SparNet: A Convolutional Neural Network for EEG Space-Frequency Feature Learning and Depression Discrimination

Xin Deng, Xufeng Fan*, Xiangwei Lv and Kaiwei Sun

Key Laboratory of Data Engineering and Visual Computing, College of Computer Science and Technology, Chongqing University of Posts and Telecommunication, Chongqing, China

Depression affects many people around the world today and is considered a global problem. Electroencephalogram (EEG) measurement is an appropriate way to understand the underlying mechanisms of major depressive disorder (MDD) to distinguish depression from normal control. With the development of deep learning methods, many researchers have adopted deep learning models to improve the classification accuracy of depression recognition. However, there are few studies on designing convolution filters for spatial and frequency domain feature learning in different brain regions. In this study, SparNet, a convolutional neural network composed of five parallel convolutional filters and the SENet, is proposed to learn EEG space-frequency domain characteristics and distinguish between depressive and normal control. The model is trained and tested by the cross-validation method of subject division. The results show that SparNet achieves a sensitivity of 95.07%, a specificity of 93.66%, and an accuracy of 94.37% in classification. Therefore, our results can conclude that the proposed SparNet model is effective in detecting depression using EEG signals. It also indicates that the combination of spatial information and frequency domain information is an effective way to identify patients with depression.

Keywords: SENet, SparNet, space-frequency domain characteristics, depression, EEG

1. INTRODUCTION

Major depressive disorder (MDD, also known as unipolar depression) is a physical disease of the brain, also known as a mental health disorder. It mainly affects the process of thought, behavior, and mood, and also can lead to the loss of interest and energy, interpersonal relationships, and job performance. According to the statistics of the World Health Organization, more than 300 million people in the world suffer from depression, and about 800,000 people die of depression every year (Belmaker and Agam, 2008; Olesen et al., 2012; Whiteford et al., 2013). Early and accurate diagnosis of depression is crucial for patients who need timely clinical treatment.

Most of the previous diagnoses of depression are based on the questionnaires as a judgment and screening tool. One of the major drawbacks of this method is that it requires experienced doctors. Therefore, finding a suitable and effective way to detect depression is an emerging research area. Currently, various physiological measurement tools are developing rapidly, such as functional magnetic resonance imaging (fMRI), electroencephalogram (EEG), and positron emission tomography (PET). Many studies attempt to

measure psychological data and develop auxiliary diagnostic methods in clinical practice (Van Der Stelt and Belger, 2007; Michel and Murray, 2012; Olbrich and Arns, 2013; Kerestes et al., 2014; de la Salle et al., 2016). The quantitative measurement of EEG signals is a neuroimaging technique with obvious practical advantages because it does not involve invasive manipulation, and it is easy to manage, well-tolerated and relatively inexpensive. In addition, the prevalence and persistence of depressive symptoms make scalp recording EEG an appropriate method to understand the underlying mechanisms of depression.

Most of the existing deep learning methods take the original data of EEG signals or transform them into frequency-domain signals as input, thus losing the spatial features between multiple brain regions and the whole brain. Liao et al. (2017) and Jiang et al. (2021) showed the effectiveness of spatial information in distinguishing depression. Considering that the irregular EEG network is one of the possible physiological symptoms of depression, the EEG activity has spatial characteristics originating from different brain regions, so the spatial EEG characteristics extracted from different brain regions can be used to identify depression. Cai et al. (2018) also found that the features in the frequency domain were more likely to distinguish patients than those in the time domain. By fusing these two features of comprehensive mixed information, it is expected to achieve richer and more accurate identification of depression. Although it is possible to explore the spatial information of the brain based on the internal structure of neural networks, there are seldom studies that have explored the spatial information of the various brain regions in depression based on the whole brain structure from EEG signals. The purpose of this study is to combine the spatial information and the frequency domain information of each brain region, and integrates them with deep learning.

The main contributions of this study are as follows. First, phase space reconstruction is used to denoise the EEG signals in the time domain and smooth the feature in the frequency domain. Second, a new model called SparNet is proposed to capture more specific information about depression in this study. It is a parallel convolutional network used to extract the features of different brain regions, and the attention mechanism module is added to the network. Third, the channels for each brain region are selected to explore the local spatial-frequency domain features. The frequency domain and the spatial features of each brain region are combined by the multi-layer parallel convolutional filter. By adding the attention mechanism, our deep learning model can assign the weights to different channels in the local brain region and also to different contribution degrees in the global brain region.

2. RELATED STUDY

All approaches to depression identification fall into two broad categories: those based on the manual features and those based on the raw data. Hosseinifard et al. (2013) used a large EEG recording dataset of 90 subjects (45 normal subjects and 45 depressed subjects) and found that in non-linear features, the correlation dimension is a powerful feature for analyzing EEG signals and identifying the depressed and the non-depressed subjects. Cai et al. (2018) through three-electrode channel acquisition, found that features were mainly concentrated in the frequency domain, and achieved the best accuracy of 79.27% with KNN. Their recent study, Cai et al. (2020b) compared KNN, DT, and SVM on the same data set. Their KNN model achieved the highest accuracy of 89.98%. Liao et al. (2017) proposed a nuclear feature filtering group common space Mode (KEFB-CSP) based on the scalp EEG signals. The signals were decomposed into each frequency band and then the spatial features were extracted by the CSP algorithm. Mumtaz et al. (2017) conducted the time-frequency decomposition of an EEG data and constructed EEG data matrix. Compared with other time-frequency methods such as STFT and EMD, the wavelet analysis has the highest classification accuracy of 87.5%. Mahato and Paul (2020) found that the average theta asymmetry of normal people was higher than that of patients with depressive people. In SVM, the classification accuracy of alpha2 and theta asymmetric combination is 88.33%. Peng et al. (2019) collected 128 electrodes of the subjects, and the research results showed that depression would affect the brain activity of almost the entire cerebral cortex, and the accuracy of 92.73% was achieved by using SVM and full frequency band features. Sun et al. (2020a) found that there were far more functional connections within hemispheres than between hemispheres. High frequency parietal occipital lobe plays an important role in depression recognition. They Sun et al. (2020b) achieved the highest classification accuracy of 82.31% by using the ReliefF feature selection method and LR classifier on the same data set. They further indicated that the functional connection feature plays an important role in depression recognition. Jiang et al. (2021) proposed an effective EEG based spatial classification detection method for depression, task-related common spatial pattern (TCSP), which significantly improved the accuracy of depression classification by using spatial information.

Although feature extraction and machine learning can effectively identify patients with depression, manual feature extraction and selection are required, which is time-consuming and laborious. There are many studies that use raw EEG data or pre-processed data as model input. Zhang et al. (2020) extracted the temporal and spatial characteristics of EEG signals by 1DCNN and added the population attention mechanism. They suggested that the combination of EEG signals and demographic factors could be better for patients with depression. Fan et al. (2020) combined CNN and LSTM to better extract time and space information. Ke et al. (2020) designed a lightweight CNN model for the online identification of patients with depression. Wan et al. (2020) proposed a convolutional neural network HybridEEGNet composed of two parallel lines for learning synchronization and regional EEG features. Seal et al. (2021) found that the right extreme value of the subjects with depression was significant, while the left extreme value of normal subjects was significant. Sharma et al. (2021) proposed a computer-aided (CAD) hybrid neural network based on EEG, that used CNN for time learning and LSTM architecture for sequence learning.

In the detection of depression, the spatial and frequency domains are two important factors, but there are no suitable neural networks to combine them together. Therefore, we propose the SpatNet neural network to combine the two features to improve the detection of depression.

3. MATERIALS AND METHODS

3.1. Participants

The dataset used in this study is from Lanzhou University. The dataset (Cai et al., 2020a) mainly includes data from patients with depression and the normal control group. Prior to the experiment, all participants signed the written informed consent. The consent and study design were approved by the Local Biomedical Research Ethics Committee of Lanzhou University Second Hospital in accordance with the World Medical Association rules. There were 48 participants, including 24 patients with depression (13 men and 11 women; 16–56 years old) and 24 healthy controls (17 men and 7 women; 18–55 years old). All patients with depression underwent a structured MINI interview, which met the diagnostic criteria for depression in the DSM-IV-based Diagnostic and Statistical Manual of Mental Disorders (DSM). Patients with MDD were selected based on their PHQ-9 (Kroenke and Spitzer, 2002), GAD-7 (Spitzer et al., 2006), and PSQI scores. Participants should be between 18 and 55 years old and have primary or higher education. For depression, inclusion criteria were MINI meeting the diagnostic criteria, patient health Questionnaire (PHQ-9) score of 5 or greater, and no psychotropic medication in the past 2 weeks.

### 3.2. Data Acquisition and Preprocessing

Subjects were asked to stay awake and still, and to reduce head and body movements and eye movements to reduce EMG and EOG, respectively, which record a 5-min resting state of closed eyes. EEG signal acquisition equipment is 128-channel HydroCel Geodesic Sensor Net. The sampling rate is 250 Hz. The reference electrode is Cz. The skin impedance of each electrode channel is kept below 70 kΩ.

The resting EEG signals are further processed using MATLAB 2021b. In the first step, the infinite impulse response digital filter IIR is used to perform 1–40 Hz band-pass filtering on the signal, and the order of the filter is set as 6. The filter can eliminate the "baseline drift" caused by low frequency noise and electrical interference from the 50 Hz-line noise. In the second step, independent principal component analysis is used to remove the EOG and EMG. Meanwhile, the integrity of channel recording signals is checked. If the invalid channels are detected, spherical interpolation is used for interpolation. The EEG signals of depressed patients and normal people are shown in **Figure 1**. In the third step, the processed time-domain signal is decomposed and reconstructed to remove the noise. A phase space reconstruction of a time-domain signal is decomposed into three signals, and then the new signal is reconstructed through the least square interpolation. The fourth step is to keep the same sample size between the subjects. The sliding window with the 2s non-overlap method is adopted for sectioning (Siuly et al., 2015), and the sample size of each subject is 148*2s. In the process of beginning and ending the experiment, the interference of brain electricity would be relatively large, so we discarded the first sample and the last sample.

