2021 IEEE 10th Data Driven Control and Learning Systems Conference
May 14-16, 2021, Suzhou, China

Deep neural network classification of EEG data in schizophrenia

Zhifen Guo[1], Lezhou Wu[2], Yun Li[3],Beilin Li[4]

1. College of Information Science and Engineering, Northeastern University, Shenyang, 110819
E-mail: 2628242884@qq.com

2. College of Information Science and Engineering, Northeastern University, Shenyang, 110819
E-mail: 614122718@qq.com

3. College of Information Science and Engineering, Northeastern University, Shenyang, 110819
E-mail: 18236884385@163.com

4. College of Information Science and Engineering, Northeastern University, Shenyang, 110819
E-mail: 1970630@stu.neu.edu.cn

2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS) | 978-1-6654-2423-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/DDCLS52934.2021.9455509

Abstract: Schizophrenia(SZ) is a disease of unknown etiology and pathogenesis and is ranked by the World Health Organization as one of the top ten diseases contributing to the global burden of disease. Studying the internal physiological differences between EEG of schizophrenia patients and normal individuals is important for diagnosing and treating schizophrenia in order to determine objective physiological diagnostic criteria. The EEG data of patients with schizophrenia were preprocessed and markers were extracted. The convolutional neural network was used to characterize the difference of distributed structure of data for classification and the classification results were given. The accuracy of the classification was 92%, and the disease classification was effectively performed using deep learning networks.

Key Words: Schizophrenia, Deep Neural Classification Network, EEG Data, Biomarker Extraction

1 Introduction

Schizophrenia is a complex, heterogeneous behavioural and cognitive syndrome that seems to originate from disruption of brain development caused by genetic or environmental factors, or both. At present, treatment mainly consists of antipsychotic drugs combined with psychological therapies, social support, and rehabilitation, but there is an urgent need for more effective treatment and service provision[1].Patients with schizophrenia present with impairments in perception, thought disorders, affective disorders, volitional and behavioral disorders cognitive intellectual disability.The causes of schizophrenia are complex and influenced by genetic and social factors. Currently, the definitive diagnosis of schizophrenia still relies on physician consultation and assessment of clinical scales and lacks objective physiological diagnostic criteria.

As of now, medical research based on deep learning is relatively mature in medical imaging. EEG is a spontaneous and rhythmic potential activity generated by the neural activity in the brain and always exists in the central system. EEG is a reflection of the activity state of the brain and an auxiliary examination method to help diagnose brain diseases[2]. Neuronal oscillations reflect the activity of neuronal aggregates involved in comprehensive cognition and can be used as a functional measure of cognitive impairment in schizophrenia[3]. In this paper, we use EEG signals to analyze the physiological differences between schizophrenia patients and normal people to find objective physiological diagnostic criteria.

Deep learning (DL) techniques have gained popularity among researchers while solving various problems [4]. The generalized feature invariant deep neural network framework is used to ensure the generalization of the model[5]. EEG signals are extracted by a series of filtering and transformation processes, including information theory markers, connectivity markers, spectral markers and evoked potential markers.The experimental process performs data processing and feature extraction of EEG, extracting a total of 21 dimensions of data in the frequency and time domains as well as computational complexity metrics, laying the foundation for deep learning interpretability in the medical field.Inputting EEG markers into the deep neural network for classification training has greatly improved the accuracy to 92%, which can assist doctors in diagnosis and treatment. This paper is divided into five parts, the second part introduces the research background and diagnosis development status of schizophrenia, the third part introduces the theoretical knowledge and construction of deep neural network, the fourth part introduces the data processing and experimental process, and finally the fifth part introduces the experimental conclusion and future work outlook.

2 Research Background and Diagnosis Development of Schizophrenia

Schizophrenia is a severe mental illness of unknown etiology, with a lifetime prevalence of about 1%, and is chronic. SZ is listed by the World Health Organization as one of the top 10 diseases contributing to the global burden of disease, placing a psychological and physical burden on patients and their families. Multidimensional analysis of diagnostic treatment is performed based on EEG features.

