Machine Intelligence Research
www.mi-research.net

19(5), October 2022, 472-484
DOI: 10.1007/s11633-022-1352-1

EEG-based Emotion Recognition Using Multiple Kernel Learning

Qian Cai[1] Guo-Chong Cui[2] Hai-Xian Wang[2,3]

[1] School of Statistics and Data Science, Nanjing Audit University, Nanjing 211815, China

[2] Key Laboratory of Child Development and Learning Science of Ministry of Education, School of Biological Science & Medical Engineering, Southeast University, Nanjing 210096, China

[3] Institute of Artificial Intelligence of Hefei Comprehensive National Science Center, Hefei 230094, China

Abstract: Emotion recognition based on electroencephalography (EEG) has a wide range of applications and has great potential value, so it has received increasing attention from academia and industry in recent years. Meanwhile, multiple kernel learning (MKL) has also been favored by researchers for its data-driven convenience and high accuracy. However, there is little research on MKL in EEG-based emotion recognition. Therefore, this paper is dedicated to exploring the application of MKL methods in the field of EEG emotion recognition and promoting the application of MKL methods in EEG emotion recognition. Thus, we proposed a support vector machine (SVM) classifier based on the MKL algorithm EasyMKL to investigate the feasibility of MKL algorithms in EEG-based emotion recognition problems. We designed two data partition methods, random division to verify the validity of the MKL method and sequential division to simulate practical applications. Then, tri-categorization experiments were performed for neutral, negative and positive emotions based on a commonly used dataset, the Shanghai Jiao Tong University emotional EEG dataset (SEED). The average classification accuracies for random division and sequential division were 92.25% and 74.37%, respectively, which shows better classification performance than the traditional single kernel SVM. The final results show that the MKL method is obviously effective, and the application of MKL in EEG emotion recognition is worthy of further study. Through the analysis of the experimental results, we discovered that the simple mathematical operations of the features on the symmetrical electrodes could not effectively integrate the spatial information of the EEG signals to obtain better performance. It is also confirmed that higher frequency band information is more correlated with emotional state and contributes more to emotion recognition. In summary, this paper explores research on MKL methods in the field of EEG emotion recognition and provides a new way of thinking for EEG-based emotion recognition research.

Keywords: Emotion recognition, electroencephalography (EEG), multiple kernel learning, machine learning, brain computer interface.

Citation: Q. Cai, G. C. Cui, H. X. Wang. EEG-based emotion recognition using multiple kernel learning. *Machine Intelligence Research*, vol.19, no.5, pp.472–484, 2022. http://doi.org/10.1007/s11633-022-1352-1

1 Introduction

Emotions are a very complicated psychological and physiological response triggered by various events, and are also considered to be a high-level brain activity. It affects and even determines people′s daily communication and decision-making process. Furthermore, emotion recognition has been recognized as a crucial and essential link of human-computer interaction, especially in advanced brain-computer interface systems[1]. Accordingly, emotion recognition technology has been widely used in recent years, particularly in medical, military and other fields, and is considered to be a necessary part of future artificial intelligence technology.

Nonphysiological signals and physiological signals are two types of body signals used for emotion recognition. Many previous studies were based on facial expression[2, 3], voice[4–6] and other nonphysiological signals because these signals are easier to obtain. However, these methods are obviously not reliable because people can intentionally conceal their real emotions and show unreal facial expressions, tone of voice, and other nonphysiological signals. Compared to these unreliable methods, emotion recognition using physiological signals is clearly more reliable because these signals cannot be deliberately hidden or controlled. These signals involve physiological characteristics such as electroencephalography (EEG)[7], electrocardiogram (ECG)[8], electromyography (EMG)[9], respiratory signals[10], pulse rate[11], skin resistance (SC)[12] and functional magnetic resonance imaging (fMRI)[13]. As a physiological signal, EEG signals reflect neural electrical oscillations in the central nervous system and are dir-

Research Article
Special Issue on Brain-inspired Machine Learning
Manuscript received March 31, 2022; accepted June 20, 2022; published online September 3, 2022
Recommended by Associate Editor Dao-Qiang Zhang
Colored figures are available in the online version at https://link.springer.com/journal/11633

ectly related to high-level cognitive processes[14], including emotion[15]. It has great research prospects in the classification and recognition tasks of emotion, tactile sensation[16] and vision[17]. In addition, EEG signals have a higher temporal resolution than other physiological signals. Therefore, EEG-based emotion recognition has greater potential. So this paper uses EEG signals to conduct emotion recognition research.

For general EEG emotion recognition methods, after acquiring emotion-related EEG signals, emotion-related features are first extracted from them and then input to the classifier for classification and identification of the subject′s emotions. For example, Mohammadi et al.[18] used wavelet entropy and frequency band energy features with K-nearest neighbor (KNN) and support vector machine (SVM) classifiers in the valence/arousal domain for the classification of emotional states. Degirmenci et al.[19] used naive Bayes (NB), SVM, and linear discriminant analysis (LDA) classifiers to explore the effectiveness of features improved with empirical mode decomposition (EMD) in emotion recognition. Yin et al.[20] proposed a new feature selection method called locally-robust feature selection (LRFS) and used classifiers such as NB, logistic regression (LR), KNN, extreme learning machine (ELM), and least squares SVM (LSSVM) to verify that their method works. With the upsurge of neural network (NN) and deep learning research, in addition to traditional machine learning methods, many researchers have applied NN to EEG-based emotion recognition[7, 21−23].

Among these classification algorithms, KNN, LDA, and NB are computationally small, but not suitable for solving nonlinear problems[24]. NNs have been widely studied in recent years due to their data-driven convenience and high accuracy, but they lack interpretability and require more empirical parameters that cannot control the learning model well[25]. In contrast, SVM has produced better performance in solving high-dimensional, nonlinear, small datasets and many other practical application problems[26]. In SVM, there is a basic and essential element used to map samples from one feature space to another, i.e., the kernel function. The classical SVM has only one kernel function, and researchers need to select the best kernel function and its parameters through experience or a large number of experimental operations. Moreover, sometimes a kernel function cannot well characterize multichannel or multitype features. Thus, the multiple kernel learning (MKL) method emerged at a historic moment.

There are many studies on MKL algorithms, and the classic methods recognized are as follows. Simple MKL (SMKL) implements a linear approach of kernel weight combination, which is an iterative algorithm proposed by Rakotomamonjy et al.[27] Generalized MKL (GMKL) is a nonlinear approach proposed by Varma and Babu[28] that solves the problem by regularizing not only the kernel combination weights but also the hyperplane weights. Group lasso regularized MKL (GLMKL) proposed by Kloft et al.[29, 30] solves a minor quadratic programming (QP) problem in each iteration to find the kernel weights and utilizes the 1-norm to regularize the kernel weights. EasyMKL proposed by Aiolli et al.[31] finds a unitary norm vector as a vector of multiple weak kernels combination to maximize the distance between positive and negative samples. Among these algorithms, EasyMKL is more accurate and robust, retaining linear time complexity while efficiently handling a large number of various kernels[31].

Multiple kernel learning SVM (MKL-SVM) can efficiently merge different kernels from multiple sources in a data-driven manner. MKL-SVM not only improves the interpretability of the decision function but also improves the classification performance[32, 33] compared to single kernel SVM. However, as far as we know, there is no related research on EEG-based emotion recognition using MKL-SVM. Therefore, we proposed an EasyMKL-SVM classifier to explore the feasibility of the MKL algorithm in EEG-based emotion recognition. We designed a tri-categorization experiment of negative, positive and neutral emotions based on the Shanghai Jiao Tong University emotional EEG dataset (SEED). Two data partition methods, random division and sequential division were used to conduct the experiment. We compared it with the conventional single kernel SVM to verify its validity. The main contributions of this paper are summarized below.

