

Retraction

Retracted: EEG Emotion Recognition Based on Temporal and Spatial Features of Sensitive signals

Journal of Electrical and Computer Engineering

Received 19 December 2023; Accepted 19 December 2023; Published 20 December 2023

Copyright © 2023 Journal of Electrical and Computer Engineering. 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This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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) Manipulated or compromised peer review

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] Z. Qu and X. Zheng, "EEG Emotion Recognition Based on Temporal and Spatial Features of Sensitive signals," *Journal of Electrical and Computer Engineering*, vol. 2022, Article ID 5130184, 8 pages, 2022.

Research Article

EEG Emotion Recognition Based on Temporal and Spatial Features of Sensitive signals

Zhihao Qu  and Xiujuan Zheng 

College of Electrical Engineering, Sichuan University, Chengdu, Sichuan 610065, China

Correspondence should be addressed to Zhihao Qu; 2019223039264@stu.scu.edu.cn

Received 15 April 2022; Accepted 17 May 2022; Published 5 December 2022

Academic Editor: Xuefeng Shao

Copyright © 2022 Zhihao Qu and Xiujuan Zheng. 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.

Currently, there are some problems in the electrocorticogram (EEG) emotion recognition research, such as single feature, redundant signal, which make it impossible to achieve high-precision recognition accuracy when used a few channel signals. To solve the abovementioned problems, the authors proposed a method for emotion recognition based on long short-term memory (LSTM) neural network and convolutional neural network (CNN) combined with neurophysiological knowledge. First, the authors selected emotion-sensitive signals based on the physiological function of EEG regions and the active scenario of the band signals, and then merged temporal and spatial features extracted from sensitive signals by LSTM and CNN. Finally, merged features were classified to recognize emotion. The method was experimented on the DEAP dataset, the average accuracy in the valence and arousal dimensions were 92.87% and 93.23%, respectively. Compared with similar studies, it not only improved the recognition accuracy, but also greatly reduced the calculation channel, which proved the superiority of the method.

1. Introduction

Emotion is a comprehensive state of human interaction with the outside world, some people are not good at expressing emotions, when the accumulation of negative emotions causes a series of mental illnesses. An efficient emotion recognition method can assist psychologists in identifying potential mental illnesses. Early studies on emotion recognition were mostly based on nonphysiological signals, such as expressions, speech, and body language, but these signals can be disguised by personal awareness and lack reliability [1]. Recently, it has been shown that EEG signals are nonfakeable and can objectively describe changes in mental states [2, 3]; so many scholars have started emotion recognition research based on EEG. Currently, the relevant methods can be divided into two categories: traditional machine learning and neural networks. In the field of machine learning, Guo et al. [4] extracted the temporal features of EEG signals, used the granger causal model and implemented emotion recognition by support vector machines (SVM). Jin et al. [5] extracted the temporal and frequency

features of the EEG signal, merged the two features through an internal cascade forest, and finally used a deep forest (DF) to classify the merged features. The final recognition accuracy in the valence and arousal dimensions reached 66.3% and 65.8%, respectively. Zhu [6] merged the temporal, frequency, and spatial features extracted from the EEG signal and used SVM to classify the merged features, achieving 71% and 70% recognition accuracy in the valence and arousal dimensions. Although the principle of machine learning is simple, easy to understand, and rapid to recognize emotion, such methods are less sensitive to the variability among features, and experimental results are usually unsatisfactory.

As neural networks have achieved increasingly significant results in several studies [7], some scholars have started to use them to analyze EEG signals. For example, Liu et al. [8] merged the multidomain linear and nonlinear features extracted from the EEG signal, and then classified the merged features via a stacked auto-encode neural network (SAE) to recognize the subject's emotion. Chao et al. [9] also proposed a recognition method based on the multidomain merged features of EEG signal in temporal, frequency, and

temporal-frequency combined with the deep belief network (DBN). Yang et al. [10] used the CNN model directly to mine information from the subjects' EEG signals and eventually achieved 90.01% and 90.65% recognition accuracy in the valence and arousal dimensions. On the other hand, Kim Y and Choi [11] analyzed EEG signals from temporal perspective, mined their temporal information using LSTM, and used fully connected layer to classify features, ultimately achieving 90.1% and 88.1% recognition accuracy. Most scholars usually use all channel signals for experiments in order to guarantee that the extracted features carry enough rich information. However, these scholars ignore the problem that not all the signals are closely related to emotions. If the signal is complex and the channel signals are redundant, the experiments will be disturbed by irrelevant information and increase the computational load. In addition, the EEG signal is a complex signal integrating temporal and spatial characteristics [12], and if only a single feature of the signal is extracted for emotion recognition, it will lead to the problem that the analysis perspective is not comprehensive enough and carries incomplete information. Most of the current studies using the idea of merging features for emotion recognition have manually extracted shallow features and then merged these shallow features. The process of such research is tedious, and the original information is easily lost during the manual extraction of features, so it is difficult to achieve better recognition results. Although neural networks can achieve end-to-end emotion recognition and extract abstract features from EEG signals [13], the time-consuming problem of processing a large number of signals also makes past scholars consider only a single feature in their research.