2.1 Background of Schizophrenia Research

The onset of schizophrenia is more than that in adolescence and early adulthood, which is characterized by incoordination between mental activities and the real environment, accompanied by obstacles in thinking, emotion, behavior, perception and other aspects, and is characterized by repeated attacks and difficult to cure[6]. Schizophrenia is mainly manifested by three core symptoms: positive symptoms, negative symptoms, and cognitive dysfunction[7]. Positive symptoms are hallucinations, delusions, excitement, etc., negative symptoms are poor thinking, emotional apathy, decreased will, etc.[8], and cognitive impairment is manifested by abnormal memory, language, attention, and executive function[9]. The research shows[10] that the number of patients with schizophrenia is increasing year by year, with 13.1 million patients in 1990 and 20.9 million in 2016. Schizophrenia has the epidemiological characteristics of "three low and three high", namely low detection rate, low visit rate, low compliance, high recurrence rate, high disability rate, and heavy disease burden[11].

Schizophrenia is a disabling mental disorder that has become an important global public health problem[11]. Studies have shown that schizophrenia can lead to different degrees of brain tissue damage and mental decline, and is the main cause of mental disability, accounting for about 83.18% of mental disability[12]. Patients are highly susceptible to relapse and worsening if their illness is prolonged, leading to loss of workforce. Most of the causes of disability are due to the misdiagnosis of patients or delayed treatment and other reasons for the increased disability rate of patients. The lifetime suicide mortality of SZ patients is 4-6%[13], as high as 50% of SZ patients attempt suicide, and as high as 13% of suicide deaths can be attributed to SZ[14].

In the present, researchers generally believe that the occurrence, development and transformation of SZ are related to the combined effects of genetic factors, social environmental factors and neurobiology, but the specific etiology and pathogenesis are difficult to determine. Lack of objective criteria for physiological diagnosis may lead to misdiagnosis and improper treatment of patients. The study of the etiology and physiopathological mechanisms of SZ is a challenging topic, and the search for objective physiological diagnostic criteria to characterize the pathogenesis and etiology of patients is the focus and difficulty of current research.

2.2 Current Development of Schizophrenia Diagnosis

Imaging and electrophysiology are the two main development directions for exploring the pathogenesis and diagnosis of SZ. Imaging refers to functional neuroimaging, which is a method that can observe the possible functional changes of treatment at clinically relevant scales[15].FMRI measures changes in blood oxygen level dependent signals in the brain[16], and is a non-invasive, radiation-free, high-spatial resolution neuroimaging technology. EEG is an electrophysiological phenomenon of brain neurons, which records the changes of postsynaptic potential in the cortical and subcortical neurons during brain activity[17]. EEG signals can be divided into spontaneous signals and evoked signals. In this study, acoustic evoked EEG signals are used for experimental research. By frequency classification, EEG can be divided into five categories: δ wave is between 0.5Hz - 4Hz, and the human is in a deep sleep state; θ wave is between 4 Hz - 8 Hz, and human is in a state of deep relaxation ; α wave is between 8 Hz - 14 Hz, and human is in a state of conscious and relax ;β wave is between 14 Hz - 40 Hz, and human brain is awake and gradually been tension ; γ wave is between 30 Hz - 100 Hz, and human is in a state of meditation.

The pathological differences between SZ and normal people were analyzed according to the rhythm changes of a certain band or a combination of several bands.

The traditional EEG signal-based emotion recognition mainly includes EEG signal acquisition, data preprocessing, emotion feature extraction and emotion classification. Extracting features from EEG signals that are highly correlated with disease types and have strong differentiation ability is an important part of the EEG signal characterization process. The current feature analysis methods used for EEG emotion recognition mainly include time domain feature analysis, statistical feature analysis, frequency domain feature analysis and nonlinear dynamics feature analysis[18].

Based on the time domain characteristic analysis of EEG, the horizontal axis of EEG is time, and the longitudinal axis is the change of EEG signal, which can intuitively reflect the signal status and the time period of change. The time-domain feature analysis mainly extracts the waveform features of EEG signals directly, including Event Related Potential (ERP), signal statistics, energy, power, higher-order over-zero analysis, Non-Stationary Index (NSI) and Fractal Dime nsion (FD), etc. Time domain waveform analysis is most commonly used in clinical practice, because it is the most intuitive method, and doctors can make judgments on some common diseases by observing the amplitude and frequency of waveform. The frequency domain characteristic analysis mainly analyzes the distribution and change of frequency band rhythm in EEG signal through the spectrum diagram. In cognitive neuroscience research, it has been found that different frequency bands of EEG signals are closely related to different brain activities. The nonlinear dynamic analysis includes many methods, among which the entropy and complexity of EEG signals have strong anti-interference ability and can achieve a faster calculation speed with only a short time series without affecting the accuracy[19, 20]

The complexity of the causes of schizophrenia makes it difficult to diagnose, and treatment methods still do not fully meet the needs of patients. Using EEG to analyze the fundamental differences between patients and normal brain electrical markers is a key step to establish objective physiological criteria.

