

## *Retraction*

# **Retracted: Mental and Emotional Recognition of College Students Based on Brain Signal Features and Data Mining**

### **Security and Communication Networks**

Received 27 June 2023; Accepted 27 June 2023; Published 28 June 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] S. An and Z. Yu, "Mental and Emotional Recognition of College Students Based on Brain Signal Features and Data Mining," *Security and Communication Networks*, vol. 2022, Article ID 4198353, 10 pages, 2022.

## Research Article

# Mental and Emotional Recognition of College Students Based on Brain Signal Features and Data Mining

Shaobo An<sup>1</sup> and Zhifen Yu<sup>2</sup> 

<sup>1</sup>School of Mathematics and Statistics, Central South University, HuNan, ChangSha 410083, China

<sup>2</sup>Middle School Affiliated to Central South University, HuNan, ChangSha 410005, China

Correspondence should be addressed to Zhifen Yu; 205138@csu.edu.cn

Received 1 December 2021; Revised 20 December 2021; Accepted 17 January 2022; Published 8 February 2022

Academic Editor: Chin-Ling Chen

Copyright © 2022 Shaobo An and Zhifen Yu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays, people pay more and more attention to the psychological situation of college students. Using data mining technology to model and analyze the collected psychological data of college students is a research hotspot in psychology and computer science. In addition, the essence of human emotional change is the higher nervous activity in the cerebral cortex. Electroencephalography (EEG) has become an important feature signal for emotion recognition because of its high time resolution and portability and practicality. Therefore, to solve the problem that the accuracy and generalization of the existing research models are not ideal, a method of college students' psychological emotion recognition based on EEG signal features and data mining is proposed. Firstly, a feature selection method based on sparse learning is used to find out a few features from the high-dimensional feature space that contribute greatly to the reconstruction of category information so as to quickly acquire a few key emotion-related features. Then, the entropy-weighted clustering algorithm is combined with sparse learning feature selection, and the local structure of heterogeneous data is divided. Experimental results show that, compared with traditional methods, the proposed method has stronger applicability and higher accuracy of five categories of emotions, which provides a valuable reference for the evaluation of depression and anxiety of college students based on brain signal characteristics.

## 1. Introduction

College students' mental health and happiness are not only vital to their own growth but also have an important contribution to society. There is evidence that the mental health of most college students is vulnerable at present. Many studies show that college students all over the world have a high rate of stress and depression. Teenagers' psychological problems are increasing gradually. More than 20% of young people have experienced psychological barriers at some point in their lives.

In recent years, with the development of medical imaging technology, emotion recognition based on central nervous system signals has become a research hotspot. The emotion recognition method based on the central nervous system refers to emotion recognition by analyzing the differences of signals sent by the brain in different emotional

states [1–4]. At present, the recognition method based on the central nervous system mainly uses EEG. The emotion recognition method based on EEG signal has become one of the hot research topics in the field of emotion computing, which has attracted the attention of many researchers at home and abroad [5–8]. Especially in recent years, with the rapid development of machine learning technology, combining EEG signals with machine learning technology to identify and analyze emotions is becoming the mainstream research method in this research field.

At present, the feature extraction methods of emotional EEG signals can be mainly divided into three kinds [9, 10]: first, feature extraction methods based on the time domain, frequency domain, and time-frequency domain; second, the feature extraction method using potential technology; the third is to use the nonlinear dynamics of the feature extraction method. Regarding the selection of EEG signals,

Jing et al. [11] pointed out that due to individual differences in emotions, different subjects may have different responses to different emotional stimulation materials. Therefore, data with large differences between samples should be excluded to enhance the accuracy and stability of model recognition. According to the research by Zheng and Lu [12], differences in EEG signals of different emotions in different bands are also relatively significant, and the band separation of EEG signals is conducive to improving the accuracy of emotion recognition. The purpose of using feature selection for emotion recognition is to find the representative and strong correlation features so that the selected features can be better studied and classified and improve the accuracy of emotion recognition.

Generally, there are two problems to be considered when constructing an emotion recognition model; one is classification accuracy, and the other is generalization ability. Especially for EEG signals, which have great differences between samples, we should pay more attention to these two problems [13–15]. At present, the algorithms selected by most related literature to build emotion recognition models are basically the support vector machine (SVM), decision tree, random forest, and neural network. For example, Mitsukura [16] segmented the EEG data of one minute with sliding windows and calculated 9 features of energy, energy ratio, and spectral entropy in alpha and beta bands on each segment. The features are input into the random forest classifier, and a good four classification accuracy of an emotional state is obtained.

Emotion recognition methods based on EEG feature signals and traditional machine learning technology [17–20] mentioned above can achieve relatively good results in general, but their shortcomings are also obvious, which are mainly reflected in two aspects: (1) most feature extraction methods of EEG signals are finely divided in all time periods and frequency domains, so the dimension of features will become very high, which makes rapid feature selection an urgent problem to be solved; (2) in many cases, the traditional emotion recognition model cannot fully reflect the hidden relationship between EEG signals and emotional states. That is to say, the more abstract, deeper, and more discriminating relationship between the EEG signal and emotional state cannot be obtained only by isomorphic data analysis, so it is difficult to make a breakthrough in classification accuracy.

