

Research Article

The Impact of Emotional Music on Active ROI in Patients with Depression Based on Deep Learning: A Task-State fMRI Study

Renzhou Gui , Tongjie Chen , and Han Nie 

The Department of Information and Communication Engineering, Tongji University, Shanghai 201804, China

Correspondence should be addressed to Tongjie Chen; chentongjie@tongji.edu.cn

Received 8 July 2019; Revised 12 September 2019; Accepted 14 September 2019; Published 20 October 2019

Academic Editor: Luis Javier Herrera

Copyright © 2019 Renzhou Gui et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuous development of science, more and more research results have proved that machine learning is capable of diagnosing and studying the major depressive disorder (MDD) in the brain. We propose a deep learning network with multi-branch and local residual feedback, for four different types of functional magnetic resonance imaging (fMRI) data produced by depressed patients and control people under the condition of listening to positive- and negative-emotions music. We use the large convolution kernel of the same size as the correlation matrix to match the features and obtain the results of feature matching of 264 regions of interest (ROIs). Firstly, four-dimensional fMRI data are used to generate the two-dimensional correlation matrix of one person's brain based on ROIs and then processed by the threshold value which is selected according to the characteristics of complex network and small-world network. After that, the deep learning model in this paper is compared with support vector machine (SVM), logistic regression (LR), k-nearest neighbor (kNN), a common deep neural network (DNN), and a deep convolutional neural network (CNN) for classification. Finally, we further calculate the matched ROIs from the intermediate results of our deep learning model which can help related fields further explore the pathogeny of depression patients.

1. Introduction

Depression is a complex mental illness in which the sufferer will continuously feel depressed, negative, and pessimistic and even have suicidal thoughts. This not only does great harm to the patient himself but also imposes a burden on his family, friends, and the people around him [1].

This disease poses great challenges to the accurate diagnosis and effective and timely treatment of the medical community. Human emotion is a huge and complex process; emotional regulation can help us develop physically and mentally and has a great positive effect on the symptoms of patients with depression. However, the human brain is composed of about 100 billion neurons, trillions of glial cells, and other small cells that maintain the function of neurons [2]. All parts of the brain are connected and interact with each other to form a huge network organization.

Functional magnetic resonance imaging (fMRI), as a neural imaging mode, can display the functional activity of specific parts of the brain in real time with high spatial and temporal resolution [3]. There are now many fMRI-based

analyses of depressive patients, which have contributed significantly to the comparison and analysis of different brain activities between normal and depressed people.

Lepping et al. [4–9] used the fMRI scanning process to conduct experiments on 19 depressed patients, MDD, and 20 never-depressed subjects, ND. They were asked to constantly listen to standardized positive- and negative-emotions music and no-music stimuli. And after the scanning, they were evaluated subjectively. The data set was published online (<https://openfmri.org/dataset/ds000171/>). The result turns out that people with depression may process emotional auditory stimuli differently depending on the type of stimulus and the emotional content of the stimulus. Helping to find the differences between research on music and emotion will improve our understanding of the complexities of the human brain. And the findings on emotional reactivity in MDD may lead to more effective and targeted treatment.

Computer science can benefit a lot from fMRI-based brain research [10]. With the development of time, more and more machine learning methods have been found to mine fMRI data, understand the deeper thinking process of the

brain [11], and find the relationship between fMRI data and related tasks [12]. By processing the different states and types of data, machine learning can be used to classify the fMRI data, including independent component analysis (ICA) [13], support vector machine (SVM) [14, 15], k-nearest neighbor (kNN) [16], Gauss naive Bayes (GNB), linear discriminant analysis (LDA) [17], logistic regression (LR) [18], autoencode [19], deep neural network (DNN), and convolutional neural network (CNN). Cai et al. [13] used the ICA method to achieve the accuracy of 72% (13/18). Afshin-Pour et al. [17] used generalized canonical correlation analysis (gCCA) and linear discriminant analysis (LDA) to realize the accuracy of 75%. Ramasubbu et al. [14] used linear SVM to process the fMRI data of resting state and classify MDD patients with major depression and the control group with 66% accuracy. Khazaei et al. [15] used the SVM method to select 264 regions of interest (ROI) [20] and corresponding signals from meta-analysis and fc mapping methods. They classified three groups: HC, AD, and MCI patients. The classification accuracy reached 88.42%. Kamonsantiroj et al. [19] used powerful combination of searchlight information mapping and decoding by autoencode to reduce the dimensionality and learn efficient codings. They achieved improvements of classification accuracy more than 20% in comparing with current neuroimaging methods.

Most of the data used in these models are relatively easy to distinguish, but the accuracy is not high enough. In recent years, deep learning has been widely used in learning and mining of fMRI data. And its powerful ability of data analysis and classification has attracted more and more attention. Zhang and Ji [21] used the CNN to classify fMRI data corresponding to 8 items, with an accuracy rate of 87.69%. Tahmassebi et al. [22] use CNN to predict relapse in heavy smoker subjects based on fMRI of resting state. They achieved 0.86 precision through the XGBoost algorithm. Sarraf and Tofghi [23, 24] used CNN to distinguish between people with Alzheimer's disease and healthy people. By using LeNet and GoogleNet frameworks, they achieved accuracy of 96.86% and 98.84%, respectively. Experiments showed CNN's shift and scale invariant features. Coupled with deep learning classification, it is proved that CNN is the most powerful method in fMRI to distinguish clinical data and health data. At the same time, current academic studies have also explored more about the brain samples of depressed patients.

The above research results are based on resting state data and have high accuracy. This is not the same as the results of the task data. For the fMRI data of task state, requiring subjects to do corresponding tasks during scanning can avoid thinking about irrelevant things. This way can better reflect the changes of the brain in response to external stimuli, as well as including the information of differences between depressed patients and normal people. But due to the complexity of fMRI data in task state, its classification accuracy is lower. And classification accuracy varies greatly according to different data. For task-state data of listening stimulus, the accuracy rate is not generally very high. Ni et al. [25] choosed SVM algorithm for classifications. Their average accuracy achieved 77.38%. Wang et al. [26] presented a novel depression disorder classification algorithm, named

weighted discriminative dictionary learning (WDDL) for fMRI data in task state, and achieved 79.31% accuracy. Based on SVM and recursive feature elimination (RFE), Chanel et al. [27] proposed a fully data-driven (voxel-based) approach to distinguish individuals with autistic spectrum disorders (ASD) from controls which can be applied to two different fMRI experiments with social stimuli (faces and bodies). Compared with the above studies that rely on resting state connectivity measurements, their method achieves accuracy between 69% and 92.3%. Chung et al. [28] presented exploratory work on the use of node-based heat kernel features applied to functional MRI data and achieved an average accuracy of 83.7%. Rosa et al. [29] used SVM to establish the fMRI data discrimination modeling framework based on sparse network and the pattern recognition which was used to distinguish MDD between normal people and patients with deep depression and finally obtained the accuracy rate of 85%. Ertugrul et al. [30] proposed a novel framework to encode the local connectivity patterns of the brain. They classified the cognitive states of Human Connectome Project (HCP) task fMRI dataset by training SVM. When the pairwise correlation between blood oxygen level-dependent (BOLD) response pairs in all regions was characterized, the classification accuracy was 77.49%.

As mentioned above, the important role of deep learning in fMRI data processing and analysis is also reflected. And the potential of machine learning in developing a brain-based diagnostic method for MDD has been confirmed. In this paper, a new form of deep learning network and large convolution kernel is proposed for four different types of fMRI data produced by depressed patients and normal people under the condition of listening to positive- and negative-emotions music, and the network structure is improved by multibranch and local residual feedback. In addition, most machine learning methods are only for classification of types; the statistics of active ROI and further analysis of the pathogenesis of MDD have not been realized which are most important to analyze the human emotion and brain decoding.

In this paper, we used the correlation matrix ROI [20] of 264 ROIs for single people's brains to perform high-precision classification and feature matching. In order to verify the availability of the new deep learning model proposed in this paper, we compare our model with five other common machine learning models, SVM, LR, kNN, a normal DNN model, and a CNN model represented by Inception-ResNet-v2 [31]. Finally, to study the pathogenesis and the operation rules of ROI of the brain in response to music stimulation, this paper used the intermediate convolution results of the trained deep learning network model to conduct statistics and explore the greater impact of ROI on depressed patients. Through the analysis and statistics of convolution features in the classifier, ROIs of high correlation with the corresponding brain regions were obtained. Finally, we compared and verified the statistical results with other studies.