3.3. The Time-Domain Denoising
3.3.1. Phase Space Reconstruction

Compared with normal EEG signals (Knott et al., 2001; Puthankattil and Joseph, 2012; Sharma et al., 2018), inhibited EEG signals have stronger stability and lower complexity. Hence, the phase space reconstruction is used to better analyze the complexity and non-stationary behavior of normal and depressed EEG signals in two-dimensional space, and the unstable noises are removed. In Sharma and Pachori (2015), Bhattacharyya and Pachori (2017), the two-dimensional diagrams of EEG signals have been used for seizure detection. The phase space reconstruction was first proposed in 1980. There are mainly two methods for phase space reconstruction: the derivative reconstruction method and the coordinate delay reconstruction method. The data set is collected in the resting state, so the signal is generally stable. The current signal feature can be

FIGURE 1 | Electroencephalogram (EEG) signals from **(A)** normal and **(B)** depressed subjects.
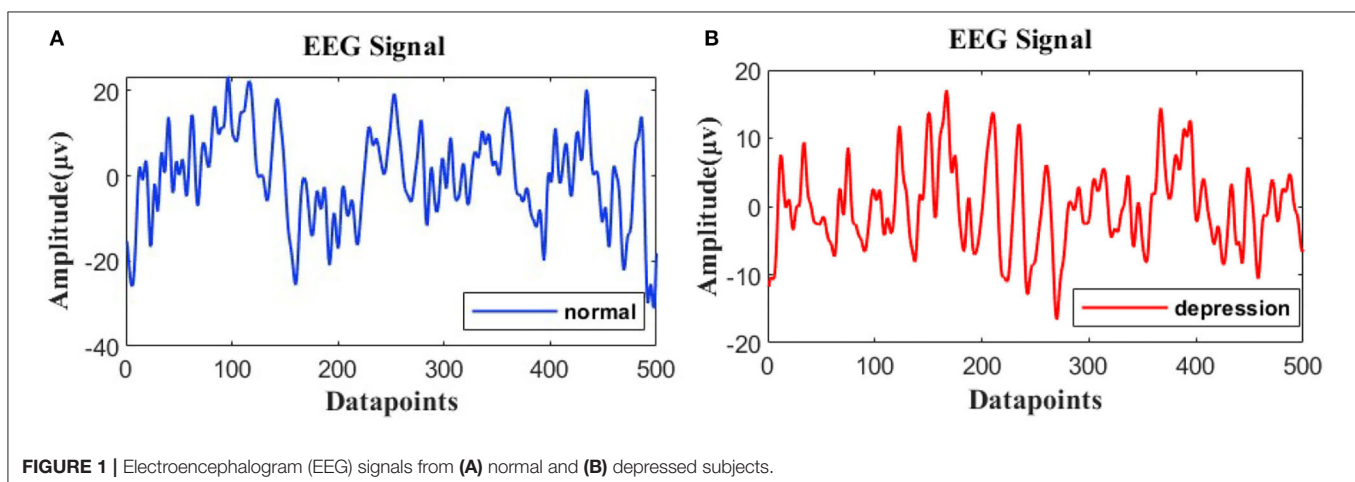
predicted by the previous signal feature and the following feature. Based on this characteristic, the coordinate delay reconstruction method is finally adopted. The generation of the PSR requires the determination of delay time t and embedding dimension d, which can be obtained by mutual information (MI) (Roulston, 1999; Bradley and Kantz, 2015) and false nearest Neighbor (FNN) (Bradley and Kantz, 2015), respectively. Supposing the time series is $x(i): i = 1, 2, \ldots, n$, the d dimensional phase space vector is constructed by different delay time t of one-dimensional time series $x(i)$, as shown in Equation (1).

$$y(i) = (x(i), \ldots, x(i + (d-1)t)), 1 \leq i \leq n - (d-1)t \quad (1)$$

The initial application of the PRS is in the chaotic time series. According to Takens' embedding theorem (Stark et al., 1997; Muldoon et al., 1998; Kukavica and Robinson, 2004), we can reconstruct a phase space from one-dimensional chaotic time series that is the same as that of the prime motorial system in the sense of topology. According to this principle, the EEG signals as the general time series can reconstruct a phase space topologically identical to the original signal by determining the delay time t and d dimensions of the phase space. Since the EEG signal is collected by the subject in the resting state with eyes closed, the brain activity is relatively stable, and the characteristics of the current signal can be determined by the characteristics of the front and rear signals. If there are interference noises at this time, the original signal will have a mutation in the waveform. We can determine and correct the reconstructed phase space.

3.3.2. Least Square Fitting

After the original EEG signal is reconstructed in the phase space, the reconstruction is selected as a three-dimensional phase space vector. Then, the linear least square method by Kiers (1997) is used to fit each point inside, because the EEG signal in the resting state would not have mutations. Thus, the current signal point could be fitted according to the characteristics of the signals before and after. Given n points $(x_i, y_i), i = 1, 2, 3, \ldots, n$. x_i is not the same, as shown in Equation (2).

$$f(x_i) = a_1 r_1(x_i) + a_2 r_2(x_i) + \cdots + a_m r_m(x_i) \quad (2)$$

$f(x)$ is closest to all the data points. We assume that the current data point is i (unknown) and all known points before and after are taken to determine the fitting function. Where, the steps to determine the coefficient a_k are as follows: First, the error function of the fitting curve and the original curve is Equation (3).

$$J(a_1, a_2, \ldots, a_m) = \sum_{i=1}^{n} \delta^2 = \sum_{i=1}^{n} [f(x_i - y_i)]^2 \quad (3)$$

For $a_1, a_2, \ldots, a_m$ minimizes J, use the necessary extreme conditions: $\partial J / \partial a_k = 0 \ (k = 1, \ldots, m)$ to get the linear Equation (4).

$$\sum_{i=1}^{n} r_j(x_i) [\sum_{k=1}^{m} a_k r_k(x_i) - y_i] = 0, (j = 1, \ldots, m) \quad (4)$$
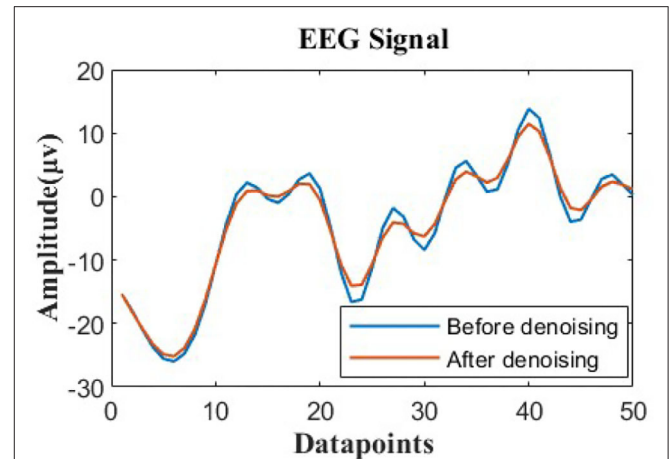

FIGURE 2 | Phase space reconstruction and the least square method were used to detect EEG signals before and after denoising in patients with depression.

$$\text{make } R = \begin{pmatrix} r_1(x_1) & \cdots & r_m(x_1) \\ \vdots & \ddots & \vdots \\ r_1(x_n) & \cdots & r_m(x_n) \end{pmatrix}, \ A = [a_1, \ldots, a_m]^T, \ Y =$$

$[y_1, \ldots, y_n]^T$, so we can rewrite this equation as Equation (5).

$$R^T R A = R^T Y \quad (5)$$

Thus, when the equation satisfies $r_1(x), \ldots, r_m(x)$ linearly independent, R column full rank, $R^T R$ invertible, there is a unique solution Equation (6).

$$A = (R^T R)^{-1} R^T Y \quad (6)$$

For the selection of function $r_k(x)$, we use the polynomial curve to better fit the EEG signal. **Figure 2** shows the comparison before and after denoising EEG signals of patients with depression by using phase space reconstruction and the least square method.

3.4. Characteristics of Feature Smoothing

Pham et al. (2015) proposed the importance of feature smoothing for emotional EEG classification. The denoised EEG signal is converted into the frequency domain signal by the fast Fourier transform. The frequency domain signal is used as the input of the neural network. Each sample is segmented according to the time sliding window. Therefore, the transformed frequency domain signals are also smoothed according to 148 time points (each sample size). The EEG signals are collected by the subjects with their eyes closed, so the changes in the EEG signals in the frequency domain are not particularly obvious. Equations (7) and (8) are used to calculate the difference between the signals at each time point and the mean value of the signals.

$$\delta_j^i = |y_j^i - E(Y)^i|, 1 \leq i \leq 40, 1 \leq j \leq 148 \quad (7)$$

$$E(Y)^i = \frac{1}{J} \sum_{j=1}^{J} y_j^i \quad (8)$$

Where Y represents the amplitude, i represents the frequency, and j represents the time point. $E(Y)$ is the mean of a sample of a subject at frequency i. δ_j^i is the error between each sample point and the mean.

$$\delta_j^i - 3\sqrt{\frac{1}{J}\sum_{j=1}^{J}(y_j^i - E(Y)^i)^2} > 0 \qquad (9)$$

Mark y_j^i as an outlier if the value of y_j^i differs from the mean by more than three standard deviations as shown in Equation (9). After the outliers are detected, the y_j^i sample points marked before are smoothed by the method of before-and-after mean interpolation. In this way, the influence of abnormal features can be further cleaned up, while maintaining the overall trend of frequency domain features. At the same time, it can reduce the risk of over-fitting during the training of the neural network.