Based on the abovementioned thinking, if we can refine the acquired signals in advance and then use the powerful information mining ability of neural networks to analyze the EEG signals from multiple perspectives, we can guarantee the recognition efficiency while achieving high accuracy recognition accuracy. Therefore, in this paper, the authors combined the knowledge of neurophysiology to select the signals that are sensitive to emotions in advance, and on this basis, we combined LSTM and CNN models to extract the deep temporal and spatial features of the sensitive signals, respectively, and merge the two features to achieve emotion recognition.

2. Data Description and Preprocessing

2.1. DEAP Dataset. In this paper, we conducted experiments using the publicly available dataset DEAP collected by M. Khateeb et al.[14]. This dataset records the subjects' emotional changes by means of video-evoked emotions, and none of the 32 subjects participating in the trial had any history of illness. Each subject watched 40 videos of 63 s duration, with the first 3 s of the video being the duration of the preparation phase and the last 60 s being the duration of the formal experiment. The data set recorded 32 EEG signals and 8 other physiological signals. In this paper, 32 EEG signals were selected for the experiment, and the sampling frequency of the signals was 128 Hz, with a total of 8064 sampling points for each channel.

All subjects rated their current emotions on four dimensions of valence, arousal, liking, and dominance on a scale of 1–9 after watching a 60-second video. In this paper, the emotion model proposed by Russell was used for emotion recognition in the valence and arousal dimensions. The label value less than 5 in both dimensions were replaced with 0 and were noted as low valence/low arousal, respectively, and the remaining label value were replaced with 1 and were noted as high valence/high arousal. High/low valence represents the positivity/negativity of emotion, while high/low arousal indicates the strength/weakness of emotion.

2.2. Data Preprocessing. To enhance the robustness of the experiments and the resistance to fitting of the network framework, we performed data expansion and baseline calibration operations on the initial signals before the start of the formal experiments. We cut each channel signal with a window of 1 s step to obtain 63 segments of 1 s signal. The first 3 segments are the baseline signals and the last 60 segments are the formal experimental signals. The input signal of the experiment is obtained by subtracting the average sampling value of the first 3 segments from the sampling value of the last 60 segments, so a single subject has a total of 40×60 experimental signals [15]. The baseline calibration equation is as follows:

$$\text{Base Avg} = \frac{\sum_1^3 \text{Base Line}_j}{3}, \quad (1)$$

$$\text{Input EEG}_i = \text{EEG}_i - \text{BaseAvg}.$$

In the above equation, Base Line_j is the reference signal 3 s before the experiment, EEG_i is the formal experiment signal, Input EEG_i is the signal after baseline calibration and was used in the formal experiment, $i = 1, 2, \dots, 60$, and $j = 1, 2, 3$.

3. Methods

3.1. Sensitive Signal Selection

3.1.1. Band Signal Selection. The EEG signal can be divided into five band signals, δ wave (1 Hz~4 Hz), θ wave (4 Hz~8 Hz), α wave (8 Hz~13 Hz), β wave (13 Hz~30 Hz), and γ wave (30 Hz~45 Hz) according to the frequency range [16]. The collectors of the DEAP dataset have prefiltered the δ wave, which is uncommon in the waking brain, so the use of this paper EEG signal contains only four band signals: θ , α , β , and γ .

The activity of band signals varies when the brain is in different states; θ wave is more common when the brain is in a state of extreme fatigue; α wave is more regular and is the most common band signals when the brain is awake; β and γ waves are more common when the human brain is in an excited state and are particularly active when the brain is performing higher functions such as emotion. It has been shown that human emotion changes trigger fluctuations in a variety of band signals [17]. Based on the above analysis, we

selected the combined signals $\alpha + \beta + \gamma$ of α , β , and γ for emotion recognition. To verify that the selected signals can achieve better emotion recognition, they also extract five groups of signals, $\theta + \alpha$, $\alpha + \beta$, $\beta + \gamma$, $\theta + \alpha + \beta$, and $\theta + \alpha + \beta + \gamma$, for comparison experiments.

3.1.2. Channel Signal Selection. The regions of the brain can be divided into five areas: frontal lobe, temporal lobe, parietal lobe, occipital lobe, and central area, each responsible for different physiological functions [18]. The frontal lobe is the main area of the brain that performs advanced functions and is responsible for the generation of thoughts and emotions, and the prefrontal area of the frontal lobe is particularly prominent when emotional mechanisms are triggered; the parietal lobe is mainly responsible for the perception of stress, pain, and other stimuli, and when the human brain is stimulated, emotions change and the posterior parietal area becomes active; the temporal lobe and occipital lobe are mainly responsible for the processing of visual stimuli and memory recall, and are not directly related to emotions; the central area is mainly responsible for the integration of spatial information from different regions, and its central location is where the most frequent spatial information processing activities occur.