Therefore, in order to solve the above two problems, this paper proposes a method of college students' psychological emotion recognition based on EEG signal features and data mining. The main work includes the following: (1) introducing the idea of sparse learning, a feature selection method of the EEG signal based on sparse learning is proposed, which can quickly, accurately, and effectively select emotion-related features from high-dimensional features and provides strong support for the subsequent construction of the emotion recognition model; (2) combining the entropy-weighted clustering algorithm with the features obtained after sparse learning, the classification model of depression and anxiety is established. By dividing the local structure of heterogeneous data, and marking and sorting the

importance of features, the accuracy and generalization of clustering are provided. Finally, aiming at the positive and negative emotion classification task, the accuracy rate of 68.35% is achieved on the experimental dataset, which proves the effectiveness of the method proposed in this paper.

The rest of the paper is organized as follows. In Section 2, a processing mode of an input EEG signal is studied in detail, while Section 3 provides the detailed feature selection method. Section 4 provides the detailed emotional identification method. Section 5 provides detailed results and discussion. Finally, the paper is concluded in Section 6.

## 2. Artifact Removal of EEG Signal

As a physiological signal, EEG reflects the information of potential changes caused by complex nerve discharges in the brain [21]. EEG artifacts seriously affect the acquisition of signals related to neural activity. Fast and efficient removal of artifacts is the premise of the subsequent emotional state interpretation. We have adopted the EEG artifact removal method based on prior information [22, 23], and the process is shown in Figure 1.

Firstly, by adding a small amount of specific artifact prior information, the problem of mismatch of artifact prior information is overcome. Then, wavelet independent component analysis (WICA) is used to separate artifacts. As a combination of wavelet transform and independent component analysis, the WICA method has the advantages of multiresolution and multidimensional analysis. Its basic idea is to map signals to the wavelet domain and then carry out independent component analysis in the wavelet domain. Based on correlation discriminant analysis, the artifact components can be automatically identified and removed. Only a small amount of artifact prior information is used to effectively separate and remove the corresponding artifact components. In the range of 1 Hz to 50 Hz, the power spectrum energy is obtained with 10 Hz as the bandwidth.

## 3. Feature Selection Method of EEG Signal Based on Sparse Learning

*3.1. Algorithm Principle.* In the traditional filtering feature selection method, when the dimension of features becomes very high, the amount of calculation will become very time-consuming. In recent years, the sparse learning algorithm [24] has shown a good application prospect in the field of signal analysis. Through proper modeling, sparse learning can be well applied to feature selection. Then, by using the fast optimization method in sparse model solving, the operation speed of feature selection in high-dimensional feature space can be effectively improved.

The purpose of sparse learning is a sparse decomposition of signals, and a few optimal signals can be selected from the dictionary to reconstruct the original signals. After multiplying high-dimensional features by sparse weighting coefficients, the process of feature selection by sparse learning can be regarded as the process of approaching category labels. The final distribution of sparse weighting coefficients

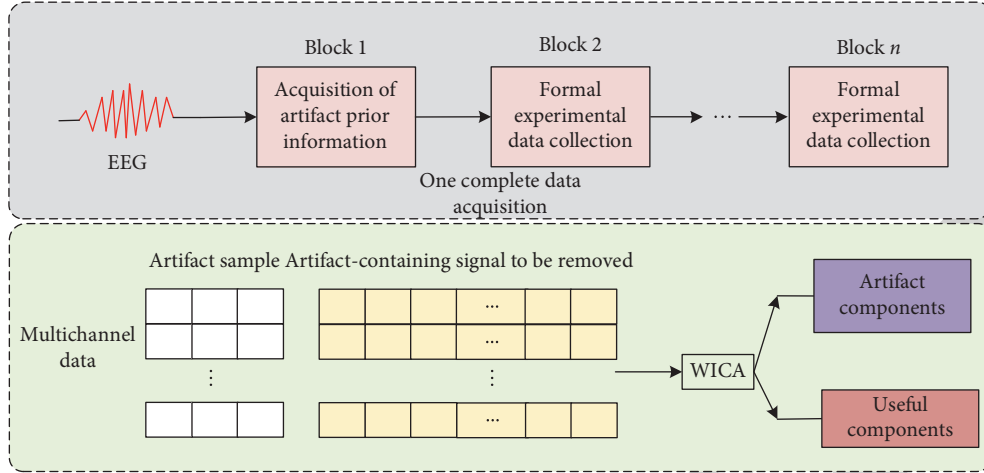


FIGURE 1: EEG artifact removal based on prior information.

shows the distribution of selected features. The overall method idea is shown in Figure 2.

The process of reconstructing category information by using high-dimensional feature space can be expressed by the following formula:

$$s_i = \mathbf{D}\alpha_i, \quad (1)$$

where  $s_i$  denotes a column vector related to class information,  $\mathbf{D} = (v_1, v_2, \dots, v_n)$  denotes a normalized feature

vector group, and  $\alpha_i$  denotes a sparse weighting coefficient. When the feature dimension is very high,  $\mathbf{D} \in \mathbf{R}^{m \times n}$  can be regarded as an overcomplete dictionary matrix [25]. If the information of  $s$  is known, the problem of feature selection, i.e., solving  $\alpha_i \in \mathbf{R}^n$ , becomes a problem of sparse learning.