3.5. Attention Module

The attention mechanism is a special structure embedded in a machine learning model, which is used to automatically learn and calculate the contribution of input data to output data. The attention mechanism is a signal processing mechanism discovered by some scientists in the study of human vision. Some practitioners in the field of artificial intelligence have introduced this mechanism into some models. At present, the attention mechanism has become one of the most widely used "components" in the field of deep learning, especially in the field of natural language processing. The classic ones are BERT, Transformer (Devlin et al., 2018; Wang et al., 2019), and other models or structures that are highly exposed in the past 2 years. In this article, we adopt the SENet (Squeeze-and-Excitation Networks) (Hu et al., 2018) module incorporated with the channel attention mechanism.

In our proposed network model, the EEG signals are converted into frequency domain features and then used as the input of the SENet module. The SENet module is originally used to process the two-dimensional images, in this article, we use it in one-dimensional signal processing. It can mainly use the global information to selectively emphasize the information features and suppress the less useful features by assigning different weight values to each channel. This is a combinatorial function of five consecutive operations: channel global average pooling (Lin et al., 2013), complete connection layer, Relu, complete connection layer, and finally Sigmoid. The sigmoid activation plays an important role as the channel weights that adapt the input specific descriptors. Due to the fully connected layer and pooling layer, the number of parameters and the computation load increased slightly. The unique structure of this extrusion and excitation network, shown in **Figure 3**, can be used with any standard architecture.
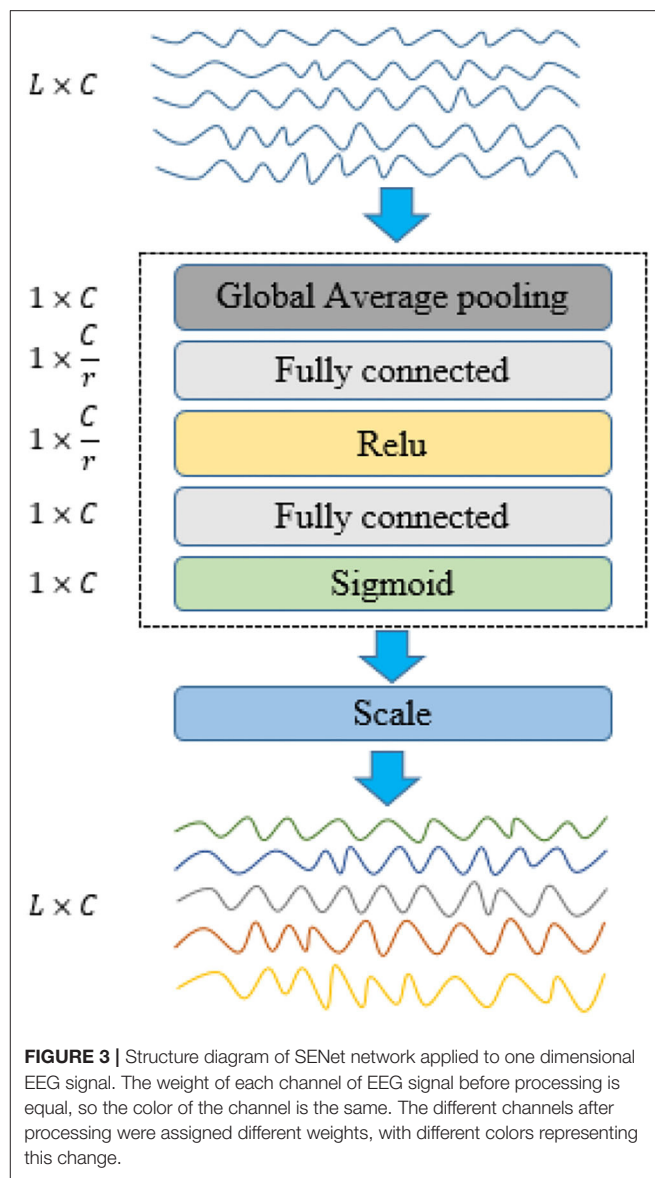

FIGURE 3 | Structure diagram of SENet network applied to one dimensional EEG signal. The weight of each channel of EEG signal before processing is equal, so the color of the channel is the same. The different channels after processing were assigned different weights, with different colors representing this change.

3.5.1. Squeeze

SENet implements compression operations through the global average pooling to generate channel statistics. Where $Z \in R^C$, and the kth element z_k of Z is calculated by Equation (10).

$$z_k = F_{sq}(u_k) = \frac{1}{L}\sum_{i=1}^{L} u_k(i), k = 1, 2, \cdots, C \qquad (10)$$

where $F_{sq}(\cdot)$ is the compression operation, and u_k is the feature on the kth channel. C is the total number of channels.

3.5.2. Excitation

The excitation operations help to capture channel dependencies and greatly reduce the number of parameters and calculations. The excitation part is mainly composed of two fully connected

layers and two activation functions, which can be written as Equation (11).

$$S = F_{ex}(Z, W) = \sigma(\gamma(Z, W)) = \sigma(W_2 \delta(W_1 Z)) \quad (11)$$

where $S = s_1, s_2, \ldots, s_C, s_k \in R^L (k = 1, 2, \ldots, L)$. $F_{ex}(\cdot)$ is the excitation operation. $W_1 \in R^{\frac{C}{r} \times C}$, $W_2 \in R^{C \times \frac{C}{r}}$, W_1, and W_2 are the weights of the two fully connected layers used for dimension reduction and dimension enhancement. Z is the fully connected input after global average pooling. r is a hyperparametric ratio, which can change the capacity and calculation cost. $\delta(x)$ is the activation function Relu used to prevent the gradient from disappearing (Gu, 2017). $\sigma(x) = \frac{1}{(1+e^{-x})}$ is a sigmoid function. Equation (12) is used to calculate the final output $\tilde{x}_k (k = 1, 2, \cdots, C)$. The output is obtained by multiplying the input channels by their respective weights.

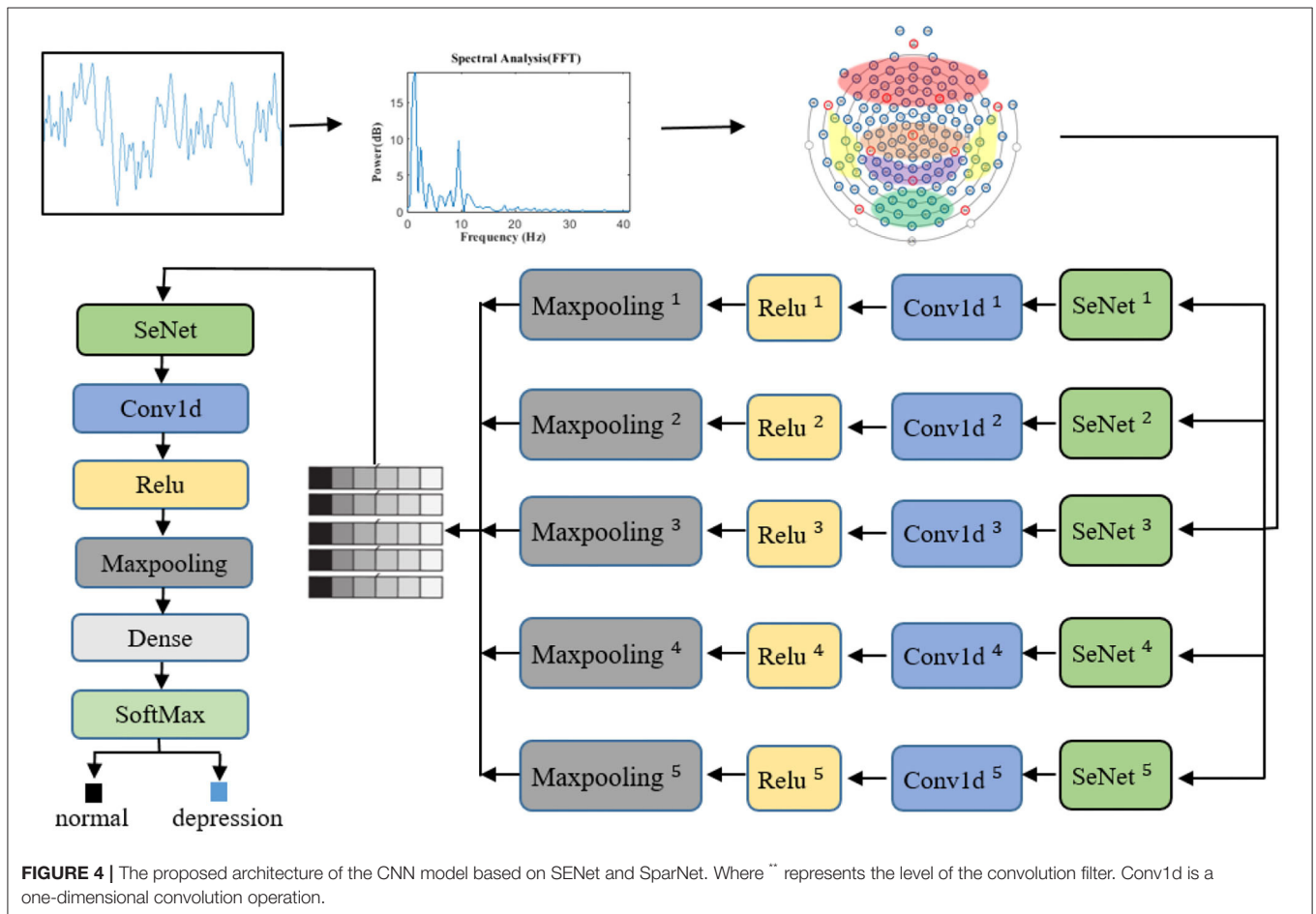$$\tilde{x}_k = F_{scale}(u_k, s_k) = u_k \cdot s_k \quad (12)$$

where $\tilde{x}_k \in R^L$ refers to the multiplication above the channel. s_k is the processed channel weight. u_k is the original eigenvector.