Based on the abovementioned analysis, the authors first selected FP1, FP2, AF3, AF4, F7, F3, Fz, F4, and F8 from the frontal lobe region for a total of 9 channels; then selected P3, PO3, P4, and PO4 from the parietal lobe region for a total of 4 channels; finally, they consider that they need to explore the emotional information in the EEG spatial domain, the Cz channel in the central area was also included in the experiment. In summary, a total of 14 channels are selected for the experiments in this paper.

3.2. Merged Feature

3.2.1. Temporal Features Extracted by LSTM. The EEG signal is nonlinear and nonstationary, and to extract features that reveal the temporal emotional information of the signal, it is necessary to correlate the information changes with the emotional fluctuations in different time periods [19].

LSTM is a new neural network which is improved on the basis of the recurrent neural network, which has a special gate structure that can selectively preserve the key information of the current sequence and combine it with the key information of the subsequent sequences. Therefore, the LSTM is able to extract features containing a large amount of temporal emotion information from the EEG signal and describe the fluctuation pattern of the EEG signal with emotion change in a global perspective [20]. The internal structure of the LSTM unit is shown in Figure 1.

In Figure 1, h_{t-1} and C_{t-1} represent the output and state of the previous cell, respectively, and the calculation formulas of the output h_t and state C_t of the current unit are shown in (2)–(6).

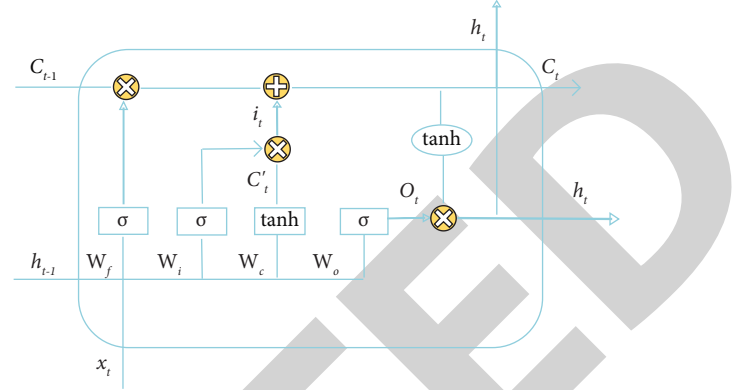


FIGURE 1: Internal structure diagram of the LSTM unit.

$$f_t = \sigma(\mathbf{W}_f \times [h_{t-1}, x_t] + b_f), \quad (2)$$

$$i_t = \sigma(\mathbf{W}_i \times [h_{t-1}, x_t] + b_i), \quad (3)$$

$$O_t = \sigma(\mathbf{W}_o \times [h_{t-1}, x_t] + b_o), \quad (4)$$

$$C_t = C_{t-1} * f_t + i_t * \tanh(\mathbf{W}_c \times x_t + b_c), \quad (5)$$

$$h_t = O_t * \tanh(C_t). \quad (6)$$

In the above formulas, x_t is the input signal, and f_t , i_t , and O_t represent the information of the forget gate, input gate, and output gate. The f , i , and O in other variables also indicate that the variable belongs to the forget gate, input gate, and output gate. \mathbf{W} and b are the weight matrix and bias of the gate, w_c and b_c are the control weight and bias of the input gate. \tanh is the hyperbolic tangent function and σ is the sigmoid activation function.

Some scholars have tried to use LSTM to analyze EEG signals, but the input signals are too redundant, resulting in the LSTM unit taking longer time in processing the information. In this paper, we have filtered the original signal in advance, and the amount of information that needs to be processed by the LSTM model is significantly reduced, which also reduces the time duration of the LSTM to extract temporal features.

3.2.2. Spatial Features Extracted by CNN. In order to describe the spatial connection between EEG signal channels, the original 1D chain EEG signal is mapped into a 2D matrix signal in this paper, and the mapped matrix structure corresponds to the electrode distribution in the cerebral cortex, and the signal is mapped as shown in Figure 2.

The CNN has the advantages of local connectivity and shared weights and has the ability to mine and integrate the local information of matrix signals [21], so it is suitable for analyzing the spatial connection of each channel in 2D matrix signals in this paper. In recent years, some scholars have tried to use the CNN to analyze the internal spatial

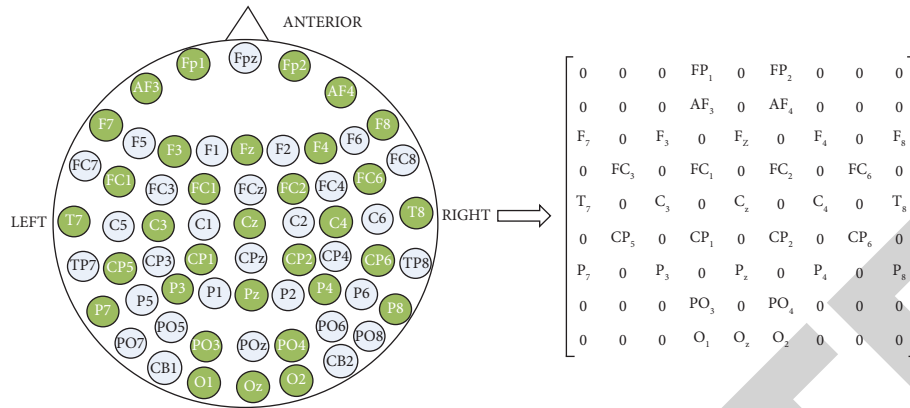


FIGURE 2: 2D spatial mapping of EEG signals.

connections of EEG signals, but most of them combine shallow multidomain features of single channels, then merge these features into matrix signals as input, and finally further mine spatial information from the merged features [22]. Such experimental results are highly dependent on the scholars' experience in feature selection, and the experiments lack robustness. The extracted spatial features cannot highlight the spatial topology between the channels and have insufficient characterization capability, leading to poor experimental results.

