sum_{j=1}^{i-1}\bar{H}_i\bar{g}_j + O\left(\alpha^2\right). \end{aligned} \tag{8}$$

where use AvgGrad to represent the expected value of the loss, and AvgGradInner to represent the inner product between different minibatches:

$$\mathrm{AvgGrad} = \mathrm{E}_{\tau,1}\left[\bar{g}_1\right]. \tag{9}$$

$$\begin{aligned} \mathrm{AvgGradInner} &= \mathrm{E}_{\tau,1,2}\left[\bar{H}_2\bar{g}_1\right] \\ &= \mathrm{E}_{\tau,1,2}\left[\bar{H}_1\bar{g}_2\right] \\ &= \frac{1}{2}\mathrm{E}_{\tau,1,2}\left[\bar{H}_2\bar{g}_1+\bar{H}_1\bar{g}_2\right] \\ &= \frac{1}{2}\mathrm{E}_{\tau,1,2}\left[\frac{\partial}{\partial\phi_1}\left(\bar{g}_1\cdot\bar{g}_2\right)\right]. \end{aligned} \tag{10}$$

where AvgGrad is the direction of reducing the task loss, and the direction of Inner is the direction of increasing the inner product, which makes the model's generalization ability stronger. The expected loss of meta-updates for the three methods can be simplified as:

$$\mathrm{E}\left[g_{\mathrm{MAM}}\right] = (1)\ \mathrm{AvgGrad} - (2(k-1)\alpha)\ \mathrm{Inner}. \tag{11}$$

$$\mathrm{E}\left[g_{\mathrm{FOMAML}}\right] = (1)\mathrm{AvgGrad} - ((k-1)\alpha)\ \mathrm{Inner}. \tag{12}$$

$$\mathrm{E}\left[g_{\mathrm{Reptile}}\right] = (k)\ \mathrm{AvgGrad} - \left(\frac{1}{2}k(k-1)\alpha\right)\mathrm{Inner}. \tag{13}$$

where It can be found that all three meta-learning algorithms can update the model by increasing the inner product of the gradients in the task while reducing the task loss, but the coefficients of the two gradients are different for different methods, which means that the model is in the same times the focus of the momentary update is different.

As mentioned above, the core of transfer learning is to use the knowledge of the source domain to improve the performance of the target domain. The feature of meta-learning is to have meta-level training objectives and use a series of meta-tasks for meta-training. Therefore, transfer learning and meta-learning are not contradictory products, they have different emphases. Meta-learning can generalize the knowledge of the training meta-task and apply it to the test meta-task, improving the performance of the test meta-task. If the relatively independent training meta-tasks are regarded as multiple source domains, and the testing meta-tasks are regarded as the target domain, this coincides with the purpose of transfer learning. This is also the simplest combination of transfer learning and meta-learning.

In the emotion recognition model, to improve the generalization ability of the model to different subjects and facilitate the model to adapt to new subjects faster, the emotion classification task of each subject is defined as a meta-task, and the above meta-learning is used. The method meta-trains the initialization parameters of the model. Specifically, this paper first randomly initializes the weight matrix of the model, that is, initializes the meta-parameter θ. Then divides the meta-tasks according to the principle of subject independence, and each subject becomes an independent meta-task. Then takes the meta-task as the unit uses Reptile's update rules and training meta-tasks to perform meta-training. Different meta-tasks alternately optimize model parameters θ to obtain meta-parameters θ_0.

Then based on meta-parameters θ_0, use a small amount of data from the target meta-task to fine-tune the model parameters. Finally, the final performance of the model is tested using target meta-task data that was not involved in training.