4. PROPOSED DEEP LEARNING SCHEME

Convolutional Neural Network is a special type of neural network which is widely used in image processing and classification tasks. It is a state-of-the-art deep learning method consisting of many stacked convolutional layers. The network consists of a convolution layer, pool layer, and final complete connection layer. The EEG signals are one-dimensional time series signals. After converting them into frequency domain signals, multi-channel one-dimensional frequency domain signals are input. Therefore, 1DCNN is used in the convolutional neural network. The features in the frequency domain can be fully combined with the spatial information between channels.

4.1. Convolution Layer

At the convolution layer, 1DCNN carries out the convolution operation on the local area of the input signal to generate the corresponding one-dimensional feature map. Different convolution kernels extract different features from the input signal respectively. Each convolution kernel detects the specific features of all positions on the input feature graph to achieve the weight allocation on the same input feature graph. The characteristics of the local connectivity and weight sharing effectively can reduce the complexity of the network and the

FIGURE 4 | The proposed architecture of the CNN model based on SENet and SparNet. Where ** represents the level of the convolution filter. Conv1d is a one-dimensional convolution operation.

number of training parameters. If the L layer is the convolution layer, Equation (13) of the one-dimensional convolution layer is

$$x_j^l = f(\sum_{i=1}^{M} x_i^{l-1} * k_{ij}^l + b_j^l) \qquad (13)$$

where k represents the convolution kernel, j represents the number of convolution kernels, and M represents the number of channels of the input $x(l-1)$ of the upper layer. b is the offset of the convolution kernel, and * stands for the convolution operation. $f(\cdot)$ is the activation function.

4.2. Activation Layer

What is done in the convolution layer of the upper layer is to process the input features in the way of convolution, i.e., to assign a weight to each pixel point. This operation is linear. But for samples, they are non-linear separable instead of linear separable. Thus, we add the activation function Relu Equation (14) here.

$$f(x) = max(0, x) \qquad (14)$$

As a non-linear factor, the activation function added to the model can make the model more expressive and better fit the data.

4.3. Pooling Layer

The downsampling stage is after the convolution layer and the number of feature graphs increases. This leads to the expansion of the data dimension, which is not conducive to calculation. Therefore, the average pooling or maximum pooling is used to process each feature map in this stage. The average pool is calculated according to the size of the predetermined pool window, and the maximum pool method selects the maximum parameter within the predetermined window range as the output value. In our study, the maximum pooling operation is adopted. The pooling kernel size is 1*2, the step size is 1, and there is no filling.

4.4. Connection Layer

After passing through the convolution layer, the data scale is channel×features, and the feature dimension needs to be straightened into one dimension. At this point, the full connection layer of the node is connected with all neuron nodes output from the feature mapping of the previous layer, and the activation function is softmax function. If the final pooling layer is l+1 and output to the full connection layer, then the output of the full connection layer is Equation (15).

$$h(x) = f(w^{l+1} \cdot x^{l+1} + b^{l+1}) \qquad (15)$$

where w represents the weight of each feature, and b represents the offset. $F(\cdot)$ represents the activation function.

4.5. Loss Function

It is used to calculate the error between the classification prediction label and the actual label. The classification cross entropy is used as the loss function, and the probability distribution is compared with the real distribution. $L_1, L_2, \ldots,$

TABLE 1 | Detailed information about the proposed SparNet deep model.

No	Names of layers	Kernel size	Parameters of layers
0	Input	—	—
1	SENet$^{(1-5)}$	—	Reduction = 2
2	Conv1D$^{(1)(3-5)}$	12*5	Stride = 1
	Conv1D$^{(2)}$	25*5	Activation = Relu
3	MaxPooling1D$^{(1-5)}$	2	Stride = 2
4	SENet	—	Reduction = 2
5	Conv1D	5*2	Stride = 1, Activation = Relu
6	MaxPooling1D	2	Stride = 2
7	Dense	—	Neurons = 2
8	Softmax	—	—

*The ** represents the level of the convolution filter: SENet$^{(1-5)}$ represents the SENet of the 1st to 5th layers.*

TABLE 2 | Channels in different brain regions are selected.

C	E36-E104,E30-E105,E41-E103,E37-E87,E42-E93,E47-E98
F	E19-E4,E22-E9,E24-E124,E27-E123,E32-E1,E33-E128
O	E70-E83,E71-E76,E69-E89,E74-E82,E73-E83,E75-E81
P	E52-E92,E60-E85,E51-E97,E67-E77,E59-E91,E72-E62
T	E58-E96,E45-E108,E114-E44,E100-E46,E102-E57,E50-E101

and L_T are represented by the one-hot encoding strategy. The loss function can be calculated as Equation (16).

$$Loss = -\sum_{i=1}^{T}\sum_{j=1}^{M} L_{i,j} * log p_{i,j} \qquad (16)$$

where T is the number of verification data samples, M is the number of classes, $p_{i,j}$ is the predicted value obtained from the fully connected layer, and $L_{i,j}$ is the true value.

4.6. SparNet

In view of the advantages of CNN, one-dimensional CNN is used to extract the spatial frequency features of EEG signals. In the proposed SparNet network architecture, each layer is directly connected to each other in feedforward mode. In **Figure 4**, the SparNet network consists of six sub-CNNs. Each sub-CNN consists of a SENet, Conv1d, Relu, and Maxpooling. Five of them are at the same level to form a network of parallel structures to operate the brain region, and the last convolutional operation performs a processing operation on the global brain. Finally, there is a fully connected layer output result. The SENet module is added to each sub-network to increase the attention mechanism between channels and brain regions. **Table 1** details each layer of the proposed SparNet network and the parameters of each layer.

First of all, the same sample size is kept between the subjects. We use the sliding window 2s non-coincidence method to slice, and the sample size of each subject is 148*2s. In this study, since the recognition features of depression patients are mainly in the frequency domain signals, we manually extracted the features in this step. The EEG signals are transferred from the time domain

to the frequency domain (1–40 Hz) by a fast Fourier transform, and the frequency domain signal is used as the input of the neural network. The channels are sorted out before entering the first layer of the model. The channels are divided into five regions according to the brain regions. At this point, the data size of each brain region is channeled *40. Before entering the convolution layer, the SENet module will adjust the weight of each channel, and then the space frequency domain features of each brain region will be extracted through convolution operation and maximum pooling. The features of five brain regions will be spliced to form a feature scale of 5*features, where 5 represents five brain regions. The SENet module then weights each brain region based on global features. After the convolution operation, the output of the full connection layer is entered. **Figure 4** shows the whole signal processing process of the network.

4.7. Characteristics in the Space Frequency Domain

Peng et al. (2019) and Jiang et al. (2021) showed that the EEG signals of depressed patients had better feature representation in the frequency domain, where absolute power and relative power in the frequency domain were of great help in identifying patients with depression. Moreover, the expression of EEG signals in the power spectrum of patients with depression is obviously different from that of healthy people. Stark et al. (1997) and Lin et al. (2013) showed that in the spatial domain of channels, the brain functional connections of patients with depression were different from those of normal people, and the accuracy of classification of patients with depression was significantly improved.

Therefore, this article aims to combine frequency domain and spatial domain to effectively identify patients with depression
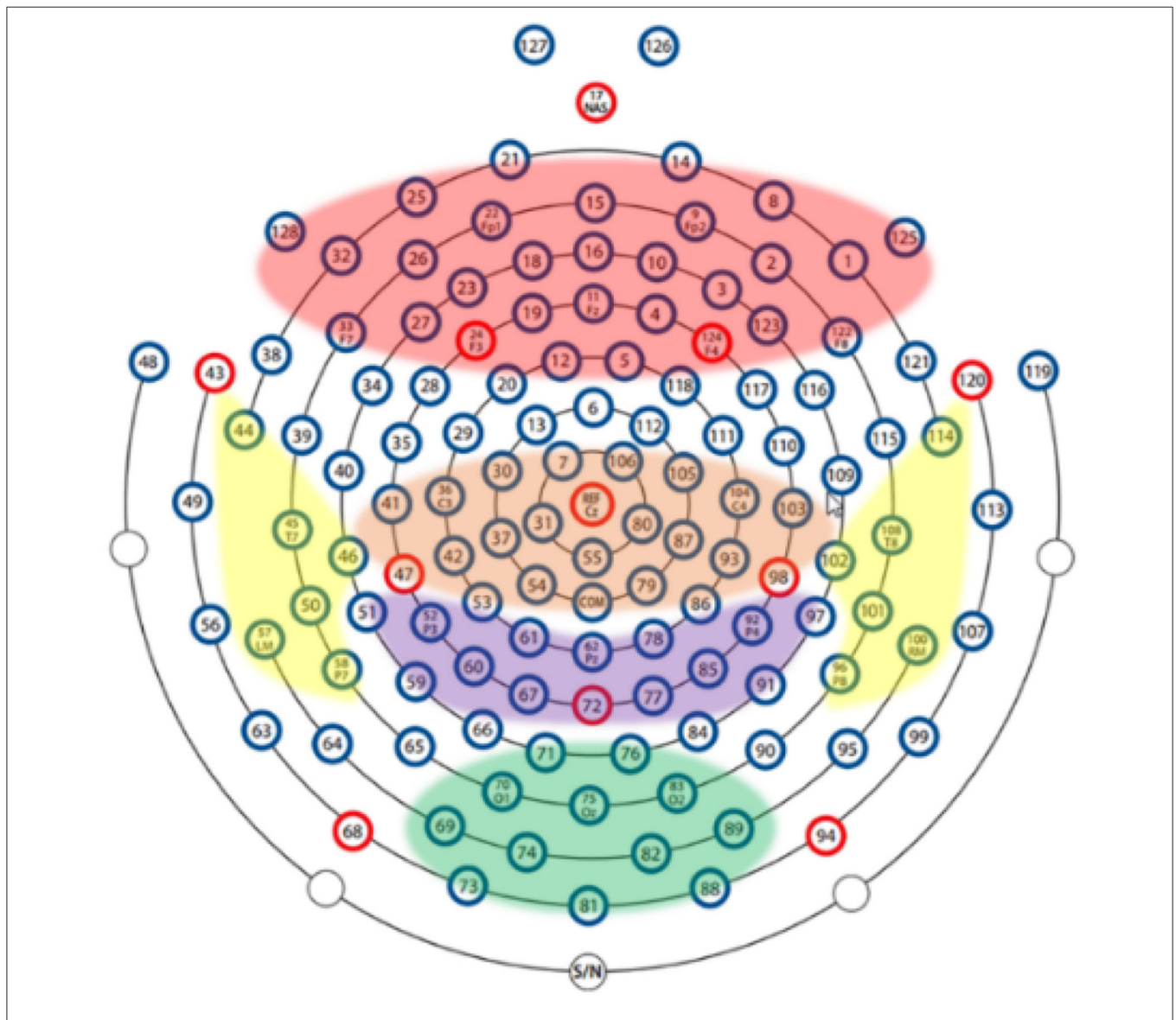

FIGURE 5 | The division of the brain between different brain regions.

in space-frequency domain. Based on the characteristics of the brain, it is divided into five brain regions. First, the local space-frequency domain is explored, and then the feature extraction of the global space-frequency domain is carried out. The brain regions are shown in **Table 2**.

