Multi-feature data fusion based on common space model and recurrent convolutional neural networks for EEG tristimania recognition used in upper limb rehabilitation exercises

Hudun Sun[1,*]

[1]Tai Chi Martial Arts Academy, Jiaozuo university, Jiaozuo, 454000, China

Abstract

Traditional tristimania recognition methods cannot accurately recognize the mood of patients, which cannot provide effective adjuvant therapy for rehabilitation. Therefore, this paper proposes a new multi-feature data fusion method for Electroencephalography (EEG) tristimania recognition. It combines common space model and recurrent convolutional neural networks to classify the tristimania group and control group. According to the phase lock value, the phase synchronization functional network between electrode channels is constructed, and the functional connection modes of two kinds under different frequency bands are analyzed. The Xception network and LSTM are used as two non-interfering parts to extract two feature matrices from EEG tristimania signals. They are fused into a single feature matrix by merge algorithm. The single feature matrix is input into the recurrent convolutional neural networks (RCNN) for feature extraction and pooling. L2 regularized Softmax function is used as the classifier to complete the training and testing of RCNN. Finally, combining the Fisher score feature selection method and the classifier dependency structure, a low dimensional and efficient feature subset is obtained. Experimental results on public tristimania data sets validate that the proposed method has better effect in terms of accuracy and PLV compared with other strategies.

*Corresponding author. Email: 352720214@qq.com

1. Introduction

Tristimania is a common mental illness, which is characterized by decreased happiness, helplessness, persistent emotional depression, pessimism and thinking retardation in daily activities [1]. Studies show that tristimania will become the second major human disease [2]. For the early diagnosis of tristimania, the current clinical diagnosis of tristimania has many obvious shortcomings, including patient denial, poor sensitivity and subjective bias, which can lead to inaccurate diagnosis results, thus affecting the timely treatment of patients and resulting in illness delay. In general, compared with language and facial expressions, physiological signals can relatively and objectively reflect the dynamic changes of the central nervous system and have been applied in various studies [3].

Electroencephalogram (EEG) contains a lot of physiological and pathological information. It is the overall reflection of the physiological activity of brain nerve cells on the surface of cerebral cortex or scalp. Compared with magnetoencephalogram (MEG) [4], functional magnetic resonance imaging (FMRI) [5], positron emission tomography (PET) [6] and other research tools, EEG is still an important tool for monitoring brain activity due to its relatively low cost, convenience of recording and not easily falsified [7]. EEG has gradually been applied in the auxiliary diagnosis of a number of diseases [8-9].

Due to the complexity of brain activity, EEG signal is not a simple superposition of brain cognitive signal and interference signal, but a nonlinear, time-varying signal. Therefore, it is necessary to obtain efficient low-dimensional EEG features. The common features include time domain, frequency domain and space domain features. Common Spatial Pattern (CSP) [10-11] is a spatial filtering algorithm under binary classification tasks, it is generally used to extract spatial features. Appel et al. proposed the CSP algorithm for EEG [12], which could extract the spatial distribution components of each class from the multi-channel EEG data. By using the diagonalization of the covariance matrix, a set of optimal spatial filters are found for projection, which maximizes the variance of the first class and minimizes the variance of the other classes, so as to obtain the feature vectors with high degree of differentiation. On the other hand, due to the successful introduction of graph theory analysis technology into the study of brain functional networks, complex network analysis has become an increasingly hot topic in the field of brain electricity in recent years. Studies have shown that the introduction of complex network model can help people better understand the relationship between brain network structures [13-14]. In the work of Thomas et al. [15], they used Phase Locking Value (PLV) to construct the brain functional network, and further discussed the task-related network differences on the motor imagination data set, thus identifying important nodes and channels. Based on the above two points, in this work, we first conduct brain network analysis and use PLV to construct phase synchronization functional networks between tristimania patients and normal controls. Then we explore the possible different connection patterns between the two categories. Finally, CSP features and network features are fused to identify tristimania.