1) This paper proposes using MKL methods for EEG emotion recognition research. We applied the EasyMKL-SVM method to EEG emotion classification and proved that it is suitable for EEG emotion recognition. This algorithm can automatically select the best combination of kernels from a set of predefined weak kernel functions in a data-driven manner, which reduces the user′s work of selecting appropriate kernel functions and optimizing parameters for different features.

2) We designed two data division methods, random partition and sequential partition, to perform three kinds of emotion classification experiments on the SEED. The former was used to verify the validity, and the latter was used to simulate real application situations. Under the two data division methods, the average classification accuracy of the three emotions reached 92.5% and 74.37%, respectively, both of which outperformed the traditional single kernel SVM method. Moreover, the MKL method proposed in this paper outperforms other state-of-the-art methods.

The remainder of this paper is structured as follows. First, the general steps and related work on EEG-based emotion recognition are presented in Section 2, followed by our proposed method in Section 3. Then, we show our experimental procedure and its results in Section 4. Finally, in Sections 4 and 5, there is a discussion of the experimental results and a general summary of the work in this paper.

2 Related work

In this section, we present the general process of EEG-based emotion recognition research and the theory of the EasyMKL algorithm.

2.1 General process of EEG-based emotion recognition

The use of physiological signals for emotion recognition, especially EEG-based methods, has received increasing attention in recent years. Fig. 1 illustrates the general framework and process of the EEG-based emotion recognition approaches. Researchers usually start their research from one or more of them, and the following is a detailed description of these steps.

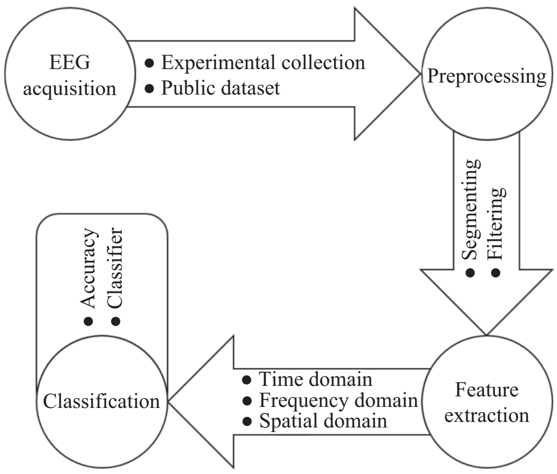


Fig. 1 General flow of EEG emotion recognition

EEG acquisition. There are two main sources of emotional EEG signals: experimental collection and public datasets. The first and most important step in collecting emotional EEG signals is to stimulate the experimental subject′s emotional state in a specific way. Stimulant materials such as music, pictures and videos are commonly used to induce emotions. To make this process more standard and objective, Frantzidis et al.[34, 35] established an international emotional picture/sound system as a benchmark for inducing the subjects to produce emotional states and label emotional categories. Emotion-related EEG signals can be collected by using EEG acquisition equipment while inducing the subjects′ emotions.

In addition to collecting EEG signals through experiments, researchers often use public emotional EEG datasets. Currently, the more commonly used datasets are SEED[36, 37] and database for emotion analysis using physiological signals (DEAP)[38]. A simple comparison of these two datasets is shown in Table 1.

As seen from Table 1, the SEED dataset is released later, and the quality of the EEG data is higher. Therefore, this paper conducts the study on the SEED, and the details of the SEED are described in Section 3.1.

Table 1 Comparison of the SEED and DEAP dataset

Name	Release time	Number of channels	Sampling rate
SEED	2015	62	1 000 Hz
DEAP	2011	32	512 Hz

Preprocessing. In the data preprocessing step, the EEG signal is usually first downsampled to reduce the amount of data. Then, bandpass filtering, blind source separation or independent component analysis[39] is used to remove artifacts and avoid the effects of artifacts.

Feature extraction. The main purpose of this step is to extract information from EEG signals that characterize the emotional state and lay the foundation for the further application of emotion recognition. The traditional features of EEG signals mainly include features from the frequency, space, and time domains[40]. In recent years, as the high complexity and nonlinearity of EEG signals have been gradually recognized, nonlinear analysis has been widely used in the analysis of EEG signals, and entropy and other complexity measures have also been heavily applied to extract the features of EEG signals[41−43].

Classification. Designing an effective sentiment classifier model and conducting classification experiments is the final step in EEG-based emotion recognition research. The quality of the classifier used is decisive for the accuracy of emotion recognition[44]. There are many classifiers that can be used to classify extracted emotional features, such as naive Bayesian (NB), SVM, decision trees, and deep learning classifiers[45].

2.2 EasyMKL-SVM

In SVM, the kernel function is used to implicitly map the samples to the feature space[46]. Researchers may hope to combine multiple possible kernels because of the various types of kernel functions and the fact that it is usually not clear which kernel is the best for the present task. Therefore, it is desirable to develop an optimized kernel function consisting of multiple kernels to fit well with the data structure[47]. To solve and optimize the problem of combining multiple kernels in SVM, the MKL-SVM method was developed.

The goal of MKL is to improve the classification performance of target kernels by combining predefined multiple kernels from different sources in a data-driven manner. We used an algorithm called EasyMKL, proposed by Aiolli and Donini[31], to solve the problem of how to efficiently combine multiple kernels. The EasyMKL algorithm can efficiently handle many different kernels and is considered to be a scalable and robust MKL algorithm[31]. The purpose of EasyMKL is to find the kernel combination that maximizes the margin between classes.

Finding the best parameters to combine a predefined set of kernel matrices is the most important thing in MKL. This is done by learning a coefficient vector $\boldsymbol{\eta}$. The combinatorial kernel is then obtained according to the following formula:

$$\boldsymbol{K}=\sum_{r=1}^{R}\boldsymbol{\eta}_r\boldsymbol{K}_r,\quad \boldsymbol{\eta}_r\geq 0.$$

EasyMKL converted the problem of learning combinatorial kernels into a minimax problem with variables $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$, i.e.,

$$\max_{\boldsymbol{\eta}:\|\boldsymbol{\eta}\|=1}\min_{\boldsymbol{\gamma}\in\Gamma}\underbrace{(1-\lambda)\boldsymbol{\gamma}^{\mathrm{T}}\widehat{\boldsymbol{Y}}\left(\sum\nolimits_{r}^{R}\boldsymbol{\eta}_r\widehat{\boldsymbol{K}}_r\right)\widehat{\boldsymbol{Y}}\boldsymbol{\gamma}+\lambda\|\boldsymbol{\gamma}\|^2}_{Q(\boldsymbol{\eta},\boldsymbol{\gamma})}.$$

It is also proposed that the vector $\boldsymbol{\eta}$ can be used as the kernel combination vector to maximize the distance between positive and negative examples. Considering the distance vector $\boldsymbol{d}(\boldsymbol{\gamma})$ with the rth entry defined as $\boldsymbol{d}_r(\boldsymbol{\gamma})=\boldsymbol{\gamma}^{\mathrm{T}}\widehat{\boldsymbol{Y}}\widehat{\boldsymbol{K}}_r\widehat{\boldsymbol{Y}}\boldsymbol{\gamma}$, the original problem can be rewritten as

$$\begin{aligned}&\min_{\boldsymbol{\gamma}\in\Gamma}\max_{\boldsymbol{\eta}:\|\boldsymbol{\eta}\|=1}Q(\boldsymbol{\eta},\boldsymbol{\gamma})=\\&\qquad\min_{\boldsymbol{\gamma}\in\Gamma}\max_{\|\boldsymbol{\eta}\|=1}(1-\lambda)\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{\gamma})+\lambda\|\boldsymbol{\gamma}\|_2^2.\end{aligned}\tag{1}$$