Firstly, the category label matrix  $\mathbf{S} = (s_1, s_2, \dots, s_k)$  is constructed;  $k$  is the number of sample categories, where  $s_i = (s_{i1}, s_{i2}, \dots, s_{im})^T$ . The specific expression is as follows:

$$s_{il} = \begin{cases} 1, & l\text{-th sample belongs to class } i, \\ 0, & l\text{-th sample does not belong to class } i \end{cases}, \quad (l = 1, 2, \dots, m) \quad (2)$$

Each column vector  $s_i$  in  $\mathbf{S}$  can be sparsely represented by a feature matrix  $\mathbf{D}$ . And finally, a sparse matrix  $\mathbf{A} = (\alpha_1, \alpha_2, \dots, \alpha_k)$  is formed by solving the sparse coefficient  $\alpha_i$ . The  $j$ -th row of the matrix  $\mathbf{A}$  represents the contribution of the  $j$ -th feature to the reconstruction of the class information of each class of samples, and the sum of the absolute values of the  $j$ -th feature can be considered as the contribution to distinguishing the  $k$ -class samples. The feature importance score  $F$  is defined as follows:

$$F_j = \sum_{l=1}^k |\alpha_{jl}|. \quad (3)$$

As the importance score of the  $j$ -th feature for classification, the larger the  $F$  value, the greater the role of the corresponding feature for classification. Then, the features can be sorted according to the calculated  $F$  value, and the features with specified dimensions can be selected.

Sparse representation can be regarded as a problem of sparse decomposition of signals in an overcomplete dictionary. The basic idea is to select some atoms in the overcomplete dictionary ( $\Phi$ ) to represent the original signal  $y$  with optimal linear weighting. This is a process of sparse weighted approximation, which can be described by the following formula:

$$\hat{\mathbf{a}} = \operatorname{argmin}_{\mathbf{a}} \|\Phi \mathbf{a} - \mathbf{y}\|_2 + \lambda \|\mathbf{a}\|_0, \quad (4)$$

where  $\lambda$  is the regularization coefficient, which determines the degree of sparsity of the algorithm.  $\hat{\mathbf{a}}$  is the required sparse representation coefficient.

**3.2. Solution of Sparse Model.** There are two kinds of commonly used sparse learning algorithms, namely, the relaxation algorithm and the greedy algorithm. The relaxation algorithm makes the problem gradually solvable by relaxing the constraints. The disadvantage of this method is that the complexity of the algorithm is high, and the complexity of the algorithm largely depends on the atomic library itself, which needs to be stored in the calculation process, so the requirement for storage space is relatively high. The greedy algorithm is to find out the supporting atoms of the signal to be reconstructed in turn through the idea of iteration. Its outstanding advantage is its fast operation speed, which has attracted extensive attention. Because the "backtracking" idea of the regularized orthogonal matching pursuit (ROMP) algorithm can better select the supporting atoms with global optimality, this paper uses this algorithm to solve the sparse model in feature selection.

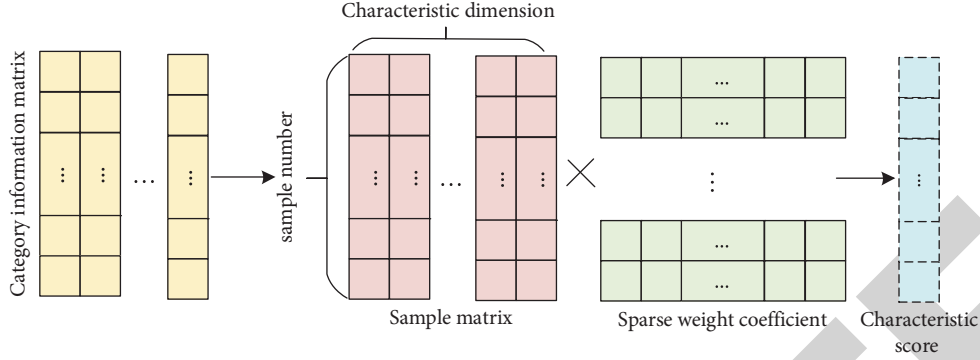


FIGURE 2: Feature selection based on sparse learning.

ROMP algorithm optimizes the selection mechanism of atoms. By adding the regularization step to realize the secondary screening of atoms, the global optimality of supporting atoms can be guaranteed to a great extent. The specific steps of the algorithm are as follows:

- (1) The initial residual  $r$  is set as signal  $y$ . Firstly,  $\Phi^T r$  is calculated, and the largest  $n$  nonzero atoms are selected in  $\Phi^T r$  to form a set  $J$ .
- (2) In the set  $J$ , the atoms are grouped according to the inner product  $\mathbf{u}$  of atoms and residuals so that any two inner products  $u_i$  and  $u_j$  in each group satisfy  $\|u_i\| \leq 2\|u_j\|$ . The set with the largest energy is selected as  $J_0$ .
- (3) the atoms in  $J_0$  are added into the supporting atom set  $X$ , the decomposition coefficient  $\hat{\mathbf{a}}$  is calculated by using the least square method on the supporting atom set, and the residual error  $\mathbf{r} = \mathbf{y} - X\hat{\mathbf{a}}$  is updated. Returning to the first step, the residuals are further decomposed until either the residuals are negligible or the number of supporting atom sets reaches a predetermined number.

#### 4. Emotional Identification Method Based on Entropy Weighted Clustering

The research steps of emotion recognition based on EEG mainly include emotion induction, EEG signal acquisition, EEG signal pretreatment, feature extraction and selection, and emotion pattern learning and classification. Each step is essential for the research of EEG-based emotion recognition methods. The process of emotion recognition based on EEG is shown in Figure 3.