B. FEATURE EXTRACTION NETWORK

This paper uses a 3D DenseNet for deep feature extraction. Different from traditional convolutional networks, DenseNet concatenates the output of the current convolutional layer with the outputs of all previous convolutional layers as the input of the next convolutional layer. Let H_l represent the lth layer convolution, then the convolution layer operation can be expressed as:

$$X_l = H_l\left([X_0, X_1, \ \mathrm{K}, X_{l-1}]\right). \tag{14}$$

In addition, to reduce the increase in computation and memory requirements caused by 3D convolution, this paper uses a pseudo-3D convolution to replace the traditional 3D convolution layer. Specifically, the three-dimensional convolution kernel of shape $k \times k \times d$ is divided into two convolution kernels of $k \times k \times 1$ and $1 \times 1 \times d$, which are equivalent to 2D convolution and 1D convolution respectively. The complexity of 3D convolution is significantly reduced, and the calculation equation of the lth convolution layer can be expressed as:

$$H_l(X) = \mathrm{Conv}^{1\times1\times d}\left(\mathrm{Conv}^{k\times k\times 1}(X)\right). \tag{15}$$

Meanwhile, the first two dimensions of the 3D representation just represent the spatial dimension, and the third dimension represents the time or frequency dimension. Similar to DenseNet, to avoid redundancy or interference caused by too many and too large feature maps, and at the same time reduce the computational cost, a transition layer is used between each Dense Block to reduce the dimension of the feature maps. Specifically, batch normalization is first performed on the input, followed by 1×1 convolution to reduce the number of feature maps, and spatial dimension average pooling to reduce the size of feature maps. To fuse high-level features, this paper uses the splicing operation to input the output of the feature extractor to the fully connected network, and finally performs multi-classification through a fully connected layer activated by the softmax function. The loss for multi-classification uses the cross-entropy loss:

$$\mathrm{L} = -\sum \log\left(y'\right). \tag{16}$$

V. EXPERIMENTAL SETUP AND RESULT ANALYSIS

This section details the experimental setup and results analysis of the meta-transfer learning-based sentiment classification method. First, this paper introduces the SEED and SEED-IV datasets and some data preprocessing operations used in the experiment. Secondly, the performance of the model is compared with the baseline model, which proves that the sentiment classification method based on meta-transfer learning proposed in this paper has better performance. Finally, through a large number of experiments, this paper determines the final structure of the network and uses the performance of the model to be further evaluated and ablation experiments.

A. DATASETS AND PREPROCESSING

To verify the effectiveness of the model proposed in this paper, this paper uses two public datasets SEED [18] and [22] for experiments. SEED dataset: A total of 15 subjects (7 males and 8 females, mean age 23.27) participated. During the experiment, 15 Chinese movie clips were selected to stimulate the subjects, respectively evoking positive, neutral, or negative emotions. The length of each video was about 4 minutes, and movie clips that evoked the same emotion were not shown continuously. There is a 5-second prompt before each video is played, a 45-second self-assessment, and a 15-second break after the video is finished. The subjects wore a 62-lead EEG helmet based on the international 10-20 EEG system for testing, and the frequency of the collected raw signals was 1000Hz. The original signal was downsampled to 200Hz and filtered using a bandpass filter from 0 to 75Hz.

SEED-IV Dataset: A total of 15 subjects (7 males and 8 females) participated. During the experiment, 72 video clips were selected to stimulate the subjects, respectively evoking happy, sad, fearful, or neutral emotions, and each video was about 2 minutes in length. There is also a 5-second prompt before each video is played, and 45 seconds of self-evaluation after the video is finished. The subjects' EEG signals were collected by the 62-channel ESI NeuroScan system, and the frequency of the collected raw signals was 1000 Hz. The original signal was first downsampled to 200Hz and subsequently filtered with a bandpass filter from 1 to 75Hz to remove artifacts. In addition, the SEED-IV dataset also provides eye movement data of the subjects so that researchers can explore the influence of eye movement data on emotions. This paper does not use eye movement data. Both datasets have extracted power spectral densities (PSD) over 5 frequency bands (δ: 1∼3Hz, θ: 4∼7Hz, α: 8∼13Hz, β: 14∼30Hz, γ: 31∼50Hz) and differential entropy (DE) features. For a random signal x, they are calculated as:

$$\mathrm{PSD} = \mathrm{E}\left[x^2\right]. \tag{17}$$

$$\mathrm{DE} = -\int_{-\infty}^{\infty} P(x)\ln(P(x))dx. \tag{18}$$

Assuming that the EEG signal obeys a Gaussian distribution $x \sim N\left(\mu, \sigma^2\right)$, the equation for calculating differential entropy can be simplified as:

$$\begin{aligned}\mathrm{DE} = &-\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} \exp\frac{(x-\mu)^2}{2\sigma^2} \\ &\times \ln\left(\frac{1}{\sqrt{2\pi\sigma}} \exp\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ = &\ \frac{1}{2}\ln 2\pi e\sigma^2.\end{aligned} \tag{19}$$

TABLE 1. Summary of SEED and SEED-IV dataset information.

