Journal of Network Intelligence ©2022 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **7**, Number **3**, August **2022**

An Improved Bi-LSTM EEG Emotion Recognition Algorithm

Shuai Ma

School of Computer and Information Engineering
Xiamen University of Technology
Xiamen,361024,China
2022031489@s.xmut.edu.cn

Jianfeng Cui*

College of Software Engineering
Xiamen University of Technology
Xiamen,361024,China
jfcui@t.xmut.edu.cn

Chin-Ling Chen*

1 School of Computer and Information Engineering,
Xiamen University of Technology, Xiamen 361024, China.
2 School of Information Engineering,
Changchun Sci-Tech University, Changchun 130600, China.
3 Department of Computer Science and Information Engineering,
Chaoyang University of Technology, Taichung 41349, Taiwan.
clc@mail.cyut.edu.tw

Weidong Xiao

College of Software Engineering
Xiamen University of Technology
Xiamen,361024,China
2013112101@xmut.edu.cn

Lijuan Liu

School of Computer and Information Engineering
Xiamen University of Technology
Xiamen,361024,China
ljliu@xmut.edu.cn

*Corresponding author: Jianfeng Cui, Chin-Ling Chen

Received April 9, 2022, revised May 11, 2022, accepted July 27, 2022.

Abstract. *EEG emotion signal feature extraction is computationally intensive and the classification accuracy of the model is not high. Therefore, it can seriously affect the overall performance of classification algorithms. In order to better apply deep learning methods to EEG emotion recognition classification. To address the above problems, this paper uses the Fast Fourier Transformation algorithm to extract richer and more complete multi-channel feature information in feature extraction. At the same time, for the problem of low classification accuracy, an improved Bi-LSTM classification model is proposed. Finally, the model proposed in this paper is verified by 32-channel EEG signal data. The experimental results show that the classification method proposed improves the accuracy of Arousal to 93.35% and Valence to 93.09%. Compared with other models, the improved Bi-LSTM model has an excellent performance in EEG emotion recognition. Meanwhile, this paper explores the brain emotion mechanism by the accuracy of the improved model for different brain region channels, showing that the brain regions with the highest emotional relevance are the parietal and frontal lobes in the brain.*
Keywords: EEG Emotion Recognition, Deep Learning, Fast Fourier Transformation, Bi-LSTM

1. **Introduction.** Emotion is a general term for a series of subjective cognitive experiences and is a psychological and physiological state produced by a combination of multiple feelings, thoughts, and behaviors. In people's daily work and life, the role of emotions is everywhere. A good emotional state is conducive to maintaining physical and mental health, while long-term negative emotions have a great impact on people's mental and physical health. Emotion recognition has important application prospects in the medical field, traffic safety, etc. [1, 2]. In the medical field [3], the emotional state of a patient has a great influence on the body. Long-term negative emotions can easily lead to depression, affect people's social functions and interpersonal communication, and even threaten life safety; for patients with cardiovascular and cerebrovascular diseases, extreme emotions such as anger and anxiety will increase the risk of disease; in the field of safe transportation [4], emotion recognition technology can be used to ensure the safety of the driver to drive the vehicle. The anger generated by the driver during driving can easily lead to road rage, which seriously affects the life safety of the driver and pedestrians on the road. It can be seen that emotions have an important impact on all aspects of human life. Therefore, it is particularly important to accurately identify emotions.

In recent years, there are two main methods of emotion recognition. One is to recognize non-physiological signals such as facial expressions, voice tones, and body postures [5]. Although non-physiological signals are easy to obtain, on some occasions, people can artificially control non-physiological signals through camouflage and other means, resulting in the inability to obtain real emotional signals and thus unable to accurately identify the real emotional state. Therefore, researchers tend to use physiological signals such as brain point signal (EEG), eye electrical signal (EOG), electrocardiogram (ECG), electromyography (EMG), and other physiological signals to conduct emotion recognition research [6, 7, 8]. Of all these physiological signals, EEG signals are of more interest to researchers. EEG signal is a method of recording brain activity using electrophysiological indicators. When the brain is active, it records the changes in electrical waves during brain activity, which is the overall reflection of the electrophysiological activity of brain nerve cells on the surface of the cerebral cortex or scalp. Compared with other physiological signals, EEG signals (EEG) can more truly and reliably reflect human emotional states [9].

Recognition and classification Model is an important part of EEG emotion recognition. The main task is to determine the EEG models corresponding to different emotional states by extracting various types of EEG features, and then classify the untrained EEG signal features. Choosing a good classification model is crucial for emotion recognition, which can effectively improve the accuracy of emotion classification. At present, the commonly used EEG emotion recognition and classification methods mainly include machine learning and deep learning. With the rapid development of artificial intelligence, machine learning has emerged as a potential classification method [10]. The machine learning method first pre-processes the EEG data and then extracts EEG features through wavelet transform [11], linear discriminant analysis[12], principal component analysis [13], and other methods. Finally, the extracted EEG features are input into the classifier to complete the classification. Commonly used classifiers include Self-Organizing Map (SOM) [14], Support Vector Machines (SVM) [15], and K-Nearest Neighbor (KNN) [16]. However, the classification model of machine learning needs to spend a lot of time on feature extraction and feature selection. When processing high-dimensional data, the amount of calculation will be very large, resulting in very slow classification speed and low classification accuracy. It is not suitable for applications that require fast forecasting. In order to overcome the shortcomings of machine learning methods, more and more scholars apply deep learning methods to the research of emotion recognition based on EEG signals. Deep learning classification models commonly used in the field of EEG emotion recognition include Convolutional Neural Network (CNN) [17], Long Short-Term Memory (LSTM) [18], and Deep Neural Network [19]. The deep learning method can perform end-to-end automatic learning for EEG signal pre-processing, feature extraction, and classification. EEG emotion recognition based on a deep learning algorithm has a higher feature dimension and shorter recognition time than artificially designed features, which can be a better assisted medical diagnosis.

In summary, for EEG emotional signal feature extraction is computationally intensive, slow and classification accuracy is not high. In this paper, we extract richer and more complete multi-channel feature information by using the FFT method to effectively improve the efficiency of feature extraction. Then an improved Bi-LSTM model is used for EEG emotion recognition classification, which has the advantage of bi-directional information, mainly in the temporal input, retaining the information of each step to dynamically adjust to the next step and improving the performance of model classification. The experimental results show that the model effectively improves the accuracy of emotion recognition classification. In addition, this paper explores the brain emotion mechanism by improving the accuracy of the model to different brain regions. The results show that the regions with the highest emotional correlation in the brain are the parietal lobe and frontal lobe.