In this paper, the original channel signal is directly mapped into a 2D matrix by the method of spatial mapping, and then the spatial features are extracted used CNN. Compared with previous methods of extracting spatial features, the used CNN to operate directly on the mapped original signal can distill the initial emotional information embedded in the original signal into significant features, which is more accurate and realistic than other models that merge shallow features as input. In addition, the 2D matrix retains the spatial connections between the channels, and the convolutional kernel of the CNN can then be used to explore the spatial connections between EEG signals from local to global. Therefore, the spatial features that were extracted have stronger reliability and integrity compared with those extracted by previous methods.

3.2.3. Merge Spatial Features and Temporal Feature. Currently, the existing methods of merging features can only extract some simple and shallow features for merging, and the information carried by the merged features is slightly shallow. Insufficient variability of features leads to low accuracy of the classifier in recognizing features. In this paper, we use the LSTM and CNN to directly extract the temporal and spatial features of the original EEG signal, then merge the extracted features, and finally classify the merged features to achieve the purpose of recognizing emotions. Compared to the shallow merged feature method, our method has the following advantages: (a) both the temporal features extracted from each time period by the LSTM through the gate structure and the spatial features extracted from the mapping matrix signals by the CNN through convolution kernel contain rich emotion information. By

merging the extracted features, the emotion information in the EEG signal can be explored more comprehensively to further improve the emotion recognition accuracy. (b) Both LSTM and CNN automatically extract features directly from the original signal without going through a manual feature extraction step, which not only increases the convenience of the method, but also avoids the loss of the original information when extracting shallow features. The flow framework of the merge feature method is shown in Figure 3.

As shown in Figure 3, the CNN first extracts the base features from the 2D EEG signal through three convolutional layers, and then uses a pooling layer to downscale the base features to obtain the spatial features. The LSTM model directly explores the temporal information of the 1D original signal, and the output of the last LSTM unit in the second layer of the LSTM model are the temporal features. The temporal and spatial features are merged, and then the fully connected layer with an activation function of SoftMax is used to classify the merged feature for recognizing emotion.

4. Experimental Results and Analysis

A total of 32 subjects participated in the experiment, and each subject had 40 original samples in the format 32×7680 , 32 is the number of channels and 7680 is the number of sampling points. The original sample was divided into 60 experimental samples in the format of 32×128 , so each subject has 2400 samples in the format 32×128 . A single experimental sample is spatially mapped to a 2D matrix format of 9×9 . The CNN input tensor format is $2400 \times 9 \times 9 \times 128$. The LSTM input tensor format is $2400 \times 32 \times 128$. We use the function to randomly divide the original data into 10 copies, of which 7 copies are used as the training set to train the network model and 3 copies are used as the test set to test the network model.

The first two layers of the LSTM model are cell layers, the number of cells is set to 256 and 128, the activation function of the two layers is tanh, and a dropout of 0.5 is set in each layer to prevent overfitting. The third layer is the fully connected layer, the activation function of this layer is SoftMax, the loss function is binary cross-entropy, the learning rate is 0.001, the number of training iterations epoch is 30, and the batch size is 32.