In order to explore the influence of various brain regions on patients with depression, we divide the brain into five parts as shown in **Figure 5**: central region, frontal lobe region, occipital lobe region, parietal lobe region, and temporal lobe region. The first layer of the SparNet network has five parallel sub-CNNs corresponding to the five brain regions. Due to the tightness between brain regions, we initially select 12 channels in each brain region to carry out space frequency domain characteristics of brain regions. In order to accurately locate the channels in each brain region, we obtained the channel location information from Luu and Ferree (2005) by referring to the 128-channel HydroCel Geodesic Sensor Net device. Each channel contains spatial information about the brain. Additionally, in order to avoid the influence of multiple brain regions on a single channel, we select 12 channels distributed in the central location of each brain region without selecting channels at the edge of the brain region. The channels are then added individually according to each brain region's contribution to the recognition of depression. The selection of preliminary channels is shown in **Table 2**.

5. RESULTS

5.1. Evaluation

There are 48 subjects in total. In order to better generalize the model, we adopt dataset division among subjects. A total of 24 sets are obtained by combining a depressed patient with healthy control, using the one-subject cross-validation (LOSOCV) to assess the generalization ability of each classification model.

In this study, accuracy, sensitivity, and accuracy based on the confusion matrix are used as the performance evaluation indexes. Sensitivity (recall rate) is defined as the percentage of patients with MDD predicted in all MDD patients (TP+ FN), and precision is defined as the percentage of healthy controls predicted in all healthy controls (TP+FP). Accuracy is defined as the percentage of correctly classified patients with MDD and healthy controls. F1 index takes into account both model accuracy and recall rate and is defined as the harmonic mean of model accuracy and recall rate.

$$Accurary = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{18}$$

$$Precision = \frac{TP}{TP + FP} \tag{19}$$

$$F_1 Score = 2 \times \frac{Precision \times Sensitivity}{Precision \times Sensitivity} \tag{20}$$
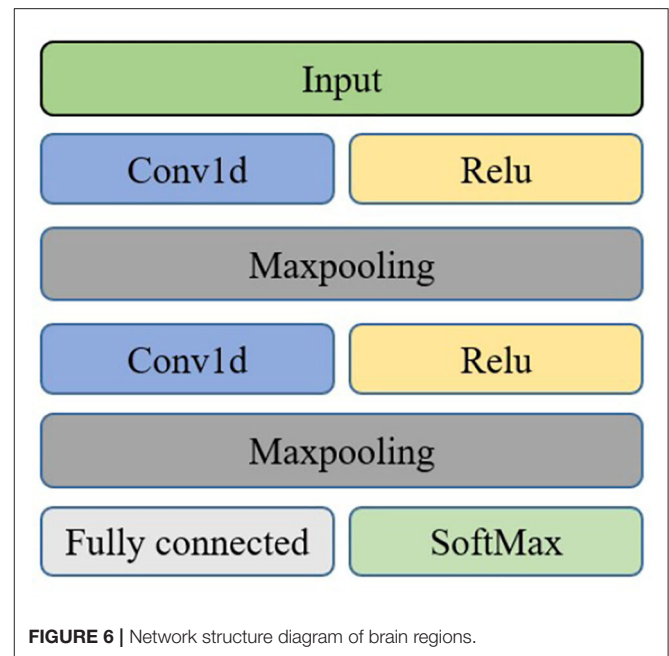

**FIGURE 6 |** Network structure diagram of brain regions.

The Receiver Operating characteristic (ROC) curve is used for evaluation. The ROC curve is widely used in binary classification evaluation, which evaluates sensitivity and specificity against several thresholds.

5.2. Partial Results

The brain is divided into five regions: frontal, central, parietal, occipital, and temporal. In order to more effectively identify people with depression, we explored the importance of different brain regions. The importance of 12 channels in each brain region is evaluated. 1DCNN is performed on each of the five brain regions, and the evaluation index is the average accuracy after each fold. The neural network model of brain regions is shown in **Figure 6**.

Each subject is included in the test set by a 24-time retention cross-validation method for each brain region. The accuracy of each brain region is shown in **Figure 7**.

As can be seen from the results, the frontal lobe has a greater contribution to the identification of patients with depression compared with other brain regions, which is consistent with the results of Jiang et al. (2021), indicating the feasibility of the frontal lobe channel detection for depression. Wan et al. (2020) also prefer the frontal lobe channel in terms of channel selection. Depression patients with low moods will lead to emotional differences from normal people. Since the frontal lobe is the main brain area for emotional processing, the frontal lobe is more important for the identification of depression patients. Thus, in order to better explore the characteristics of the global space frequency domain, the number of channels in the frontal lobe is increased to 25 channels, E19-E4, E22-E9, E24-E124, E27-E123, E32-E1, E33-E128, E15, E11, E1 - E32, E27 - E123, E13 - E112, E29 - E111, E28-117, the E6, E121 - E38, E34 - E36, combined with the cerebral cortex region occupied by the frontal lobe.