2. **Related Work.** CNN and LSTM are the two most commonly used deep learning neural network models. CNN can be applied in deep learning for visualization [20]. LSTMs are scalable models that can efficiently solve multiple learning problems associated with sequential data [21]. Li et al. [22] used discrete wavelets to extract features from EEG, and then combined them with long-term short-term memory neural network to use the extracted features to train CNN, and the accuracy rate in both Arousal and Valence reached 72%. Choi et al. [23] used the raw EEG data to be directly input into the LSTM model for classification, and the experimental results showed that the accuracy of Arousal was 74.65%, and the accuracy of Valence was 78%. Zhan et al. [24] used STFT to extract power spectral features from four frequency bands and designed a shallow deep parallel CNN inspired by the Mobilenet technique for learning spatial representations from labeled frames. The experimental results show that the accuracy rates of Arousal and Valence are 84.97% and 82.95%, respectively. Zhong et al. [25] proposed an attention mechanism-based salient region extraction method, aiming to enhance the network's ability to generate representations by modeling the interdependencies between the networks

and finally classify them through CNN. Experimental The results show that Arousal has an accuracy of 68.5% and Valence has an accuracy of 66.23%. Three datasets were used in the research work of Cimtay et al. [26]: DEAP, SEED, and LUMED. The experiment directly uses the original EEG signal to train the CNN model, and the accuracy rate on the DEAP dataset is 72.81%. Sharma et al. [27] decomposed EEG signal memory DWT and used the LSTM-based deep learning method to classify emotional signals, and the accuracy of the proposed algorithm reached 82.01%. Cimtay et al. [28] used two types of CNN models, one-path CNN, and two-path, using a 4-fold cross-validated CNN model. However, they train directly with raw EEG signals without feature selection, and the classification accuracy can reach 91.5%. The research overview of EEG signal emotion recognition is shown in Table 1.

All of the above research results are worthy of study. Since EEG is a non-smooth and non-linear random signal, but the above research work did not enable a richer and more complete extraction of multi-channel EEG feature information in the data pre-processing stage, which in turn led to a low classification accuracy of the classification model designed in the study for emotion recognition. Therefore, this paper constructs a new combinatorial model combining FFT with an improved Bi-LSTM, a combination of methods that has not been done before. The combined model uses the FFT algorithm in the data preprocessing stage, which can transform the signal pass in the EEG time domain to the frequency domain, reflecting the characteristics of the EEG signal in the frequency change and more intuitive understanding of the characteristics of each frequency in the EEG. The result is a richer and more complete multi-channel EEG feature extraction of the EEG signal. At the same time, in the design of the classification model, the temporal dynamics of the EEG signal, which is crucial to the recognition of emotion recognition, is taken into account. In this paper, an improved Bi-LSTM emotion classification model is designed. This model, compared with the traditional LSTM model, has forward and backward factors to jointly determine the results of emotion classification, which makes the accuracy of classification effectively improved.

Table 1. Overview of EEG Signal Emotion Recognition Research.

Reference	Year	Dataset	Feature Extraction	Classifier	Performance
Li et al.[22]	2016	DEAP	DWT	CNN	Arousal : 72% Valence : 72%
Choi et al.[23]	2018	DEAP	Raw Data	LSTM	Arousal : 74.65% Valence : 78%
Zhan et al.[24]	2019	DEAP	STFT	CNN	Arousal : 84.97% Valence : 82.95%
Zhong et al.[25]	2020	DEAP	DE	CNN	Arousal : 68.5% Valence : 66.23%
Cimtay et al.[26]	2020	DEAP	Raw Data	CNN	72.81%
Sharma et al.[27]	2020	DEAP	DWT	CNN	82.01%
Cimtay et al.[28]	2020	DEAP	Raw Data	CNN	91.5%

3. **Materials and Methods.** The overall process of EEG-based emotion recognition classification starts with collecting EEG data and then cleaning the raw EEG data using signal pre-processing techniques. Next, these cleaned raw data are segmented in the EEG feature extraction step. Finally, these extracted EEG features are used to train the model to obtain better classification results. The overall flowchart of EEG-based emotion recognition classification is shown in Figure 1 below.

Figure 1. Overall flow chart of EEG emotion recognition and classification.

3.1. **DEAP Database.** The DEAP dataset [29] is a large, open-source dataset that contains multiple physiological signals with emotional assessments for detecting human emotional states. First, 1000 music clips were selected to stimulate the subjects emotionally, and then 40 music videos were selected from the 1000 videos according to the subjects' evaluations as the official stimulus source for this dataset. So this dataset detected and recorded EEG, ECG, EMG, EEG, electromyography evoked by 32 subjects (16 males and 16 females) watching 40 music videos with different emotional tendencies of about 1 minute in duration 40 channels of mixed physiological signals, such as electrooculography, etc. After watching the video, the subjects used a continuous value of 1-9 to evaluate the video in terms of Arousal, Valence, Dominance, and Liking. The evaluation value is determined by Small to large indicating that each indicator is from negative to positive, from weak to strong. The EEG signals in each experiment formed a matrix of size $32\times32\times40\times8064$ according to the format of [tester, EEG channel, test stimulus, EEG signal data]. A summary of the DEAP dataset is shown in Table 2.

Table 2. Synopsis of the DEAP dataset.

Types of datasets	Multimodal dataset
No. of participant	32
No. of EEG channel	32
Data collection method	Showing one-minute-long excerpts of music videos
No. of a used data collection resource	40 music videos
Sampling rate	128HZ
Rating values	Continuous scale 1-9
Rating scales	Arousal, Valence, Dominance, and Liking

During the acquisition of the DEAP dataset, 32 EEG-related electrodes out of 128 electrodes were selected for acquisition, with a signal sampling frequency of 512 Hz. Since the correlation between similar EEG signals is relatively strong, using all 128 electrodes would result in an overlap of the acquired signals. The data set is referenced to the international standard location of 32 electrodes, which are evenly distributed in the brain space and contain information from each brain region. The correspondence between each electrode and channel number is shown in Table 3, and the distribution of electrodes is shown in Figure 2.

3.2. **EEG Signal Preprocessing.** In this paper, EEG signals on 32 channels from 32 subjects viewing 40 videos in the DEAP dataset were used as the experimental dataset, where EEG signals were first downsampled to 128 Hz in order to collect accurate data content between 0-48 Hz. EMG and EEG data were then removed from the downsampled data. Incremental waves in the signal are also separated from the analysis process using a band-pass filter. Finally, the EEG signal is de-common averaged to improve the correlation between the subject's subjective emotional evaluation and the EEG signal in the data. Each recorded EEG data was divided into 60s segments and a 3s pretest baseline was removed, and the baseline signal was subtracted from the 63s signal to obtain the stimulus-related signal changes.

Table 3. EEG channels used in DEAP dataset.