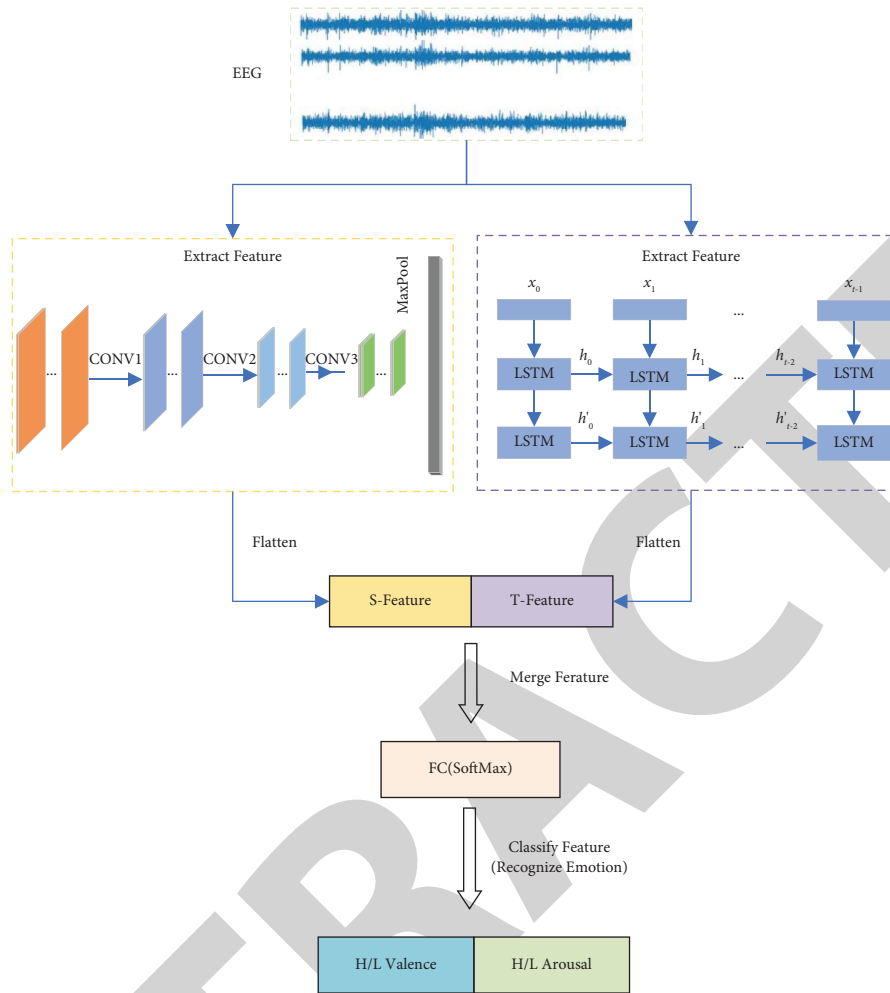


FIGURE 3: Flow chart of the merge feature method.

The CNN model has five layers, the first three layers are convolutional, the fourth layer is pooling, and the fifth layer is fully connected. The settings of dropout, loss function, learning rate, epoch, and batch size in the CNN model are the same as those of the LSTM model, and the step size of the convolutional layer is set to 1. Other specific parameters of each layer are shown in Table 1. The hardware devices are Intel(R) Core (TM) i9-9900K CPU and NVIDIA Ge Force RTX 2080 SUPER, the software environment is Python 3.6, and the neural network framework is tensorflow2.0.

4.1. Band Selection Experiment. The authors conducted a total of six sets of band selection experiments using the LSTM model in the valence and arousal dimensions, and the input samples were all original 1D signals. We use the average recognition accuracy of all subjects as the evaluation index, the experimental results are shown in Figure 4.

As can be seen from Figure 4, the experiment with $\alpha + \beta + \gamma$ signals achieved 90.45% and 91.68% recognition accuracy in the valence and arousal dimensions, respectively, which are 0.87% and 1.26% higher than those with $\theta + \alpha + \beta + \gamma$ signals, respectively, and 1%–3% in both dimensions compared to the four groups of experiments with

$\theta + \alpha$, $\alpha + \beta$, $\beta + \gamma$, and $\theta + \alpha + \beta$ signals of improvement. The abovementioned results indicate that after filtering out the θ wave, which is weakly correlated with emotion, the filtered $\alpha + \beta + \gamma$ combination signal can more accurately describe subjects' emotional changes compared with other signals, and the emotional signal-to-noise ratio of the EEG signal has been improved. All experiments in this paper were conducted using $\alpha + \beta + \gamma$ signals.

4.2. Channel Selection Experiment. In this section, we use the CNN model for experiments, and the main experiment uses the selected 14-channel signal as the input, and the comparison experiment selects the 32-channel signal as the input. Finally, the average recognition accuracy of 32 subjects and the average experiment duration of a single subject are used as evaluation metrics. The experimental results are shown in Figure 5.

It can be seen from Figure 5 that after reducing the experiment channels based on the knowledge of brain functional regions, no matter in the dimension of valence or arousal, there is no big difference in the accuracy of most subjects in the two types of experiments. Compared to the 32-channel experiment, two subjects in the 14-channel

TABLE 1: CNN structure parameters.

Layers	Convolution kernel size	Filter size	Number of convolution kernels/filters	Activation function
1	3×3	—	32	RELU
2	3×3	—	64	RELU
3	3×3	—	128	RELU
4	—	3×3	1	—
5	—	—	—	SoftMax

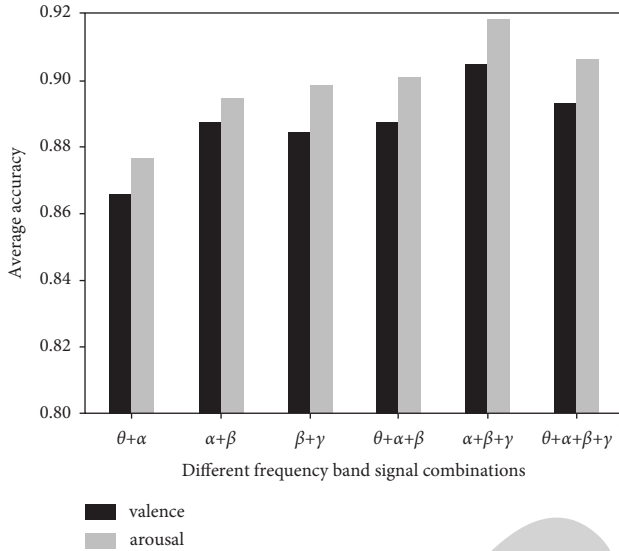


FIGURE 4: Band selection experiment results.