**4.1. Representation of Sparse Learning Features.** In order to reduce the spatial dimension of the EEG feature data obtained by solving the sparse model while retaining the features that are helpful for clustering, we use the sparse fractional feature representation. First, the input dataset is transformed with the objective of minimizing the L1 norm, assuming the set  $\{x_i\}_{i=1}^n$ ,  $x_i \in R^d$ . Let  $X = [x_1, x_2, \dots, x_n] \in R^{d \times n}$  and the sparse fractional coefficient for  $x$  be solved, as shown in the following formula:

$$\min_{S_i} \|S_i\|_1, \quad (5)$$

$$\text{s.t. } x_i = X' s_i,$$

where  $X'$  is the matrix without column  $i$ .

The objective function of the sparse fractional coefficient is

$$S(r) = \frac{\sum_{i=1}^n (x_{ir} - (Xs_i)_r)^2}{\text{Var}(X(r))}, \quad (6)$$

where  $\text{Var}(X(r))$  represents the accumulated difference of the  $r$  dimension feature.

Finally, the feature with the minimum objective function value  $S(r)$  is selected as the clustering feature representation in this paper so that the unnecessary clustering features are removed.

**4.2. Entropy Weighted Clustering Algorithm.** Data mining refers to the process of revealing the implicit and potential value information from a large number of fuzzy, incomplete, and noisy complex data [26]. Among them, clustering analysis is a widely used data mining method at present, which can be regarded as a process of dividing data object sets. Recently, Fathian and Jafarianmoghaddam [27] proposed an optimization scheme for ad hoc networks based on the entropy-weighted clustering algorithm. Entropy is a measure of uncertainty that arises from the measurement of "disorder" in a physical thermodynamic system. After selecting appropriate features as the representation of clustering features, the clustering analysis process is started. For the construction of the emotional tendency recognition model, the main purpose is to calculate the clustering center matrix so as to get the membership of EEG signals to different emotional categories.

Therefore, if the object subjected to entropy-weighted clustering is set as  $X = \{x_1, x_2, \dots, x_N\} \subset R^D$ , and the degree of membership expression is set as  $u_{ij}^{(1)}$ , the clustering center matrix is expressed as

$$v_{ik} = \frac{\sum_{j=1}^n u_{ij}^2 x_{jk}}{\sum_{j=1}^n u_{ij}^2}. \quad (7)$$

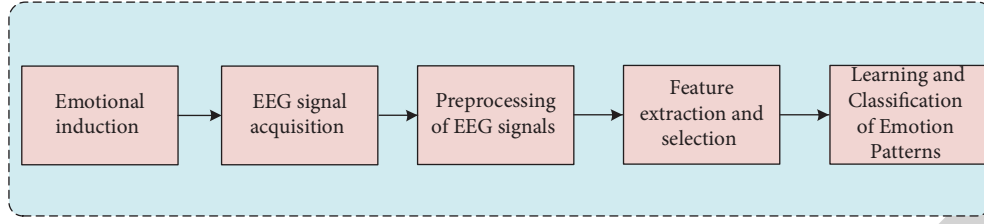


FIGURE 3: Process of emotion recognition based on EEG.

The datasets  $U = [u_1, u_2, \dots, u_n]$  and  $V = [v_1, v_2, \dots, v_n]$  are given, and the calculation of  $w_{ik}$  is set as

$$w_{ik} = \frac{\exp\left(-\sum_{j=1}^n u_{ij}^2 (x_{jk} - v_{ik})^2 / \gamma\right)}{\sum_{s=1}^d \exp\left(-\sum_{j=1}^n u_{ij}^2 (x_{js} - v_{is})^2 / \gamma\right)}. \quad (8)$$

Combining the above formula with the objective clustering function (6), the following formula is obtained:

$$\begin{aligned} \psi(w_{ik}) = & \sum_{i=1}^c \sum_{j=1}^n u_{ij}^2 \sum_{k=1}^d w_{ik} (x_{jk} - v_{ik})^2 + \gamma \sum_{i=1}^c \sum_{j=1}^n w_{ik} \log w_{ik} \\ & - \sum_{i=1}^c \lambda_i^w \left( \sum_{k=1}^d w_{ik} - 1 \right). \end{aligned} \quad (9)$$

The partial derivatives of formula (9) are calculated for  $w_{ik}$  and  $\lambda_i^w$ , respectively, and the result is made equal to 0; then,

$$\frac{\partial \psi(w_{ik})}{\partial w_{ik}} = \sum_{j=1}^n u_{ij}^2 (x_{jk} - v_{ik})^2 + \gamma (\log w_{ik} + 1) - \lambda_i^w = 0. \quad (10)$$

$$\frac{\partial \psi(w_{ik})}{\partial \lambda_i^w} = \sum_{k=1}^d w_{ik} - 1 = 0. \quad (11)$$

Combining formula (10) and formula (11), we can get

$$w_{ik} = \frac{\exp\left(-\sum_{j=1}^n u_{ij}^2 (x_{jk} - v_{ik})^2 / \gamma\right)}{\sum_{s=1}^d \exp\left(-\sum_{j=1}^n u_{ij}^2 (x_{js} - v_{is})^2 / \gamma\right)}. \quad (12)$$

Therefore, the above formula (12) is a necessary condition of formula (5). Then, the entropy-weighted membership representation calculation method is defined as follows:

$$u_{ij} = u_{ij}^I + u_{ij}^{II}, \quad (13)$$

$$u_{ij}^I = \frac{1}{\sum_{s=1}^C \left[ \left( \sum_{k=1}^d w_{ik} (x_{jk} - v_{ij})^2 \right) / \left( \sum_{k=1}^d w_{sk} (x_{jk} - v_{sk})^2 \right) \right]}, \quad (14)$$

$$u_{ij}^{II} = \alpha \frac{N_i - N_j}{\sum_{k=1}^d w_{ik} (x_{jk} - v_{ik})^2}. \quad (15)$$