Channels No.	Electrode Name	Channels No.	Electrode Name
1	Fp1	17	Fp2
2	AF3	18	AF4
3	F3	19	Fz
4	F7	20	F4
5	FC5	21	F8
6	FC1	22	FC6
7	C3	23	FC2
8	T7	24	Cz
9	CP5	25	C4
10	CP1	26	T8
11	P3	27	CP6
12	P7	28	CP2
13	PO3	29	P4
14	O1	30	P8
15	Oz	31	PO4
16	Pz	32	O2

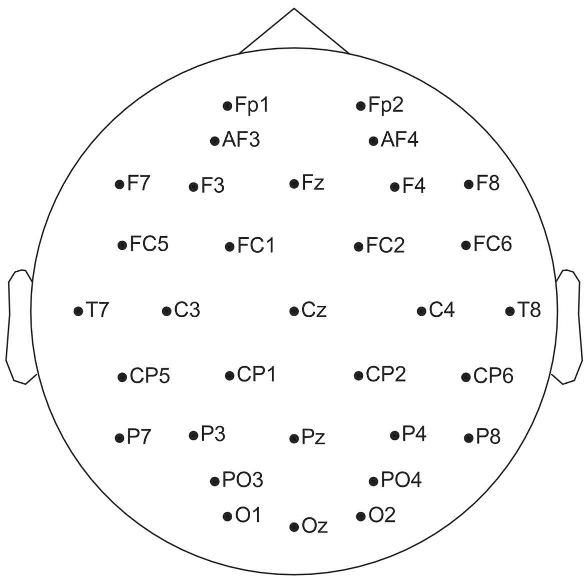

Figure 2. 32-channel EEG sensor location.

3.3. **Feature Extraction.** In the process of EEG-based emotion recognition research, feature extraction is mainly to reduce the dimension of EEG data to extract emotion-related features, which are used to study the emotional state of the subjects. As the key link of emotion recognition, the performance of the emotion recognition model is directly determined by the merit of features, and it is crucial to extract features with good representativeness and high correlation with emotion. Feature extraction is a way to compress the dimensionality of feature space by finding the most effective features for classification recognition from a large number of feature data, i.e., to obtain a set of "few but precise" features with a low probability of classification error. These features are then used for classification, which can improve the accuracy of classification. There are three main types of features for EEG signals: time-domain features, frequency domain features, and time-frequency domain features [30].

There are many different methods for feature extraction of current EEG signals, which include fast Fourier transform (FFT) [31], wavelet transform (WT)[32] and time-frequency distribution (TFD)[33], etc. FFT is a fast algorithm for discrete Fourier transform (DFT), which is a fast Fourier transform. It reduces the computational time as well as the computational complexity by calculating the coefficients making the computation easier to perform the DFT in an iterative manner. It also reduces the rounding error calculation associated with it. Since the Fourier transform can convert the time domain signal to the frequency domain signal and also decompose the brain point signal into frequency components, the sequence of the Fourier transform can be better calculated. Therefore, the most popular FFT algorithm is used in this paper. The algorithm is imported into a pre-processed EEG dataset through the Python module of PyEEG, which is formatted in NumPy. In this paper, the EEG features are reduced from (40,40,8064) to the final dimension of (42480,70) by using the FFT algorithm for feature extraction, which speeds up the training and provides better accuracy. The FFT feature extraction method is shown in Figure 3.

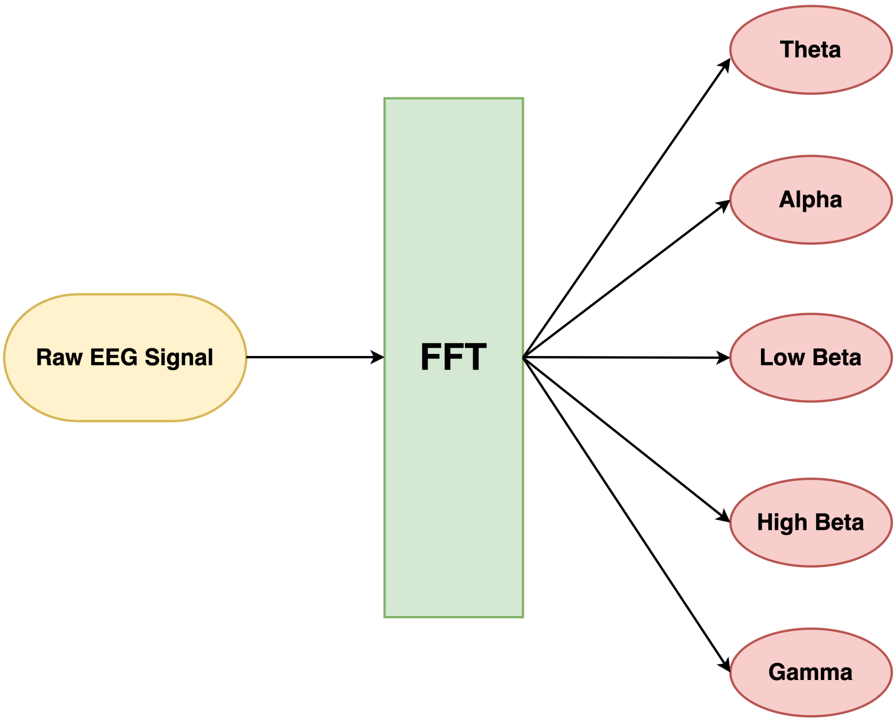

FIGURE 3. FFT feature extraction method.

The FFT time-domain analysis shows the variation of EEG waveform with time, while the frequency domain analysis shows the variation of EEG waveform with frequency. The main idea of frequency domain analysis is to transform the signal in the time domain to the frequency domain by some algorithm, reflecting the characteristics of the signal with frequency so that the distribution of each rhythm in the EEG can be observed more intuitively. The frequency-domain analysis divides the EEG signal into delta band (0-4Hz), theta band (4-8Hz), alpha (8-13Hz), beta (13-25Hz), and gamma band (25-50Hz) for feature extraction.

In Equation 1, the length of the EEG emotional signal sequence is set to be $N(N = 2M)$. When M is not an integer, several zero values can be added to the signal sequence. where $X(K)$ for $k = 0, 1, 2 \dots N - 1$ is the coefficient of the discrete Fourier and $x_i(n)$is the time domain input signal.

$$X(K) = \sum_{i=0}^{N-1} x_i(n) W_N^{kn} (\mathrm{n} = 0, 1, 2, \ldots; \mathrm{k} = 0, 1, 2, \ldots \mathrm{N} - 1), W_{N=} e^{\frac{-j2\pi ik}{N}} \tag{1}$$

Secondly, the 2FFT algorithm of decimation by time is used for the sequence after $X(K)$. First, the brain point emotional signal sequence is divided into two groups according to the parity of n, as shown in Equation 2 and 3.

$$X(2r) = x_1(r) \left(n = 0, 1, \ldots, \left(\frac{N}{2} \right) - 1 \right) \tag{2}$$

$$X(2r+1) = x_2(r) \left(n = 0, 1, \ldots, \left(\frac{N}{2} \right) - 1 \right) \tag{3}$$

Then, Equation 2 and 3 are respectively substituted into formula 1 and simplified to obtain formulas 4 and 5.

$$X(K) = X_1(K) = W_N^k X_2(k) \left(n = 0, 1, \ldots, \left(\frac{N}{2} \right) - 1 \right) \tag{4}$$

$$X\left(k + \frac{K}{2} \right) = X_1(K) - W_N^k X_2(K) \left(k = 0, 1, \ldots, \left(\frac{N}{2} \right) - 1 \right) \tag{5}$$