experiment improved more significantly, with the subject numbered 21 having a 7.05% increase in recognition accuracy in the valence dimension and the subject numbered 7 having a 5.87% increase in recognition accuracy in the arousal dimension. In addition, the average time spent by a single subject in the 14-channel experiment was 30 s, which was 48.3% less than the average time spent in the 32-channel experiment (58 s). The abovementioned results show that the channel signals selected by combining the knowledge about brain regions can not only achieve higher recognition accuracy, but also improve the recognition efficiency. The subsequent experiments in this paper were conducted using the selected 14-channel signals.

4.3. Merge Feature Experiment. The authors used the LSTM model and CNN model to extract temporal and spatial features of sensitive signals, respectively, and merge the two types of features, and finally use the fully connected layer with SoftMax function to classify the merged features. In order to verify the superiority of the merged features, we also use the CNN model and LSTM model with the same parameters to extract temporal and spatial features, respectively, for comparison experiments. The experimental results are shown in Figure 6. The triangular symbols in Figure 6 represent the mean and the horizontal lines in Figure 6 represent the median.

As can be seen from Figure 6, the experiments using merged features achieve 92.87% and 93.23% recognition accuracy in valence and arousal dimension, respectively, which

are 2.64% and 2.05% higher than the results of the experiments using only temporal features and 1.10% and 1.12% higher than the results of the experiments using only spatial features. The above single feature experiment and the merged feature experiment use the same network structure parameters, but the experimental results of the merged features are better, indicating that the information carried by the temporal and spatial features they extracted can complement each other. In addition, while the merged feature method improves the experimental accuracy, the median of the subjects' accuracy also improves, and the upper and lower edge values of the box plot are close to the median. It indicates that the method not only improves the accuracy of the experiment, but also further enhances the robustness and stability of the experiment.

4.4. Comparison with Similar Studies. In addition to the comparative experiments designed in this paper, experiments on emotion recognition using different methods on the same data set in recent years are selected for comparison with this paper. The experiment results are shown in Table 2: As can be seen from Table 2, their method not only significantly improves the experimental accuracy, but also uses fewer channels compared to other existing methods, which significantly reduces the computational load. In Table 2, their proposed method has the most significant improvement compared to the method proposed by Jin et al. [5], the recognition accuracy of valence and arousal dimensions improved by 26.57% and 27.43%, respectively. Guo et al. [4], Jin et al. [5], Zhu et al. [6], Liu and Qiao [8], and Chao et al. [9] also used merged features for emotion recognition, but the recognition effect is not as good as that of their method. The reason is that these scholars extracted the features manually and lost some of the original information in the feature extraction process, resulting in poor abstraction of the merged features, which are not easily recognized by the classifier. In addition, the authors directly use the LSTM and CNN to extract different deep features from the original signal and merge them, and the operations of extracting features are performed for the signal data, which can be used to analyze the EEG signals collected in different scenes. Compared with the emotion recognition methods of Kim and Choi [11] and Bozhkov and Georgieva [12] using single deep features, the authors' method also showed significant improvement in valence and arousal dimension and used fewer channel signals.

The abovementioned results show that merged features are more representative and more comprehensive in the study of signals compared to single features and it is easier to achieve better recognition results.