FIGURE 7 | The accuracy of different brain regions in identifying people with depression.
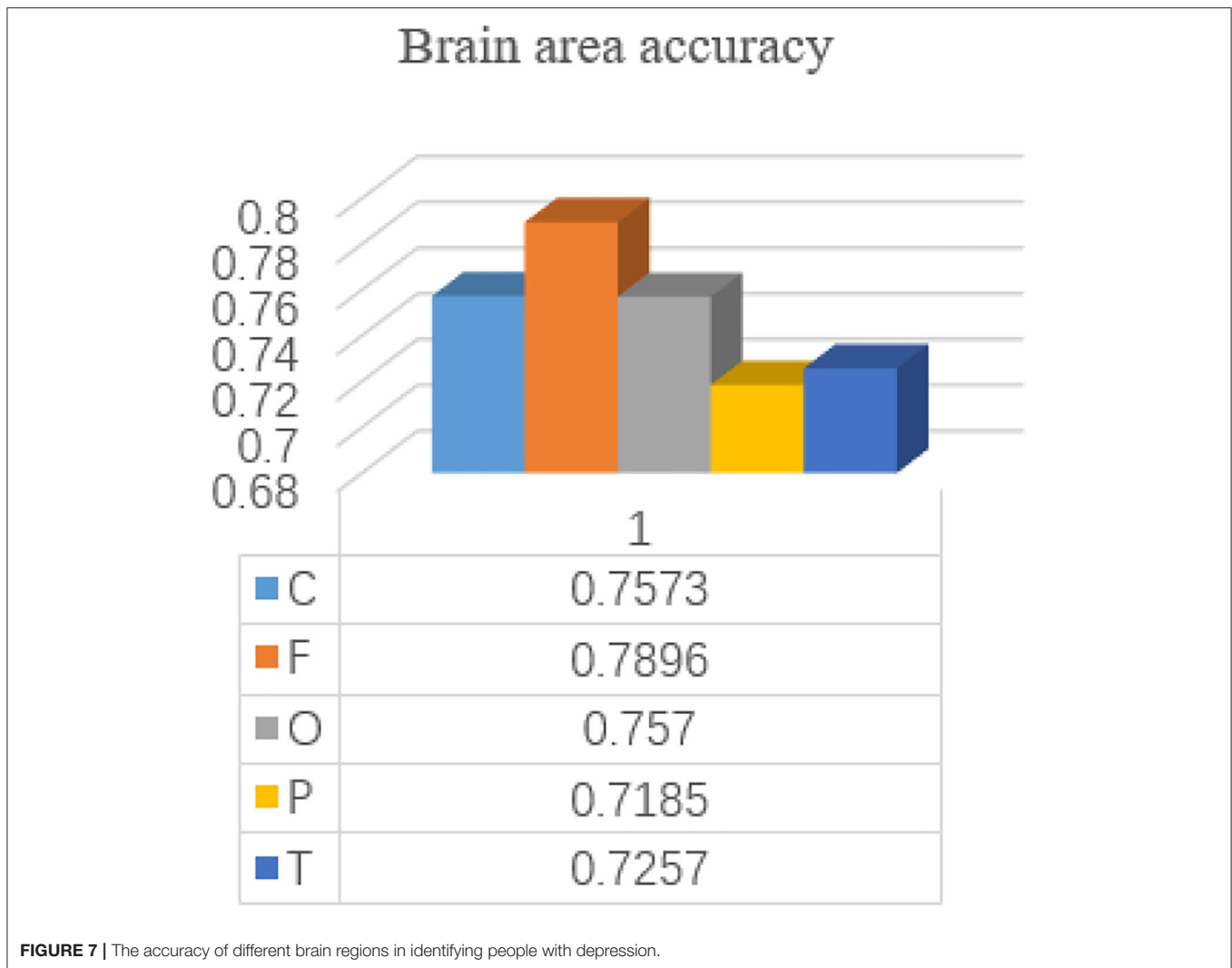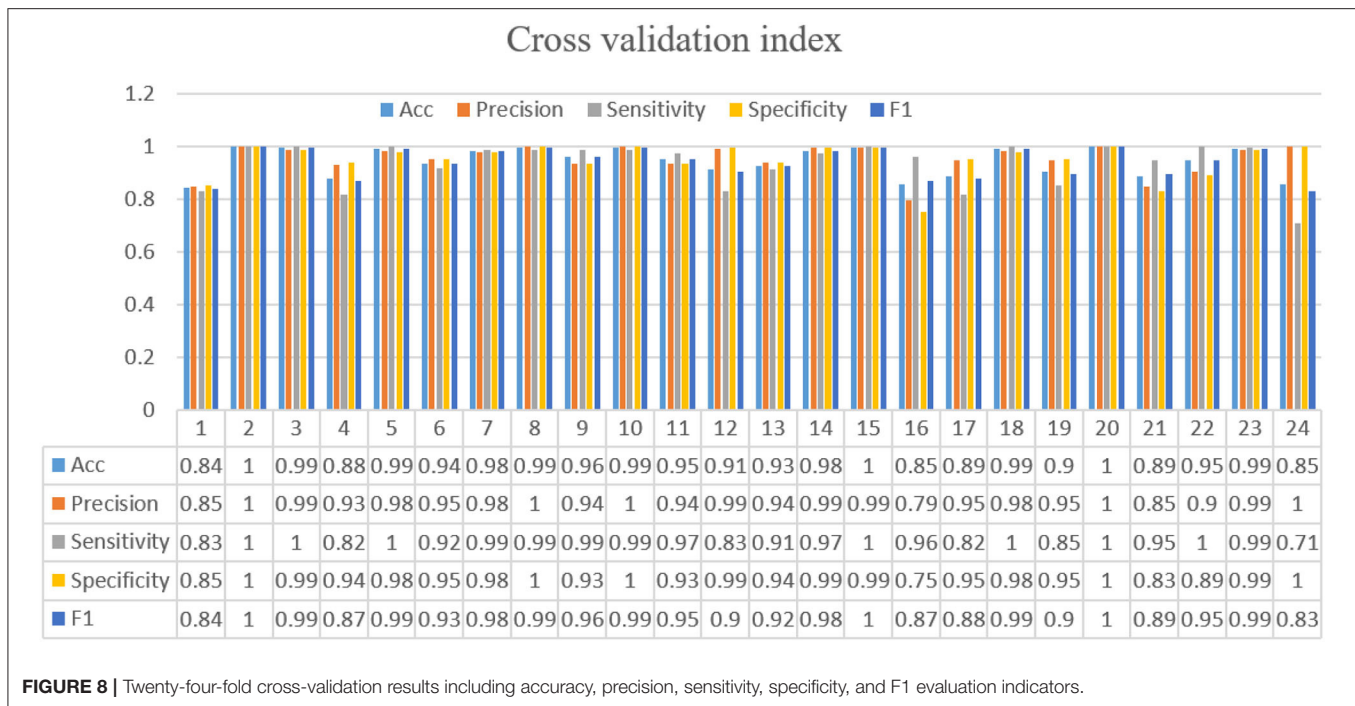
5.3. Global Results

The software Matlab 2020b is used to preprocess the data, and the EEG data are segmented with a 2s window length. After that, the time domain signal is converted to the frequency domain as the characteristic input of the model. The space frequency domain features of the EEG signal will be extracted by a two-layer convolution operation. In addition, in the convolution operation, each channel will be adjusted according to the parameters of a convolution kernel, and will not be affected by the parameters of other channels. Therefore, the channel grouping order has no significant impact on performance.

For training the model, the batch size is 8, and each network is trained with 50 epochs. The cross entropy is selected as the loss function. In the optimization stage, the RMS algorithm is selected to obtain better results and a shorter running time. If the loss function is verified not to improve after 10 consecutive epochs, the early stop criterion is used. **Figure 8** shows the calculation results for each fold. The ROC and AUC values of the above methods are shown in **Figure 9** for each fold. Due to the differences between the subjects, the minimum AUC

area obtained is 0.829, and there are 16-fold AUC areas over 0.95, which also demonstrates the effectiveness of the model. The results of cross-validation are statistically analyzed. The mean values of F1, Acc, Precision, Sensitivity, Specificity are 0.947, 0.953, 0.937, 0.951, and 0.942, respectively. The standard deviation of F1, Acc, Precision, Sensitivity, Specificity are 0.054, 0.056, 0.082, 0.063, 0.056, respectively.

The results of 24-folds are combined and analyzed, and the evaluation indexes obtained are shown in **Table 3**, in which the accuracy rate reaches 94.37%. At the same time, the ROC and AUC values of the above methods are shown for the overall results, including the ROC and AUC values of the individual categories of normal people and depressed patients, as shown in **Figure 10**. The AUC area of the overall data is 0.9682. All the evaluation indicators decrease when the input characteristics in SparNet are time domain signals. The results are shown in **Table 4**, named "With Time Domain". Thus, the frequency domain feature is more effective than the time domain feature.

For the proposed model, in order to verify the effectiveness of the attention mechanism in identifying patients with depression,

FIGURE 8 | Twenty-four-fold cross-validation results including accuracy, precision, sensitivity, specificity, and F1 evaluation indicators.

six SENet modules are removed from the model and the comparison before and after the results are shown in **Table 4**.

It can be seen from the results that adding an attention mechanism into the model can improve the learning ability of the model. The accuracy rate of depression identification is improved by 3.13%, and all the evaluation indicators are also improved.

In this article, after converting the EEG signals into the frequency domain, the smoothing operation is performed on the frequency domain features. The comparative experiments are conducted before and after feature smoothing, and the experimental results are shown in **Table 4**.

It can be seen from the results that the accuracy is improved by 4 % using the smoothed features as the input. Furthermore, we can observe that all the evaluation indicators are improved at the same time. The calculation cost of the network is calculated. The parameter number of the SparNet network is 8,725, and the network parameter without the SENet module is 455. The number of network parameters without smooth is 8,725, mainly to illustrate the necessity of feature smoothing for neural networks. The SparNet parameters are mainly contributed by the SENet module.

6. DISCUSSION

Screening for depression is very important for early diagnosis and treatment. However, the previous diagnosis of depression is confined to a manual questionnaire survey and feature extraction, which is subject to many limitations. For example, a questionnaire survey required experienced doctors, while feature extraction required a lot of manpower to find the characteristics of relevant indicators. The deep learning methods can overcome this limitation and can be used anywhere without highly trained experts. In this study, we use feature smoothing and deep learning for the automatic detection of patients with MDD and healthy controls with good performance. The attention mechanism is combined with 1DCNN, and the spatial-frequency domain features are extracted by brain regions. The accuracy rate of our model reaches 94.37%.

The main innovation of this study is to make full use of the spatial and frequency domain characteristics of the brain. We also try to smooth the input of the neural network. Experimental results show that the frequency domain characteristics of the input smoothing processing can effectively improve the identification accuracy of patients. In addition, due to the characteristics of the convolutional neural network, the brain is divided into different regions for feature extraction of parallel structures.

It can be seen from the final results that adding the SE module into the 1DCNN neural network can make the model have higher accuracy. This is because the convolutional layer provides a powerful feature fusion technology, although the weights between channels are unified by default. The SE module can better highlight the importance of different channels. At the same time, although there has been a feature of converting the EEG signal into the frequency domain as the input, the segmentation of brain regions according to the brain's structure can make the spatial information of brain regions better interpreted. The results for each brain region and the final results for the whole brain region show that although each brain region is helpful in identifying depression, the characteristics of the whole brain are better. It could be thought that a large amount of space-frequency information is lost for a single brain region and cannot get a significant effect.
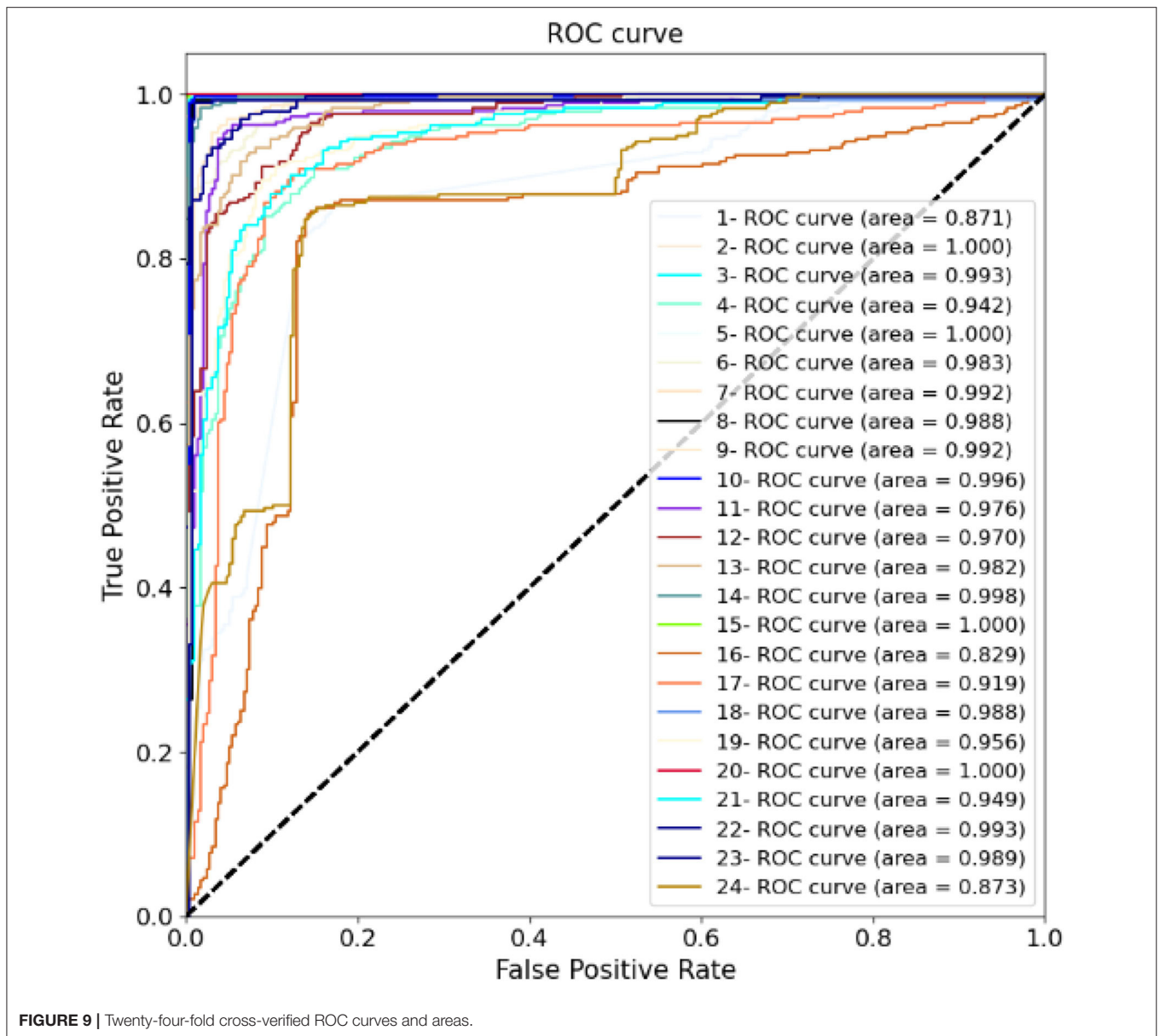
FIGURE 9 | Twenty-four-fold cross-verified ROC curves and areas.