In the above formula,  $N_j$  is calculated as follows:

$$N_j = \alpha \frac{\sum_{s=1}^C \left[ 1 / \left( \sum_{k=1}^d w_{sk} (x_{jk} - v_{sk})^2 \right) N_s \right]}{\sum_{s=1}^C \left[ 1 / \sum_{k=1}^d w_{sk} (x_{jk} - v_{sk})^2 \right]}. \quad (16)$$

## 5. Experiment and Result Analysis

**5.1. Data Collection.** A total of 80 college students aged from 18 to 25 were recruited in this experiment. Different emotions were induced by external stimuli in the form of pictures. Since the extraction of EEG signals is generally achieved by placing electrodes on the scalp, the signals are very weak, so hardware amplification by an amplifier is required after the EEG signals are extracted. The experimental data acquisition used GTEC's 16-channel G-USBamp system. Table 1 shows the specific settings of the acquisition parameters, and Figure 4 shows the placement locations of 16-channel EEG electrodes.

The letters in Figure 4 indicate where the electrodes are located. F represents the frontal lobe, C represents the center, T represents the temporal lobe, P represents the parietal lobe, and O represents the occipital lobe.

**5.2. Emotional Data Analysis.** For holistic analysis of the collected data, a depression scoring baseline was set. Those with the score below 53 points are defined as no depression, with the score between 53 and 62 as mild depression, the score between 62 and 72 as moderate depression, and the score above 72 as severe depression. The pie chart of the depressive mood data is shown in Figure 5. Health accounted for 70%, mild depression accounted for 16%, moderate depression accounted for 9%, and severe depression accounted for 5%. Generally speaking, most students are healthy. Similarly, anxiety was scored on a 50-point basis. The pie chart of depression data is shown in Figure 6. The EEG correlation between positive and negative emotions (depression and anxiety) is shown in Figure 7.

**5.3. Data Mining Performance Verification of Entropy-Weighted Clustering.** The main hardware configuration of the computer is 2.6GHZ CPU, 4G memory, and 500G hard disk. In order to save the initialization time cost of the algorithm, the dataset is firstly normalized, and the initial entropy-weighting coefficient is set.

First, the number of datasets is set to 15; that is, the initial target clustering center is 15, and the initial two-dimensional view is shown in Figure 8.

TABLE 1: EEG signal acquisition parameter settings.

Parameter name	Parameter value
Amplifier	16-channel G-USBamp system
Sampling frequency	512 Hz
High pass filter cutoff frequency	0.1 Hz
Low-pass filter cutoff frequency	60 Hz
Trap/notch frequency	50 Hz
Electrode position distribution	16 locations in 10–20 international standard system (as shown in Figure 4)
Grounding electrode position	Forehead
Reference electrode position	Right earlobe
Electrode material	Ag/AgCl
Recording software	g.Recorder

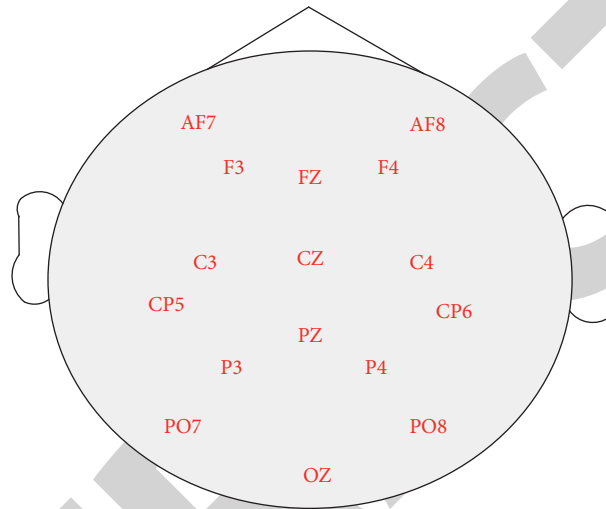


FIGURE 4: Schematic diagram of electrode position distribution.

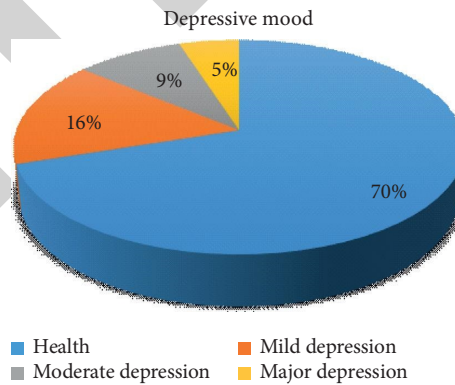


FIGURE 5: Depressed emotion pie chart.

After entropy-weighted clustering, the number of target clustering centers is reduced to three. In order to verify the convergence rate of the algorithm, the number of iterations of the algorithm was set to 2, 3, 9, and 16 times, respectively. The clustering results after 9 iterations are shown in Figure 9.

When the number of iterations was set to 2, 3, 9, and 16, the clustering centers became 9, 8, 6, and 4, respectively. The more iterations, the closer the number of target clustering centers was to 3. The number of iterations has a nonlinear

relationship with the number of clustering centers. In order to reach the target clustering center number 3, the iteration is continued. The experimental results show that when the number of iterations is 17, the number of clustering centers becomes three, as shown in Figure 10. It can be seen that the number of clustering centers converges stably, and the clustering centers no longer change with the increase of iteration times, which verifies the generalization of the clustering algorithm.