2. Materials and Methods

The overall data preprocessing and deep learning classification process in this paper are shown in Figure 1. The fMRI data are

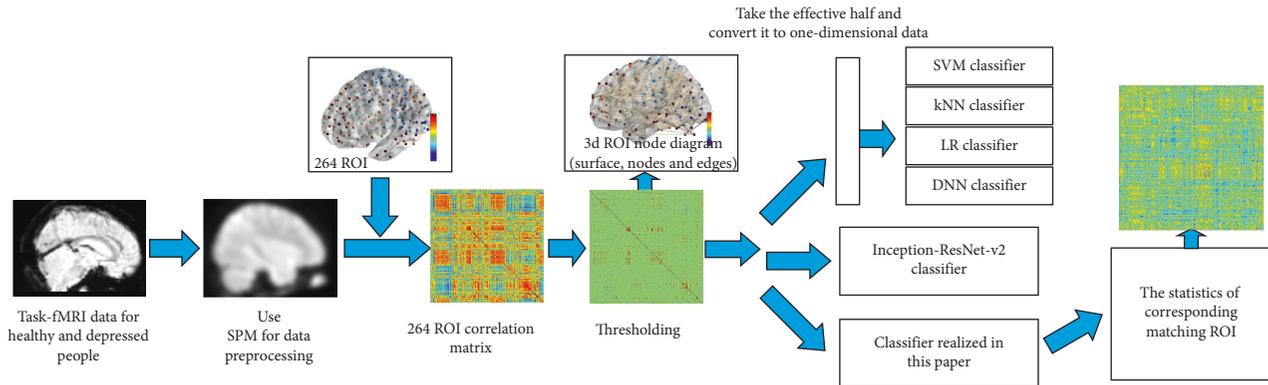


FIGURE 1: Entire process of data processing in this paper.

first preprocessed to obtain correlation matrixes of 264 ROI [20], which are then processed by thresholds and classified by different machine learning models. After classification, the trained convolution kernels were used to calculate the statistics of matched ROI and the ROI positions with obvious difference between depressed patients and normal people.

2.1. Experimental Materials and Preprocessing. The experimental data of Lepping et al. [4–9] are used in this paper; their study was approved by the Human Subject Committee of the University of Kansas Medical Center and found that stimulation with music was more active in the corresponding brain regions than stimulation without music. The data can be obtained from the OpenfMRI database. Its accession number is ds000171. Scanning was conducted on a 3 Tesla Siemens Skyra scanner (Siemens, Erlangen, Germany) with TR/TE/flip angle = 3 s/25 ms/90 degree, field of view (FOV) = 220 mm, matrix = 64×64 , slice thickness = 3 mm, 0 mm skip, and in-plane resolution = 2.9×2.9 mm. In addition to this, the 3-dimensional anatomical data of size $176 \times 256 \times 256$ with a voxel size of 1 mm^3 have been acquired. Figure 2 shows slices of the brain from one patient in all three axes for both the anatomical (Figure 2(a)) and the functional (Figure 2(b)) representations.

In the process of fMRI scans, 19 patients with MDD (11 women, nine men, with an average age of 34.15) and 20 never-depressed subjects (ND) (11 female, eight men, average age of 28.5) listen to the no-music and the music stimuli of positive and negative emotions. Participants in the MDD group were all experiencing a current depressive episode at the time of scanning, determined by screening for research purposes using the SCIDI/NP. It was evaluated subjectively after scanning. Participants ranged from 18 to 59 in age, and the MDD participants in the study all experienced the current depressive episode when tested.

Each subject was tested five times in this experiment, including 3 times of music-type stimulation and 2 times of no-music stimulation. A total of $19 \times 5 + 20 \times 5 = 195$ groups were tested, among which $19 \times 3 + 20 \times 3 = 117$ groups were music-stimulation experiments (data adopted in this paper). Each test experienced the process as shown in Figure 3. In addition to the pure tones, two positive music stimuli and two negative music stimuli were experienced, and the results of $(2 + 2) \times 117 = 468$ groups were obtained.

Statistical parametric mapping (SPM12) (<https://www.fil.ion.ucl.ac.uk/spm/>) was used to preprocess the data which included slice of time correction, motion correction, normalization of standard space, and smooth filtering.

After calculating the time series of 264 brain regions, the toolboxes of DPARSFA [32] were used to generate the correlation coefficient matrix for different states and objects and the corresponding labels including normal people with positive music, normal people with negative music, depressed patient with positive music, and depressed patient with negative music.

In the study of this paper, there are 39 objects, each with 12 samples. Therefore, there are a total of 468 sample data. These data can be divided into data for 228 depressed patients and 240 normal subjects according to the type of subjects. It can also be divided into data for 234 positive music stimuli and 234 negative music stimuli according to the type of emotion of the music. A total of 468 correlation matrixes are generated, 80% of which are used as training data and 20% are used as test data.

2.2. Threshold Processing. The human brain is a complex three-dimensional network where each ROI can be viewed as a node of the network. Each value of the correlation coefficient matrix can be considered as the weight of each node pair in the undirected complex network. That is to say, the correlation coefficient matrix of the fMRI data is similar to the undirected network structure. This is a complex network approach to studying the interactions between human brain functional regions.

Brain anatomical networks are sparse and complex and have economical small-world properties [33]. The network has a lot of noise and weak connections with little correlation. Through processing by threshold, weak connections and noise edges can be removed to ensure that the two connected nodes have a high similarity in time behavior. Different values of the threshold may affect the statistical characteristics and topology of the brain network constructed [2, 34]. The value of the choice should be as high as possible, while network connectivity (no isolated brain regions) ensures the relative integrity of the network and the small world of the network [35, 36].

Therefore, the ideal threshold can be obtained through the constraints of complex networks and small-world networks. The main statistical characteristics of complex

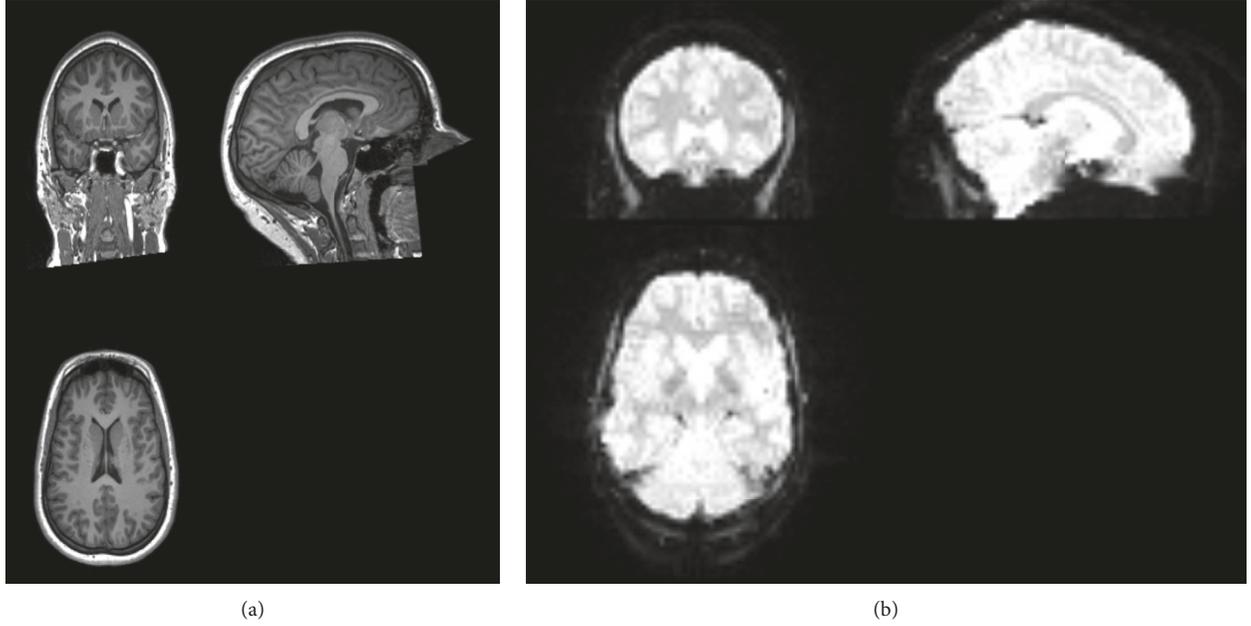


FIGURE 2: (a) Anatomical and (b) functional slices of the brain.

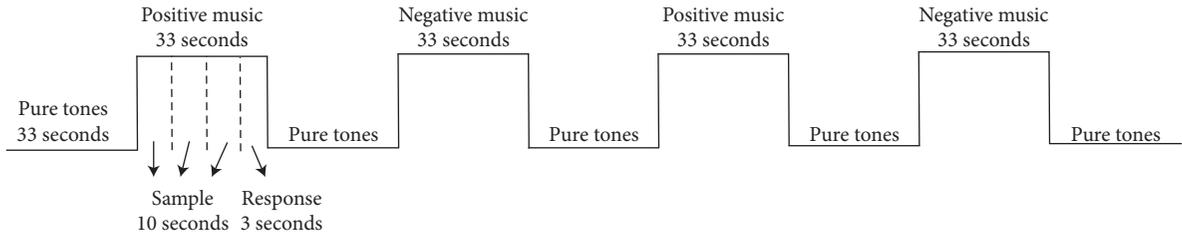


FIGURE 3: Test time process of the subjects.

networks [37, 38] include network average value $\langle k \rangle$, average clustering coefficient C_{net} , and average path length L_{net} . C_{net} is the coefficient of the degree of clustering among vertices which can represent the degree of interconnection between the adjacent points of ROI in the brain network. L_{net} represents the average distance between all pairs of nodes, and it describes the average degree of separation (how small the network is) between intermediate nodes of the brain network.