Multichannel EEG signals often contain a lot of redundant information. The redundancy feature will adversely affect the performance of the algorithm, resulting in high computational complexity, over-fitting and low efficiency. Therefore, more representative features need to be screened from the eigenmatrix. In practical applications, there are a variety of feature selection strategies, such as Pearson Correlation (PC) [16], reliefF algorithm [17] and Fisher Score (F-Score) [18]. F-score essentially measures the discrimination of a feature by calculating the ratio of the inter-class and intrclass variances of a single feature. The higher F-Score denotes the better distinguishment ability. This paper combines the F-Score method and the classifier dependent structure to realize feature selection, which not only ensures the recognition ability of the selected feature subset and model, but also saves the calculation amount.

With the development of deep learning, computer-assisted classification and diagnosis of tristimania can help reduce the misdiagnosis rate of clinical diagnosis and help patients get correct treatment as soon as possible. For example, Acharya et al. [18] used Convolutional Neural Network (CNN) to classify 15 healthy subjects and 15 tristimania subjects for the right hemisphere of the brain, and achieved 96% and 93.5% respectively. Li et al. [20] used EEG feature transform to identify tristimania. They used autoregressive models and Hjorth algorithm to extract power spectral density and activity as the original features. At the same time, the integration learning and deep learning are used to deal with these features. The performance of Support Vector Machine (SVM) and CNN classifier is tested on five and total frequency bands respectively. The results showed that EEG signal could be used as a reliable indicator for the identification of tristimania, and the best accuracy was 89.02% when integrated model and power spectral density feature were used. The accuracy with deep learning method and activity features was 84.75%. Zhang et al. [21] combined EEG data with demographic information including gender and age into one-dimensional CNN and obtained a classification accuracy of 75.29%. CNN [22] is an efficient model widely used in the field of pattern recognition. However, since the state of CNN only depends on the current input, it is not suitable for dealing with timing problems. However, the Long short-term Memory network (LSTM) [23] and Gated Recurrent Unit (GRU) [24] cannot achieve parallel computing operations due to their structural characteristics, resulting in inferior computing efficiency to CNN. Therefore, this paper proposes to apply recurrent convolutional network (RCNN)[25] and common space model into tristimania recognition of EEG signals. This model not only has the characteristics of RCNN processing time series, but also retains the characteristics of parallel computation of CNN.

In this paper, our main contributions are as follows: CSP algorithm is used to extract features and brain functional connectivity matrix is constructed according to PLV to extract brain network features. The features are visible, which is convenient for feature selection and method verification. The combination of F-Score and classifier is used to optimize the extracted features, and the best feature collection with high distinction is

obtained. The detailed process is shown in figure 1. Experiments are conducted on tristimania data sets to evaluate our proposed strategy, and the results demonstrate the effectiveness of our approach.

The following chapters are arranged as follows. The section 2 describes the proposed EEG tristimania recognition in detail including the extracted features, the feature selection method, and the RCNN. The section 3 introduces the data set, related experiments and results analysis. Section 4 gives the summary for this paper.
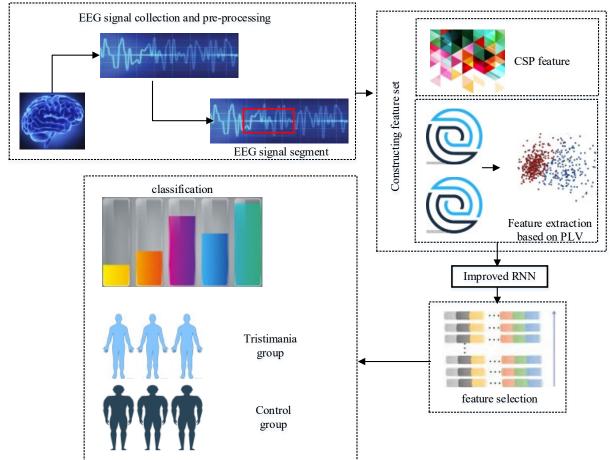

Figure 1. The framework of proposed method

2. Proposed EEG tristimania recognition

2.1. Common spatial pattern algorithm

The common spatial pattern (CSP) algorithm is an efficient spatial filtering algorithm, whose goal is to create the optimal common space filter. CSP is a spatial filtering feature extraction algorithm for two classification tasks. It can extract spatial components of each category from multi-channel brain-machine interface data. The basic principle of CSP is to find a set of optimal spatial filters for projection by diagonalization of matrix, so as to maximize the variance difference between the two kinds of signals and obtain feature vectors with high discrimination. As shown in figure 2, CSP obtains the feature vector with the most discriminating force by diagonalizing the covariance matrix of the two types of tasks simultaneously.
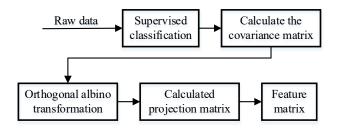
Figure 2. Diagram of CSP feature extraction method