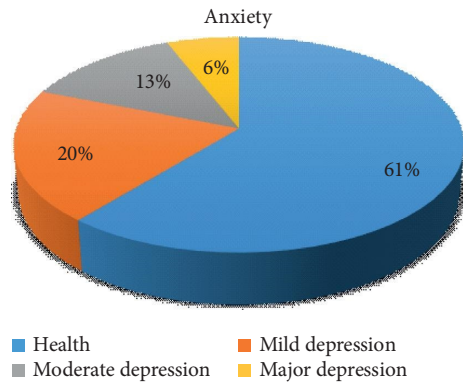


FIGURE 6: Anxiety emotion pie chart.

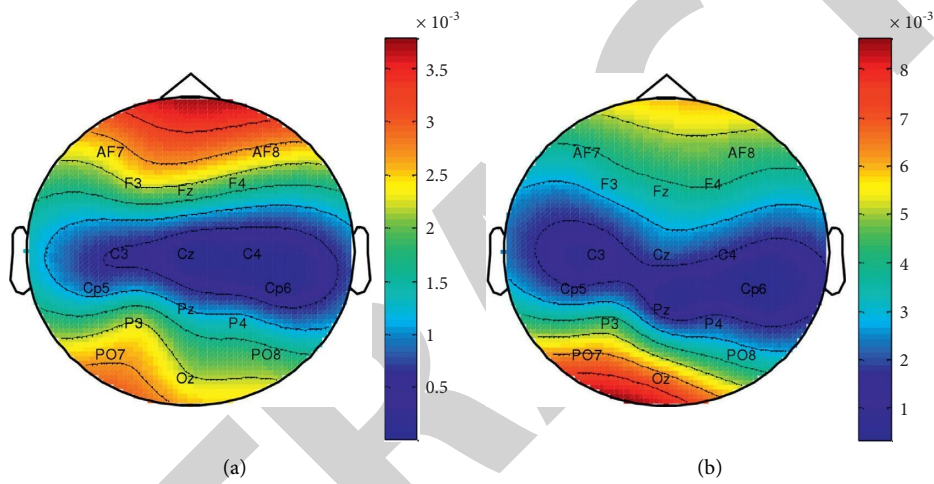


FIGURE 7: EEG correlation between positive and negative dichotomies of emotion. (a) Depressed emotion. (b) Anxiety emotion.

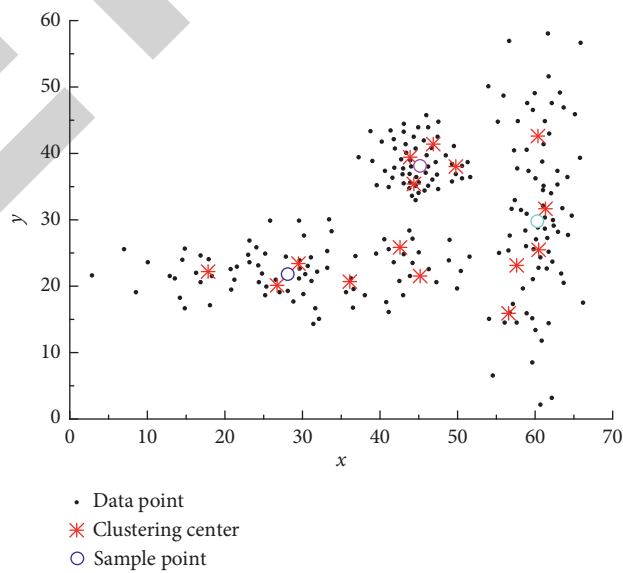


FIGURE 8: Two-dimensional view distribution of initial cluster center.



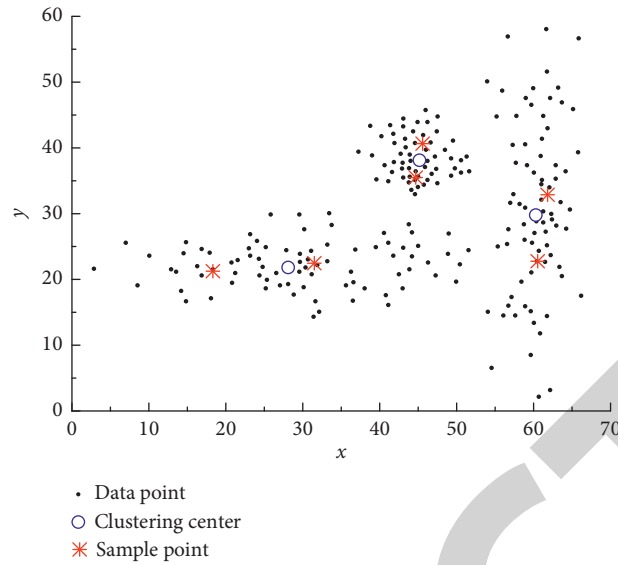


FIGURE 9: Clustering results after 9 iterations.