The calculation process is shown as follows:

$$\begin{aligned}
 \langle k \rangle &\geq \frac{1}{N} \sum_{i=1}^N k_i, \\
 C_{\text{net}} &= \frac{1}{N} \sum_{i=1}^N C_i = \frac{1}{N} \sum_{i=1}^N \frac{2E_i}{k_i(k_i - 1)}, \\
 L_{\text{net}} &= \frac{1}{N} \sum_{i=1}^N \frac{\sum_{j=1, j \neq i}^N d(v_i, v_j)}{N - 1}.
 \end{aligned} \tag{1}$$

For the undirected network, the maximum number of possible edges between k_i nodes is $k_i(k_i - 1)$, while the actual number of edges is E_i . Therefore, we define

$C_i = 2 \times (E_i / (k_i(k_i - 1)))$ as the clustering coefficient of node v_i . And the $d(v_i, v_j)$ is the shortest distance between v_i and v_j .

To maintain the network's small-world characteristics, the average degree of brain function network $\langle k \rangle$ should be greater than $\ln N$ and N is the number of network nodes. When the small world of the network is satisfied, the clustering coefficient C_{rand} and average characteristic path length L_{rand} of the random network of the same size (the same number of nodes and edges) should be satisfied:

$$\begin{aligned}
 \gamma &= \frac{C_{\text{net}}}{C_{\text{rand}}} \gg 1, \\
 \lambda &= \frac{L_{\text{net}}}{L_{\text{rand}}} \sim 1,
 \end{aligned} \tag{2}$$

where $C_{\text{rand}} = \langle k \rangle / N$ and $L_{\text{rand}} = \ln N / \ln k$. By continuously setting the threshold size and calculating whether each statistical feature of the corresponding network after each processing by threshold meets the above complex network constraints, the value of optimal matching threshold will be found.

The number of network nodes in this paper is 264, so the network average degree $\langle k \rangle$ cannot be smaller than $\ln 264$ which is 5.5759. Different thresholds and corresponded values

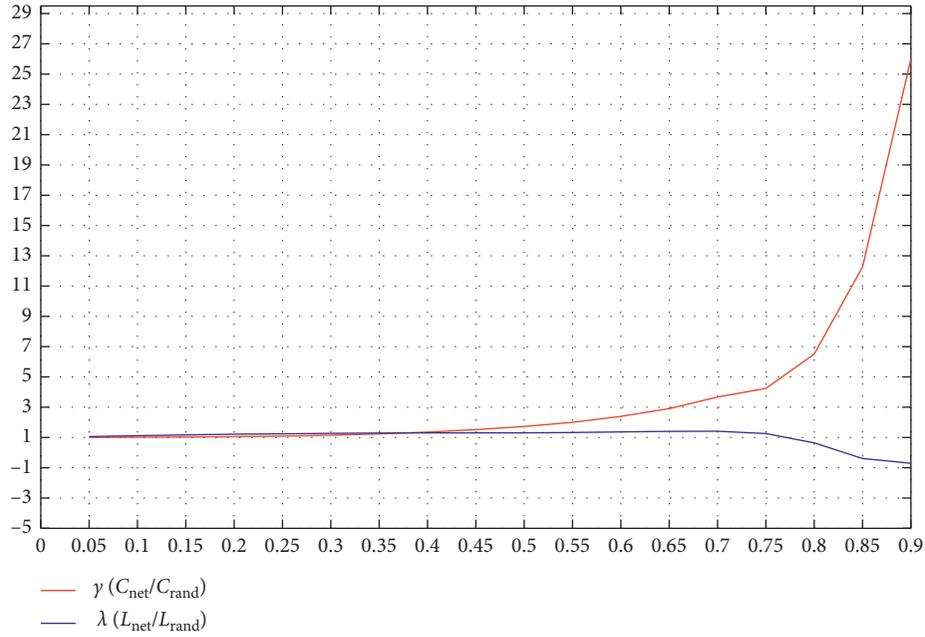


FIGURE 4: Variation curve of γ and λ under different thresholds.

are shown in Figure 4. According to this statistics, when the threshold is 0.75, the average network degree is about 26.3820, which is much larger than 5.5759. And the average clustering coefficient $C_{net} = 0.4238$ and the average characteristic path length $L_{net} = 2.1439$. At this time, $C_{rand} = 0.0999$ and $L_{rand} = 1.70379$. $\gamma = C_{net}/C_{rand} \approx 4.240891517 \gg 1$ and $\lambda = L_{net}/L_{rand} \approx 1.258315441 \sim 1$. These calculations result in a corresponding small-world feature and functional connectivity for the corresponding brain network. At the threshold of 0.75, the result and distribution of the network characteristic value of one sample are shown in Figure 5. And the effect diagram of the correlation coefficient matrix processed by the threshold value is shown in Figure 6.

2.3. Deep Learning Models. Nowadays, CNN plays an important role in the field of deep learning of images. The convolution kernel is locally sensitive to the input space and can distinguish between displacement, scaling, and other forms of distortion of the image object and content, thereby better mining the strong local spatial correlation in the image.

However, for our correlation coefficient matrix, there is no correlation between positions in the matrix, as shown in Figure 7. Position (I, J) represents the correlation coefficient of the Ith ROI and the Jth ROI and has no correlation with the correlation coefficients of position (K, M). However, the feature matching function of convolution can be applied well. Due to the calculation of the two-dimensional convolution, the closer the feature of the matching image is to the convolution kernel, the higher the calculation result will be. At the same time, through continuous training, each convolution kernel is also a more approximate matching feature as shown in Figure 8. Therefore, in order to match ROI with important

characteristics in the coefficient matrix, large convolution kernel of the same size as the matrix can be used to achieve matching. In this way, we can not only classify the data but also get those ROIs that contribute more to the classification results and explore significantly different ROI positions between depressed and normal people.

In 2014, the GoogleNet [31, 39] network won the 2014 ImageNet champion. Because of the efficient Inception network structure, it has a deeper network structure but high computational efficiency. Drawing on the idea of net in net, these filters have different feeling fields and achieve high classification accuracy.

The basic structure of the inception module is shown in Figure 8(a). This module can be used not only in the network structure of CNN but also in multilayer perceptions to achieve the same effect. The convolution modules with different sizes and dimensions can be replaced by perceptron modules with different widths and depths to realize DNN in the common deep learning network, as shown in Figure 9(b).

In this paper, drawing on the Inception structure of DNN form and the idea of residual networks (ResNet) [40], we proposed a deep learning network with multibranch and local residual feedback for task-state fMRI data, as shown in Figure 10. This network has 17 layers, and the total number of parameters is 114159902. The network adopts ReLU as activation function and cross entropy as loss function to train the network. The model also prevents overfitting by Dropout with 0.5 rate and L2 regularization.

In the network structure, the convolutional layer's convolution kernel is the same size as the input correlation coefficient matrix, which fully utilizes the CNN feature matching capability. After convolution, 1,500 feature matching results are obtained and then sent to the two DNN's Inception-ResNet structures. Finally, after the full