The specific process is as follows:

(1) Assuming that X_1 and X_2 are two types of data samples, respectively. Firstly, it calculates the covariance matrix R_1 and R_2 of the two types of data:

$$R_i = \frac{X_i X_i^T}{trace(X_i X_i^T)}, i = 1,2 \tag{1}$$

Where $trace(\cdot)$ denotes the sum of elements on the diagonal of the matrix, namely the trace of the matrix. Then the mixed space covariance matrix is obtained:

$$R_c = \overline{R_1} + \overline{R_2} \tag{2}$$

Where $\overline{R_i}$ is the mean covariance matrix of the two types of tasks.

(2) Applying principal component analysis (PCA) method to carry out orthogonal albino transformation and diagonalization processing for positive definite matrix R_c at the same time. Then, eigenvalue decomposition is performed for R:

$$R_c = U_c \Lambda_c U_c^T \tag{3}$$

Where, U_c represents the eigenvector matrix. Λ_c is the diagonal matrix formed by descending order of eigenvalues. After transformation, the albino value matrix can be obtained as follows:

$$P = \frac{1}{\sqrt{\Lambda_c}} U_c^T \tag{4}$$

(3) Constructing space filter. Apply P to R_1 and R_2 and obtain:

$$S_1 = PR_1 P^T, \quad S_2 = PR_2 P^T \tag{5}$$

S_1 and S_2 have the common eigenvectors. The eigenvector matrix B and the diagonal matrix Λ_1, Λ_2 satisfy the following conditions.

$$S_1 = P\Lambda_1 P^T, \quad S_2 = P\Lambda_2 P^T \tag{6}$$

$$\Lambda_1 + \Lambda_2 = I \tag{7}$$

Where I is the unit matrix.

(4) Calculating the projection matrix.

Since the sum of the eigenvalues of the two kinds of matrices is 1, then one class has the largest eigenvalue while the other class has the smallest value. The transformation of the albino EEG to the eigenvectors corresponding to the maximum eigenvalues in Λ_1, Λ_2 is optimal for separating the variances in the two signal matrices. Therefore, matrix B can be used to realize the classification of the two types of problems. The spatial filter, i.e. the projection matrix W, can be obtained as:

$$W = (B^T P)^T \tag{8}$$

(5) Feature $Z_{M \times N}$ can be obtained by filtering the original tristimania EEG through the constructed corresponding filter W.

$$Z_{M \times N} = W_{M \times N} \times X_{M \times N} \tag{9}$$

It can select the front and back m rows of $Z_{M \times N}$ (2m<M) features as the raw input data.

2.2. Construction of PLV phase synchronization brain functional network

Cognitive process not only involves the activity of various brain regions, but also involves the information transmission and interaction between different functional regions. In essence, connection patterns can mine the interactions between relevant brain regions, and network attributes are network patterns that reflect the transmission and processing of information. There are many network construction methods, where brain synchronization activity analysis, which captures brain synchronization activity in the form of synchronization measurement matrix, has attracted much attention in recent years. Studies have shown that EEG signals have phase synchronization across brain regions, but not amplitude synchronization. Therefore, the synchronization measurement based on phase information is more discerning than that based on amplitude information [26, 27]. In this paper, we use phase lock value (PLV) to measure the phase synchronization information between the two brain regions. Assuming that the instantaneous phase of the two signals x(t) and y(t) are $\phi_x(t)$ and $\phi_y(t)$, PLV is defined as:

$$PLV = \left| \frac{1}{N} \sum_{j=0}^{N-1} e^{i\Delta\phi(t)} \right| \tag{10}$$

Where N is the number of samples. $\Delta\phi(t) = \phi_x(j\Delta t) - \phi_y(j\Delta t)$ represents the phase difference between the signals x(t) and y(t). Δt indicates the sampling period. j is the j-th sample point. Gupta et al. [28] used network indicators (mean path length, global efficiency, local efficiency) of amplitude square coherent

networks for conducting experiments, but the classification accuracy was not high. Their study showed that global differences between different brain functional networks were less than local differences for EEG data. Therefore, four global attributes and one local attribute are selected as attribute indicators for complex network analysis in the experiment, and Brain Connectivity Toolbox [29] is used to extract network features in this paper.