Maximizing the above function $Q(\boldsymbol{\eta},\boldsymbol{\gamma})$ has a simple analytic solution $\boldsymbol{\eta}^*$ that can be expressed as

$$\boldsymbol{\eta}^*=\frac{\boldsymbol{d}(\boldsymbol{\gamma})}{\|\boldsymbol{d}(\boldsymbol{\gamma})\|_2}.$$

This solution can be substituted into a min-max problem as follows:

$$\min_{\boldsymbol{\gamma}\in\Gamma}Q(\boldsymbol{\eta}^*,\boldsymbol{\gamma})=\min_{\boldsymbol{\gamma}\in\Gamma}(1-\lambda)\|\boldsymbol{d}(\boldsymbol{\gamma})\|_2+\lambda\|\boldsymbol{\gamma}\|_2^2.$$

It can be seen that the regularized minimizer of the 2-norm of the distance vector will be the optimal $\boldsymbol{\gamma}$. EasyMKL further simplifies the problem by transforming the minimization of the 1-norm of the corresponding distance vector into minimizing an upper bound, thus obtaining:

$$\begin{aligned}&\min_{\boldsymbol{\gamma}\in\Gamma}(1-\lambda)\|\boldsymbol{d}(\boldsymbol{\gamma})\|_1+\lambda\|\boldsymbol{\gamma}\|_2^2=\\&\qquad\min_{\boldsymbol{\gamma}\in\Gamma}(1-\lambda)\boldsymbol{\gamma}^{\mathrm{T}}\widehat{\boldsymbol{Y}}\left(\sum_{r}^{R}\widehat{\boldsymbol{K}}_r\right)\widehat{\boldsymbol{Y}}\boldsymbol{\gamma}+\lambda\|\boldsymbol{\gamma}\|_2^2.\end{aligned}\tag{2}$$

The resulting minimization problem is the replacement of the kernel matrix in the kernel optimization of the margin distribution (KOMD) problem with the sum of kernels. Then, EasyMKL uses the 1-norm upper bound to modify the original problem and changes the optimal solution $\boldsymbol{\eta}^*$ of the initial problem (1) with a new $\boldsymbol{\eta}^*$:

$$\boldsymbol{\eta}^*=\boldsymbol{\eta}^*\frac{\|\boldsymbol{d}(\boldsymbol{\gamma})\|_1}{\|\boldsymbol{d}(\boldsymbol{\gamma})\|_2}.$$

Therefore, the problem to be solved becomes

$$\begin{aligned}&\min_{\boldsymbol{\gamma}\in\Gamma}(1-\lambda)\boldsymbol{\eta}^*\boldsymbol{d}(\boldsymbol{\gamma})+\lambda\|\boldsymbol{\gamma}\|_2^2=\\&\qquad\min_{\boldsymbol{\gamma}\in\Gamma}(1-\lambda)\boldsymbol{\eta}^*\frac{\|\boldsymbol{d}(\boldsymbol{\gamma})\|_1}{\|\boldsymbol{d}(\boldsymbol{\gamma})\|_2}\boldsymbol{d}(\boldsymbol{\gamma})+\lambda\|\boldsymbol{\gamma}\|_2^2.\end{aligned}\tag{3}$$

From this new formulation, it can be seen that compared to the original formula, only a coefficient $\|\boldsymbol{d}(\boldsymbol{\gamma})\|_1/\|\boldsymbol{d}(\boldsymbol{\gamma})\|_2$ is added after the original optimal solution $\boldsymbol{\eta}^*$. It can be concluded that the number of kernels constrains this coefficient according to Holder's inequality:

$$1\leq\frac{\|\boldsymbol{d}(\boldsymbol{\gamma})\|_1}{\|\boldsymbol{d}(\boldsymbol{\gamma})\|_2}\leq\sqrt{R}.$$

This coefficient tends to 1 if $\boldsymbol{d}(\boldsymbol{\gamma})$ is very sparse, while it tends to $\sqrt{R}$ if the values in $\boldsymbol{d}(\boldsymbol{\gamma})$ are similar. That is, to solve the minimum problem, EasyMKL uses the proposed formula to promote sparse solutions of the $\boldsymbol{d}(\boldsymbol{\gamma})$, also known as the distance vector. Then, $\boldsymbol{\eta}^*=\boldsymbol{d}(\boldsymbol{\gamma})/\|\boldsymbol{d}(\boldsymbol{\gamma})\|_2$, the weight vector, will be sparser than the original problem's solution when solving the problem in (3). Obviously, the EasyMKL minimum problem in (2) has the same solution as the problem in (3).

3 Method

In this paper, an SVM classifier based on the MKL algorithm EasyMKL was proposed to adapt to different features and achieve better performance. We call it EasyMKL-SVM. The whole process of our novel method will be described in detail in this section, and the flowchart is shown in Fig. 2.

3.1 Dataset preparation

The emotion recognition experiments in this paper are based on a free, publicly available emotional EEG dataset, the SEED[36, 37]. This dataset is an EEG signal dataset provided by Shanghai Jiao Tong University in 2015 specifically for emotion analysis and has been used by many researchers as a benchmark dataset. It contains 62 channels of EEG signals acquired from 15 subjects, and it performed 3 experiments for each subject at different times. In each experiment, 15 Chinese film clips were selected as stimuli. In one session, the subject had 5 s to get a hint, 45 s for self-assessment and 15 s to calm down. Zheng and Lu[37] used movie clips in their experiments to induce three types of emotional states in their subjects: negative, neutral and positive. The 62-channel EEG signals in the SEED were sampled at a rate of 1 000 Hz and acquired according to the international 10–20