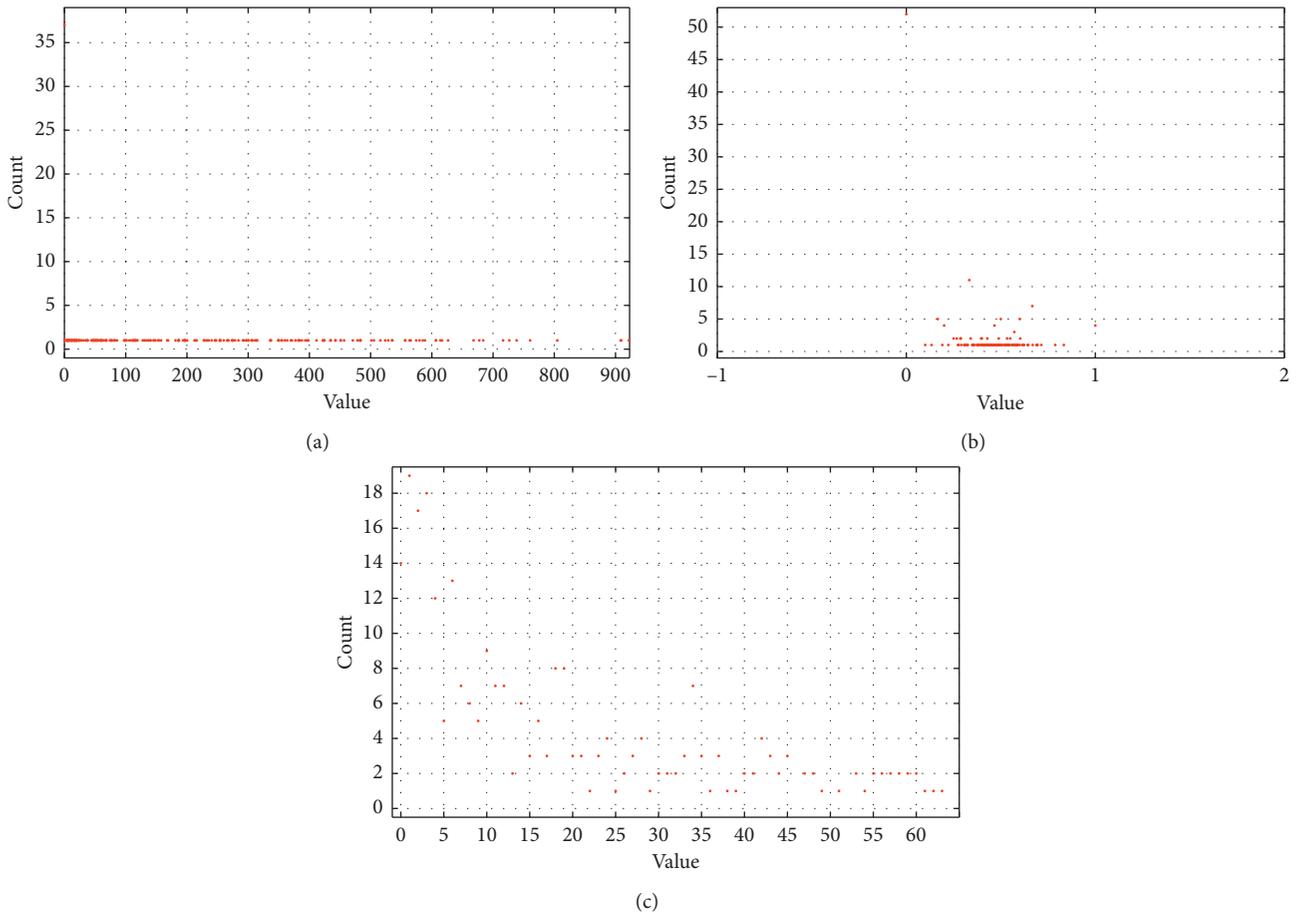


FIGURE 5: Change curve of an fMRI-related characteristic value at the threshold of 0.75. (a) Betweenness centrality distribution. (b) Clustering coefficient distribution. (c) Degree distribution.

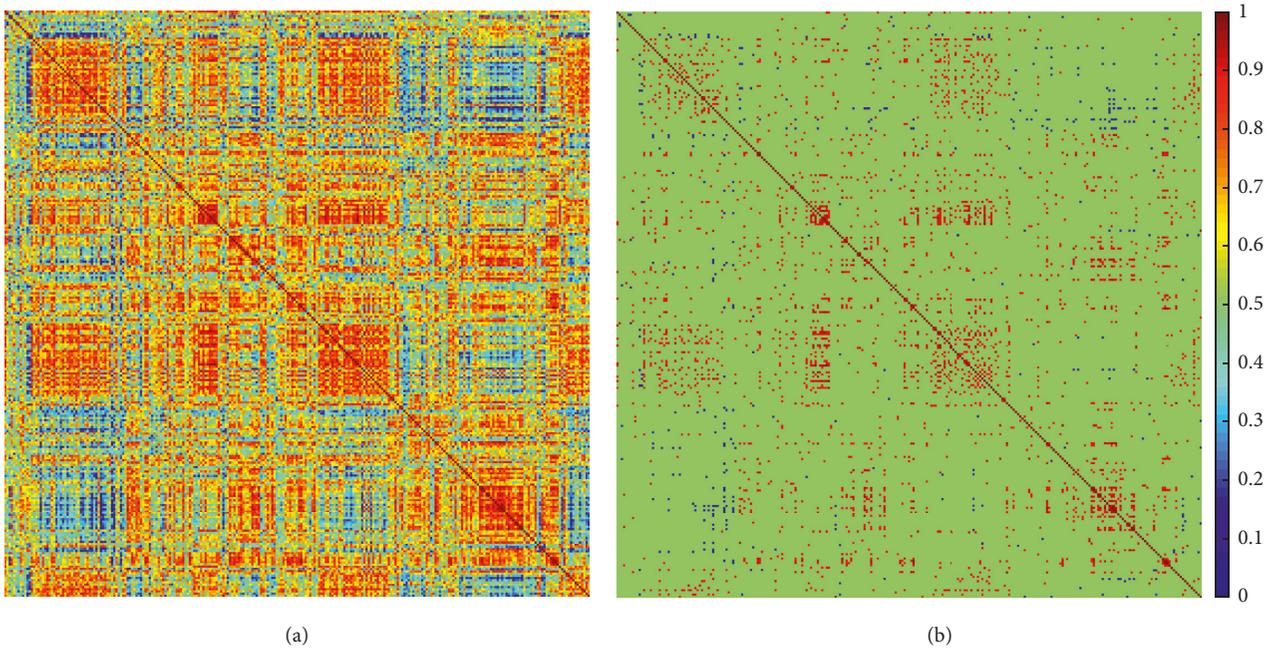


FIGURE 6: Correlation matrix rendering. (a) Before threshold processing. (b) After threshold processing.

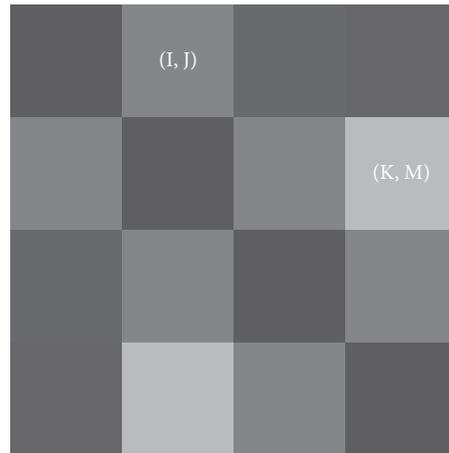


FIGURE 7: Two different points in the correlation coefficient matrix.

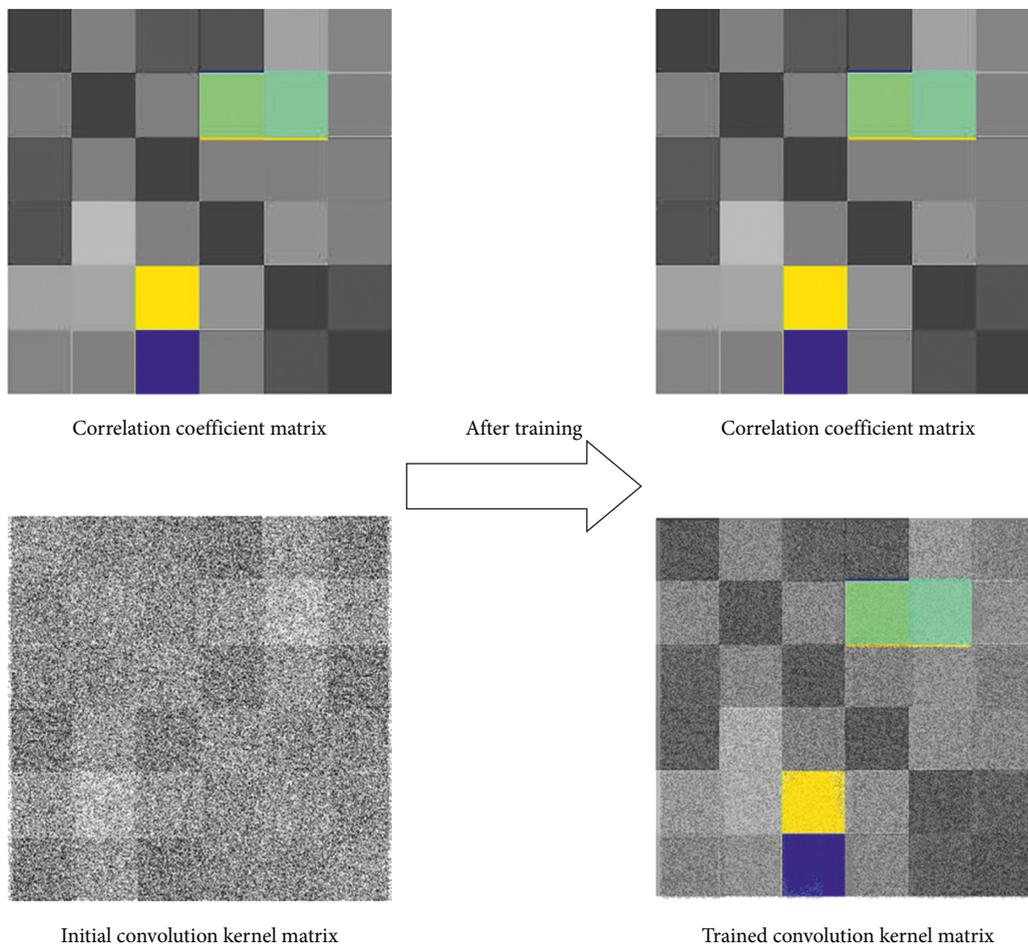


FIGURE 8: Matrix of convolution kernel before and after training.

link layer, the classification output is obtained by SoftMax layer. Such a network not only ensures sufficient depth to make the network have good nonlinearity but also transmits errors to each layer to better train the entire network.