The clustering coefficient is the separation degree of brain functions, which is the ratio of the connections number between a node and its neighbors to the maximum number of possible connections. The clustering coefficient of node i is defined as:

$$C_i = \frac{\sum_{k \neq i} \sum_{l \neq i, l \neq k} c_{ik} c_{il} c_{kl}}{\sum_{k \neq i} \sum_{l \neq i, l \neq k} c_{ik} c_{il}} \quad (11)$$

Where c_{ij} is the connection weight between the two nodes i and j of the adjacencies matrix.

Characteristic path length L refers to the minimum number of edges connecting two nodes in the network. The average characteristic path length is the average of the shortest path length among all network nodes. For a weighted network, the length of the edge between nodes i and j is expressed by the reciprocal of the edge weight.

$$L = \frac{1}{(1/N(N-1)) \sum_{i=1}^{N} \sum_{j \neq i}^{N} (1/L_{ij})} \quad (12)$$

Where N is the node in the weighted network. L_{ij} is the edge number of the shortest path between any two nodes i and j.

Global efficiency Ge measures the propagation of information in the entire network, which is the inversely proportional to the length distance of the shortest path between any two nodes in the network.

$$Ge = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq i}^{N} L_{ij}^{-1} \quad (13)$$

Local efficiency measures the capability of a regional specialization network by looking at the connectivity of each sub-network. It provides an estimation of the fault tolerance of the network. Let G_i represent the sub-network of all adjacent nodes containing the i-th node. The average efficiency of sub-net G_i:

$$Le_i = \frac{1}{N_{G_i}(N_{G_i} - 1)} \sum_{i,k \in G_i} \frac{1}{L_{j,k}} \quad (14)$$

Where, N_{G_i} denotes the node number in the sub-graph G_i. Mediation centrality is a common network local attribute, it is used to measure the centrality of nodes in a graph:

$$b_i = \sum_{m \neq i \neq n} \frac{\sigma_{mn}(i)}{\sigma_{mn}} \quad (15)$$

$\sigma_{mn}(i)$ indicates the shortest path from node m to node n through node i. σ_{mn} is the shortest path length from node m to node n.

2.3. RCNN-based feature extraction

The network model in this paper consists of Xception [30], LSTM and CNN-Fusion. The innovation is the introduction of a recurrent network, which can extract the features of EEG signals, and build a feature Fusion network (CNN-Fusion). The recurrent convolutional neural network model is shown in figure 3. After L2 regularization and other adjustments, it is suitable for fine-grained classification of EEG signal features and can exert the best performance. Firstly, the sample data sets are input into Xception network and LSTM respectively. Then the two eigenmatrices obtained by the two networks are simultaneously input into the Merge layer of CNN-Fusion. Through this layer, features are fused and merged into a single feature matrix. Then the merged eigenmatrices are convolved. Finally, the category with the largest probability of species is output from the output layer to complete the classification prediction.
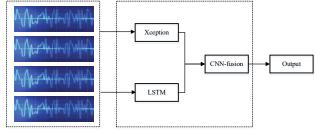

Figure 3. Structure of RCNN

Xception and LSTM

Xception is a convolutional neural network that uses separable convolution combined with local connections

and weight sharing. Firstly, the sample is convolved by the 3D filter of the convolution layer to produce the feature map. Then, the weighted average sum of the local regions of the feature map is carried out, and the bias is increased. A new feature map is obtained by using a nonlinear activation function and feature fusion at the pooling layer. Finally, the feature map is output by the output layer. The network can not only reduce the number of parameters, but also reduce the probability of over-fitting. Its advantage lies in that the low-level features of EEG signal samples are first obtained through convolution processing in the convolution layer, and then the high-level features of the samples are obtained through continuous convolution, which enables effective feature extraction.

The LSTM contains forgetting gate, input gate and output gate [32]. The forgetting gate reads the output of the upper cell and the input of the current cell. The output values between 0 and 1 are given to each neuron, and the neuron decides to choose 0 or 1. The input gate generates a vector through the activation function for updating network data. Output gate Sigmoid layer outputs information through the activation function processing. As a kind of temporal recurrent neural network, this network is often used for the connection of experimental statements and the association between the features of samples. In this study, it is used for feature extraction of EEG signal samples.