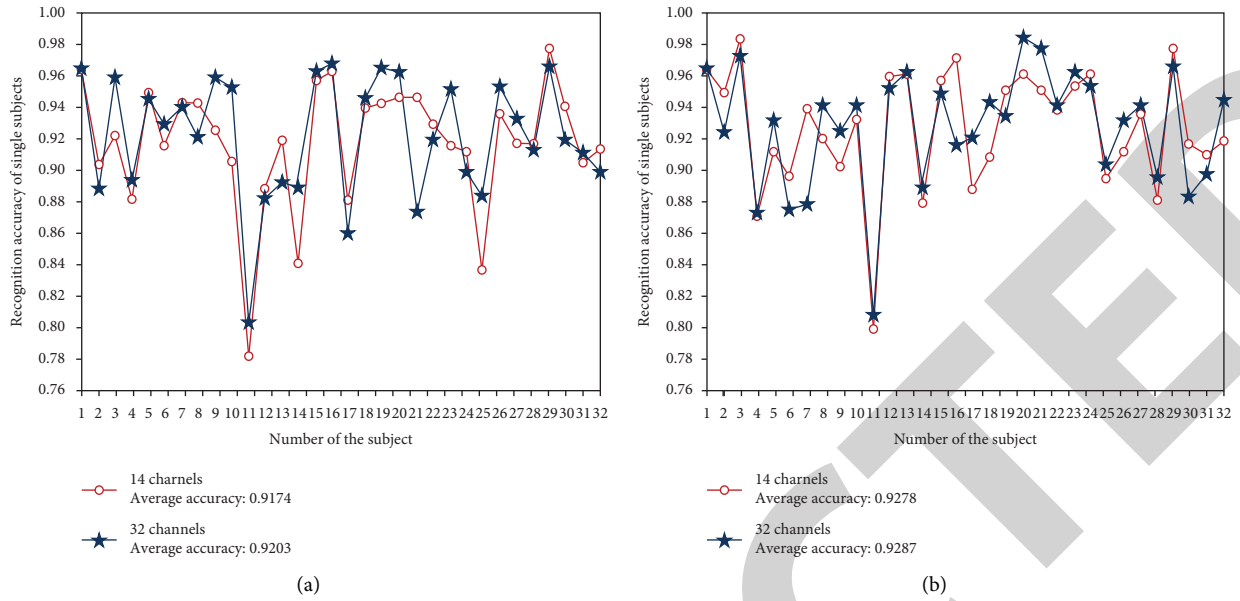


FIGURE 5: The results of the channel selection experiment.

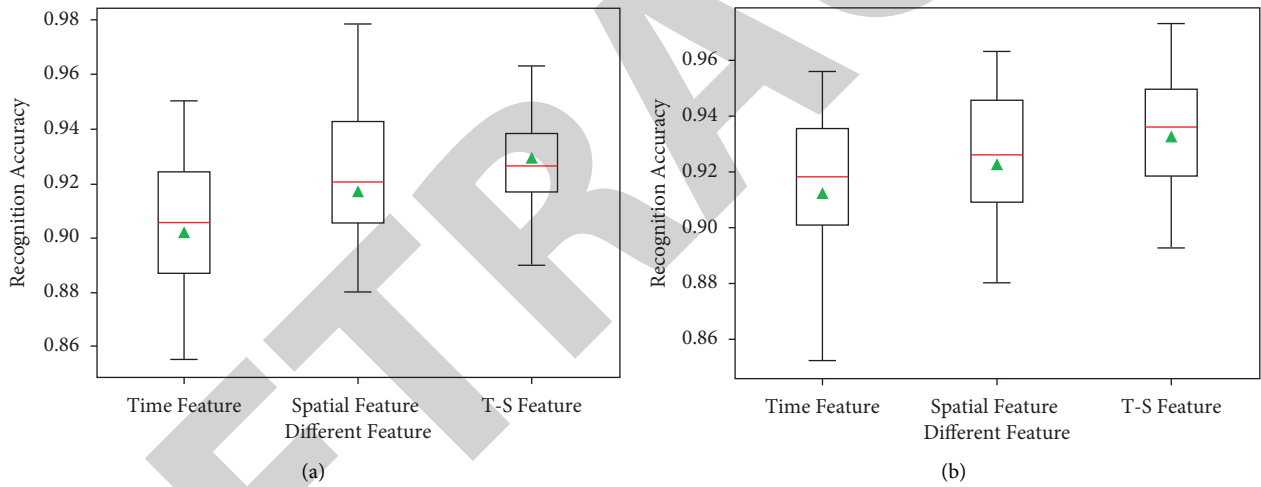


FIGURE 6: The experiment results of single feature and merge feature.

TABLE 2: Comparison of similar experiment results

Authors (year)	Features	Classifier	Valence (%)	Arousal (%)	Channel number
Guo et al. [4] (2018)	Temporal	SVM	87.15	86.6	14
Jin et al. [5] (2018)	Frequency + temporal	DF	66.30	65.80	32
Zhu [6] (2018)	Temporal + frequency + spatial	SVM	71.00	70.00	32
Liu and Qiao [8] (2021)	Frequency + temporal-frequency + nonlinear	SAE	80.30	81.50	32
Chao et al. [9] (2018)	Temporal + frequency + temporal-frequency + spatial	DBN	70.15	75.92	32
Yang et al. [10] (2019)	Spatial	FC (SoftMax)	90.01	90.65	32
Kim and Choi [11] (2021)	Temporal	FC (SoftMax)	90.11	88.10	32
The authors'	Temporal + spatial	FC (SoftMax)	92.87	93.23	14

5. Conclusions

In this paper, we combined the knowledge of neurophysiology to select the EEG signals from the perspective of frequency bands and channels, and on this basis, they extracted and merged the deep features of multiple domains, and finally classified the merged features for recognizing emotions. The following conclusions were drawn from the experimental results: (1) the signals selected according to the activity of the frequency band signals in different scenes carried more refined emotional information and reduced the interference of redundant information to the experiment. (2) The channel signals selected according to the brain regions and their responsible physiological functions could not only improve the accuracy of the experiment, but also reduce the computational load of the experiment. (3) The merged features carried more comprehensive information and enhanced characterization ability, which can improve the recognition accuracy. In the future, they would apply the method to multiple data sets for cross-data EEG emotion recognition research to enhance the generalizability of the method. It is hoped that the method will be applied to real-life emotion recognition as soon as possible.