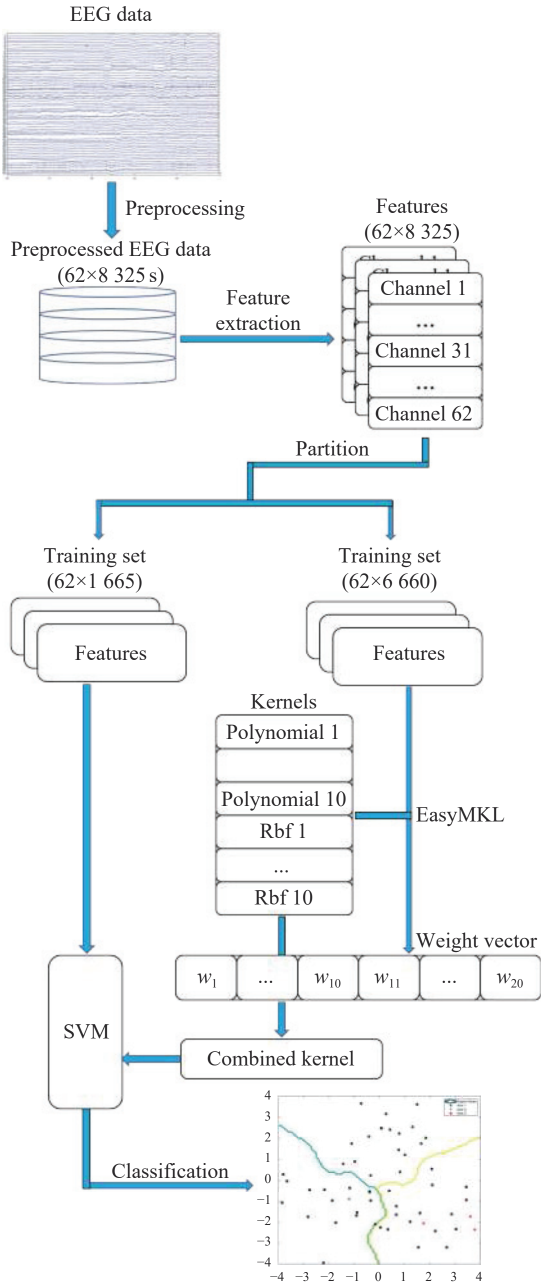


Fig. 2   Flowchart of the proposed method

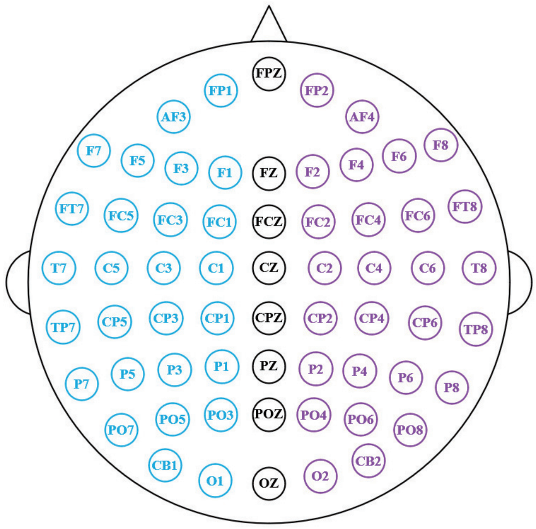


Fig. 3   Electrode position arrangement of the electrode caps in the SEED

system[37] using the electric source imaging (ESI) neuroscan system. The specific electrode position is shown in Fig. 3.

3.2 Data preprocessing

The SEED provides processed EEG data that have been sampled down to 200 Hz and bandpass filtered to 0−75 Hz. The duration of each trial in the SEED was different, and the shortest was 185 s. According to the research in [48], when watching emotionally stimulating videos, the emotional activity of the subjects will first increase and then decrease, which means that the selection of the time period will affect the accuracy of the classification task. Therefore, this paper chooses the last 185 s of each trial to conduct emotion recognition research with reference to the method of [48]. For trials that are longer than 185 s, considering that it takes some time to elicit emotions, the last 185 s was selected. Finally, according to the recommended settings in [37], the data were segmented with a time window of 1 s without overlap.

3.3 Feature extraction

The frequency domain features suggested by the SEED and their combined features in space were used for this study. We extracted the power spectral density (PSD), differential entropy (DE), rational asymmetry (RASM), differential asymmetry (DASM), asymmetry (ASM), and differential causality (DCAU)[36, 37] from the five frequency bands (delta: 1–3 Hz, theta: 4–7 Hz, alpha: 8–13 Hz, beta: 14–30 Hz, and gamma: 31–50 Hz) of the EEG data as features. These features that were used for comparison tests are defined as follows:

PSD. PSD is a feature that describes the variation of signal power with frequency, which can be obtained by dividing the square of the amplitude-frequency characteristic after discrete Fourier transfer (DFT) by the length of the EEG sequence. It was calculated as follows:

$$PSD = \frac{|F(X)|^2}{N}$$

where X is the EEG signal of length N and $F(X)$ is the DFT of X. We calculated the PSD features in each EEG channel for 5 frequency bands in each sample. In the end, a 62 (channels)×185 (durations)×5 (frequency bands) feature vector can be obtained for each trial.

DE.

$$h(X) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) \mathrm{d}x = \frac{1}{2}\log\left(2\pi \mathrm{e}\sigma^2\right)$$

where X is a time series following a Gaussian distribution $N(\mu, \sigma^2)$. For a fixed-length EEG sequence, DE has been shown to be equal to the logarithmic energy spectrum (ES) of a specific frequency band[49], where ES is the average energy of the EEG signal in this frequency band. In this study, the DE features of the five previously mentioned frequency bands were calculated, and the final feature vector size for each trial was the same as that of PSD, which was 62×185×5.

RASM, DASM and ASM. Twenty-seven pairs of hemispherically asymmetric electrodes are given in Table 2. The difference in DE features on these electrode pairs is defined as DASM, the ratio is defined as RASM, and the combination of both is defined as ASM. RASM, DASM and ASM can be expressed as

$$DASM = h\left(X_i^{left}\right) - h\left(X_i^{right}\right)$$

$$RASM = \frac{h\left(X_i^{left}\right)}{h\left(X_i^{right}\right)}$$

$$ASM = \begin{bmatrix} DASM \\ RASM \end{bmatrix}$$

where i represents the label of the electrode pair in Table 2 and $h(X)$ is the same as defined in the DE part.

Table 2 The 27 pairs of hemispherical asymmetric electrodes

Pair number	Left	Right	Pair number	Left	Right
1	FP1	FP2	15	PO5	PO6
2	F1	F2	16	CB1	CB2
3	FC1	FC2	17	F5	F6
4	C1	C2	18	FC5	FC6
5	CP1	CP2	19	C5	C6
6	P1	P2	20	CP5	CP6
7	PO3	PO4	21	P5	P6
8	O1	O2	22	PO7	PO8
9	AF3	AF4	23	F7	F8
10	F3	F4	24	T7	T8
11	FC3	FC4	25	TP7	TP8
12	C3	C4	26	P7	P8
13	CP3	CP4	27	PO7	PO8
14	P3	P4			

In the end, the DASM and RASM feature vectors obtained from each trial have the same size, 27×185×5. The ASM feature represents the total asymmetric features, so its feature vector size is 54×185×5.

DCAU. Similar to RASM, the DCAU feature was defined as the ratio of the DE features on the 23 pairs of frontal-posterior electrodes shown in Table 3. DCAU can be expressed as

$$DCAU = \frac{h\left(X_i^{frontal}\right)}{h\left(X_i^{posterior}\right)}$$

where i is the pair number in Table 3 and $h(X)$ is the same as defined in the DE part. The final DCAU feature vector size of each trial is 23×185×5.

Table 3 The 23 pairs of frontal-posterior electrodes

Pair number	Frontal	Posterior	Pair number	Frontal	Posterior
1	FPZ	OZ	13	F6	P6
2	FP1	O1	14	F8	P8
3	FP2	O2	15	FT7	TP7
4	AF3	CB1	16	FC5	CP5
5	AF4	CB2	17	FC3	CP3
6	F7	P7	18	FC1	CP1
7	F5	P5	19	FCZ	CPZ
8	F3	P3	20	FC2	CP2
9	F1	P1	21	FC4	CP4
10	FZ	PZ	22	FC6	CP6
11	F2	P2	23	FT8	TP8
12	F4	P4			

3.4 Partition

After obtaining the features introduced in the previous section, we used two different partitioning strategies to divide them into training data and test data. Random division was used to verify the effectiveness of the EasyMKL-SVM method, while sequential division was used to simulate actual application situations.

Random division. Like general machine learning methods, one-fifth of the data from one subject were randomly selected for testing, and the remaining four-fifths of the data from the same experiment were used for model training.

Sequential division. One-fifth of the data from one subject were selected for testing; i.e., the data of nine adjacent trials in one subject were used as a test set, and the data of the remaining trials from the same subject were used for model training.

3.5 Kernel selection

We used ten polynomial kernels and ten radial basis function (RBF) kernels with different parameters to construct the kernel vector and learned the weight vector corresponding to the kernel vector on the training set by the EasyMKL method. The parameter of the polynomial kernel function is the degree of the highest order term, and we set it to an arithmetic sequence from 1 to 10. We set the parameter gamma of the RBF kernel to a geometric sequence from 10^{-5} to 10^4. Specifically, the basis kernel vector constructed in this paper contains {Poly (degree = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10); Rbf (gamma = 0.000 01, 0.000 1, 0.001, 0.01, 0.1, 1, 10, 100, 1 000, 10 000)}, a total of 20 commonly used and more effective kernel functions.