3 Construction and Training of Deep Neural Classification Network

The best deep neural classification network type, the number of hidden layers of the classification network and the

number of hidden layers of the network are selected to obtain a higher accuracy. Subjects' data were collected according to inclusion and exclusion criteria, and data preprocessing and data feature extraction were performed. The extracted EEG markers were input into the network for training and obtained the accuracy of the deep neural classification network .

3.1 Construction of Deep Neural Classification Network

Compared with the shallow learning of machine learning, deep learning has outstanding generalization ability to deal with complex problems. Deep learning is representational learning, that is, the automatic generation of useful features from data. Deep learning network provides a multilayer nonlinear network structure, realizes complex function approximation, and represents different representations of input data from low level to high level[21].The structure diagram of deep neural classification network is shown in Figure 1.

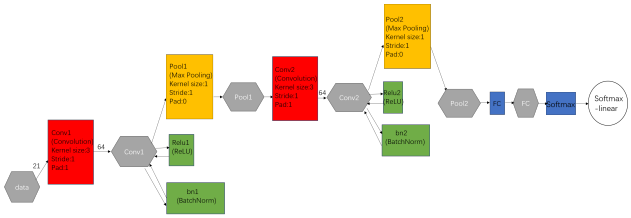

Figure 1: Deep neural classification network

The accuracy of the classification results and the generalization of the model are closely related to the structure of the deep neural classification network. The construction of deep neural classification network is mainly considered and weighed from the following three points. Firstly, the appropriate autoencoder is determined, that is, the network structure corresponding to data format, data length or data volume is selected to represent the characteristic model. Secondly, the number of hidden layers of the model is determined, that is, the depth of deep neural network and the number of hidden layer nodes. The number of hidden layers and nodes has great influence on the performance of the network model. If there are too many hidden layers in the network, problems such as gradient disappearance and gradient explosion may occur. If the number of hidden layers of the network is too small it may lead to problems such as excessive training errors. The number of hidden nodes must be less than N-1 (N is the number of training samples), otherwise the systematic error of the network model tends to zero because it is independent of the characteristics of the training samples. The number of training samples must be more than the connection weight of the network model, which is generally 2-10 times. Finally, the activation function and loss function of each layer are determined. The activation function is also called nonlinear function. The function of activation function is to transform linear model into nonlinear model to deal with nonlinear problems. With different activation functions, the model can also learn different forms of classification. The loss function measures the output loss of training samples, and the optimal extremum of the loss function is solved in the process of deep network back propagation.

The autoencoder was proposed by Kramer et al.[22] in 1991. It is a neural network with a special structure, whose input and output tensors have exactly the same shape. Autoencoder is an unsupervised form in which an exact copy of the input is generated in the form of encode-decoding. The working logic of the autoencoder is shown in Figure 2. The raw data is input into the defined autoencoder to characterize the input, and the data with the same dimensional characteristics as the input is output through the decoder.

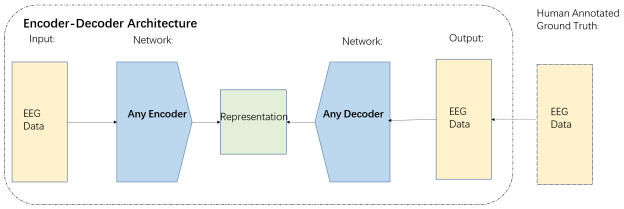

Figure 2: Self-encoder model

In this experiment, a convolution-based autoencoder is used. The implicit layer of the autoencoder is composed of convolution layer and deconvolution layer. Convolutional Neural Network (CNN) is a deep neural network with a convolutional structure, whose structural composition is convolution-pooling-full connection. The working structure of Convolutional Neural Network is shown in Figure 3. The working principle of convolution is to use the convolution kernel for feature extraction. The size and step size of the convolution kernel can be set independently and whether to expand and scan the input data to extract features. Therefore, the convolution kernel is also called the feature extractor. Convolutional neural network based on the above working mode has the advantages of translation invariance. The calculation formula of the j_{th} feature map at the l layer of the convolution layer is:

$$x_j^l = f(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l) \quad (1)$$

M_j refers to the mapping set, k is the convolution kernel, and b_j is the bias parameter. In essence, * is to make the convolution kernel k perform the convolution operation on all the associated feature maps of the $l-1$ layer. Function f is a nonlinear activation function.

The dimension formula of the output matrix of convolution operation is as follows:

$$outputsize = (N - F + 2 * pad)/stride + 1 \quad (2)$$

Where N is the size of the input data of the previous layer; F is the size of the convolution kernel; pad is the dimensionality of the increment of the input data of the previous layer; stride is the step size.

The nonlinear activation function uses Softmax function in the output layer of this experiment. The calculation formula of activation function is as follows:

$$a_i^L = \frac{e^{Z_i^L}}{\sum_{j=1}^{n_L} e^{Z_j^L}} \quad (3)$$

Where e is a natural number,$Z_i^L$is the input of class i in the L^{th} layer.
The calculation formula of logarithmic likelihood function is as follows:

$$J(W,b,a^L,y) = -lna_i^L \tag{4}$$

For class i, the gradient is calculated as:

$$\begin{aligned}
&\frac{\partial J(W,b,a^L,y)}{\partial W_i^L} = \frac{\partial J(W,b,a^L,y)}{\partial a_i^L}\frac{\partial a_i^L}{\partial Z_i^L}\frac{\partial Z_i^L}{\partial W_i^L} \\
&= -\frac{1}{a_i^L}\frac{(e^{Z_j^L})(\sum_{j=1}^{n_L} e^{Z_j^L}) - e^{Z_i^L}e^{Z_i^L}}{(\sum_{j=1}^{n_L} e^{Z_j^L})^2}a_i^{L-1} \\
&= -\frac{1}{a_i^L}(\frac{e^{Z_i^L}}{\sum_{j=1}^{n_L} e^{Z_j^L}} - \frac{e^{Z_i^L}}{\sum_{j=1}^{n_L} e^{Z_j^L}}\frac{e^{Z_i^L}}{\sum_{j=1}^{n_L} e^{Z_j^L}})a_i^{L-1} \\
&= -\frac{1}{a_i^L}a_i^L(1-a_i^L)a_i^{L-1} = (a_i^L - 1)a_i^L
\end{aligned} \tag{5}$$

$$\frac{\partial J(W,b,a^L,y)}{\partial b_i^L} = a_i^L - 1 \tag{6}$$

Where i is the class serial number of the training sample.

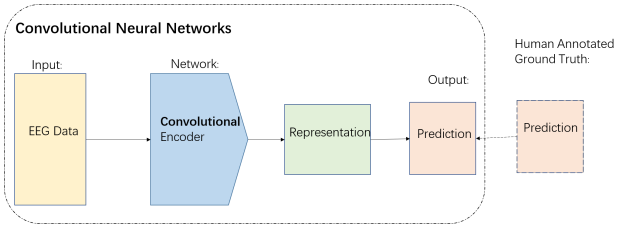

Figure 3: Convolutional neural network

Pooling operation does not change the number of channels. Max pooling is adopted in this experimental study, and the maximum value in the data area is selected as the pooled value. The main function of pooling is down sampling. Reduce the number of parameters to reduce computational redundancy; Implement invariance includes translation invariance, rotation invariance, etc; Realize nonlinearity ; Expand the field of perception.
The role of the full connection layer is to act as a classified part of the entire network. Each node of the full connection layer is connected to all nodes of the upper layer, that is, the local features obtained by convolution are assembled into a complete graph by weight matrix, and then classified by weight. The working structure of the full connection layer is shown in Figure 4. The calculation formula of the full connection is as follows:

$$\begin{bmatrix} a1 \\ a2 \\ a3 \end{bmatrix} = \begin{bmatrix} W11 & W12 & W13 \\ W21 & W22 & W23 \\ W31 & W32 & W33 \end{bmatrix} * \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} + \begin{bmatrix} b1 \\ b2 \\ b3 \end{bmatrix} \tag{7}$$