	SEED	SEED-IV
EEG montage number	62	62
Kind of emotion	Positive Neutral Negative	Happy Neutral Sad Fear
Sampling Rate	200HZ	200HZ
Filter	0∼75HZ	1∼75HZ
Number of subjects	15	15
Number of scenarios	3	3
Number of video clips	15	24
Frequency band	δ : 1∼3Hz θ : 4∼7Hz α : 8∼13Hz β : 14∼30Hz γ : 31∼50Hz	δ : 1∼3Hz θ : 4∼7Hz α : 8∼13Hz β : 14∼30Hz γ : 31∼50Hz

Subsequently, all features are smoothed using the moving average method or linear dynamic system (LDS) method. The frequency-domain feature mainly used in this paper is the differential entropy feature after linear dynamic system smoothing. Details about the dataset and data processing are shown in Table 1.

B. EVALUATION METRICS

For EEG emotion recognition tasks, there are usually two experimental settings: cross-subject and non-cross-subject. Due to the large emotional differences between different subjects, the performance of the same model in the two experimental settings also often varies widely. This paper focuses on using transfer learning methods to reduce the differences between subjects, so the experimental setting of this paper adopts a cross-subject experimental setting, and uses K-fold cross-validation, leave-one-subject-out (LOSO) organization data, that is, each fold uses all the data of one subject for testing, and all the data of the remaining subjects as the training set traverses all subjects in this way and cross-validates K subjects. The final result is the average performance of the model across all subjects. The emotion recognition task studied in this paper is a multi-classification problem, so this paper reports its accuracy (ACC), F1-score (F1-score), and Kappa (Cohen's kappa) for all experimental results to evaluate the classification performance of the model:

$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN}. \tag{20}$$

$$\text{F1 - score} = \frac{2TP}{2TP + FP + FN}. \tag{21}$$

$$\text{Kappa} = \frac{p_0 - p_e}{1 - p_e}, \tag{22}$$

where TP is the number of samples that predict positive classes as positive classes, TN is the number of samples that predict negative classes as negative classes, FP is the number of samples that predict negative classes as positive classes, and FN is the number of samples that predict positive classes as negative classes The number of samples in the class. p_0 is the observed agreement rate, and p_e is the expected agreement rate.

C. BASELINE METHODS AND EXPERIMENTAL SETTINGS

To verify the advanced nature of the proposed method, seven baseline models are selected for comparison. The baseline models used are as follows:

(1) SVM [38]: Support Vector Machine, a traditional machine learning method, determines the decision boundary by solving the hyperplane with the largest margin for the training samples.

(2) RF [39]: Random Forest is an ensemble learning method that integrates multiple decision trees through the idea of ensemble learning, and multiple decision trees are classified through a voting mechanism.

(3) MLP [40]: Multilayer Perceptron, which is a multi-layer perceptron, a most basic feed-forward artificial neural network, which realizes the weight update of the neural network by back-propagating the error.

(4) 3D-CNN [41]: an emotion classification model built using 3D convolutional neural networks, which constructs a 3D EEG topographic map in the time domain, and uses 3D convolution to capture the spatiotemporal information of EEG signals.

(5) ResLSTM [28]: Residual LSTM, that is, an LSTM neural network combined with a residual structure, captures the temporal information of EEG signals through long short-term memory.

(6) CDCN [42]: Channel-fused Dense Convolutional Network, which is a dense convolutional network of channel fusion, which uses the two-dimensional DenseNet to classify the EEG feature matrix.