3.6 Classification

Through the kernel vector and its weight vector obtained by the EasyMKL algorithm, a combined kernel can be calculated and then combined with the SVM and the test set to complete the classification experiment. In this study, we conducted three-category (negative, neutral, and positive) emotion recognition experiments for each subject and performed five-fold cross-validation. The results of the classification experiments are shown in detail in Section 4.

4 Experiments and results

The experimental results were divided into two parts, random division and sequential division according to the different data partitioning strategies mentioned in Section 3.4. In the following, we will demonstrate the results.

4.1 Random division

First, we follow the traditional machine learning process to conduct classification experiments to verify the validity of the proposed EasyMKL-SVM method. Four-fifths of the data were randomly selected from each subject's data to train the model parameters, and the rest were input as a test set into the resulting model for classification. A five-fold cross-validation was then performed, and its average accuracy was used to evaluate the classification performance.

Fig. 4 shows the classification accuracies of the six features (ASM, DASM, DCAU, DE, PSD, and RASM) extracted on the five frequency bands mentioned earlier. It can be seen that the DE feature eventually achieved the best classification performance in all four frequency bands except delta. The highest accuracy of 99.72% was achieved for the DE feature in the gamma band on the 15th subject. After that, we used the DE feature as an example to compare the impact of different frequency bands on classification performance. The average classification accuracy of the DE features of the five frequency bands on 15 subjects is shown in DE of Fig. 4. We can see clearly that the features in the high frequency bands perform better than the features in the low frequency bands. Among them, DE features extracted from gamma bands have the highest average classification accuracy, reaching 92.25%.

As a control, the traditional single kernel SVM with an RBF kernel was used for the experiment. Fig. 5 shows the comparison between the traditional SVM and the EasyMKL-SVM method for the average classification accuracy of the DE features of the 15 subjects in five frequency bands. We can see that the EasyMKL-SVM proposed by us performs better than the traditional SVM, especially in the low-frequency DE features.

4.2 Sequential division

The random division method mentioned above is mainly used to verify the validity, which does not fit well with the practical application scenarios. We usually use the previously collected EEG data to train the model in practical applications and input and classify the newly collected data before performing subsequent operations. In other words, the test data are continuous, so we used the sequential division method to conduct experiments to simulate the actual application. We sequentially selected the data of three consecutive trials in an experiment to test the classification performance and the data of the remaining twelve trials from the same experiment to train the model. Then, we still performed a five-fold cross-validation and evaluated the performance with its average accuracy.

The content of Fig. 6 is similar to that of Fig. 4 and shows the classification results of the sequentially divided data. Similar to the result of the random division results, the DE features perform best in almost all five frequency bands. We can also see that the performance of the higher frequency bands is better than that of the lower bands. The highest average classification accuracy of 74.37% was provided by the DE features of the gamma band and was significantly lower than that of random division.

5 Discussions

From the above experimental results of classification using our method for different features in different frequency bands, we have some important findings. In this section, we will discuss these findings in detail.

This paper used two data partition methods, random division and sequential division, and their results are shown in Fig. 4 and Fig. 6, respectively. For randomly divided data, the average classification accuracy can reach 92.25%, while for sequentially divided data, the average accuracy can reach 74.37%. From Figs. 4 and 6, we can also see that DE features have the best classification per-

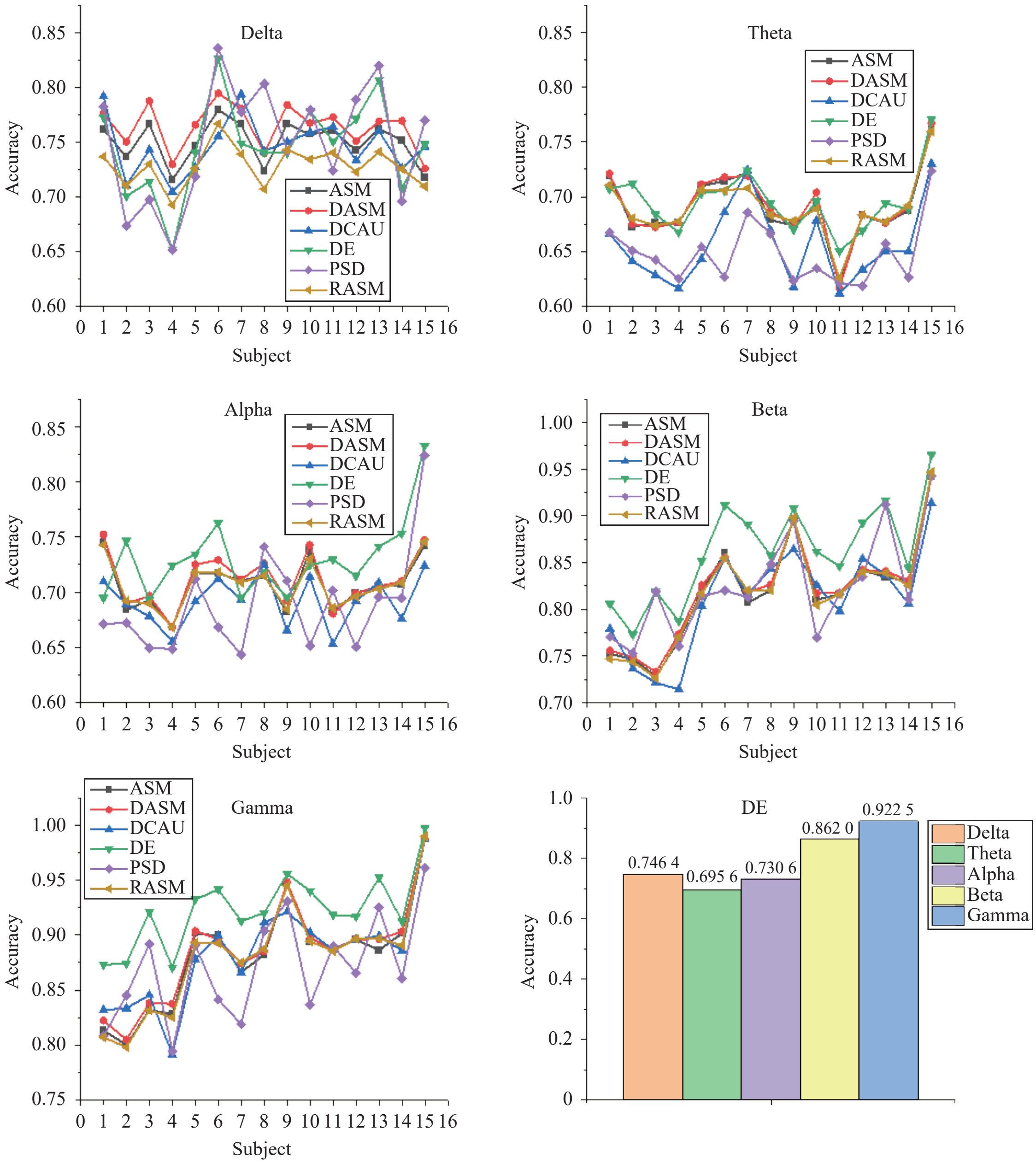


Fig. 4 Classification results of six features of the random data division method on five frequency bands

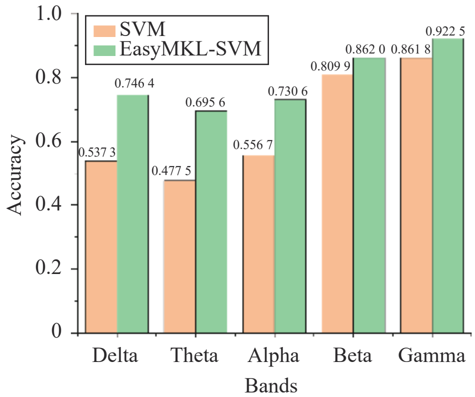


Fig. 5 Classification results of DE features in five bands on traditional SVM and EasyMKL-SVM

formance in both data division methods compared to other features. Compared with the PSD feature and the features that simply utilize the spatial position information of the electrode (DASM, RASM, ASM, and DCAU), the DE feature performs better in almost all five frequency bands. Therefore, performing simple mathematical operations on the features on the symmetrical electrodes cannot effectively integrate the spatial information of the EEG signals to obtain features with better performance. Further research on how to make better use of EEG spatial information is needed.