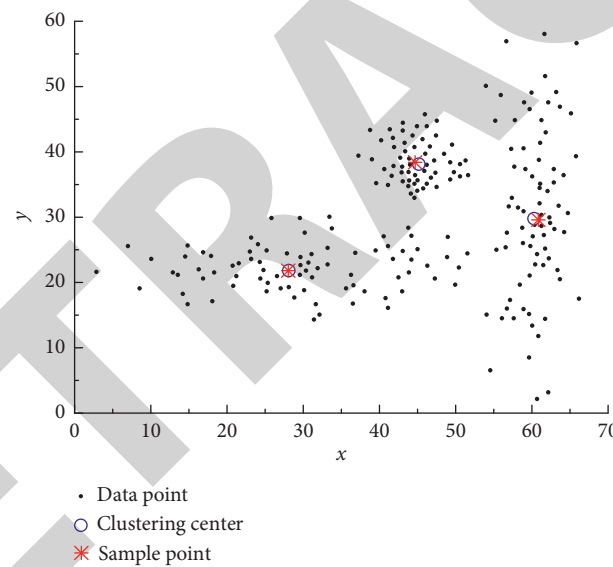


FIGURE 10: Clustering results after 17 iterations.

**5.4. Comparative Analysis of Emotion Recognition.** The proposed algorithm is compared with other algorithms to illustrate the effectiveness of the proposed algorithm. The classification of emotions here refers to the positive and negative (depression and anxiety) categories of tasks. The classification accuracy rate of the two emotional classifications is calculated and compared. This study was compared with DEAP BaseLine [28] and CNN [29]. The DEAP BaseLine is the baseline result of the DEAP database publisher's classification of emotions using traditional machine learning techniques. The average classification accuracy of different methods with the different number of features is shown in Figure 11.

As can be seen from Figure 11, regardless of the number of feature dimensions, the classification accuracy of the

proposed method is significantly higher than the existing DEAP BaseLine and CNN methods, reaching the highest accuracy of 68.35%. The classification results of the four methods were ranked as ours, CNN, and DEAP BaseLine. This is because the combination of entropy-weighted clustering algorithm and sparse learning feature selection can extract the local correlation of heterogeneous data and obtain more discriminating abstract features, thus obtaining the higher accuracy of sentiment classification. When the feature dimension is about 300, the proposed method can achieve the highest average, while the classification results of other methods are almost unchanged. It should be noted that when the feature number is greater than 700, the classification accuracy of the proposed method is no longer changed. This is because all the

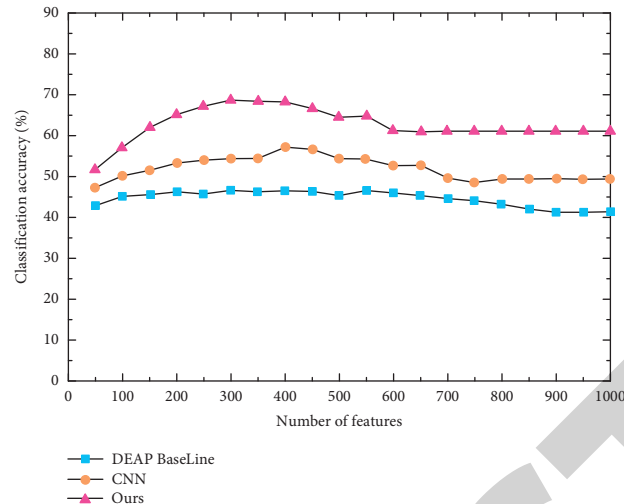


FIGURE 11: Average classification accuracy of different methods with different feature numbers.

features corresponding to the sparse weights have been used, and no more new features will be generated thereafter.

## 6. Conclusions

This paper proposes a method of college students' psychological emotion recognition based on sparse learning and entropy-weighted clustering. Sparse learning is mainly used to achieve the feature selection of EEG signals, which provides strong support for the subsequent construction of an emotional recognition model. At the same time, the entropy-weighted clustering algorithm and the features obtained after sparse learning are combined to establish a classification model for depression and anxiety, and better clustering accuracy and generalization are obtained. Under the condition of different feature dimensions, the classification accuracy of the proposed method is higher than that of the existing methods, and the highest accuracy reaches 68.35%. The experimental results of positive and negative dichotomies of emotions verify the effectiveness of the proposed method. In future research, we will focus on the finer spatiotemporal correlation extraction of electroencephalogram signals and try to identify emotions using nonlinear dynamic characteristics.

## Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] R. J. Huster, S. Enriquez-Geppert, and C. F. Lavalée, "Electroencephalography of response inhibition tasks," *International Journal of Psychophysiology Official Journal of the International Organization of Psychophysiology*, vol. 87, no. 3, pp. 217–233, 2016.