The number of fully connected input nodes is n, and the number of output nodes is N. The back propagation of the full connection layer calculates the partial derivatives of the three parameters, and the gradient from the upper layer to the layer is known $\frac{\partial loss}{\partial a}$.Firstly, the partial derivative of the output a_i with respect to the input x_j is obtained as follows:

$$\frac{\partial a_i}{\partial x_j} = \frac{\sum_j^n w_{ij} * x_j}{\partial x_j} = w_{ij} \tag{8}$$

The partial derivative of the loss function with respect to x is obtained by the chain rule as follows:

$$\frac{\partial loss}{\partial x_k} = \sum_j^N \frac{\partial loss}{\partial a_j}\frac{\partial a_j}{\partial x_k} = \sum_j^N \frac{\partial loss}{\partial a_j} * w_{jk} \tag{9}$$

We know $\frac{\partial a_i}{\partial w_{ij}} = x_j$,so:

$$\frac{\partial loss}{\partial w_{kj}} = \frac{\partial loss}{\partial a_k}\frac{\partial a_k}{\partial w_{kj}} = \frac{\partial loss}{\partial a_k} * x_j \tag{10}$$

It can be deduced from the above:$\frac{\partial a_i}{\partial b_i} = 1$, that is, the partial derivative of the loss function with respect to the bias coefficient is equal to the partial derivative with respect to the upper layer.

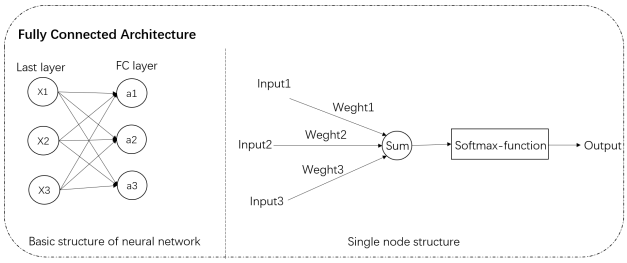

Figure 4: Full-connected network node diagram

3.2 Data Acquisition and Processing

During data collection, experimental volunteers need to be selected according to the following exclusion criteria for inclusion. Specifically, the criteria for selecting typical patients:
1)Patient inclusion criteria:

- Meet the DSM-5 diagnostic criteria for BD-D or MDD.
- HAMD-17 scale $\geq$ 17 points.
- The Young Mania Scale < 8 points.

Patient exclusion criteria:

- Patient exclusion criteria.
- Previous or current suffering from other mental diseases other than BD-D or MDD, mental retardation.
- Secondary depressive episode caused by brain organic psychosis or other diseases.
- Any neurodegenerative diseases, brain injuries or cerebrovascular diseases and other brain organic diseases.

2)Inclusion criteria for healthy control group:

- No history of mental illness, SAS (Self-Rated Anxiety Scale) and SDS (Self-Rated Depression Scale) excluded anxiety and depression.

- HAMD-17 scale <7 points.

Exclusion criteria for healthy control group:

- A positive family history of mental illness.
- Any neurological and other mental disorders.
- A history of head trauma.

A total of 81 people were included in the database, including 49 patients and 32 healthy people. 64 channels of EEG acquisition equipment were used to collect data from patients. The subjects were stimulated into three conditions: (1) the object pressed the button once every 1-2 seconds to provide a sound at 1000 Hz, 80 dB sound pressure level, and there was zero delay (key tone) between the press and the start tone, (2) the task stopped after 100 sounds were transmitted. Preserving the chronological sequence of tones for playback (playback tones), (3) the subject presses the button at roughly the same rate, and no sound appears (individual buttons). Data is collected 100 times in a row for each object, about 5 seconds each time, and 1024 times per second. The EEG signals were continuously digitized at a frequency of 1024Hz by fast Fourier transform (FFT), and the average earlobe electrode was referenced offline. Digital low-pass filtering was performed between 0.5Hz and 15Hz. Too low and too high frequencies are not very informative for this experimental study as noise filtered out. EEG data were then divided into a period of 3000 milliseconds, and the EEG data of each subject were divided into 280-300 acquisition records. Lock to the time when the button is pressed and make baseline correction from -600 to -500ms. After the above data preprocessing, data feature extraction is performed next. Eliminate irrelevant data, extract the 21-dimensional electrical markers including information theory markers, connectivity markers, spectrum markers (δ, θ, α, β, γ) and evoked potential markers. The 21-dimensional markers are shown in Table 1, in which the normalized processing of the five frequency band data of the spectrum markers can accelerate the calculation speed and improve the accuracy.