Fig. 5 shows the comparison results of the traditional SVM and EasyMKL-SVM on randomly divided data. The results show that the classification performance of

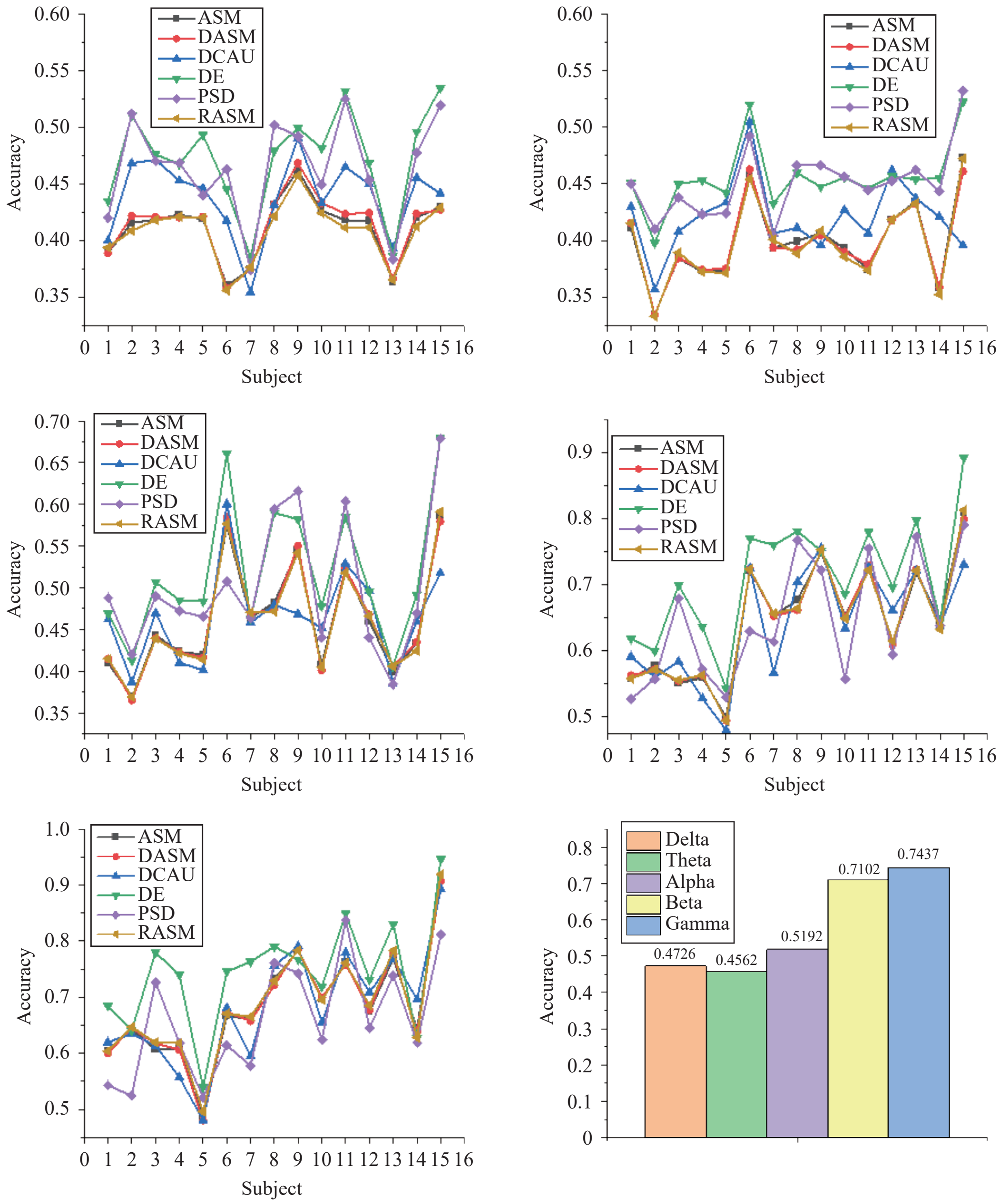


Fig. 6 Classification results of six features of the sequential data division method on five frequency bands

EasyMKL-SVM is better than that of the traditional SVM, and EasyMKL-SVM has more obvious advantages in the low frequency band features. This shows that EasyMKL-SVM is more adaptable to features because the method uses a combined kernel composed of multiple kernel functions. Based on this, an effective multidomain features fusion method for EEG emotion recognition may be found in the future. To more clearly demonstrate the superiority of the proposed EasyMKL-SVM method, we present the average classification accuracy of multiple single kernel methods in sequentially divided data in Table 4. Here, the kernel function is used for combination in the proposed EasyMKL-SVM method, and the accuracy is obtained by averaging a five-fold cross-validation using the DE features of the gamma bands. Similar to random division, the classification performance of the combined kernel outperforms the performance of any of the single kernels that comprise it. In summary, these results indicate that the proposed method is effective and that MKL deserves further investigation in EEG-based emotion recognition.

After that, we compared the classification performance of DE features in different frequency bands. Figs. 4 and 6 show the results of the two data division methods. It can be found that the performance of the beta band and gamma band features is better than other lower fre-

Table 4 Comparison of the classification performance of different kernel functions

Kernel function	Accuracy	Kernel function	Accuracy
Linear	72.04%	Poly2	73.20%
Rbf 0.0001	59.98%	Poly3	73.74%
Rbf 0.001	59.98%	Poly4	71.18%
Rbf 0.01	60%	Poly5	70.43%
Rbf 0.1	65.6%	Poly6	70.39%
Rbf1	73.02%	Poly7	70.35%
Rbf10	72.93%	Poly8	70.33%
Rbf100	48.98%	Poly9	70.33%
Rbf1000	33.37%	Poly10	70.32%
Poly1	72.88%	Combined kernel	**74.37%**

quency band features, which indicated that compared with lower frequency bands, the information in higher frequency bands contributed more to emotion recognition. The effectiveness of the beta and gamma bands in emotion recognition was also proven in [50, 51].

It is worth noting that the classification accuracy differs significantly under the two data division methods. For the random division method, the highest average classification accuracy reached 92.25%, while the sequential division method was only 74%. After analysis, the possible reasons are as follows. Since the experimental sample data are divided by manual segmentation, each trial will obtain 185 samples. Under the random division method, a part of the data in almost every trial will be divided into the training set, which may cause data leakage and lead to a certain degree of overfitting, making the model look accurate, while the performance in actual application will not be consistent. Therefore, the results under randomly divided data are only useful for the comparison of the performance between different methods, and the evaluation of the actual application of the method should be based on the results of sequentially divided data. This suggests that emotional EEG signals are complicated. Even if the same emotion is induced by different movie clips in the same experiment, different EEG signals may be generated due to different audio-visual stimuli. This shows that further research is needed on essential features that can characterize emotions to eliminate the influence of other information in EEG signals.