TABLE 3 | Performance values obtained when testing the SparNet model using EEG data.

	Class	Predicted depression	Normal	Acc	Precision	Sensitivity	Specificity	F1
Actual	Depression	3,327	225	0.9437	0.9375	0.9507	0.9366	0.9440
	Normal	175	3,377		0.9500	0.9366	0.9507	0.9432

Table 5 is a comparison of results on the same data set. It can be seen that the accuracy of this study is higher than that of traditional machine learning methods for linear and non-linear extraction. Additionally, compared with Peng et al. (2019), we use fewer channels to achieve better accuracy and obtain the best results in this dataset in automated detection of patients with depression and healthy controls, which could provide better solutions for future clinical applications. Compared with Zhang et al. (2020), the combination of multi-parallel 1DCNN and SE modules in this study tests the importance of space-frequency information and other advantages. However, the main shortcoming of this study is the data size to train the network, which can be overcome by simplifying the deep model. Our future goals
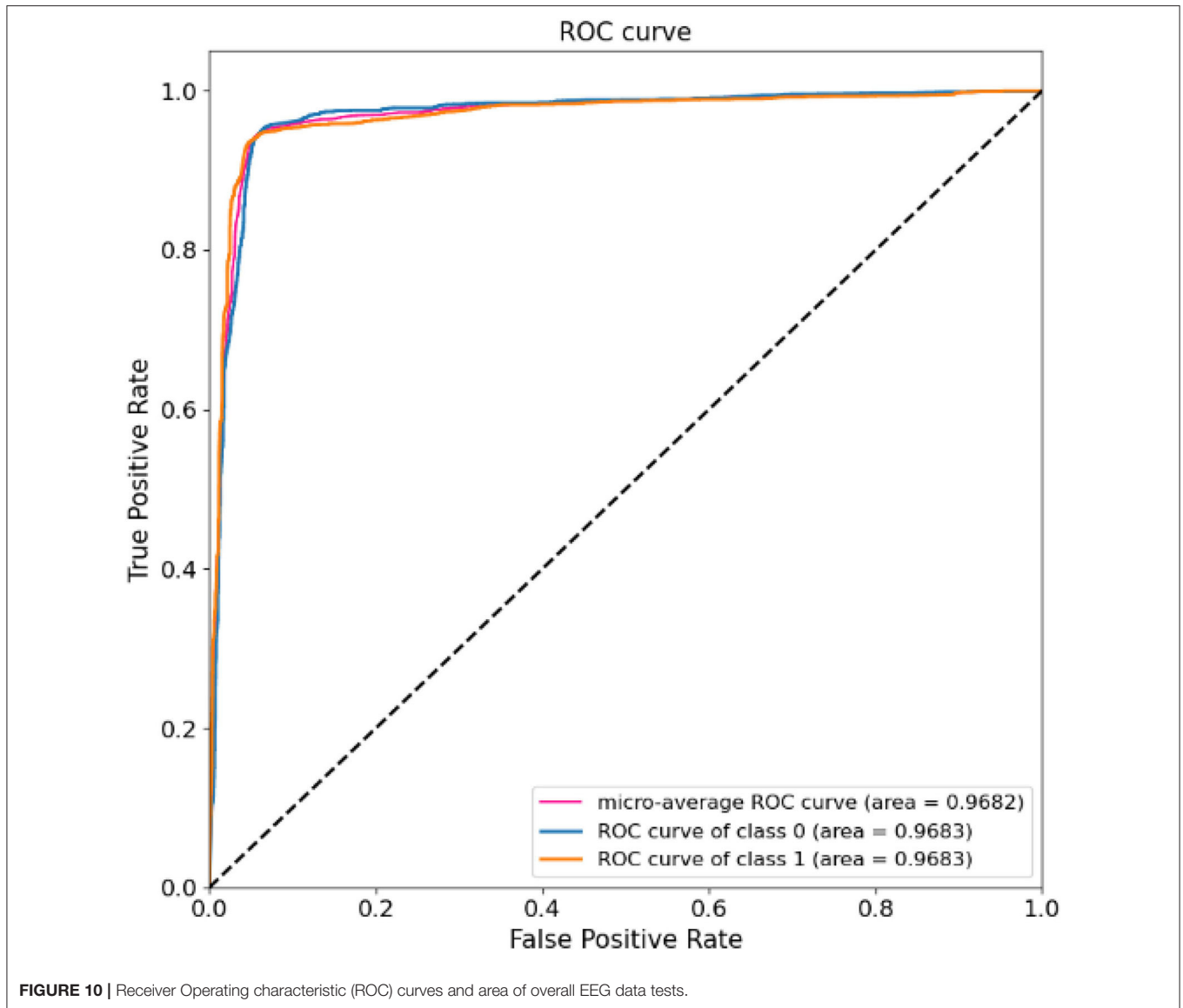
FIGURE 10 | Receiver Operating characteristic (ROC) curves and area of overall EEG data tests.

TABLE 4 | Comparison of results before and after adding SENet module.

	Acc	Precision	Sensitivity	Specificity	F1
SparNet	0.9437	0.9375	0.9507	0.9366	0.9440
With time domain	0.8851	0.9189	0.8607	0.9104	0.9130
Without the SENet	0.9124	0.8953	0.9341	0.8907	0.9143
Without smooth	0.9037	0.9099	0.9042	0.9104	0.9070

TABLE 5 | The confounding matrix and evaluation index are used to compare the classification results of SparNet and correlation methods in the same data set.

References	Methods	Channels number	Classification methods	Accuracy (%)
(Sun et al., 2020b)	Linear features, non-linear features, PLI	16 Channels	ReliefF, LR	82.31
(Peng et al., 2019)	PLI	128 Channels	Kendall rank correlation coefficient+SVM	92.73
(Sun et al., 2020a)	Linear features, non-linear features, PLI, network measures	128 Channels	C4.5, BFDT, LR	84.18
(Zhang et al., 2020)	Time domain feature	3 Channels	1DCNN	75.29
SparNet	Frequency domain feature	73 Channels	1DCNN	**94.37**

The bold value is the accuracy derived from the SparNet.

are to further expand the experimental space, collect more samples, and apply the developed methods to other types of EEG data.

7. CONCLUSION

The main study in this article is using the novel neural network called SparNet based on the EEG signals to identify depressed people or not. First, the denoising method of the phase space reconstruction is used to denoise and clean the data. Second, the input features are smoothed before the frequency domain features are input into the model. Third, a new model called SparNet is proposed to extract the space-frequency domain features of the local brain regions and the whole-brain. Finally, the cooperating of the attentional mechanisms to the model improves the identification accuracy of the patients with depression. Compared with other methods, the proposed model can obtain a better classification performance. From the results of the local brain regions, it can be seen that the frontal lobe plays a better role in the identification of patients with depression. From the results of the global brain region, it can be seen that the combination of the spatial features and the frequency domain features can effectively improve the accuracy of depression identification. The combination of features of the different brain regions may be the focus of future research. The methods and findings of this study may contribute to the wider application of the diagnosis of deep depression in clinical applications and neuroscience research.

REFERENCES

Belmaker, R. H., and Agam, G. (2008). Major depressive disorder. *N. Engl. J. Med.* 358, 55–68. doi: 10.1056/NEJMra073096

Bhattacharyya, A., and Pachori, R. B. (2017). A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform. *IEEE Trans. Biomed. Eng.* 64, 2003–2015. doi: 10.1109/TBME.2017.2650259

Bradley, E., and Kantz, H. (2015). Nonlinear time-series analysis revisited. *Chaos Interdiscip. J. Nonlinear Sci.* 25, 097610. doi: 10.1063/1.4917289

Cai, H., Gao, Y., Sun, S., Li, N., Tian, F., Xiao, H., et al. (2020a). Modma Dataset: A Multi-modal Open Dataset for Mental-disorder Analysis. New York, NY: Cornell University.