Next, the extracted 21-dimension data is preprocessed. Based on the Batch Normalization(BN) algorithm, it normalizes or standardizes each dimension, converting the input data into numbers between 0 and 1. The purpose of this operation is to speed up the training rate, increase the learning rate and improve the accuracy. Avoid distribution data deviation; Stay away from the derivative saturation region.

4 Experimental Process and Analysis

The experimental process consists of three steps: input data into the network, training data and network model, and test set to verify the training results. Firstly, the EEG data of 81 labeled subjects were scrambled and divided into training set and test set in a ratio of 4:1. A total of 18560 data were collected from the training set and 4641 from the test set. Each data was 22 dimensions, including 21 dimensions of EEG markers and 1 dimension of label data. The data is output to the training network successively in the form of 3*7 and one label bit.

Deep neural classification networks are mainly composed of convolutional autoencoders. The 21-dimensional data is first input into the convolutional neural network. After the [3,3,64] convolutional kernel scan, the 64-dimensional data samples are output after nonlinearization by the ReLU activation function. After the Pooling and BN operations, the 64-dimensional data samples are output by the same operation transformation of the next layer of the convolutional network. Next, input two layers of full connection layer to carry out data sample classification. The working structure of the experimental deep neural classification network is shown in Figure 5.

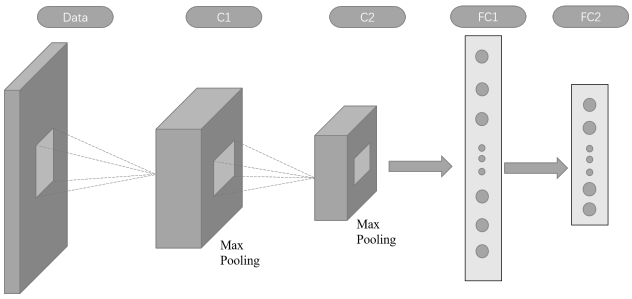

Figure 5: Deep neural network connection diagram

The batch-size is set to 128 for training, and the accuracy of multi-GPU training results reaches 92%.

5 Conclusions

In this experimental study, EEG data from patients with schizophrenia and normal subjects were preprocessed and markers were extracted and input into the network for training. Brain markers were extracted from the EEG data of 81 subjects and 49 patients and 32 healthy subjects. The results of the 21-dimensional EEG data input network showed good generalization ability. The type and layer number of the deep neural classification network established by the accuracy analysis experiment of training results are suitable for the input data type. It can be used as objective physiological diagnostic criteria for diagnosis and follow-up treatment. The theory of obvious classification effect is consistent with the practice.

In the future research and work, we will continue to optimize the network structure and look for data sets with larger data volume and more perfect data. To find the best number of layers and hidden layer nodes to improve the accuracy of classification results; The contribution rate of each dimension in the 21-dimension data to the classification results was calculated step by step.

REFERENCES

[1] Owen MJ,et al.Schizophrenia.Lancet.Published online January 14,2016.

[2] Yin Hongyi. Deep learning based EEG analysis of schizophrenia [J]. Microcomputers and Applications, 2016,35 (6) : 54-57.

[3] Junfeng Sun.Abnormal Dynamics of EEG Oscillations in Schizophrenia Patients on Multiple Time Scales[J]. IEEE Trans on Biomedical Enginnering,2017,61(6),1756-1764.