Since the time period division and selection method used in this paper referred to [48], it was compared with its results. Qing et al.[48] conducted emotion recognition experiments on the SEED using DE features and a soft-voting classifier and achieved a three-category accuracy of 71.41% for positive, neutral, and negative emotions when using the last 185 s data of each trial, while the proposed EasyMKL-SVM method in this paper achieved an average classification accuracy of 74.37%. This shows that our method has better performance, and the MKL method has certain research value in the field of EEG emotion recognition.

6 Conclusions

In summary, the EasyMKL-SVM method we proposed has better performance and better adaptability to different features than traditional SVM and other methods. Through experiments, we found that simple mathematical operations on the features on the symmetrical electrodes cannot effectively fuse the spatial information of the EEG signals to achieve better performance. It was confirmed that information in higher frequency bands contributed more to emotion recognition based on EEG signals. In the future, we will continue to study the EEG-based emotion recognition method based on MKL and devote ourselves to researching a method that can effectively fuse multidomain features to extract the essential characteristics of emotional EEG signals.

Acknowledgements

This work was supported by National Natural Science Foundation of China (No. 62176054), and University Synergy Innovation Program of Anhui Province, China (No. GXXT-2020-015).

References

[1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, J. G. Taylor. Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, vol. 18, no. 1, pp. 32–80, 2001. DOI: 10.1109/79.911197.

[2] N. Mehendale. Facial emotion recognition using convolutional neural networks (FERC). *SN Applied Sciences*, vol. 2, no. 3, Article number 446, 2020. DOI: 10.1007/s42452-020-2234-1.

[3] S. Minaee, M. Minaei, A. Abdolrashidi. Deep-emotion: Facial expression recognition using attentional convolutional network. *Sensors*, vol. 21, no. 9, Article number 3046, 2021. DOI: 10.3390/s21093046.

[4] D. M. Schuller, B. W. Schuller. A review on five recent and near-future developments in computational processing of emotion in the human voice. *Emotion Review*, vol. 13, no. 1, pp. 44–50, 2021. DOI: 10.1177/1754073919898526.

[5] Y. M. Huang, K. X. Tian, A. Wu, G. B. Zhang. Feature fusion methods research based on deep belief networks for speech emotion recognition under noise condition. *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 5, pp. 1787–1798, 2019. DOI: 10.1007/s12652-017-0644-8.

[6] J. H. Tao, J. Huang, Y. Li, Z. Lian, M. Y. Niu. Correction

to: Semi-supervised ladder networks for speech emotion recognition. *International Journal of Automation and Computing*, vol. 18, no. 4, Article number 680, 2021. DOI: 10.1007/s11633-019-1215-6.

[7] Y. Q. Yin, X. W. Zheng, B. Hu, Y. Zhang, X. C. Cui. EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. *Applied Soft Computing*, vol. 100, Article number 106954, 2021. DOI: 10.1016/j.asoc.2020.106954.

[8] P. Sarkar, A. Etemad. Self-supervised learning for ecg-based emotion recognition. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, Barcelona, Spain, pp. 3217–3221, 2020. DOI: 10.1109/ICASSP40776.2020.9053985.

[9] V. Kehri, R. Ingle, S. Patil, R. N. Awale. Analysis of facial EMG signal for emotion recognition using wavelet packet transform and SVM. *Machine Intelligence and Signal Analysis*, M. Tanveer, R. B. Pachori, Eds., Singapore: Springer, pp. 247–257, 2019. DOI: 10.1007/978-981-13-0923-6_21.

[10] Q. Zhang, X. X. Chen, Q. Y. Zhan, T. Yang, S. H. Xia. Respiration-based emotion recognition with deep learning. *Computers in Industry*, vol. 92–93, pp. 84–90, 2017. DOI: 10.1016/j.compind.2017.04.005.

[11] L. Shu, Y. Yu, W. Z. Chen, H. Q. Hua, Q. Li, J. X. Jin, X. M. Xu. Wearable emotion recognition using heart rate data from a smart bracelet. *Sensors*, vol. 20, no. 3, Article number 718, 2020. DOI: 10.3390/s20030718.

[12] D. Ayata, Y. Yaslan, M. Kamaşak. Emotion recognition via random forest and galvanic skin response: Comparison of time based feature sets, window sizes and wavelet approaches. In *Proceedings of Medical Technologies National Congress*, IEEE, Antalya, Turkey, 2016. DOI: 10.1109/TIPTEKNO.2016.7863130.

[13] S. Huang, W. Shao, M. L. Wang, D. Q. Zhang. fMRI-based decoding of visual information from human brain activity: A brief review. *International Journal of Automation and Computing*, vol. 18, no. 2, pp. 170–184, 2021. DOI: 10.1007/s11633-020-1263-y.

[14] L. M. Ward. Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences*, vol. 7, no. 12, pp. 553–559, 2003. DOI: 10.1016/j.tics.2003.10.012.

[15] J. A. Coan, J. J. B. Allen. Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological Psychology*, vol. 67, no. 1–2, pp. 7–50, 2004. DOI: 10.1016/j.biopsycho.2004.03.002.

[16] J. Jin, Z. M. Chen, R. Xu, Y. Y. Miao, X. Y. Wang, T. P. Jung. Developing a novel tactile P300 brain-computer interface with a cheeks-stim paradigm. *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 9, pp. 2585–2593, 2020. DOI: 10.1109/TBME.2020.2965178.

[17] Y. Yu, Y. D. Liu, E. W. Yin, J. Jiang, Z. T. Zhou, D. W. Hu. An asynchronous hybrid spelling approach based on EEG-EOG signals for Chinese character input. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 6, pp. 1292–1302, 2019. DOI: 10.1109/TNSRE.2019.2914916.

[18] Z. Mohammadi, J. Frounchi, M. Amiri. Wavelet-based emotion recognition system using EEG signal. *Neural Computing and Applications*, vol. 28, no. 8, pp. 1985–1990, 2017. DOI: 10.1007/s00521-015-2149-8.

[19] M. Degirmenci, M. A. Ozdemir, R. Sadighzadeh, A. Akan. Emotion recognition from EEG signals by using empirical mode decomposition. In *Proceedings of Medical Technologies National Congress*, IEEE, Magusa, Cyprus, 2018. DOI: 10.1109/TIPTEKNO.2018.8597061.

[20] Z. Yin, L. Liu, J. N. Chen, B. X. Zhao, Y. X. Wang. Locally robust EEG feature selection for individual-independent emotion recognition. *Expert Systems with Applications*, vol. 162, Article number 113768, 2020. DOI: 10.1016/j.eswa.2020.113768.

[21] T. F. Song, W. M. Zheng, P. Song, Z. Cui. EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 532–541, 2020. DOI: 10.1109/TAFFC.2018.2817622.

[22] H. Chao, L. Dong, Y. L. Liu, B. Y. Lu. Emotion recognition from multiband EEG signals using CapsNet. *Sensors*, vol. 19, no. 9, Article number 2212, 2019. DOI: 10.3390/s19092212.

[23] Y. Cimtay, E. Ekmekcioglu. Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset EEG emotion recognition. *Sensors*, vol. 20, no. 7, Article number 2034, 2020. DOI: 10.3390/s20072034.

[24] E. Iáñez, J. M. Azorín, A. Úbeda, E. Fernández, J. L. Sirvent. LDA-based classifiers for a mental tasks-based brain-computer interface. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, Istanbul, Turkey, pp. 546–551, 2010. DOI: 10.1109/ICSMC.2010.5642018.

[25] L. Guo, Y. X. Wu, L. Zhao, T. Cao, W. L. Yan, X. Q. Shen. Classification of mental task from EEG signals using immune feature weighted support vector machines. *IEEE Transactions on Magnetics*, vol. 47, no. 5, pp. 866–869, 2011. DOI: 10.1109/TMAG.2010.2072775.