Finally, assuming that the length of the EEG emotional signal is limited, the formula of its FFT is shown in Equation 6, where $W_{N=} e^{\frac{-j2\pi ik}{N}}$ denotes the rotation factor.

$$X_1(k) = \sum_{r=0}^{\frac{N}{2}-1} x_1(r) W_{N/2}^{kr} \quad X_2(k) = \sum_{r=0}^{\frac{N}{2}-1} x_2(r) W_{N/2}^{kr} \tag{6}$$

3.4. **An Improved Bi-LSTM Classification Model.** LSTM is a special kind of recurrent neural network (RNN). The structural defects of the traditional conventional recurrent neural network lead to problems such as gradient disappearance due to excessive sequence length [34]. To address these problems, the LSTM improves the structure of the original network and adds forgetting gates, input gates, and output gates, thus enabling better results in handling long-term information [35].The input of LSTM includes not only the current input X_t, but also the output H_{t-1} of the previous moment H_t and the cell state C_t of the hidden layer, which together constitute the real input at this time, and the output H_t and the cell state C_t are calculated to obtain the real input at this time. The output H_t and the cell state C_t are calculated and used as the input for the next moment. The forgetting gate is responsible for deciding whether to save or discard unnecessary information, the input gate is used to update the cell state, and the output gate is used to obtain the output H_t from the cell state C_t and the computation at this time. The structure diagram of LSTM is shown in Figure 4.

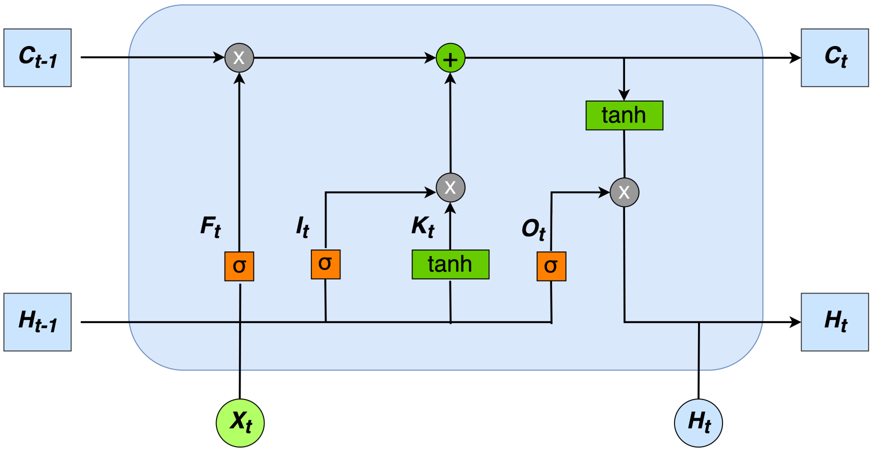

FIGURE 4. Structure diagram of LSTM.

The calculation Equations 7, 8, 9, 10, and 11 of each gate of LSTM are as follows:

$$F_t = \sigma\left(w_f\left[H_{t-1}, X_t\right] + b_f\right) \tag{7}$$

$$I_t = \sigma\left(w_i\left[H_{t-1}, X_t\right] + b_i\right) \tag{8}$$

$$O_t = \sigma\left(w_o\left[H_{t-1}, X_t\right] + b_o\right) \tag{9}$$

$$K_t = \tanh\left(w_k\left[H_{t-1}, X_t\right] + b_k\right) \tag{10}$$

$$C_t = F_t * C_{t-1} + I_t * K_t \tag{11}$$

In the formula, and are the sigmoid functions and tanh functions respectively, is the cell state information of the memory cell, w and b are the weight and bias terms respectively, and represents the forgetting gate, input gate, output gate and cell state respectively, and is the accumulation of information at the current moment.

The output of the unidirectional LSTM is only related to the input of the current moment and the input of the last moment, but there are some specific scenarios where the output of the current moment should also be related to the input of the future moment, and such scenarios require a joint forward LSTM and an inverse LSTM. compared to the unidirectional LSTM, the bidirectional LSTM is better at solving long-time gradient explosions than the unidirectional LSTM. The network structure of the bidirectional long short-term memory neural network is shown in Figure 5.

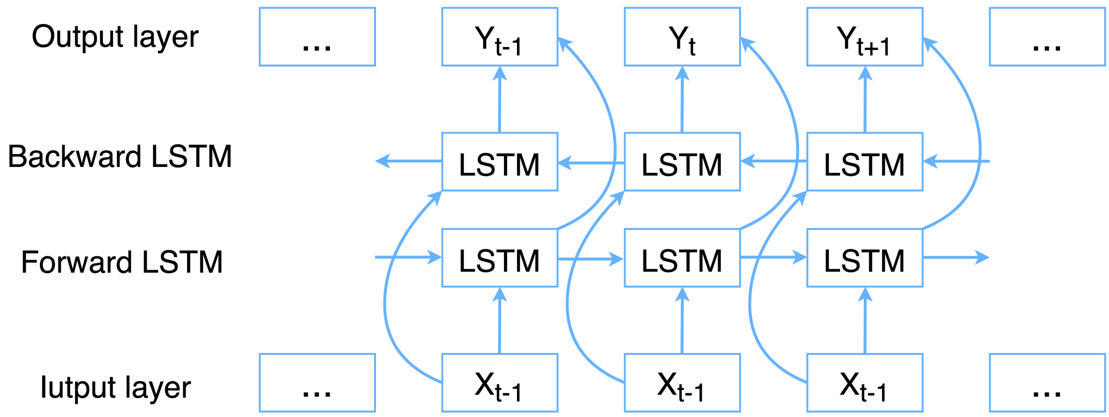

FIGURE 5. Structure diagram of Bi-LSTM.

The output of each time t is determined by the forward LSTM output and the reverse LSTM output. The forward LSTM propagation and the reverse LSTM propagation are similar to the one-way LSTM propagation. The update Equation 12 of the forward LSTM network, the update Equation 13 of the reverse LSTM network, and the Equation 14 of the output of the superimposed LSTM network of two different directions.

$$\overrightarrow{h_t} = W_{X\overrightarrow{h_t}} x_t + W_{\overrightarrow{h_t}\overrightarrow{h_t}} \overrightarrow{h_{t-1}} + b_{\overrightarrow{h_t}} \tag{12}$$

$$\overleftarrow{h_t} = W_{X\overleftarrow{h_t}} x_t + W_{\overleftarrow{h_t}\overleftarrow{h_t}} \overleftarrow{h_{t-1}} + b_{\overleftarrow{h_t}} \tag{13}$$

$$y_t = W_{\vec{h}y} \vec{h} + W_{\overleftarrow{h}y} \overleftarrow{h} + b_y \tag{14}$$