Data Availability

The tagged datasets used to support the findings of this study are publicly available dataset and can be downloaded from the following website <https://www.eecs.qmul.ac.uk/mmv/datasets/deap/download.html>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Funding

No fund support

References

- [1] W. M. Vanderlind, J. Everaert, and J. Joormann, "Positive Emotion in Daily Life: Emotion Regulation and depression," *Emotion*, vol. 22, 2021.
- [2] J. Zhang, Z. Yin, P. Chen, and S. Nichele, "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review," *Information Fusion*, vol. 59, pp. 103–126, 2020.
- [3] Y. Höller, E. Trinka, G. Kalss, G. Schiepek, and R Michaelis, "Correlation of EEG spectra, connectivity, and information theoretical biomarkers with psychological states in the epilepsy monitoring unit A pilot study," *Epilepsy & Behavior*, vol. 99, Article ID 106485, 2019.
- [4] J. L. Guo, F. Fang, and W. Wang, "EEG emotion recognition based on sparse group lasso-granger causality feature," *Pattern Recognition and Artificial Intelligence*, vol. 31, no. 10, pp. 941–949, 2018.
- [5] Y. X. Jin, Y. Luo, and Y. Yu, "Research on EEG emotion recognition based on deep forest," *Softw. Guide*, vol. 18, no. 7, pp. 53–55, 2019.
- [6] H. Y. Zhu, *Research of Emotion Recognition Based on Multi-Domain EEG Features and Integration of Feature selection*, Jilin University, Changchun, China, 2021.
- [7] C. Zhang, Y. Guo, and M. Li, "Review of development and application of artificial neural network models," *Comput Eng Appl*, vol. 57, pp. 57–69, 2021.
- [8] P. Liu and X. Y. Qiao, "EEG emotion recognition based on stacked auto-encoder neural network," *Journal of Experimenting Technology*, vol. 35, no. 2, pp. 145–151, 2021.
- [9] H. Chao, H. Zhi, L. Dong, and Y. Liu, "Recognition of emotions used multichannel EEG data and DBN-GC-based ensemble deep learning framework," *Computational intelligence and neuroscience*, vol. 2018, Article ID 9750904, 11 pages, 2018.
- [10] H. Yang, J. Han, and K. Min, "A multi-column CNN model for emotion recognition from EEG signals," *Sensors*, vol. 19, no. 21, p. 4736, 2019.
- [11] Y. Kim and A. Choi, "Emotion Classification based on EEG signals with LSTM deep learning method," *Journal of the Korea Industrial Information Systems Research*, vol. 26, no. 1, pp. 1–10, 2021.
- [12] L. Bozhkov and P. Georgieva, "Overview of Deep Learning Architectures for EEG-Based Brain imaging," in *Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, IEEE, Rio de Janeiro, Brazil, July 2018.
- [13] H. Chao, L. Dong, and Y. Liu, "Improved deep feature learning by synchronization measurements for multi-channel EEG emotion recognition," *Complexity*, vol. 2020, Article ID 6816502, 15 pages, 2020.
- [14] M. Khateeb, S. M. Anwar, and M. Alnowami, "Multi-domain feature fusion for emotion classification using DEAP dataset," *IEEE Access*, vol. 9, Article ID 12134, 2021.
- [15] M. Y. Shen, *Research on Emotion Classification Methods Based on Neural Network and EEG signals*, Jilin University, Changchun, China, 2021.
- [16] B. Garcia-Martinez, A. Martinez-Rodrigo, R. Alcaraz, and A. Fernandez-Caballero, "A Review on Nonlinear Methods Using Electroencephalographic Recordings for Emotion Recognition," *IEEE Transactions on Affective Computing*, vol. 12, no. 3, pp. 801–820, 2021.
- [17] F. Demir, N. Sobahi, S. Siuly, and A. Sengur, "Exploring Deep Learning Features for Automatic Classification of Human Emotion Using EEG Rhythms," *IEEE Sensors Journal*, vol. 21, no. 13, Article ID 14923, 2021.
- [18] L. Tong, J. Zhao, and W. Fu, "Emotion Recognition and Channel Selection Based on EEG Signal," in *Proceedings of the 2018 11th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, pp. 101–105, IEEE, Changsha, China, September 2018.
- [19] X. Li, *Research of Emotion Recognition Based on Multi-Channel Brain Neural signals*, Tianjin University, Tianjin, China, 2019.
- [20] H. Jiang, R. Jiao, Z. Wang, T. Zhang, and L. Wu, "Construction and Analysis of Emotion Computing Model Based on LSTM," *Complexity*, vol. 2021, Article ID 8897105, 12 pages, 2021.
- [21] B. Pan and W. Zheng, "Emotion recognition based on EEG using generative adversarial nets and convolutional neural network," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 2520394, 11 pages, 2021.
- [22] S. Hwang, K. Hong, G. Son, and H. Byun, "Learning CNN features from DE features for EEG-based emotion recognition," *Pattern Analysis and Applications*, vol. 23, no. 3, pp. 1323–1335, 2020.