For the optimization algorithm, there are many algorithms for the optimization of the parameters of the

deep learning model, including stochastic gradient descent (SGD), root mean square propagation (RMSprop), adaptive moment estimation (Adam), and Nesterov adaptive moment estimation (Nadam) [41]. Through continuous algorithmic improvements and targeted modifications, they can achieve better performance for

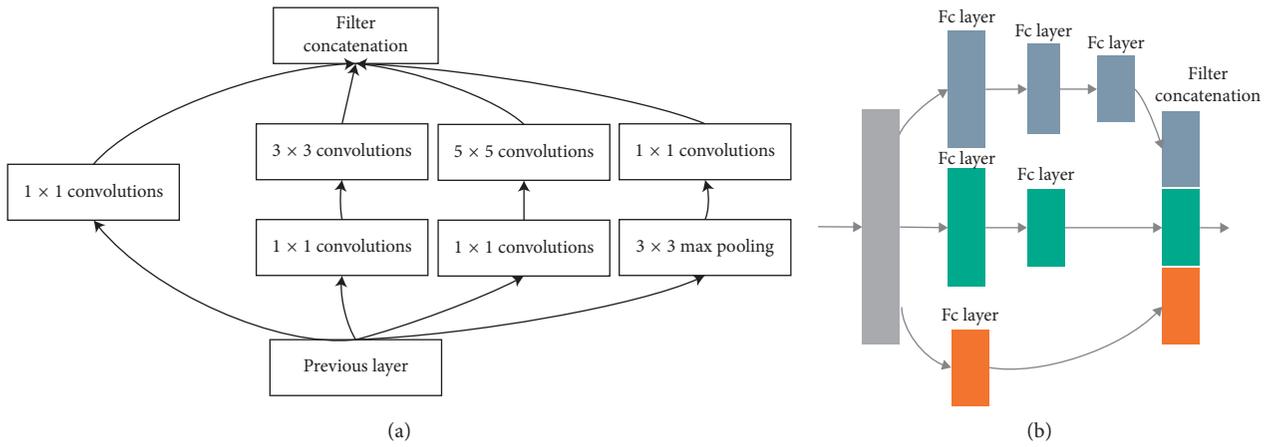


FIGURE 9: Basic structure of the inception module. (a) The original form. (b) Modified form.

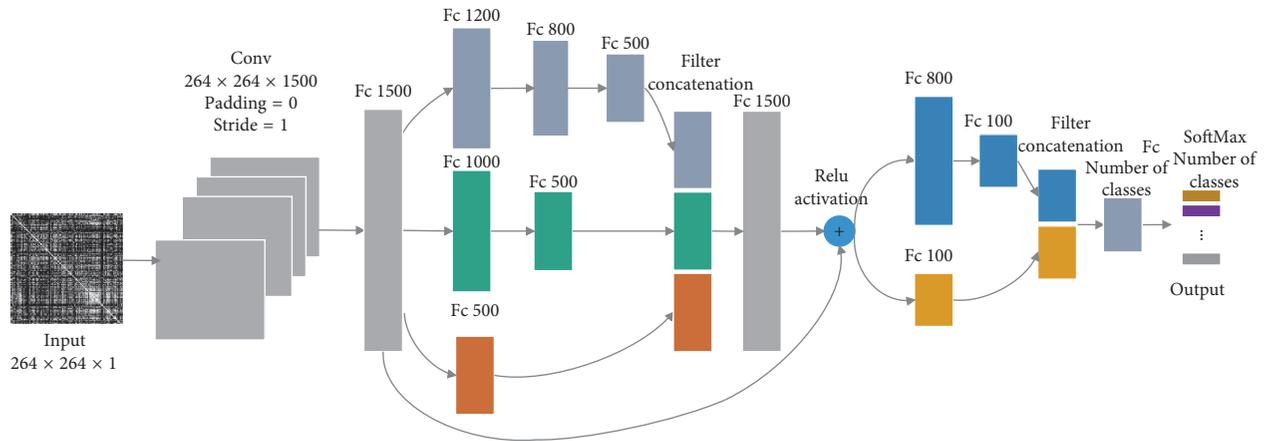


FIGURE 10: Deep network model in this paper.

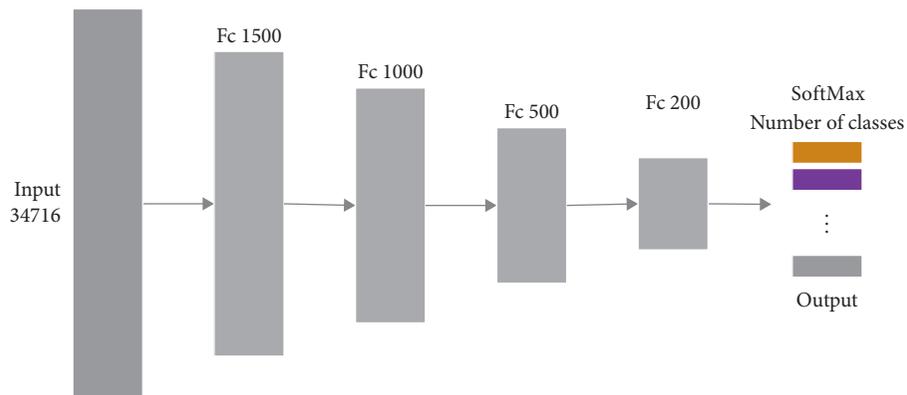


FIGURE 11: DNN model.

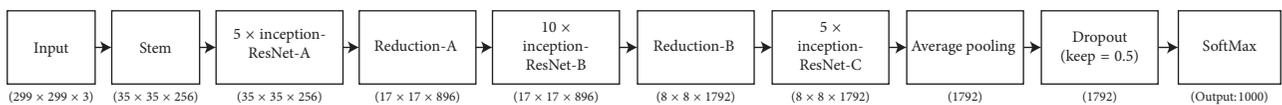


FIGURE 12: Inception V4 network model.

TABLE 1: Classification scores.

Type of music stimulus	Classification score	kNN	SVM	LR	DNN	GoogleNet	Model in this paper
Positive and negative music	Accuracy	0.5213	0.7979	0.7872	0.9149	0.8617	0.9468
	Recall	0.2979	0.7660	0.7021	0.8936	0.8085	0.8936
	AUC	0.5430	0.8918	0.8628	0.9701	0.9393	0.9887
	F1 score	0.3836	0.7912	0.7674	0.9130	0.8539	0.9438
Positive music	Accuracy	0.3617	0.6596	0.5532	0.8298	0.6808	0.9361
	Recall	1.0	0.8462	0.8462	0.4615	1.0	0.8461
	AUC	0.6538	0.8303	0.7217	0.7805	0.9253	0.9864
	F1 score	0.4643	0.5789	0.5116	0.6	0.6341	0.88
Negative music	Accuracy	0.5532	0.7447	0.6596	0.8298	0.8298	0.8936
	Recall	0.36	0.68	0.56	0.92	0.84	0.8
	AUC	0.7	0.8327	0.7636	0.9655	0.94	0.9909
	F1 score	0.4615	0.7391	0.6363	0.8519	0.84	0.8889