In summary, this paper combines the advantages of Bi-LSTM to solve the gradient explosion and the advantages of LSTM to eliminate the gradient disappearance and proposes an improved Bi-LSTM model with one Bi-LSTM layer, four LSTM layers and two Dense layers in the model architecture, and the model is shown in Figure 6. In this paper, the feature extraction is performed by using the FFT algorithm for 32-lead EEG data, and then the EEG features are input into the first Bi-LSTM layer, and the input layer EEG features of the Bi-LSTM are $I = [l, m, n]$, where l denotes the batch size,m denotes the number of channels, and n denotes the number of single-channel features. the EEG features are classified into the bi-directional LSTM for independent calculation. The output state is calculated as shown in Equation 15. h_1 represents the output state of the forward LSTM, h_2 represents the output state of the backward LSTM, W_1 is the forward output weight matrix, and W_2 is the backward output matrix.

$$Y_t = \sigma\left(W_1 h_1 + W_2 h_2\right) \tag{15}$$

There are 128 cells in the Bi-LSTM layer (256 in total), which will then replicate the first LSTM layer in the network, supplying the input sequence input to the first and back-propagating it to the second. To prevent overfitting of the data, we set the value of dropout to 0.3. The next 4 layers consist of 4 LSTM layers, and the value of dropout is set uniformly to 0.2 for all 4 LSTM layers, and the activation function is ReLU. where the first layer is a 256-neuron LSTM layer, the second and third layers are 64-neuron LSTM layers of 0.5, and the fourth layer is a 32-neuron LSTM layer. Then input to a 32-neuron Dense layer and a 16-neuron Dense layer, the association between features is extracted through the two Dense layers by nonlinear transformation, the multi-channel EEG sentiment features are extracted by using the modeling ability of LSTM on sequences, and finally the classification task is completed by Softmax activation function.

4. **Results.**

4.1. **Experiment platform.**
The experiments in this paper use the Python 3.8 programming language and the Pytorch deep learning framework. The operating system is 64-bit Linux, the CPU is Intel(R) Xeon(liceLake) Platinum 8369R @ 2.90GHz, the GPU is NVIDIA A10, and the video RAM is 24 GB.

4.2. **Evaluation method.**
In order to evaluate the performance of the model classification, this paper evaluates the classification effect of the model by *Acc* (Accuracy). Accuracy is the most intuitive evaluation index to measure the classification effect, and the formula of accuracy is shown in Equation 16, where TP (True Positive) refers to the number of results correctly predicted as positive samples, FP (False Positive) refers to the number of results correctly predicted as positive samples. The number of results incorrectly predicted as positive samples, TN (True Negative) is the number of results correctly predicted as negative samples, and FN (False Negative) is the number of results incorrectly predicted as negative samples. Take Arousal class in EEG as an example, where TP is the number of samples that will be correctly classified as Arousal class, TN

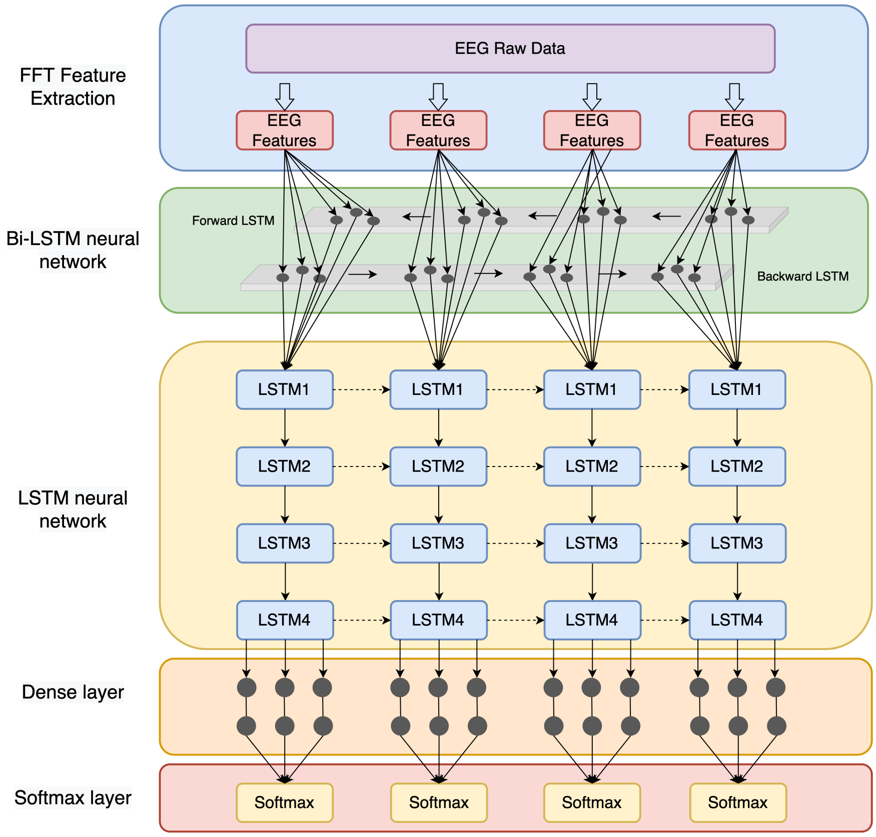

Figure 6. An Improved Bi-LSTM Model.

is the number of Arousal samples that are incorrectly classified as other samples, FP denotes data samples that are not Arousal class samples but are classified as Arousal class, and FN denotes data samples that are not Arousal class and are not classified as Arousal class data samples.

$$ACC = \frac{TP + TN}{TP + TP + FP + FN} \tag{16}$$

4.3. **Feature Extraction and Classification Result.** In this paper, feature extraction is performed by using the FFT algorithm. In the process of feature extraction, when the EEG feature extraction is too intensive, the classification model will suffer from an overfitting problem because the features are too similar. When the EEG feature extraction is too sparse, the classification model will suffer from under-learning due to the small number of features, and all of the above situations will lead to poor classification accuracy. In order to improve the classification accuracy, we need to choose the appropriate window size and Batch size, where window size refers to the length of a sliding data sequence and Batch size refers to the number of data samples crawled in one training. Therefore, we compare different values of the hyperparameters' window size and Batch size in the algorithm to determine the specific values of the hyperparameters and find the best accuracy rate. The accuracy rates of different hyperparameters are shown in Table 4. step size is uniformly set to 16, and when the window size is 128 and Batch size is 128, the accuracy rate of Arousal is 81.99% and that of Valence is 82.22%. when Batch size is 256, the accuracy rate of Arousal is the accuracy rate of Arousal is 83.68% and the accuracy rate of Valence is 84.72%. When the window size is 256 and the Batch size is 128, the accuracy rate of Arousal is 87.23% and the accuracy rate of Valence is 86.75%. When the Batch size is 256, the accuracy rate of Arousal is 88.29% and the accuracy rate of Valence is 88.17%. When the window size is 512 and the Batch size is 128, the accuracy of Arousal is 91.9% and the accuracy of Valence is 90.01%. Batch size is 256, the accuracy of Arousal is 93.35% and the accuracy of Valence is 93.09%. Batch size is 512, the accuracy of Arousal is 91.46% and the accuracy of Valence is 90.77%. The comparison graph of the accuracy of different hyperparameters is shown in Figure 7. It can be seen that when the EEG feature selection is too sparse, the classification model will be under-learning due to the small number of features, which will lead to a low accuracy of classification. When the EEG features are selected too densely, the classification model appears to be overfitted because the EEG signal features are too similar, which leads to a low classification accuracy. Therefore, the model performs the optimal classification performance when the window size is 512 and the batch size is 256. Figure 8 represents the accuracy and loss rate variation of the 32-channel EEG emotion classification under the optimal performance.