(7) ACRNN [43]: Attention-based Convolutional Recurrent Neural Network, which is an attention-based convolutional recurrent neural network, which captures more discriminative spatiotemporal information through channel-dimensional attention and self-attention mechanisms.

At the same time, for a fairer comparison, for the baseline model, the LOSO strategy is first used to train the data of the source subjects in the experiment, and then the same data partition strategy as the meta-transfer learning method proposed in this paper is used. The model is fine-tuned on the target subject's data and finally tested with other data on the target subject.

The GPU server cluster used has a total of 4 GPU computing nodes, using the Ubuntu operating system, each node has two 8-core 16-thread processors and 4 discrete graphics cards that support CUDA. The server cluster is configured with the OpenPAI cluster management platform to facilitate the use of multiple users. Each time, the Docker image is specified in units of tasks, and the running code is submitted. The detailed configuration of each node is shown in Table 2.

In the research process, the notebook is used to perform simple data preprocessing and experimental code debugging

TABLE 2. Detailed configuration table of GPU server node for experiment.

	Configuration
Processor	2×Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHZ(8C 16T)
Memory	128GB
Graphic Card	4×Nvidia RTX 2080 8G
Hard Disk	4.0TB
Operating system	Ubuntu 16.04 (Linux 4.4.0
Software Environment	OpenPAI 0.17.0, Docker, CUDA 10.1, Python 3.7, Tensorflow 1.15.0, Pytorch 1.3, etc.

TABLE 3. Results compare with baseline model on SEED dataset.

	Directly-transfer			Fine-tuning or meta-transfer		
	ACC	F1-score	Kappa	ACC	F1-score	Kappa
RF	0.5329	0.4865	0.2991	0.6566	0.6471	0.4848
SVM	0.5312	0.4683	0.2956	0.6928	0.6872	0.5387
MLP	0.5510	0.4808	0.3272	0.7005	0.6738	0.5495
3D-CNN	0.6059	0.5632	0.4083	0.7415	0.7383	0.6126
LSTM	0.5897	0.5563	0.4097	0.7441	0.7128	0.6158
CDCN	0.6313	0.6027	0.4470	0.6931	0.6680	0.5396
ACRNN	0.6689	0.6666	0.5257	0.7628	0.7387	0.6444
Ours	0.6886	0.6780	0.5341	0.7752	0.7720	0.6629

TABLE 4. Results compare with baseline model on SEED-IV dataset.

	Directly-transfer			Fine-tuning or meta-transfer		
	ACC	F1-score	Kappa	ACC	F1-score	Kappa
RF	0.3472	0.2965	0.1340	0.4737	0.4511	0.2917
SVM	0.4109	0.3477	0.2094	0.4849	0.4626	0.3085
MLP	0.4435	0.3582	0.2464	0.5066	0.3974	0.3283
3D-CNN	0.4785	0.4160	0.2776	0.5410	0.5070	0.3839
ResLSTM	0.4707	0.4337	0.3103	0.5110	0.4609	0.3482
CDCN	0.4645	0.4260	0.2850	0.5450	0.5211	0.3928
ACRNN	0.4646	0.4626	0.2847	0.4920	0.4619	0.2962
Ours	0.4920	0.4530	0.3231	0.6120	0.5793	0.4879

on the dataset first. Then large-scale experimental verification and testing are performed using a GPU server cluster. When implementing the model, it mainly uses two deep learning frameworks, Keras API based on Tensorflow and Pytorch. Both are well-known open-source software libraries in the field of deep learning. Both are easy to learn, easy to operate and have a relatively complete software ecosystem. At the same time, both of them have relatively perfect GPU acceleration functions, and the code runs more efficiently.

D. COMPARATIVE ANALYSIS WITH BASELINE METHODS

The performances of the model in this paper and the baseline models on the SEED and SEED-IV datasets are shown in Table 3 and Table 4, respectively. In the case of the direct transfer model, the traditional machine learning methods SVM and RF have the worst classification performance because it is difficult to capture the complex feature patterns of physiological signals. MLP, as the most basic neural network model, can capture a small number of time-frequency features, so the performance is better than traditional machine learning methods. CNN and RNN can better extract deep temporal or spatial features, so the final effect of 3D-CNN, CDCN, and ResLSTM models can achieve better than MLP models. On the one hand, the high-precision feature extractor used in the meta-learning emotion recognition

TABLE 5. The performance of the meta-learned emotion recognition model.