CNN-Fusion

The CNN-Fusion is shown in figure 4. It consists of Merge layer, convolution layer (Convl~Conv4), pooling layer (Pool1~Pool4), full connection layer (FC), output layer. As can be seen from figure 4, CNN-Fusion structure first changes from two feature matrices to a single feature matrix, then obtains advanced features of thyroid ultrasound images through multi-layer convolution and pooling. Finally, it outputs detection results according to the specific ultrasound image classifiers. In the Merge layer, the size of the convolution kernel is set as 5×5, that is, the sliding window performs nonlinear operations with the image region of 5×5 and the convolution kernel each time. The input is the output feature matrix of Xception and LSTM.

The size of the convolution kernel in the convolution operation is set to 5×5. That is, the sliding window performs nonlinear operations on the 5×5 image region and the convolution kernel each time. So the input y_j^k of the Conv layer is:

$$y_j^k = \sum_{i \in Z^+} x_i^{k-1} \cdot w_{ij}^k + b_j^k \qquad (16)$$

Where w_{ij}^k is the convolution kernel. x_i^{k-1} is the input feature graph. b_j^k is offset. i and j are the i-th input and j-th neuron, respectively. k is the number of convolution layer and Z^+ is a positive integer. After the convolution kernel performs the convolution operation, it needs to output the feature graph through the activation function. So the output of Conv is:

$$Y_j^k = f(y_j^k) \qquad (17)$$

Where $f(\cdot)$ is the activation function ReLU. Y_j^k is the eigenmatrix. After the convolution layer, the obtained features are down-sampled through the maxpooling layer (feature mapping layer) to reduce the size of input data and obtain the local optimal value [32], which is represented as:

$$x_j^k = f[\alpha_j^k p(Y_j^{k-1}) + b_j^k] \qquad (18)$$

Where $p(\cdot)$ is the max-pooling function. α is the weight coefficient. Pooling reduces the computational complexity to some extent.
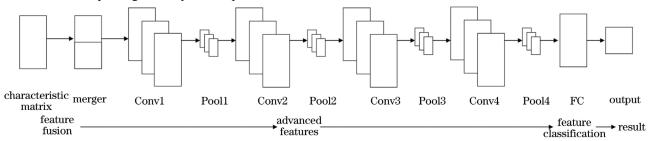

Figure 4. CNN-Fusion structure

Activate function optimization selection

After the convolution operation of the convolution kernel in the convolution layer, it is necessary to output the feature graph through an activation function. Moreover,

the activation function can effectively improve the feature learning ability of the network. Common activation functions are Sigmoid [33] and ReLU [34] functions. Where, the Sigmoid function is:

$$S(x) = \frac{1}{1+\exp(-x)} \qquad (19)$$

ReLU function is:

$$f(x) = \max(0, x) \qquad (20)$$

Where x is the neuron input. Figure 5 shows the activation function curve. As can be seen from the figure, ReLU has the advantage of gradient unsaturation compared with Sigmoid. When $x > 0$, $R(x) = x$. When $x < 0$, R(x) is set to 0, so that it can reduce the gradient dispersion. ReLU function can simplify the calculation process and it is more suitable for this deep network model. Therefore, the activation function of the convolution layer is selected as ReLU.
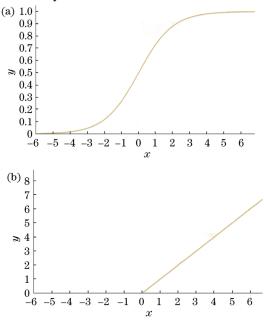

Figure 5. Activation function plots. (a) Sigmoid; (b) ReLU

L2 regularization optimization

In this paper, Softmax is used for optimization. Because Softmax is more sensitive, the noise points in the samples can easily cause the change of Softmax function value, which tends to overfit the model. In view of the adverse effects of noise points in samples on the Loss function of Softmax, L2 regularization is introduced to obtain the compression coefficient and reduce the model parameters. Over-fitting can be avoided to some extent.