- [2] A. Reches, R. R. Nir, and M. J. Shram, "A novel electroencephalography-based tool for objective assessment of network dynamics activated by nociceptive stimuli," *European Journal of Pain*, vol. 20, no. 2, pp. 250–262, 2016.
- [3] V. K. Kimiskidis, "Transcranial magnetic stimulation (TMS) coupled with electroencephalography (EEG): biomarker of the future," *Revue Neurologique*, vol. 20, no. 6, pp. 123–126, 2016.
- [4] T. Sara, N. Chris, and P. Agnès, "Older people's experiences of mobility and mood in an urban environment: a mixed methods approach using electroencephalography (EEG) and interviews," *International Journal of Environmental Research and Public Health*, vol. 14, no. 2, pp. 151–157, 2017.
- [5] T. Heunis, C. Aldrich, and V. P. J. De, "Recent advances in resting-state electroencephalography biomarkers for autism spectrum disorder—a review of methodological and clinical challenges," *Pediatric Neurology*, vol. 11, no. 3, pp. 28–37, 2016.
- [6] B. Tobias and T. Myron, "Induction of spreading depression in the ischemic hemisphere following experimental middle cerebral artery occlusion: effect on infarct morphology," *Journal of Cerebral Blood Flow and Metabolism*, vol. 16, no. 2, pp. 202–213, 2016.
- [7] U. J. Chaudhary, D. W. Carmichael, and R. Roman, "Mapping preictal and ictal haemodynamic networks using video-electroencephalography and functional imaging," *Brain*, vol. 135, no. 12, pp. 3645–3663, 2017.
- [8] E. Albornoz and D. Milone, "Emotion recognition in never-seen languages using a novel ensemble method with emotion profiles," *IEEE Transactions on Affective Computing*, vol. 8, no. 99, pp. 21–29, 2017.
- [9] J. C. Castillo, Á. Castro-González, and A. Fernández-Caballero, "Software architecture for smart emotion recognition and regulation of the ageing adult," *Cognitive Computation*, vol. 8, no. 2, pp. 1–11, 2016.
- [10] A. Russo, P. Jayakar, and M. Lallas, "The diagnostic utility of 3D electroencephalography source imaging in pediatric epilepsy surgery," *Epilepsia*, vol. 57, no. 1, pp. 24–31, 2016.
- [11] C. Jing, B. Hu, and W. Yue, "Subject-independent emotion recognition based on physiological signals: a three-stage

- decision method," *BMC Medical Informatics and Decision Making*, vol. 17, no. 3, pp. 24–31, 2017.
- [12] W. L. Zheng and 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. 22–31, 2015.
- [13] L. D. Bai, J. Q. Guo, and T. Y. Xu, "Emotional monitoring of learners based on EEG signal recognition," *Procedia Computer Science*, vol. 174, no. 3, pp. 364–368, 2020.
- [14] A. H. Masruroh, E. M. Imah, and E. Rahmawati, "Classification of emotional state based on EEG signal using AMGLVQ - sciencedirect," *Procedia Computer Science*, vol. 157, pp. 552–559, 2019.
- [15] S. B. Wankhade and D. D. Doye, "Deep learning of empirical mean curve decomposition-wavelet decomposed EEG signal for emotion recognition," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 28, no. 1, pp. 153–177, 2020.
- [16] Y. Mitsukura, "EEG signal processing for real applications," *Journal of Signal Processing*, vol. 20, no. 1, pp. 1–7, 2016.
- [17] S. Liu, X. Wang, and M. Liu, "Towards better analysis of machine learning models: a visual analytics perspective," *Visual Informatics*, vol. 1, no. 1, pp. 48–56, 2017.
- [18] M. Kohli, L. M. Prevedello, R. W. Filice, and J. R. Geis, "Implementing machine learning in radiology practice and research," *American Journal of Roentgenology*, vol. 208, no. 4, pp. 1–7, 2017.
- [19] X. Lin, X. Wang, and L. Li, "Intelligent detection of edge inconsistency for mechanical workpiece by machine vision with deep learning and variable geometry model," *Applied Intelligence*, vol. 50, no. 7, pp. 2105–2119, 2020.
- [20] T. Kim, I. Y. Jung, and Y. C. Hu, "Automatic, location-privacy preserving dashcam video sharing using blockchain and deep learning," *Human-centric Computing and Information Sciences*, vol. 10, no. 1, pp. 1–23, 2020.
- [21] M. Fujiu, J. R. Toleikis, and J. A. Logemann, "Glossopharyngeal evoked potentials in normal subjects following mechanical stimulation of the anterior faucial pillar," *Electroencephalography and Clinical Neurophysiology Evoked Potentials*, vol. 92, no. 3, pp. 183–195, 2016.
- [22] P. M. Rossini, G. Deuschl, V. Pizzella, F. Tecchio, and C. H. Lücking, "Topography and sources of electromagnetic cerebral responses to electrical and air-puff stimulation of the hand," *Electroencephalography and Clinical Neurophysiology*, vol. 100, no. 3, pp. 229–239, 2016.
- [23] J. N. Acharya and A. J. Hani, "American clinical neurophysiology society guideline 2: guidelines for standard electrode position nomenclature," *American Journal of EEG Technology*, vol. 56, no. 4, pp. 245–252, 2016.
- [24] Y. Yan, C. Li, and S. Meng, "Emotion recognition based on sparse learning feature selection method for social communication," *Signal Image and Video Processing*, vol. 13, pp. 1253–1257, 2019.
- [25] T. Kujirai and T. Yokota, "Greedy action selection and pessimistic Q-value updating in multi-agent reinforcement learning with sparse interaction," *SICE Journal of Control Measurement and System Integration*, vol. 12, no. 3, pp. 76–84, 2019.
- [26] M. D. Leoni, W. M. Aalst, and M. Dees, "A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs," *Information Systems*, vol. 56, no. 3, pp. 235–257, 2016.
- [27] M. Fathian and A. R. Jafarianmoghaddam, "New clustering algorithms for vehicular ad-hoc network in a highway communication environment," *Wireless Networks*, vol. 21, no. 8, pp. 2765–2780, 2015.
- [28] S. Koelstra, "DEAP: a database for emotion analysis using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
- [29] M. A. Li, J. F. Han, and J. F. Yang, "Automatic feature extraction and fusion recognition of motor imagery EEG using multilevel multiscale CNN," *Medical, & Biological Engineering & Computing*, vol. 59, no. 10, pp. 2037–2050, 2021.