Table 4. Accuracy of features with different hyperparameters.

Window size	Step size	Batch size	Arousal	Valence
128	16	128	81.99%	82.22%
128	16	256	83.68%	84.72%
256	16	128	87.23%	86.75%
256	16	256	88.29%	88.17%
512	16	128	91.9%	90.01%
512	16	256	93.35%	93.09%
512	16	512	91.46%	90.77%

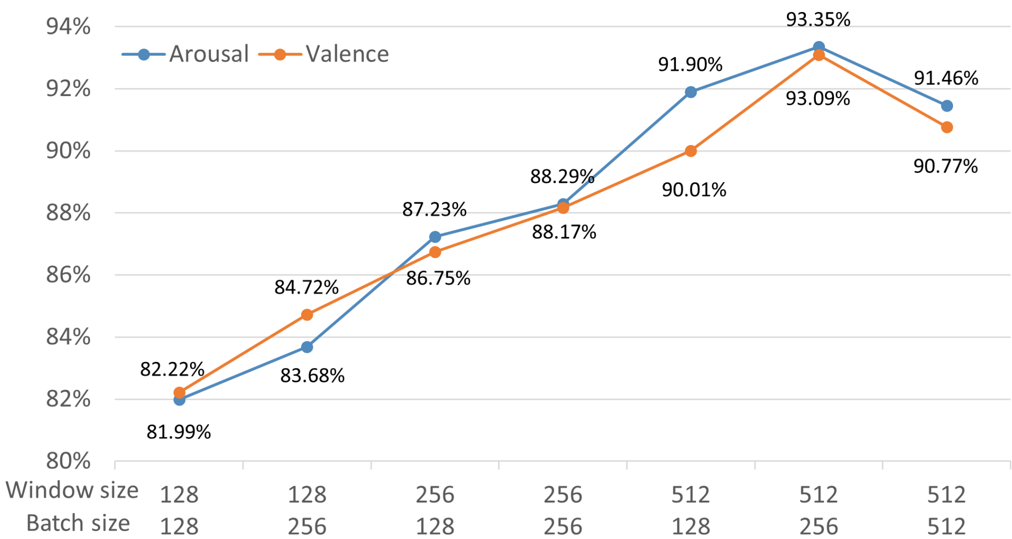

Figure 7. The comparison graph of the accuracy of different hyperparameters.

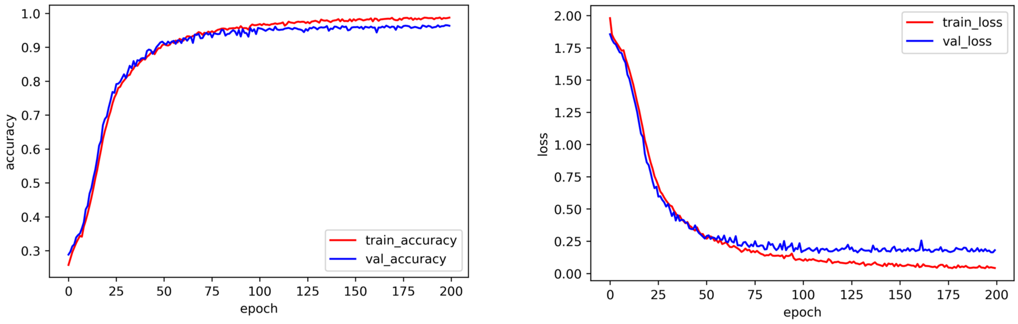

FIGURE 8. Accuracy and loss rate graphs for 32-channel data.

This paper summarizes some previous classification results for EEG-based emotion recognition. choi et al.[23] used the raw EEG data to classify through LSTM. The experimental results showed that the accuracy of Arousal classification was 74.65% and the accuracy of Valence classification was 78%.Zhan et al.[24] used STFT to extract power spectrum features from four frequency bands and designed a Mobilene technique inspired by shallow depth parallel CNN for learning spatial representation from labeled frames. The experimental results showed 84.97% and 82.95% accuracy for Arousal and Valence, respectively. Liu et al. [32] used the original dataset for classification using BDAE and achieved 85.2% and 80.5% accuracy for Arousal and Valence, respectively. Yang et al.[33] performed feature extraction by using DE followed by CNN The accuracy of performing classification Arousal and Valence were 90.24% and 89.45%, respectively. Yang et al.[34] finally improved the results to 91.03% and 90.8% by combining PSD and ED for feature extraction and performing classification by CNN. In this paper, the FFT algorithm is used in the data preprocessing stage to transform the signal in the EEG time domain to the frequency domain by the FFT algorithm to complete a richer and more complete multi-channel EEG feature extraction of the EEG signal. Meanwhile, the design of the classification model takes into account the temporal dynamics in the EEG signal which is crucial for the recognition of emotion recognition. In this paper, an improved Bi-LSTM emotion classification model is designed, which improves the accuracy to 93.35% for Arousal classification and 93.09% for Valence, compared with the traditional LSTM model, which jointly determines the results of emotion classification by the forward and backward factors. Table 5 and Figure 9 summarize the comparison of the existing literature results.

TABLE 5. Comparison of the classification effect of this paper with other methods.

Authors	Feature Extraction	Classifier	Performance	
			Arousal	Valence
Choi et al.[23]	Raw Data	LSTM	74.65%	78%
Zhan et al.[24]	STFT	CNN	84.97%	82.95%
Liu et al.[32]	Raw Data	BDAE	85.2%	80.5%
Yang et al.[33]	DE	CNN	90.24%	89.45%
Yang et al.[34]	PSD,ED	CNN	91.03%	90.80%
Our method	FFT	Bi-LSTM	93.35%	93.09%

In order to better explore the EEG regions with the highest emotional correlation. In this paper, different EEG regions are selected as input to the model for classification.Table

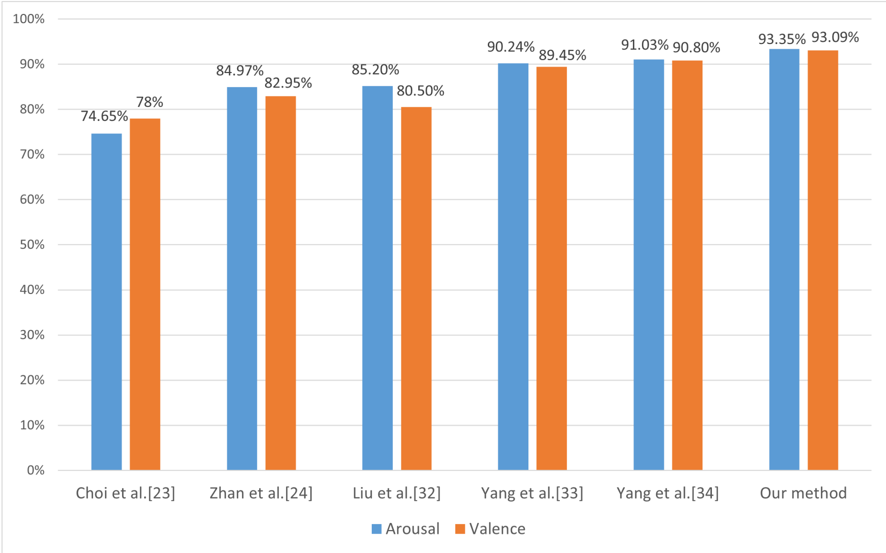

FIGURE 9. Performance comparison between relevant approaches.