After L2 regularization is introduced, the Softmax function is:

$$H(\theta) = J(\theta) + \lambda_1 \|\theta\|_2^2 \qquad (21)$$

Where, $H(\theta)$ is the Softmax with regularization. $\|\theta\|_2$ is the rule item of L2 norm [35]. λ_1 is the adjustable attenuation coefficient. In this paper, by adjusting the value of λ_1, the weight of large values is "weakened", which is helpful to improve the classification of EEG signal features. By introducing L2 regularization into Softmax, the network model can effectively improve the classification effect when classifying EEG signal features.

2.4. F-score feature selection

In general, the correlation between multiple features leads to feature collinearity, which further induces redundant information. Redundant features adversely affect the performance of recognition algorithms. In this study, we propose a new strategy to combine CSP features with brain network features. Simply lumping all the features together may be a sub-optimal operation. Therefore, it is very important to select an optimal subset through a reasonable feature selection strategy to screen out the irrelevant and redundant features from the feature vectors. In this study, we combine the F-score feature discrimination criterion and use the classifier-dependent structure to realize feature selection to reduce the influence of feature multi-collinearity [36]. The recognition ability of the selected feature subset and the classifier is improved and the computation is saved. F-score essentially measures the discrimination of a feature by calculating the ratio of the inter-class and intr-class variances of a single feature. If a feature is discriminative, then the variance between the feature and samples of the same category should be as small as possible, while the variance between the feature and samples of different categories should be as large as possible, so as to facilitate classification and prediction. We give each feature a score, and the F-Score of the i-th feature is defined as follows:

$$F(i) = \frac{\sum_{k=1}^{c}(\overline{x}_{k,i} - \overline{x}_i)^2}{\sum_{j=1}^{n_k}(\overline{x}_{k,j,i} - \overline{x}_{k,i})^2 \sum_{k=1}^{c}\frac{1}{n_k-1}} \qquad (22)$$

Where $\overline{x}_i$ and $\overline{x}_{k,i}$ are the mean value of the i-th feature in the whole data set and the mean value of the i-th

feature in the k-th class, respectively. $x_{k,j,i}$ is the i-th feature of the j-th sample in the k-th class. n_k is the sample number in the class k . The higher F-score denotes the better discriminant ability, indicating that the corresponding features have stronger discriminant ability. As shown in figure 6, we sort the original feature set in descending order according to F-Score value. The packaging model of RCNN and three comparison classifiers are used to select the best special collection according to the classification performance. The feature selection steps are shown in algorithm 1. This method improves the accuracy of subsequent operations to a certain extent, and F-Score method has a small amount of calculation, high accuracy, strong operability, and it saves the operation time.
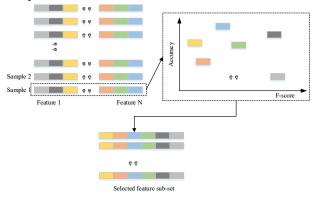

Figure 6. Flow chart of feature selection

Algorithm 1. F-score feature selection method

Input: feature sets of training sample $X = [X_1, X_2, \cdots, X_N]$

Output: optimal feature subset F_{opt} .

Step 1. All F-Score values of feature set X are calculated and sorted in descending order to obtain the sorted feature set F.

Step 2. Initializing the feature subset F_{set} . Acc_{opt} is defined as the best classification result and initialized to zero. Initializing the optimal feature subset F_{opt} to empty.

Step 3. The feature vector with the highest F-score value is selected from F. Adding the selected vector to F_{set} . Let

---

n be the feature number of a single sample in F_{set} .

Step 4. The classification accuracy μ is obtained by 10-fold cross-validation. If $\mu \geq Acc_{opt}$, Acc_{opt} is replaced by μ and F_{opt} is updated.

Step 5. Repeat step 2 to step 4 until n>N.

3. Experiment results

3.1. Data sets and preprocessing

EEG data sets are collected from 69 recruited volunteers. The volunteers are all right-handed undergraduates and have obtained informed consent before the experiment. They complete two tristimania assessments, CES-D [37] and PROMIS [38], and write the questionnaires. After strict screening and matching, the total score of each participant in the two questionnaires is calculated according to the score of the two questionnaires. The 30 students with the highest scores are assigned to the high-risk tristimania group and the 30 students with the lowest scores to the low-risk control group. Table 1 describes the collected data sets.