Subject	Directly-transfer			Fine-tuning or meta-transfer		
	ACC	F1-score	Kappa	ACC	F1-score	Kappa
Subject 1	0.7675	0.7584	0.6508	0.6435	0.6194	0.5208
Subject 2	0.7089	0.7033	0.5623	0.5313	0.4378	0.3599
Subject 3	0.7379	0.7415	0.6078	0.4789	0.4728	0.3030
Subject 4	0.7190	0.7160	0.5780	0.6042	0.6035	0.4719
Subject 5	0.7674	0.7583	0.6512	0.6265	0.6222	0.5027
Subject 6	0.7619	0.7554	0.6421	0.8130	0.8078	0.7492
Subject 7	0.9000	0.8982	0.8499	0.5236	0.4917	0.3559
Subject 8	0.9184	0.9173	0.8774	0.6974	0.6972	0.5967
Subject 9	0.6714	0.6782	0.5069	0.5814	0.5772	0.4423
Subject 10	0.7811	0.7777	0.6714	0.6275	0.6145	0.5026
Subject 11	0.8289	0.8290	0.7438	0.5673	0.4526	0.4101
Subject 12	0.6082	0.6138	0.4122	0.6469	0.6540	0.5252
Subject 13	0.8151	0.7944	0.7234	0.6901	0.6735	0.5861
Subject 14	0.7120	0.7106	0.5697	0.6435	0.6426	0.5228
Subject 15	0.9314	0.9308	0.8971	0.5070	0.3281	0.4694
Mean	0.7753	0.7722	0.6629	0.6121	0.5891	0.4785
STD	0.0883	0.0868	0.1324	0.0831	0.1018	0.1139

method proposed in this paper conducts comprehensive modeling of the features of EEG signals and can capture the multi-dimensional information of EEG signals more comprehensively. Therefore, in the direct transfer model. In this case, the model proposed in this paper can still achieve optimal performance. In addition, it is worth mentioning that the computational complexity of the proposed model is the lowest compared with other methods.

However, under the training settings of fine-tuning or meta-learning, the performance of various methods has improved, which indicates that subject differences clearly exist, and the impact of subject differences can indeed be mitigated by transfer learning methods. Improve the accuracy of the model. At the same time, the meta-learning method proposed in this paper is better than all the fine-tuning effects of other models. On the one hand, it shows that the meta-learning method can effectively improve the generalization performance of the model. On the other hand, it also shows that the effect of the meta-transfer learning method is better than the traditional one. The fine-tuning method is more effective.

E. MODEL TESTING AND ABLATION EXPERIMENTS

To explore the effectiveness of each module of the model, this paper conducts ablation experiments on the SEED-IV dataset. Specifically, this paper uses a subject-independent way to conduct experiments, using 15-fold cross-validation and LOSO scheme to divide the data set, that is, using the data of 14 subjects as training data each time (source data), and the data of the remaining 1 subject is used as validation data (target data). A small number of samples (about 20%) are used in the validation data to adjust the model, and the rest of the samples are used as the test set. Finally, this paper reports the average of the ACC, F1-score, and Kappa values of the 15-fold cross-validation classification results to evaluate the model. The final performance of the meta-learned emotion recognition model is shown in Table 5.

The same emotion recognition model will show a great difference in performance on the data of different subjects,

TABLE 6. Meta-Learning ablation experiment results.

model	ACC / STD	F1-score / STD	Kappa / STD
directly-transfer	0.4920/0.0809	0.4531/0.1038	0.3232/0.1066
Few-Sample	0.5616/0.1059	0.5345/0.1205	0.4119/0.1428
Full-Sample	0.6121/0.0831	0.5891/0.1018	0.4785/0.1139

while the meta-learning model proposed in this paper is subject to the independent situation of the three emotion classifications. It can achieve an accuracy of 77.5% and a standard deviation of 8.8% on the SEED dataset of 2000, and an accuracy of 61.2% and 8.3% on the SEED-IV dataset of four sentiment classifications. The model performance is relatively stable.