6 describes the accuracy of Arousal and Valence classification for different channel numbers. When all channels are selected for input to the model classification, the accuracy of Arousal and Valence is not very high because the features are too redundant and the training time increases. When 15 channels are selected as Fp1, Fp2, AF3, AF4, F7, F3, Fz, F4, F8, FC1, FC2, P3, P4, PO3, PO4, the accuracy rate is higher, and the accuracy rate of Arousal is 95.76%, Valence The accuracy rate is 96.22%. This shows that the emotional characteristics of the parietal and frontal lobes of the brain are more obvious. The 15 channels in the parietal and frontal lobes are shown in Figure 10. The graphs of accuracy and loss rate changes for 15-channel EEG emotion classification are shown in Figure 11.

TABLE 6. Accuracy of different channels.

Channel selection	Arousal	Valence
32 channels	93.35%	93.09%
Fp1,Fp2,AF3,AF4,F7,F3,Fz,F4,F8,FC1,FC2,P3,P4,PO3,PO4	95.76%	96.22%

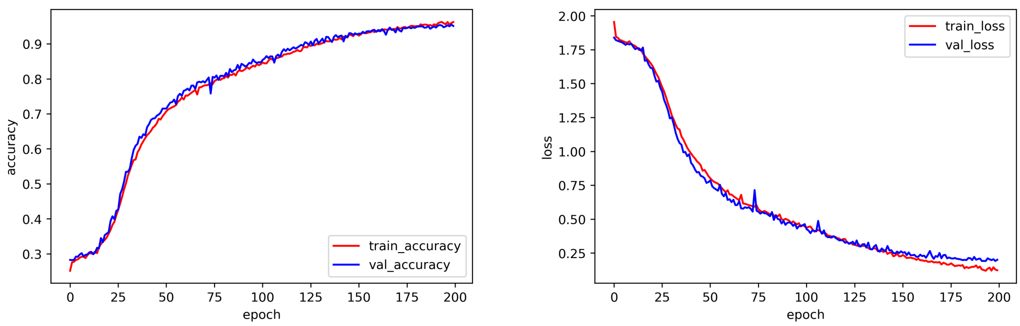

FIGURE 11. Accuracy and loss rate graphs for 15-channel data.

5. **Conclusions.** In this paper, we propose an improved Bi-LSTM emotion recognition model to classify EEG emotions in response to the problems of slow feature extraction

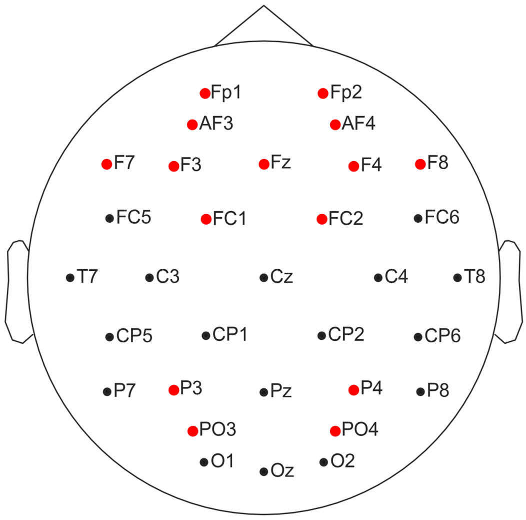

FIGURE 10. 15-channel EEG sensor location.

and low classification accuracy in EEG. This paper first extracts the EEG signal features by FFT algorithm, which can reduce the computation time and computation complexity of feature extraction, and transforms the signal in the EEG time domain to the frequency domain by FFT algorithm to complete a richer and more complete multichannel EEG feature extraction of the EEG signal. The extracted features are then fed into the improved Bi-LSTM model for training and classification. Finally, the proposed model is validated by 32 channels of EEG signal data. The experimental results show that the classification model proposed in this paper improves the accuracy of Arousal and Valence to 93.35% for Arousal and 93.09% for Valence. In terms of brain regions, the accuracy of EEG signals of 32 channels and 15 channels were selected for comparison, and the experimental results showed that the accuracy was higher when 15 channels were selected, and the accuracy of Arousal was 95.76% and that of Valence was 96.22%, which showed that the emotional correlation between the parietal and frontal lobes of the brain was higher and the emotional characteristics were more obvious. In the future, we hope to apply the designed model to different datasets and improve the accuracy of emotion recognition. We also want to design a lightweight and highly accurate emotion recognition system. For example, doctors can use this system to better assess the mood of depressed patients and find appropriate treatment plans.

Acknowledgment. This work was partly supported in part by the National Natural Science Foundation of China (No.62103345), Fujian Provincial Natural Science Foundation of China (No. 2020J02160), Xiamen Youth Innovation Fund Project (No. 3502Z20206076), Xiamen Major Science and Technology Project (No. 3502Z20221024) and The High-Level Talents Research Launched Project of Xiamen University of Technology (No. YKJ19012R).

REFERENCES

[1] A. Kallipolitis, M. Galliakis, A. Menychtas, I. Maglogiannis, "Affective analysis of patients in home-care video-assisted telemedicine using computational intelligence," *Neural Computing and Applications*, vol. 32, no. 23, pp. 17125-17136, 2020.

[2] Z. Halim, M. Rehan, "On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning," *Information Fusion*, vol. 53, pp. 66-79, 2020.

[3] H.Y. Huang, Q.Y. Xie, J.H. Pan, Y.B. He, Z.F. Wen, R.H. Yu, Y.Q. Li, "An EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness," *IEEE Transactions on Affective Computing*, vol. 99, pp. 1–1, 2019.

[4] W.H. Liu, J.H. Qian, Z.W. Yao, X.T. Jiao, J.H. Pan, "Convolutional two-stream network using multi-facial feature fusion for driver fatigue detection," *Future Internet*, vol. 11, no. 5, pp. 115, 2019.

[5] L. Kessous, G. Castellano, G. Caridakis, "Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis," *Journal on Multimodal User Interfaces*, vol. 3, no. 1, pp. 33-48, 2010.

[6] Z.k. Gao, Y.L. Li, Y.X. Yang, X.M. Wang, N. Dong, H-D. Chiang, "A GPSO-optimized convolutional neural networks for EEG-based emotion recognition," *Neurocomputing*, vol. 380, pp. 225-235, 2020.