Table 1. Dataset description

	high-risk group	low-risk group
Number	30	30
Gender(female/male)	24/6	23/7
age	19~22	19~22
Sum of score	35~63	20~33

The EEG acquisition device uses a 64-electrode international standard EEG cap with a sampling frequency of 500Hz. Finally, 62 effective EEG channels are selected for further experimental study. The electrode positions are shown in figure 7.
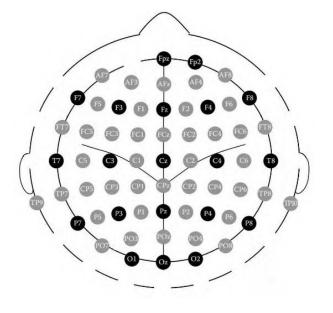
Figure 7. The position of 62 EEG Electrodes

FCz is used as a reference when recording online, and the global average reference is re-referenced when offline. The ElectroOculoGraphy (EOG) of a channel is obtained by placing an electrode above the right eye, and the impedance between all electrodes is kept below 10KΩ. The EEG and EOG signals are filtered by band pass filter (0.05~100 Hz). Independent component analysis (ICA) is used to correct the ocular artifacts.

In this paper, 25000 data points are intercepted from the collected data of 60 volunteers, and the sample size is expanded by piecewise method. The resulting sample format is 63×500 per sample. Among them, 63 represents the number of channels adopted in the experiment, and 500 represents a sample with 500 sampling points. In addition, all the data are processed by wavelet denoising. They are standardized by z-Score before input.

3.2. Results of PLV phase synchronization analysis

Resting-state functional connections between brain regions can provide more information about the pathophysiology of depression [39]. This paper uses graph theory to quantify brain networks. A brain network can be expressed as a graph G=(V,E). Using EEG electrodes as network nodes. Each lead is treated as a node. The relationships between channels represent edges. In order to better reflect the closeness of brain connections, we construct and visualize phase

synchronization functional networks based on PLV matrices. Since PLV ranges from 0 to 1, the PLV value can be used as a synchronization measure to represent connection strength in a weighted network analysis.

Studies have shown that the PLV value of the tristimania group and the control group is greatly affected by the frequency bandwidth, so the signal needs to be filtered to the narrow bandwidth of interest for analysis [31]. Furthermore, Fingelkurts et al. [40] found that tristimania affected almost the entire cerebral cortex, with relatively large oscillating components in a wide frequency range (0.5~30Hz). In addition, EEG signals of patients with tristimania have different oscillations in frequency bands δ, θ, α and β, so this paper uses bandpass filter to filter signals into these four bands for analysis. The functional connectivity matrix describes the connectivity between all pairs of channels in a network in the form of a two-dimensional matrix. The average results of the PLV matrix for all subjects in the tristimania and control groups in the four bands are shown in figure 8.
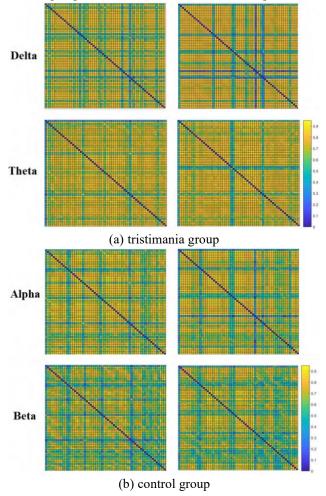

(a) tristimania group

(b) control group

Figure 8. The PLV matrixes of the tristimania and control groups.

If the value of PLV is close to 1, the signals are more synchronized. As you can see, the high-risk tristimania group and the low-risk control group show different connectivity patterns. The result of β band is not obvious, and the connection degree is relatively low. The θ and α bands show greater connection strength. Previous studies on different synchrony measurement methods have shown that the functional connectivity between tristimania patients and the normal control group shows significant differences in the θ and α bands. In order to show the characteristics of the network more intuitively, in addition to using graph theory to take the five network features as the evaluation index of the network, this study also conducts t-test analysis on PLV matrix and five network parameters respectively. The results show that there are significant differences in different attributes under four frequency bands (P<0.05), the specific values are shown in table 2. Compared with the other two bands, the network connection of θ and α bands has a more significant statistical difference.