To explore meta-learning and its performance under a reduced sample size, this paper compares model direct transfer, few-shot meta-learning, and full-shot meta-learning. Among them, only 5% of the target data is used to adjust the model during few-shot learning. As shown in Table 6, when the model is directly transferred, the accuracy rate of the model is only 49.2%, while the meta-learning model after adaptation with a small number of samples can reach 53.6%. This accuracy rate can also be compared with the current more advanced models, an accuracy of 61.2% can be achieved after increasing the amount of training data. This shows that the meta-learning strategy in the model does improve the generalization performance of the model. In addition, the meta-learning with few samples also achieves satisfactory performance under the condition of consuming the least resources, indicating that the meta-learning strategy is an efficient learning method.

VI. CONCLUSION

Accurate identification of emotions is of great significance for everyone to manage their emotions, enrich psychological theories, and treat mental illnesses. Emotion recognition based on EEG signals can better reflect human psychological and emotional changes. Compared with facial expression pictures, it is less affected by environmental factors when taking pictures, which is more conducive to extracting the depth features of EEG. Many researchers have studied EEG-based emotion recognition methods, but the proposed models will have obvious performance degradation in subject-independent settings. Some researchers apply transfer learning methods to emotion recognition to improve performance when subjects are independent, but they still require a lot of data dependencies.

To solve the problem of data dependence when subjects are independent, this paper proposes a meta-transfer learning framework for emotion recognition, which integrates the idea of meta-learning and transfer learning to improve the performance of subjects when they are independent and reduce the need for the dependency of training data. At the same time, this paper evaluates the model using two public datasets, SEED, and SEED-IV, and the results show that the meta-transfer learning method outperforms the baseline model in terms of accuracy, F1-score, and Kappa value. In addition, to verify the effectiveness of each component of the model, ablation experiments were carried out, and the experimental results confirmed the effectiveness of these components. Compared with traditional methods, the meta-transfer learning framework used in this paper reduces the dependence on training subjects and can achieve rapid adaptation to unknown subjects with only a small sample size. In addition, the meta-transfer learning used in this paper is used. The framework can also provide a reference for other problems in the context of transfer learning. Emotion classification using EEG signals can be divided into two types. One is to use multi-channel EEG signals for emotion recognition. The other is to use single channel EEG signals for emotion recognition. The proposed method uses multi-channel EEG signals for high-precision emotion recognition. However, the use of single channel EEG signals helps to propose methods for the deployment of wearable devices.

REFERENCES

[1] Z. Jia, Y. Lin, J. Wang, Z. Feng, X. Xie, and C. Chen, ''HetEmotionNet: Two-stream heterogeneous graph recurrent neural network for multi-modal emotion recognition,'' in *Proc. 29th ACM Int. Conf. Multimedia*, Oct. 2021, pp. 1047–1056.

[2] Z. Jia, Y. Lin, X. Cai, H. Chen, H. Gou, and J. Wang, ''SST-EmotionNet: Spatial-spectral-temporal based attention 3D dense network for EEG emotion recognition,'' in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 2909–2917.

[3] A. Vulli, P. N. Srinivasu, M. S. K. Sashank, J. Shafi, J. Choi, and M. F. Ijaz, ''Fine-tuned DenseNet-169 for breast cancer metastasis prediction using FastAI and 1-cycle policy,'' *Sensors*, vol. 22, no. 8, p. 2988, Apr. 2022.

[4] K. K. Sahoo, I. Dutta, M. F. Ijaz, M. Wozniak, and P. K. Singh, ''TLEFuzzyNet: Fuzzy rank-based ensemble of transfer learning models for emotion recognition from human speeches,'' *IEEE Access*, vol. 9, pp. 166518–166530, 2021.

[5] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, ''Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM,'' *Sensors*, vol. 21, no. 8, p. 2852, 2021.