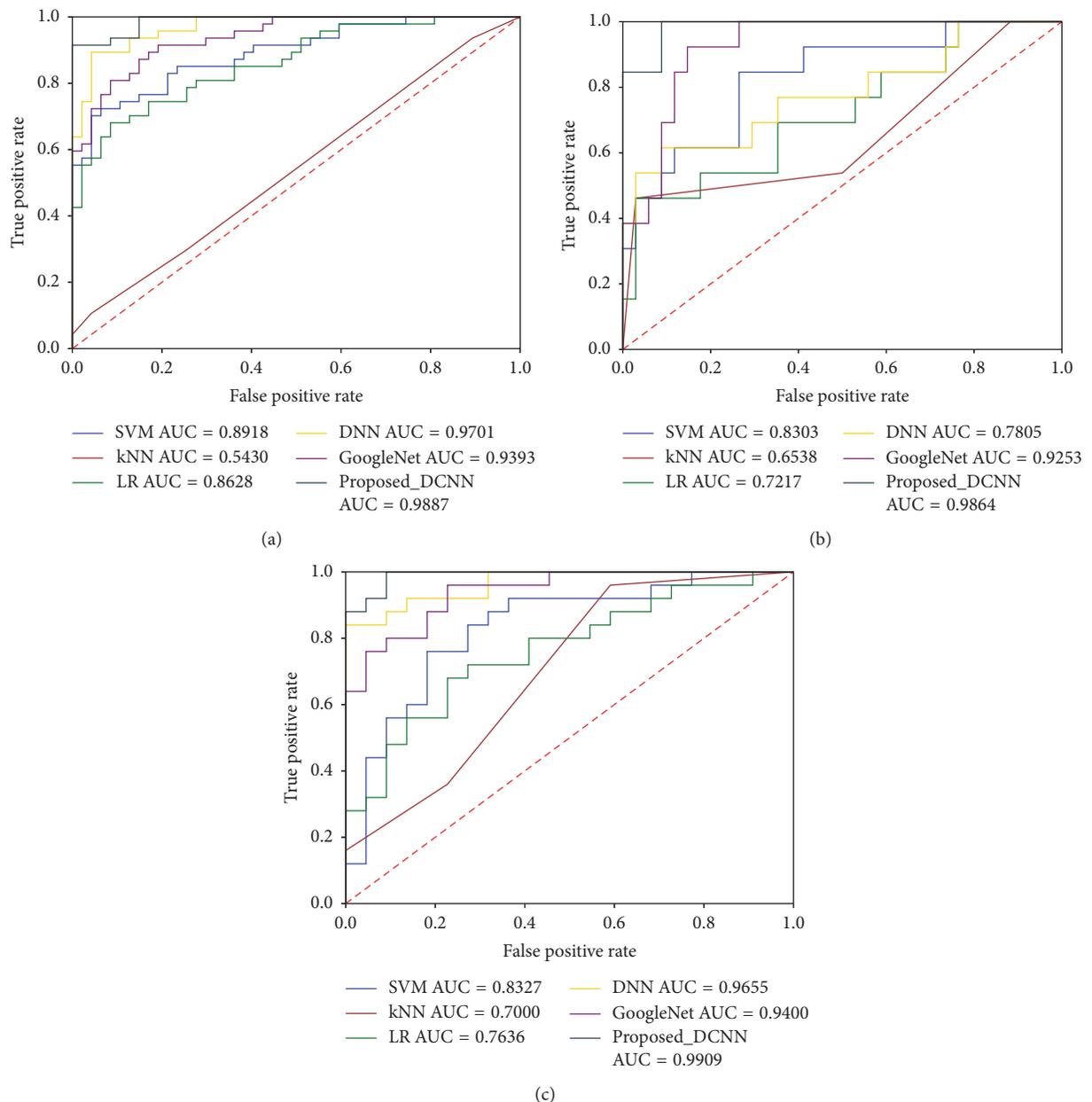


FIGURE 13: ROC curves for classification of both positive and negative music using six models. (a) Under both positive and negative music. (b) Only under positive music. (c) Only under negative music.

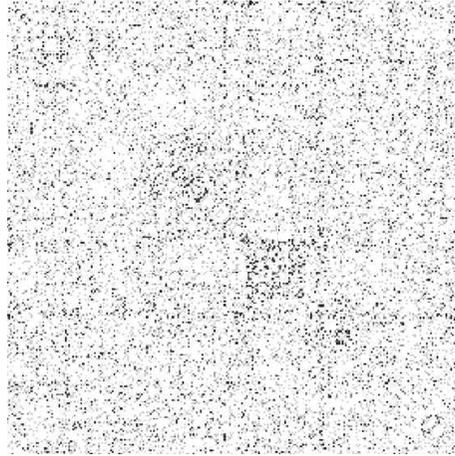


FIGURE 14: Graph of the result of the cumulative calculation of 1,500 convolution matrices (black for high statistics and white for low statistics). Each convolution matrix is thresholded with a threshold of 0.75 before being accumulated to remove the effects of noise.

TABLE 2: ROI correlation statistical results.

ROI#	Brain area	Hemisphere/BA	Stereotaxic coordinates			Statistical value
			x	y	z	
1	Inferior occipital gyrus	L/17	-25	-98	-12	146
142	Lingual gyrus	L/17	-12	-95	-13	
178	Middle frontal gyrus	L/6	-23	11	64	104
185	Posterior lobe, pyramis	R/*	35	-67	34	
222	Thalamus	R/*	6	-24	0	80
226	Thalamus	L/*	-5	-28	-4	
159	Cuneus	R/19	15	-77	31	76
182	Superior frontal gyrus	L/11	-21	41	-20	
177	Inferior parietal lobule	L/40	-53	-49	43	73
204	Inferior parietal lobule	R/40	55	-45	37	

TABLE 3: ROI statistical results.

ROI#	Brain area	Hemisphere/BA	Stereotaxic coordinates			Statistical value
			x	y	z	
142	Lingual gyrus	L/17	-12	-95	-13	392
2	Fusiform gyrus	R/18	27	-97	-13	389
185	Posterior lobe, pyramis	R/*	35	-67	-34	383
182	Superior frontal gyrus	L/11	-21	41	-20	366
184	Posterior lobe, pyramis	R/*	17	-80	-34	349
127	Precuneus	R/19	28	-77	-32	343
141	Fusiform gyrus	R/18	17	-91	-14	327
159	Cuneus	R/19	15	-77	31	306
178	Middle frontal gyrus	L/6	-23	11	64	283
222	Sublobar, thalamus	R/*	6	-24	0	281

different models. Despite the superior training results, it has been found that the adaptive optimization methods such as Adam, Adagrad, or RMSprop are less general than the stochastic gradient descent (SGD) [42]. Keskar and Socher et al. [42] proposed a hybrid strategy that firstly trained the model by using adaptive methods and then switched to SGD when appropriate, and the result showed this method was capable of closing the generalization gap between SGD and Adam on a majority of the tasks. Therefore, through the comparison of experimental

results, we choose a simple hybrid strategy combining Adam algorithm and SGD algorithm to provide optimization of model parameters, that is, first use Adam training network to reach a stable value and then train it again by using SGD.

In order to compare with the simple models, we used SVM (LR) and kNN model to classify the data. And as the comparisons of not using convolutional layers and using small convolutional layers, we designed a simple DNN and adopted a famous CNN, Inception-ResNet-v2. In the DNN

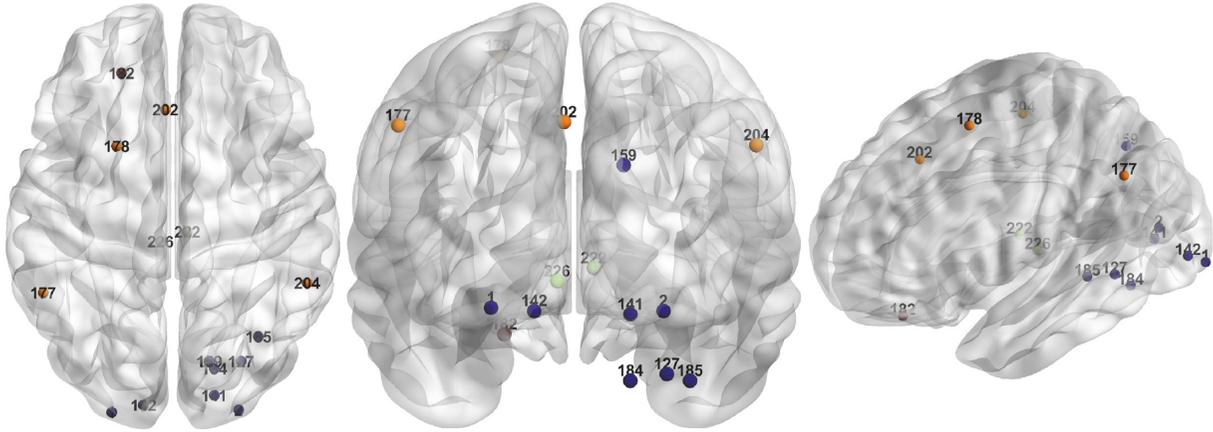


FIGURE 15: Statistical ROI distribution in the brain.

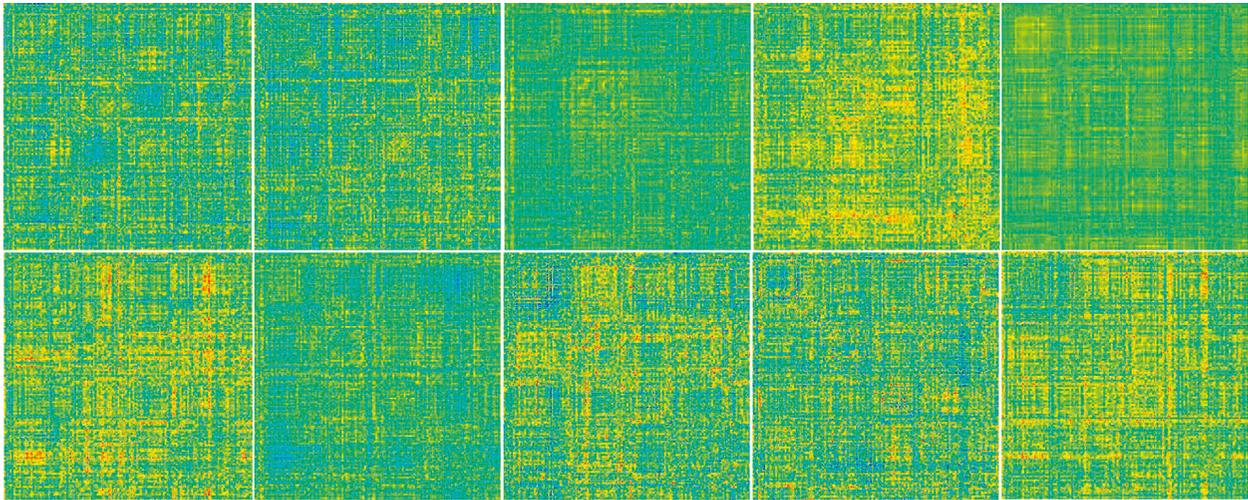


FIGURE 16: Convolution renderings of the first layer of the network model.

model, a total of 4 layers of hidden neural network are adopted to achieve the classification, as shown in Figure 11. The input layer contains 34,716 input data. The number of neurons in the four hidden layers is, respectively, 1500, 1000, 500, and 200 in order, the coefficient of Dropout in each layer is 0.5, the attenuation coefficient of L2 regularization is 0.001, the last layer removes Dropout and regularization, and the corresponding result is output by SoftMax according to the number of the classification.