Table 1: Classification of Brain Markers

Abbreviations	Marker Name	Concept of Class
$PE\theta$	Permutation entropy	Information theory
K	Kolmogorov complexity	Information theory
$wSMI\theta$	Weighted symbolic mutual information	Connectivity
α	Alpha PSD	Spectral
$\|\alpha\|$	Normalized alpha PSD	Spectral
β	Beta PSD	Spectral
$\|\beta\|$	Normalized beta PSD	Spectral
δ	Delta PSD	Spectral
$\|\delta\|$	Normalized delta PSD	Spectral
γ	Gamma PSD	Spectral
$\|\gamma\|$	Normalized gamma PSD	Spectral
θ	Theta PSD	Spectral
$\|\theta\|$	Normalized theta PSD	Spectral
MSF	Median power frequency	Spectral
SE90	Spectral entropy 90	Spectral
SE95	Spectral entropy 95	Spectral
SE	Spectral entropy	Spectral
CNV	Contingent negative variation	Evoked
P1	Short-latency auditory potential to the first sound	Evoked
P3a	Mid-latency auditory potential to the first sound	Evoked
P3b	Mid-latency auditory potential to the first sound	Evoked

[4] D. Goularas and S. Kamis. Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data. 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML),2019,12-17.

[5] A. Zhou, Y. Cui and T. Jiang. Multisite Schizophrenia Classification Based on Brainnetome Atlas by Deep Learning. 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), Nanjing, China,2018,451-455.

[6] Wang Jun,Cui Xiao, Zhou Cuixia, et al. Cognitive behavioral therapy: a review [J]. Rehabilitation Theory & Practice,2013,19(9):834-838.

[7] Shmukler AB,Gurovich IY,Agius M,et al. Long-term trajectories of cognitive deficits in schizophrenia: A critical overview[J].Eur Psychiat,2015,30(8):1002-1010.

[8] Ghoreishi A,Shajari Z. Reviewing the Dissociative Symptoms in Patients With Schizophrenia and their Association With Positive and Negative Symptoms[J]. Iran J Psychiatry Behav Sci,2014,8(1):13-18.

[9] Sun Zhengyu,Gu Lihong,Li Lin.Journal of Psychiatry,2017,30(5):394-400.

[10] Charlson FJ,Ferrari AJ,Santomauro DF,et al.Global Epidemiology and Burden of Schizophrenia:Findings From the Global Burden of Disease Study 2016[J].Schizophr Bull,2018,44(6):1195-1203.

[11] Guan Lili,Du Lizhe,Ma Hong.Int J Mental Health,2012,26(12):913-919.

[12] Huang Z S. Analysis of the assessment results of 428 cases of mental disability. The 14th National Academic Conference of Behavioral Medicine, Hangzhou,2012.

[13] N. C. Hettige, A. Bani-Fatemi, I. Sakinofsky, et al. A biopsychosocial evaluation of the risk for suicide in schizophrenia[J]. CNS Spectrums, 2018, 23(4): 253-263.

[14] C. B. Caldwell and I. I. Gottesman. Schizophrenics kill themselves too:a review of risk factors for suicide[J].Schizophrenia Bulletin,1990,16(4): 571-589.

[15] T. Sharma. Insights and treatment options for psychiatric disorders guided by functional MRI[J]. The Journal of Clinical Investigation, 2003, 112(1): 10-18

[16] C. C Abbott, A. Jaramillo, C.E Wilcox, et al. Antipsychotic drug effects in schizophrenia: a review of longitudinal FMRI investigations and neural interpretations[J].Current Medicinal Chemistry, 2013, 20(3): 428-437.

[17] M. X. Cohen.Where does EEG come from and what does it mean?[J].Trends in Neurosciences,2017,40(4):208-218.

[18] Ran MS,Weng X,Chan CL,et al.Different outcomes of never-treated and treated patients with schizophrenia:14-year follow-up study in rural China[J].Br J Psychiat,2015,207(6):495-500.

[19] Zhang Guanhua,Yu Minjing,Chen Guo,et al.A review of EEG features for emotion recognition [J].Science China(Information Sciences),2019,49(9):1097-1118.

[20] Hu Jianfeng, Wang Taotao. Analysis of Driving Fatigue Detection Based on Fuzzy Entropy of EEG Signals [J]. China Safety Science Journal, 2018(4): 13-18.

[21] Boersma M, Smit D, de Bie H, et al. Network Analysis of Resting State EEG in the Developing Young Brain: Structure Comes With Maturation [J]. Human Brain Mapping, 2011, 32(3):413-425.

[22] Kramer M A.Nonlinear principal component analysis using autoassociative neural networks [J].AIChE Journal,1991,37(2):233-243.