Cai, H., Han, J., Chen, Y., Sha, X., Wang, Z., Hu, B., et al. (2018). A pervasive approach to EEG-based depression detection. *Complexity* 2018, 5238028. doi: 10.1155/2018/5238028

Cai, H., Qu, Z., Li, Z., Zhang, Y., Hu, X., and Hu, B. (2020b). Feature-level fusion approaches based on multimodal EEG data for depression recognition. *Inform. Fusion* 59, 127–138. doi: 10.1016/j.inffus.2020.01.008

de la Salle, S., Choueiry, J., Shah, D., Bowers, H., McIntosh, J., Ilivitsky, V., et al. (2016). Effects of ketamine on resting-state EEG activity and their relationship to perceptual/dissociative symptoms in healthy humans. *Front. Pharmacol.* 7, 348. doi: 10.3389/fphar.2016.00348

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* Minneapolis, MN.

Fan, Y., Yu, R., Li, J., Zhu, J., and Li, X. (2020). "EEG-based mild depression recognition using multi-kernel convolutional and spatial-temporal feature," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (Seoul: IEEE), 1777–1784. doi: 10.1109/BIBM49941.2020.9313499

Gu, D. (2017). *3D densely connected convolutional network for the recognition of human shopping actions* (Ph.D. thesis). Université d'Ottawa/University of Ottawa, Ottawa, ON, Canada.

Hosseinifard, B., Moradi, M. H., and Rostami, R. (2013). Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Comput. Methods Prog. Biomed.* 109, 339–345. doi: 10.1016/j.cmpb.2012.10.008

Hu, J., Shen, L., and Sun, G. (2018). "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7132–7141. doi: 10.1109/CVPR.2018.00745

Jiang, C., Li, Y., Tang, Y., and Guan, C. (2021). Enhancing EEG-based classification of depression patients using spatial information. *IEEE Trans. Neural Syst. Rehabil. Eng.* 29, 566–575. doi: 10.1109/TNSRE.2021.3059429

Ke, H., Chen, D., Shah, T., Liu, X., Zhang, X., Zhang, L., et al. (2020). Cloud-aided online EEG classification system for brain healthcare: a case study of depression evaluation with a lightweight CNN. *Softw. Pract. Exp.* 50, 596–610. doi: 10.1002/spe.2668

Kerestes, R., Davey, C. G., Stephanou, K., Whittle, S., and Harrison, B. J. (2014). Functional brain imaging studies of youth depression: a systematic review. *Neuroimage Clin.* 4, 209–231. doi: 10.1016/j.nicl.2013.11.009

Kiers, H. A. (1997). Weighted least squares fitting using ordinary least squares algorithms. *Psychometrika* 62, 251–266. doi: 10.1007/BF022 95279

Knott, V., Mahoney, C., Kennedy, S., and Evans, K. (2001). EEG power, frequency, asymmetry and coherence in male depression. *Psychiatry Res. Neuroimaging* 106, 123–140. doi: 10.1016/S0925-4927(00)00080-9

Kroenke, K., and Spitzer, R. L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatr. Ann.* 32, 509–515. doi: 10.3928/0048-5713-20020901-06

Kukavica, I., and Robinson, J. C. (2004). Distinguishing smooth functions by a finite number of point values, and a version of the takens embedding theorem. *Phys. D Nonlinear Phenomena* 196, 45–66. doi: 10.1016/j.physd.2004.04.004

Liao, S.-C., Wu, C.-T., Huang, H.-C., Cheng, W.-T., and Liu, Y.-H. (2017). Major depression detection from EEG signals using kernel eigen-filter-bank common spatial patterns. *Sensors* 17, 1385. doi: 10.3390/s17061385

Lin, M., Chen, Q., and Yan, S. (2013). *Network in Network*. New York, NY: Cornell University.

Luu, P., and Ferree, T. (2005). *Determination of the HydroCel Geodesic Sensor Nets? Average Electrode Positions and Their 10-10 International Equivalents*. Electrical Geodesics, Inc.

Mahato, S., and Paul, S. (2020). Classification of depression patients and normal subjects based on electroencephalogram (EEG) signal using alpha power and theta asymmetry. *J. Med. Syst.* 44, 1–8. doi: 10.1007/s10916-019-1486-z

Michel, C. M., and Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage* 61, 371–385. doi: 10.1016/j.neuroimage.2011.12.039

Muldoon, M. R., Broomhead, D. S., Huke, J. P., and Hegger, R. (1998). Delay embedding in the presence of dynamical noise. *Dyn. Stabil. Syst.* 13, 175–186. doi: 10.1080/02681119808806259

Mumtaz, W., Xia, L., Mohd Yasin, M. A., Azhar Ali, S. S., and Malik, A. S. (2017). A wavelet-based technique to predict treatment outcome for major depressive disorder. *PLoS ONE* 12, e0171409. doi: 10.1371/journal.pone.0171409

Olbrich, S., and Arns, M. (2013). EEG biomarkers in major depressive disorder: discriminative power and prediction of treatment response. *Int. Rev. Psychiatry* 25, 604–618. doi: 10.3109/09540261.2013.816269

Olesen, J., Gustavsson, A., Svensson, M., Wittchen, H.-U., Jönsson, B., CDBE2010 Study Group, et al. (2012). The economic cost of brain disorders in Europe. *Eur. J. Neurol.* 19, 155–162. doi: 10.1111/j.1468-1331.2011.03590.x

Peng, H., Xia, C., Wang, Z., Zhu, J., Zhang, X., Sun, S., et al. (2019). Multivariate pattern analysis of EEG-based functional connectivity: a study on the identification of depression. *IEEE Access* 7, 92630–92641. doi: 10.1109/ACCESS.2019.2927121

Pham, T. D., Tran, D., Ma, W., and Tran, N. T. (2015). "Enhancing performance of EEG-based emotion recognition systems using feature smoothing," in *International Conference on Neural Information Processing* (Switzerland: Springer), 95–102. doi: 10.1007/978-3-319-26561-2_12

Puthankattil, S. D., and Joseph, P. K. (2012). Classification of EEG signals in normal and depression conditions by ANN using RWE and signal entropy. *J. Mech. Med. Biol.* 12, 1240019. doi: 10.1142/S0219519412400192

Roulston, M. S. (1999). Estimating the errors on measured entropy and mutual information. *Phys. D Nonlinear Phenomena* 125, 285–294. doi: 10.1016/S0167-2789(98)00269-3

Seal, A., Bajpai, R., Agnihotri, J., Yazidi, A., Herrera-Viedma, E., and Krejcar, O. (2021). Deprnet: a deep convolution neural network framework for detecting depression using EEG. *IEEE Trans. Instrument. Meas.* 70, 1–13. doi: 10.1109/TIM.2021.3053999

Sharma, G., Parashar, A., and Joshi, A. M. (2021). DepHNN: a novel hybrid neural network for electroencephalogram (EEG)-based screening of depression. *Biomed. Signal Process. Control* 66, 102393. doi: 10.1016/j.bspc.2020.102393

Sharma, M., Achuth, P., Deb, D., Puthankattil, S. D., and Acharya, U. R. (2018). An automated diagnosis of depression using three-channel bandwidth-duration localized wavelet filter bank with EEG signals. *Cogn. Syst. Res.* 52, 508–520. doi: 10.1016/j.cogsys.2018.07.010

Sharma, R., and Pachori, R. B. (2015). Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions. *Expert Syst. Appl.* 42, 1106–1117. doi: 10.1016/j.eswa.2014.08.030

Siuly, S., Kabir, E., Wang, H., and Zhang, Y. (2015). Exploring sampling in the detection of multicategory EEG signals. *Comput. Math. Methods Med.* 2015, 576437. doi: 10.1155/2015/576437

Spitzer, R. L., Kroenke, K., Williams, J. B., and Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the gad-7. *Arch. Internal Med.* 166, 1092–1097. doi: 10.1001/archinte.166.10.1092

Stark, J., Broomhead, D. S., Davies, M. E., and Huke, J. (1997). Takens embedding theorems for forced and stochastic systems. *Nonlinear Anal. Theory Methods Appl.* 30, 5303–5314. doi: 10.1016/S0362-546X(96)00149-6

Sun, S., Chen, H., Shao, X., Liu, L., Li, X., and Hu, B. (2020a). "EEG based depression recognition by combining functional brain network and traditional biomarkers," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (Seoul: IEEE), 2074–2081. doi: 10.1109/BIBM49941.2020.9313270

Sun, S., Li, J., Chen, H., Gong, T., Li, X., and Hu, B. (2020b). *A Study of Resting-state EEG Biomarkers for Depression Recognition*. New York, NY: Cornell University.

Van Der Stelt, O., and Belger, A. (2007). Application of electroencephalography to the study of cognitive and brain functions in schizophrenia. *Schizophr. Bull.* 33, 955–970. doi: 10.1093/schbul/sbm016

Wan, Z., Huang, J., Zhang, H., Zhou, H., Yang, J., and Zhong, N. (2020). HybridEEGNet: a convolutional neural network for EEG feature learning and depression discrimination. *IEEE Access* 8, 30332–30342. doi: 10.1109/ACCESS.2020.2971656

Wang, Z., Ma, Y., Liu, Z., and Tang, J. (2019). *R-transformer: Recurrent Neural Network Enhanced Transformer*. New York, NY: Cornell University.

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., et al. (2013). Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. *Lancet* 382, 1575–1586. doi: 10.1016/S0140-6736(13)61611-6

Zhang, X., Li, J., Hou, K., Hu, B., Shen, J., and Pan, J. (2020). "EEG-based depression detection using convolutional neural network with demographic attention mechanism," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (Montreal, QC: IEEE), 128–133. doi: 10.1109/EMBC44109.2020.9175956