The original GoogleNet Inception-ResNet-v2 network, as shown in Figure 12, consists of 10 parts, including three Inception modules and the corresponding Reduction module. The final output has no full link layer, and the final output is obtained through averaging pooling layer and SoftMax layer. Although this network has strong network and learning ability, it is easy to overfit the current data set due to the complexity of the network and cannot easily see the features learned by convolutional network.

The operating environment is as follows: the CPU is Intel(R) Core(TM) i57300hq, GPU is NVIDIA GeForce GTX 1050Ti, memory is 8G, video memory is 4G, operating system is Windows10 64 bit operating system, and we installed CUDA

10.0. Matlab2014a is used for preprocessing, Python3.6 for deep learning, and Pytorch as a deep learning tool. The deep learning is accelerated by GPU, and the optimization algorithm adopts Adam with the learning rate of 0.001 and SGD with the learning rate of 0.005; the exponential attenuation rate of the first-order moment estimation is 0.999, the exponential attenuation rate of the second-order moment estimation is 0.999, and the size of batch of each training batch is 5. The epoch times of training include 500 times when using Adam and 20 times when using SGD.

3. Experimental Results

3.1. Deep Learning and Classification Result. In this paper, we compare five other common machine learning models, such as SVM, LR, kNN, a normal DNN model, and a CNN model represented by GoogleNet Inception-ResNet-v2 with the model proposed in this paper. First, fMRI data were used to generate the correlation coefficient matrix, and then the threshold method was used for processing according to the characteristics of complex network and small-world network.

Table 1 shows the classification scores including accuracy, recall, AUC, and F1-score of the validation data set for the classification of normal people and depressed patients under positive-music and negative-music stimuli. The ROC curves corresponding to each model are shown in Figure 13. It can be found from the Table 1 that when listening to positive music, the classification accuracy was much higher than that when only listening to negative music, indicating that negative music stimulated depressed patients more than normal people. When the two types of data are trained together, the classification accuracy is further improved. Our highest classification accuracy of the model proposed in this paper reached 94.68% when listening to all types of music, 93.61% when listening to positive music, and 89.36% when listening to negative music.

3.2. Convolutional Layer Statistics and the Corresponding Active ROIs. As shown in Figure 14, after training the model, we extract the convolutional layers of the model that trains all music types and accumulate the matrixes formed by all convolutional layers. The data of this convolutional layer are used as the characteristic matching result of the correlation coefficient matrix of fMRI, and the active ROIs are counted under the current emotional music stimulation.

According to the result, the maximum statistical value is in the first row, the 142nd column in the position of the statistical matrix, and the statistical value is 146. And the corresponding ROI is the correlation between the first ROI and the 142nd ROI. The ranking of other first five statistics and corresponding ROI are shown in Table 2. In addition, separate statistics are conducted for each ROI, that is, the statistical sum of the matrix columns exceeding the threshold number, and the coordinate table values corresponding to 264 ROI. The top 10 are shown in Table 3.

From the data statistics of these two tables, we can see that the same regions include 142, 159, 178, 182, 185, 222, and other similar regions, and these regions are the features of matching in deep learning. And these important ROIs are shown in Figure 15.

4. Discussion and Conclusions

The purpose of this paper is not only to propose a model that can improve the classification accuracy of the correlation coefficient matrix of the resting state fMRI but also to use the influence of emotional music to explore the active influence of depression on the brain through this deep learning model. By statistics and analysis of the characteristics of positive ROI, it can help related fields to further explore the causes of depression patients.

As shown in Table 1, our model achieves the best classification accuracy in different classification results. This classification accuracy is also a high classification result of the current research of task state data [24, 25, 28, 30]. This confirms the validity of the deep learning model, which allows for further statistical analysis of the feature-matched convolutional layer. The feature of our deep learning model is using convolution kernel of the same size as the

correlation coefficient matrix to realize feature matching. There are no spatial attributes between the adjacent data in the correlation matrix, but the CNN feature matching and the ability of convolution can be utilized.

Through deep learning, we can find the active ROIs matched by two-dimensional convolution (Tables 2 and 3). Many of them are consistent with current research results. Gourgouvelis et al. [43] show some areas of decreased activity associated with depression, including ROI (36, 88, 11) (in the Brodmann 19 area) and ROI (24, 73, 13) (in the Brodmann 18 area). Brodmann area 18 and Brodmann area 19 (the areas where ROI (27, 97, 13), ROI (28, 77, 32), ROI (17, 91, 14), and ROI (15, 77, 31) are counted in this paper) form a visually relevant cortex that also indirectly affects human psychological emotions. Harvey et al. [44] and Whalley et al. [45] mentioned the comparative effect of Brodmann area 40 and Brodmann area 6 in normal and depressed patients which also accounted for a large proportion. The supramarginal gyrus part of Brodmann area 40 (the area where ROI (48, 25, 27) and ROI (55, 45, 37) are counted in this paper) is the region in the inferior parietal lobe that is involved in phonology and reading of meaning. In addition, we also found the 182nd ROI (21, 41, 20) with significant difference in our result which is located in Brodmann area 11. Brodmann area 11 is part of the prefrontal cortex which helps solve cognitive functions related to thinking and perception, such as problem solving and emotion management.

These results reflect that the combination of convolution with large kernel and DNN onto small data sets is a good way to find the features of the correlation coefficient matrix and improve the classification accuracy. From the perspective of music, it is of great significance and research value to explore the ROI characteristics of normal people and depressed patients in fMRI data and reflect which part of ROI affects the mood of depressed patients. Deep analysis of the brain mechanism of depressed patients is more conducive to solving the condition of depressed patients and reducing their harm to themselves and society. The next step of this paper is to further analyze the correlation matrix and feature matrix generated by convolution and help the actual situation of depressed patients.

Appendix

Visualization of Our Trained Model

In Figure 14, we show a cumulative calculation of the 1,500 convolution matrices. In fact, each convolution kernel has its own characteristics. We randomly selected 10 renderings corresponding to the convolution kernel, as shown in Figure 16, and the range of color mapping is consistent with Figure 6.

Data Availability

For original fMRI data, please visit <https://www.openfmri.org/dataset/ds000171/>. The processed data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

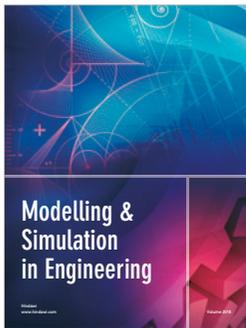
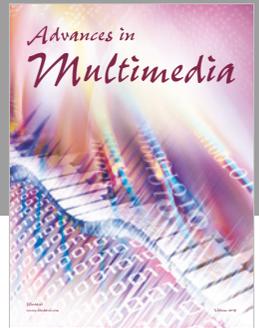
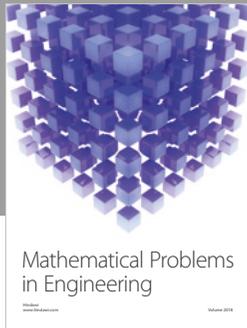
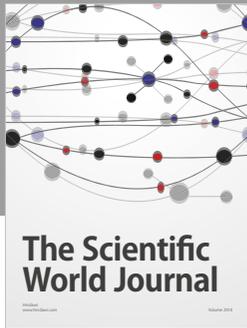
Acknowledgments

This work was supported by National Natural Science Foundation of China (61271351 and 41827807).