[6] M. F. Ijaz, M. Attique, and Y. Son, ''Data-driven cervical cancer prediction model with outlier detection and over-sampling methods,'' *Sensors*, vol. 20, no. 10, p. 2809, May 2020.

[7] A. Nichol, J. Achiam, and J. Schulman, ''On first-order meta-learning algorithms,'' 2018, *arXiv:1803.02999*.

[8] Z. Jia, Y. Lin, Y. Liu, Z. Jiao, and J. Wang, ''Refined nonuniform embedding for coupling detection in multivariate time series,'' *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 101, no. 6, Jun. 2020, Art. no. 062113.

[9] Z. Jia, X. Cai, Y. Hu, J. Ji, and Z. Jiao, ''Delay propagation network in air transport systems based on refined nonlinear Granger causality,'' *Transportmetrica B, Transp. Dyn.*, vol. 10, no. 1, pp. 586–598, Dec. 2022.

[10] Z. Li, J. Wang, Z. Jia, and Y. Lin, ''Learning space-time-frequency representation with two-stream attention based 3D network for motor imagery classification,'' in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2020, pp. 1124–1129.

[11] Z. Jia, J. Ji, X. Zhou, and Y. Zhou, ''Hybrid spiking neural network for sleep electroencephalogram signals,'' *Sci. China Inf. Sci.*, vol. 65, no. 4, pp. 1–10, Apr. 2022.

[12] Z. Jia, Y. Lin, J. Wang, X. Ning, Y. He, R. Zhou, Y. Zhou, and L.-W.-H. Lehman, ''Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification,'' *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1977–1986, 2021.

[13] X. Cai, Z. Jia, M. Tang, and G. Zheng, ''BrainSleepNet: Learning multivariate EEG representation for automatic sleep staging,'' in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2020, pp. 976–979.