[7] G. Chen, X. Zhang, Y. Sun, J. Zhang, "Emotion feature analysis and recognition based on reconstructed EEG sources," *IEEE Access*, vol. 8, pp. 11907-11916, 2020.

[8] Z. Wang, X. Zhou, W. Wang, C. Liang, "Emotion recognition using multimodal deep learning in multiple psychophysiological signals and video," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 4, pp. 923-934, 2020.

[9] M.H. Black, N.T.M. Chen, K.K. Iyer, O.V. Lipp, S. Bölte, M. Falkmer, T. Tan, S. Girdler, "Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography," *Neuroscience & Biobehavioral Reviews*, vol. 80, pp. 488-515, 2017.

[10] M.E. Wu, J.H. Syu, C.M. Chen, "Kelly-Based Options Trading Strategies on Settlement Date via Supervised Learning Algorithms," *Computational Economics*, vol. 59, no. 4, pp. 1627-1644, 2022.

[11] N. Ji, L. Ma, H. Dong, X.J. Zhang, "EEG signals feature extraction based on DWT and EMD combined with approximate entropy," *Brain sciences*, vol. 9, no. 8, pp. 201, 2019.

[12] R. Atangana, D. Tchiotsop, G. Kenne, L.C. DjoufackNkengfack, "EEG signal classification using LDA and MLP classifier," *Health Informat*, vol. 9, no. 1, pp. 14-32, 2020.

[13] M.A. Rahman, M.F. Hossain, M. Hossain, R. Ahmmed, "Employing PCA and t-statistical approach for feature extraction and classification of emotion from multichannel EEG signal," *Egyptian Informatics Journal*, vol. 21, no. 1, pp. 23-35, 2020.

[14] E.C. Djamal, D.P. Gustiawan, D. Djajasasmita, "Significant variables extraction of post-stroke EEG signal using wavelet and SOM kohonen," *Telkomnika*, vol. 17, no. 3, pp. 1149-1158, 2019.

[15] F.P. George, I.M. Shaikat, P.S. Ferdawoos, M.Z. Parvez, J. Uddin, "Recognition of emotional states using EEG signals based on time-frequency analysis and SVM classifier," *International Journal of Electrical & Computer Engineering* , vol. 9, no. 2, pp. 2088-8708, 2019.

[16] M. Li, H.P. Xu, X.W. Liu, S.F. Lu, "Emotion recognition from multichannel EEG signals using K-nearest neighbor classification," *Technology and health care*, vol. 26, no. S1, pp. 509-519, 2018.

[17] S.J. Oh, J.Y. Lee, D.K. Kim, "The design of CNN architectures for optimal six basic emotion classification using multiple physiological signals," *Sensors*, vol. 20, no. 3, pp. 866, 2020.

[18] X.B. Du, C.X. Ma, G.H. Zhang, J.Y. Li, Y.K. Lai, G.Z. Zhao, X.M. Deng, Y.J. Liu, H.G. Wang, "An efficient LSTM network for emotion recognition from multichannel EEG signals," *IEEE Transactions on Affective Computing*, pp. 1-1, 2020.

[19] A. Al-Nafjan, M. Hosny, Y. Al-Ohali, A. Al-Wabil, "Recognition of affective states via electroencephalogram analysis and classification," *International Conference on Intelligent Human Systems Integration. Springer, Cham*, vol. 722, pp. 242-248, 2018.

[20] L.A. Tawfeeq , S.S. Hussein, M.J. Mohammed, S.S. Abood, "Abood S S. Predication of Most Significant Features in Medical Image by Utilized CNN and Heatmap," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 12, no. 4, pp. 217-225, 2021

[21] 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, pp. 106954, 2021.

[22] X. Li, D.W. Song, P. Zhang, Y.X. Hou, B. Hu, "Emotion recognition from multi-channel EEG data through convolutional recurrent neural network," *2016 IEEE international conference on bioinformatics and biomedicine (BIBM)*, pp. 352-359, 2016.

[23] E.J. Choi, D.K. Kim, "Arousal and valence classification model based on long short-term memory and deap data for mental healthcare management," *Healthcare informatics research*, vol. 24, no. 4, pp. 309-316, 2018.

[24] Y. Zhan, M.I. Vai, S. Barma, S.H. Pun, J.W. Li, P.U. Mak, "A computation resource friendly convolutional neural network engine for EEG-based emotion recognition," *2019 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, pp. 1-6, 2019.

[25] X.L. Zhong, Z. Yin, J.H. Zhang, "Cross-subject emotion recognition from EEG using convolutional neural networks," *2020 39th Chinese Control Conference (CCC)*, pp. 7516-7521, 2020.

[26] 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, pp. 2034, 2020.

[27] R. Sharma, R.B. Pachori, P. Sircar, "Automated emotion recognition based on higher order statistics and deep learning algorithm," *Biomedical Signal Processing and Control*, vol. 58, pp. 101867, 2020.

[28] Y. Cimtay, E. Ekmekcioglu, S. Caglar-Ozhan, "Cross-subject multimodal emotion recognition based on hybrid fusion," *IEEE Access*, vol. 8, pp. 168865-168878, 2020.

[29] 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, 2011.

[30] D. Nath, M. Singh, D. Sethia, D. Kalra, S. Indu, "An efficient approach to eeg-based emotion recognition using lstm network," *2020 16th IEEE international colloquium on signal processing & its applications (CSPA)*, pp. 88-92, 2020.

[31] M. Akin, "Comparison of wavelet transform and FFT methods in the analysis of EEG signals," *Journal of medical systems*, vol. 26, no. 3, pp. 241-247, 2002.

[32] A.S. Al-Fahoum, A.A. Al-Fraihat, "Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains," *International Scholarly Research Notices*, vol. 2014, no. 7, 2014.

[33] B. Boashash, G. Azemi, J.M. O'Toole, "Time-frequency processing of nonstationary signals: Advanced TFD design to aid diagnosis with highlights from medical applications," *IEEE signal processing magazine*, vol. 30, no. 6, pp. 108-119, 2013.

[34] W.C. Kong, Z.Y. Dong, Y.W. Jia, D.J. Hill, Y. Xu, Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841-851, 2017.

[35] S. Kumar, A. Damaraju, A. Kumar, S. Kumari, C.M. Chen, "LSTM Network for Transportation Mode Detection," *Journal of Internet Technology*, vol. 22, no. 4, pp. 891-902, 2021.

[36] W. Liu, W.L. Zheng, B.L. Lu, "Emotion recognition using multimodal deep learning," *International conference on neural information processing*, pp. 521-529, 2016.

[37] Y. Yang, Q.F. Wu, Y.Z. Fu, X.W. Chen, "Continuous convolutional neural network with 3D input for EEG-based emotion recognition," *International Conference on Neural Information Processing*, vol. 11307, pp. 433-443, 2018.

[38] Y.L. Yang, Q.F. Wu, M. Qiu, Y.D. Wang, X.W. Chen, "Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network," *2018 international joint conference on neural networks (IJCNN)*, pp. 1-7, 2018.