References

- [1] L. Andrade, J. J. Caraveo-Anduaga, P. Berglund et al., "The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys," *International Journal of Methods in Psychiatric Research*, vol. 12, no. 1, pp. 3–21, 2003.
- [2] S. Luo, *Study on Functional Connectivity of Resting Brain Based on fMRI*, Guilin University of Electronic Technology, Guilin, China, 2014.
- [3] E. Li, "The application of functional magnetic resonance imaging in cognitive neuroscience," *Computer Science and Exploration*, vol. 2, no. 6, pp. 589–600, 2008, in Chinese.
- [4] R. J. Lepping, R. A. Atchley, E. Chrysikou et al., "Neural processing of emotional musical and nonmusical stimuli in depression," *PLoS One*, vol. 11, no. 9, Article ID e0156859, 2016.
- [5] R. J. Lepping, R. A. Atchley, and C. R. Savage, "Development of a validated emotionally provocative musical stimulus set for research," *Psychology of Music*, vol. 44, no. 5, pp. 1012–1028, 2015.
- [6] R. J. Lepping, R. A. Atchley, L. E. Martin et al., "Limbic responses to positive and negative emotionally evocative music: an fMRI study," in *Proceedings of the Society for Neuroscience Annual Meeting 2012*, New Orleans, LA, USA, October 2012.
- [7] R. J. Lepping, R. A. Atchley, L. E. Martin et al., "The effect of musical experiences and musical training on neural responses to emotionally evocative music and no-musical sounds," in *Proceedings of the Social and Affective Neuroscience Society Annual Conference 2013*, San Diego, CA, USA, April 2013.
- [8] R. J. Lepping, R. A. Atchley, L. E. Martin et al., "Musical and no-musical sounds evoke different patterns of neural activity: an fMRI study," in *Proceedings of the Psychonomic Society Annual Meeting*, Minneapolis, MN, USA, November 2012.
- [9] R. J. Lepping, R. A. Atchley, T. M. Patrician et al., "Music to my ears: neural responsiveness to emotional music and sounds in depression," in *Proceedings of the Society of Biological Psychiatry Annual Scientific Convention 2013*, San Francisco, CA, USA, May 2013.
- [10] G. Atluri, A. MacDonald III, K. O. Lim, and V. Kumar, "The brain-network paradigm: using functional imaging data to study how the brain works," *Computer*, vol. 49, no. 10, pp. 65–71, 2016.
- [11] M. J. Patel, A. Khalaf, and H. J. Aizenstein, "Studying depression using imaging and machine learning methods," *NeuroImage: Clinical*, vol. 10, no. 4, pp. 115–123, 2016.
- [12] A. Tahmassebi, A. H. Gandomi, M. H. J. Schulte, A. E. Goudriaan, S. Y. Foo, and A. Meyer-Baese, "Optimized Naive-Bayes and decision tree approaches for fMRI smoking cessation classification," *Complexity*, vol. 2018, Article ID 2740817, 24 pages, 2018.
- [13] J. Cai, L. Zheng, L. Li et al., "A comparative study of fMRI on epileptic focus detection by independent component analysis," *Journal of Clinical Radiology*, vol. 31, no. 10, pp. 1370–1374, 2012, in Chinese.
- [14] R. Ramasubbu, M. R. G. Brown, F. Cortese et al., "Accuracy of automated classification of major depressive disorder as a function of symptom severity," *NeuroImage: Clinical*, vol. 12, pp. 320–331, 2016.
- [15] A. Khazaei, A. Ebrahimzadeh, and A. Babajani-Feremi, "Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease," *Brain Imaging and Behavior*, vol. 10, no. 3, pp. 799–817, 2016.
- [16] A. Panat, A. Patil, and R. Chauhan, "Feature extraction and classification of FMRI images," *International Journal of Electrical Electronics & Data Communication*, vol. 2, no. 10, 2014.
- [17] B. Afshin-Pour, S.-M. Shams, and S. Strother, "A hybrid LDA+gCCA model for fMRI data classification and visualization," *IEEE Transactions on Medical Imaging*, vol. 34, no. 5, pp. 1031–1041, 2015.
- [18] S. Ryali, K. Supekar, D. A. Abrams, and V. Menon, "Sparse logistic regression for whole-brain classification of fMRI data," *Neuroimage*, vol. 51, no. 2, pp. 752–764, 2010.
- [19] S. Kamonsantiroj, P. Charoenvorakiat, and L. Pipanmaekaporn, "Learning representation for fMRI data analysis using autoencoder," in *Proceedings of the IIAI International Congress on Advanced Applied Informatics*, pp. 560–565, Kumamoto, Japan, July 2016.
- [20] J. D. Power, A. L. Cohen, S. M. Nelson et al., "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.
- [21] Z. Zhang and J. Ji, "fMRI data classification method based on convolutional neural network," *Pattern Recognition and Artificial Intelligence*, vol. 30, no. 6, pp. 549–558, 2017, in Chinese.
- [22] A. Tahmassebi, A. H. Gandomi, I. McCann, M. H. J. Schulte, A. E. Goudriaan, and A. Meyer-Baese, "Deep learning in medical imaging: fMRI big data analysis via convolutional neural networks," in *Proceedings of the Practice and Experience on Advanced Research Computing*, pp. 1–4, ACM, Pittsburgh, PA, USA, July 2018.
- [23] S. Sarraf and G. Tofghi, "Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data," in *Proceedings of the Future Technologies Conference*, Vancouver, Canada, November 2017.
- [24] S. Sarraf and G. Tofghi, "DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI," 2016, <https://www.biorxiv.org/content/10.1101/070441v4>.
- [25] H. Ni, W. Wang, C. Pang, and J. Yu, "Classifying judging states from fMRI data of visual recognition task," in *Proceedings of the International Workshop on Complexity & Data Mining*, pp. 47–50, Washington, DC, USA, December 2012.
- [26] X. Wang, Y. Ren, Y. Yang, W. Zhang, and N. N. Xiong, "A weighted discriminative dictionary learning method for depression disorder classification using fMRI data," in *Proceedings of the IEEE International Conferences on Big Data and Cloud Computing*, pp. 618–623, Chengdu, China, July 2016.
- [27] G. Chanel, S. Pichon, L. Conty, S. Berthoz, C. Chevallier, and J. Grèzes, "Classification of autistic individuals and controls using cross-task characterization of fMRI activity," *NeuroImage: Clinical*, vol. 10, pp. 78–88, 2016.
- [28] A. W. Chung, E. Pesce, R. P. Monti, and G. Montana, "Classifying HCP task-fMRI networks using heat kernels," in *Proceedings of the 2016 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, Trento, Italy, June 2016.

- [29] M. J. Rosa, L. Portugal, T. Hahn, and J. Mourao-Miranda, "Sparse network-based models for patient classification using fMRI," in *Proceedings of the 2013 3rd International Workshop on Pattern Recognition in Neuroimaging*, vol. 105, pp. 66–69, Shenzhen, China, October 2013.
- [30] I. O. Ertugrul, M. Ozay, and F. T. Y. Vural, "Encoding the local connectivity patterns of fMRI for cognitive task and state classification," *Brain Imaging & Behavior*, vol. 13, no. 4, pp. 1–12, 2019.
- [31] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," 2016, <https://arxiv.org/abs/1602.07261>.
- [32] Y. Chao-Gan and Z. Yu-Feng, "DPARSF: a MATLAB toolbox for pipeline data analysis of resting-state fMRI," *Frontiers in Systems Neuroscience*, vol. 4, p. 13, 2010.
- [33] S. Achard and E. Bullmore, "Efficiency and cost of economical brain functional networks," *PLoS Computational Biology*, vol. 3, no. 2, p. e17, 2007.
- [34] M. Ke, H. Shen, and D. Hu, "Analysis of complex networks of resting state brain function based on fMRI," *Journal of National University of Defense Technology*, vol. 32, no. 1, pp. 147–151, 2010.
- [35] C. Stam, B. Jones, G. Nolte, M. Breakspear, and P. Scheltens, "Small-world networks and functional connectivity in Alzheimer's disease," *Cerebral Cortex*, vol. 17, no. 1, pp. 92–99, 2006.
- [36] M. P. V. D. Heuvel, C. J. Stam, M. Boersma, and H. E. Hulshoff Pol, "Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain," *Neuroimage*, vol. 43, no. 3, pp. 528–539, 2008.
- [37] S. Xue, Y. Tang, J. Li et al., "A brain function network construction method based on fMRI data," *Computer Application Study*, vol. 27, no. 11, pp. 4055–4057, 2010, in Chinese.
- [38] X. Chen, J. Jiao, M. Ke et al., "Brain network modeling and functional connectivity characteristics under resting state," *Journal of Lanzhou University of Technology*, vol. 36, no. 5, pp. 88–92, 2010.
- [39] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, Boston, MA, USA, June 2015.
- [40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June–July 2016.
- [41] A. Tahmassebi, "iDeepLe: deep learning in a flash," in *Proceedings of the Disruptive Technologies in Information Sciences*, Orlando, FL, USA, May 2018.
- [42] N. S. Keskar and R. Socher, "Improving generalization performance by switching from Adam to SGD," 2017, <https://arxiv.org/abs/1712.07628>.
- [43] J. Gourgouvelis, P. Yilder, and B. Murphy, "Exercise promotes neuroplasticity in both healthy and depressed brains: an fMRI pilot study," *Neural Plasticity*, vol. 2017, Article ID 8305287, 13 pages, 2017.
- [44] P.-O. Harvey, P. Fossati, J.-B. Pochon et al., "Cognitive control and brain resources in major depression: an fMRI study using the n-back task," *Neuroimage*, vol. 26, no. 3, pp. 860–869, 2005.
- [45] M. G. Whalley, M. D. Rugg, and C. R. Brewin, "Autobiographical memory in depression: an fMRI study," *Psychiatry Research: Neuroimaging*, vol. 201, no. 2, pp. 98–106, 2012.



Hindawi

Submit your manuscripts at
www.hindawi.com