[14] Z. Jia, X. Cai, and Z. Jiao, "Multi-modal physiological signals based squeeze-and-excitation network with domain adversarial learning for sleep staging," *IEEE Sensors J.*, vol. 22, no. 4, pp. 3464–3471, Feb. 2022.
[15] F. Bahari and A. Janghorbani, "EEG-based emotion recognition using recurrence plot analysis and k nearest neighbor classifier," in *Proc. 20th Iranian Conf. Biomed. Eng. (ICBME)*, Dec. 2013, pp. 228–233.
[16] X. W. Wang, D. Nie, and B. L. Lu, "EEG-based emotion recognition using frequency domain features and support vector machines," in *Proc. Int. Conf. Neural Inf. Process.* Berlin, Germany: Springer, 2011, pp. 734–743.
[17] Y. Bengio, I. Goodfellow, and A. Courville, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2016.
[18] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification," in *Proc. 6th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, Nov. 2013, pp. 81–84.
[19] L. I. Goldfischer, "Autocorrelation function and power spectral density of laser-produced speckle patterns," *J. Opt. Soc. Amer.*, vol. 55, no. 3, pp. 247–253, Mar. 1965.
[20] Y. Liu and O. Sourina, "Real-time fractal-based valence level recognition from EEG," in *Transactions on Computational Science XVIII*. Berlin, Germany: Springer, 2013, pp. 101–120.
[21] Y. P. Lin, C. H. Wang, and T. P. Jung, "EEG-based emotion recognition in music listening," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1798–1806, Jul. 2010.
[22] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Trans. Auton. Mental Develop.*, vol. 7, no. 3, pp. 162–175, Sep. 2015.
[23] A. Al-Nafjan, M. Hosny, A. Al-Wabil, and Y. Al-Ohali, "Classification of human emotions from electroencephalogram (EEG) signal using deep neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 9, pp. 419–425, 2017.
[24] W.-L. Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, "EEG-based emotion classification using deep belief networks," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, Jul. 2014, pp. 1–6.
[25] Y. Yang, Q. M. J. Wu, W.-L. Zheng, and B.-L. Lu, "Eeg-based emotion recognition using hierarchical network with subnetwork nodes," *IEEE Trans. Cogn. Devel. Syst.*, vol. 10, no. 2, pp. 408–419, Jun. 2017.
[26] R. Fourati, B. Ammar, and C. Aouiti, "Optimized echo state network with intrinsic plasticity for EEG-based emotion recognition," in *Proc. Int. Conf. Neural Inf. Process.* Cham, Switzerland: Springer, 2017, pp. 718–727.
[27] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Emotion*, vol. 8, no. 10, pp. 355–358, 2017.
[28] J. Ma, H. Tang, W.-L. Zheng, and B.-L. Lu, "Emotion recognition using multimodal residual LSTM network," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 176–183.
[29] J. Li, Z. Zhang, and H. He, "Hierarchical convolutional neural networks for EEG-based emotion recognition," *Cogn. Comput.*, vol. 10, no. 2, pp. 368–380, 2018.
[30] T.-P. Jung and T. J. Sejnowski, "Utilizing deep learning towards multi-modal bio-sensing and vision-based affective computing," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 96–107, Jan. 2022.
[31] Y. Li, L. Wang, W. Zheng, Y. Zong, L. Qi, Z. Cui, T. Zhang, and T. Song, "A novel bi-hemispheric discrepancy model for EEG emotion recognition," *IEEE Trans. Cognit. Develop. Syst.*, vol. 13, no. 2, pp. 354–367, Jun. 2021.
[32] T. Zhang, X. Wang, X. Xu, and C. L. P. Chen, "GCB-Net: Graph convolutional broad network and its application in emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 379–388, Jan. 2022.
[33] X. Chai, Q. Wang, Y. Zhao, Y. Li, D. Liu, X. Liu, and O. Bai, "A fast, efficient domain adaptation technique for cross-domain electroencephalography(EEG)-based emotion recognition," *Sensors*, vol. 17, no. 5, p. 1014, May 2017.
[34] E. J. Candès, X. Li, and Y. Ma, "Robust principal component analysis," *J. ACM JACM*, vol. 58, no. 3, pp. 1–37, 2011.
[35] X. Chai, Q. Wang, Y. Zhao, X. Liu, O. Bai, and Y. Li, "Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition," *Comput. Biol. Med.*, vol. 79, pp. 205–214, Dec. 2016.
[36] Z. Yin and J. Zhang, "Cross-session classification of mental workload levels using EEG and an adaptive deep learning model," *Biomed. Signal Process. Control*, vol. 33, pp. 30–47, Mar. 2017.
[37] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 1126–1135.
[38] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, Jun. 1999.
[39] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[40] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation*. San Diego, CA, USA: California Univ. San Diego La Jolla Institute of Cognitive Science, 1985.
[41] Y. Zhao, J. Yang, J. Lin, D. Yu, and X. Cao, "A 3D convolutional neural network for emotion recognition based on EEG signals," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–6.
[42] Z. Gao, X. Wang, Y. Yang, Y. Li, K. Ma, and G. Chen, "A channel-fused dense convolutional network for EEG-based emotion recognition," *IEEE Trans. Cognit. Develop. Syst.*, vol. 13, no. 4, pp. 945–954, Dec. 2021.
[43] W. Tao, C. Li, R. Song, J. Cheng, Y. Liu, F. Wan, and X. Chen, "EEG-based emotion recognition via channel-wise attention and self attention," *IEEE Trans. Affect. Comput.*, early access, Sep. 22, 2020, doi: 10.1109/TAFFC.2020.3025777.

HENGYAO TANG received the M.Eng. degree from the Wuhan University of Technology, in 2007. Currently, he is an Associate Professor at the School of Computer, Huanggang Normal University. His current research interests include image processing, database theory, and technologies.

GUOSONG JIANG received the Ph.D. degree in engineering from the Huazhong University of Science and Technology, in 2009. Currently, he is a Professor at the School of Computer, Huanggang Normal University. His current research interests include network storage and media big data processing.

QINGDONG WANG received the Ph.D. degree in engineering from the Huazhong University of Science and Technology, in 2007. Currently, he is a Senior Engineer at the School of Computer, Huanggang Normal University. His current research interests include high speed sampling and digital signal processing, application of the high Internet of Things technology, and big data analysis and mining.