[26] M. J. Abdi, S. M. Hosseini, M. Rezghi. A novel weighted support vector machine based on particle swarm optimization for gene selection and tumor classification. *Computational and Mathematical Methods in Medicine*, vol. 2012, Article number 320698, 2012. DOI: 10.1155/2012/320698.

[27] A. Rakotomamonjy, F. R. Bach, S. Canu, Y. Grandvalet. SimpleMKL. *Journal of Machine Learning Research*, vol. 9, pp. 2491–2521, 2008.

[28] M. Varma, B. R. Babu. More generality in efficient multiple kernel learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ACM, Montreal, Canada, pp. 1065–1072, 2009. DOI: 10.1145/1553374.1553510.

[29] M. Kloft, U. Brefeld, S. Sonnenburg, A. Zien. Non-sparse regularization and efficient training with multiple kernels. [Online], Available: https://arxiv.org/abs/1003.0079v1, 2010.

[30] Z. L. Xu, R. Jin, H. Q. Yang, I. King, M. R. Lyu. Simple

and efficient multiple kernel learning by group lasso. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, Haifa, Israel, pp. 1175–1182, 2010.

[31] F. Aiolli, M. Donini. EasyMKL: A scalable multiple kernel learning algorithm. *Neurocomputing*, vol. 169, pp. 215–224, 2015. DOI: 10.1016/j.neucom.2014.11.078.

[32] W. Samek, A. Binder, K. R. Müller. Multiple kernel learning for brain-computer interfacing. In *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Osaka, Japan, pp. 7048–7051, 2013. DOI: 10.1109/EMBC.2013.6611181.

[33] X. O. Li, X. Chen, Y. N. Yan, W. S. Wei, Z. J. Wang. Classification of EEG signals using a multiple kernel learning support vector machine. *Sensors*, vol. 14, no. 7, pp. 12784–12802, 2014. DOI: 10.3390/s140712784.

[34] C. A. Frantzidis, C. Bratsas, C. L. Papadelis, E. Konstantinidis, C. Pappas, P. D. Bamidis. Toward emotion aware computing: An integrated approach using multichannel neurophysiological recordings and affective visual stimuli. *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 589–597, 2010. DOI: 10.1109/TITB.2010.2041553.

[35] P. J. Lang, M. M. Bradley, B. N. Cuthbert. International Affective Picture System (IAPS): Technical Manual and Affective Ratings, Technical Report No. 3, NIMH Center for the Study of Emotion and Attention, USA, pp. 39–58, 1997.

[36] R. N. Duan, J. Y. Zhu, B. L. Lu. Differential entropy feature for EEG-based emotion classification. In *Proceedings of the 6th International IEEE/EMBS Conference on Neural Engineering*, IEEE, San Diego, USA, pp. 81–84, 2013. DOI: 10.1109/NER.2013.6695876.

[37] W. L. Zheng, B. L. Lu. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 162–175, 2015. DOI: 10.1109/TAMD.2015.2431497.

[38] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras. DEAP: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012. DOI: 10.1109/T-AFFC.2011.15.

[39] X. Chen, X. Y. Xu, A. P. Liu, S. Lee, X. Chen, X. Zhang, M. J. McKeown, Z. J. Wang. Removal of muscle artifacts from the EEG: A review and recommendations. *IEEE Sensors Journal*, vol. 19, no. 14, pp. 5353–5368, 2019. DOI: 10.1109/JSEN.2019.2906572.

[40] R. Jenke, A. Peer, M. Buss. Feature extraction and selection for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 327–339, 2014. DOI: 10.1109/TAFFC.2014.2339834.

[41] C. Zhang, H. Wang, R. R. Fu. Automated detection of driver fatigue based on entropy and complexity measures. *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 168–177, 2014. DOI: 10.1109/TITS.2013.2275192.

[42] V. S. Vijith, J. E. Jacob, T. Iype, K. Gopakumar, D. G. Yohannan. Epileptic seizure detection using non linear analysis of EEG. In *Proceedings of International Conference on Inventive Computation Technologies*, IEEE, Coimbatore, India, 2016. DOI: 10.1109/INVENTIVE.2016.7830193.

[43] R. C. Guido. A tutorial review on entropy-based handcrafted feature extraction for information fusion. *Information Fusion*, vol. 41, pp. 161–175, 2018. DOI: 10.1016/j.inffus.2017.09.006.

[44] K. H. Kim, S. W. Bang, S. R. Kim. Emotion recognition system using short-term monitoring of physiological signals. *Medical and Biological Engineering and Computing*, vol. 42, no. 3, pp. 419–427, 2004. DOI: 10.1007/BF02344719.

[45] X. W. Wang, D. Nie, B. L. Lu. Emotional state classification from EEG data using machine learning approach. *Neurocomputing*, vol. 129, pp. 94–106, 2014. DOI: 10.1016/j.neucom.2013.06.046.

[46] K. R. Muller, S. Mika, G. Ratsch, K. Tsuda, B. Scholkopf. An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, vol. 12, no. 2, pp. 181–201, 2001. DOI: 10.1109/72.914517.

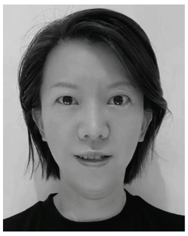
[47] X. Chen, Z. J. Wang. Pattern recognition of number gestures based on a wireless surface EMG system. *Biomedical Signal Processing and Control*, vol. 8, no. 2, pp. 184–192, 2013. DOI: 10.1016/j.bspc.2012.08.005.

[48] C. M. Qing, R. Qiao, X. M. Xu, Y. Q. Cheng. Interpretable emotion recognition using EEG signals. *IEEE Access*, vol. 7, pp. 94160–94170, 2019. DOI: 10.1109/ACCESS.2019.2928691.

[49] L. C. Shi, Y. Y. Jiao, B. L. Lu. Differential entropy feature for EEG-based vigilance estimation. In *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Osaka, Japan, pp. 6627–6630, 2013. DOI: 10.1109/EMBC.2013.6611075.

[50] M. Soleymani, M. Pantic, T. Pun. Multimodal emotion recognition in response to videos. *IEEE Transactions on Affective Computing*, vol. 3, no. 2, pp. 211–223, 2012. DOI: 10.1109/T-AFFC.2011.37.

[51] W. L. Zheng, J. Y. Zhu, B. L. Lu. Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 417–429, 2019. DOI: 10.1109/TAFFC.2017.2712143.



Qian Cai received the B. Sc. and M. Sc. degrees in mathematics from Anhui University, China in 2000 and 2003, respectively. Currently, she is with School of Statistics and Data Science, Nanjing Audit University, China.

Her research interests include statistical pattern recognition and data science.

E-mail: caiq@nau.edu.cn

ORCID iD: 0000-0001-7255-6321

Guo-Chong Cui received the B.Sc. degree in biomedical engineering from Yanshan University, China in 2019. He is a master student in biomedical engineering at Department of Biomedical Engineering, School of Biological Science & Medical Engineering, Southeast University, China.

His research interests include EEG-based emotion recognition, machine learning and neural information processing.

E-mail: gc_cui@seu.edu.cn (Corresponding author)

ORCID iD: 0000-0002-1929-0154

Hai-Xian Wang received the B.Sc. and M.Sc. degrees in statistics and the Ph.D. degree in computer science from Anhui University, China in 1999, 2002 and 2005, respectively. Currently, he is with Key Laboratory of Child Development and Learning Science of Ministry of Education, School of Biological Science & Medical Engineering, Southeast University, China.

His research interests include biomedical signal processing, brain-computer interfaces, and machine learning.

E-mail: hxwang@seu.edu.cn (Corresponding author)

ORCID iD: 0000-0001-8220-9737