Table 2. The p values of different features for tristimania group and control group under different bands

	Δ	θ	α	β
PLV	0.026	0.019	0.020	0.031
Clustering coefficient	0.015	0.011	0.009	0.017
Average path	0.024	0.016	0.016	0.033
Global efficiency	0.023	0.023	0.021	0.025
Local efficiency	0.017	0.015	0.013	0.021
Intermediate centrality	0.019	0.012	0.014	0.019

Due to the complex and sparse functional brain network, it is necessary to set the threshold value in the process of network construction. By comparing with the set threshold, the less correlated edges can be removed and the connections of brain functional networks can be clearer. In this paper, the threshold is set to 0.75, and the connection strength with PLV value greater than 0.75 is retained to visualize the brain network.

3.3. Comparison results

The experiment is implemented by Python3.7 programming, Keras framework, Tensorflow. Experiment environment is Inter(R) Core(TM) i5-1035G1 processor, 16G memory, 64-bit Windows10 system.

Five 10-fold cross-validation scheme is used to evaluate the classification performance of the data set. Table 3 shows the specific number of different features. The whole classification process includes the following steps:

(1) dividing the whole data set into 10 subsets, selecting 9 subsets as the training set and the rest as the test set.
(2) Selecting the optimal feature combination according to F_{opt}.
(3) Using the optimal features obtained from the training set to retrain the classifier.
(4) Estimating the classification results with the optimal features of the test set.
(5) Repeat the previous steps 5 times to calculate the average classification accuracy.

Table 3. Number of features used for tristimania recognition

Type of feature	Number of figure
CSP	27
Clustering coefficient	63
Average path	1
Global efficiency	1
Local efficiency	63
Intermediate centrality	63
CSP+network feature	220

For RCNN classifier, the input sequence is subjected to the action of one-dimensional convolution kernel to obtain the sequence feature sequence, and finally the full connection layer is introduced. The probability of belonging to a certain category at the current moment is calculated by Softmax. All activation functions use ReLU, the convolution kernel size is set to 3, and the learning rate is set to 0.002. As for the comparison model, the non-deep learning model SVM is adopted in this paper, and the Gaussian kernel is selected with the penalty factor

C=2. For the deep learning model, the Multilayer Perceptron (MLP) and Bi-LSTM are used to conduct a binary comparison experiment. They both use the Rmsprop optimizer to minimize the cross entropy loss function. To prevent neural network over-fitting, it use Dropout of size 0.3 after the last fully connection layer with batch size of 100.

The four classifiers are used to select the best feature F_{opt} combining with F-Score respectively. Then the average training results are obtained, as shown in figure 9. The accuracy of the proposed model reaches 89.67% in the full frequency band. On the whole, the recognition accuracy of the proposed strategy is significantly improved compared with other models in both single and full frequency bands. The classification accuracy of α band is higher than that of θ band. In the θ band, the classification accuracy of MLP and LSTM is about 7% and 9% higher than that of SVM, respectively. When RCNN is used for classification, the accuracy is improved by about 12.5%. In the α band, the accuracy of MLP classification is improved by about 8% compared with SVM, the accuracy of LSTM classification is improved by about 9%, and that of RCNN is improved by about 14%.
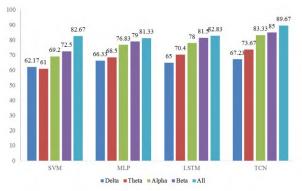

Figure 9. Classification results with different classifiers (%)

It should be noted that the use of different feature selection methods may lead to some differences in the selected optimal feature subset, resulting in different recognition accuracy. For example, F-Score uses discriminability as a criterion to select the optimal feature subset, while maximum correlativity-Minimum redundancy (mRMR) selects good features according to the maximum statistical dependence criterion of mutual information. ReliefF is a typical Filter feature selection algorithm. Therefore, this paper further applies the feature

selection method in the literature to carry out the comparison experiment, and RCNN is adopted in all the classifiers. The results are shown in table 4. Although different feature selection methods select different optimal feature subsets, the results of F-score combining with RCNN model strategy proposed in this paper are better than other methods. We also select other three state-of-the-art multi-feature data fusion methods to make comparison and verify the effectiveness of the proposed method including DDMFI [41], AS-Net [42] and MSFN [43]. The results are shown in table 5. It also demonstrates that the proposed method achieves the better results than other methods.

Table 4. Comparison results of different feature selection methods/%

Method	AC
mRMR	79.28
ReliefF	83.44
F-score	89.78

Table 5. Comparison results with other methods/%

Method	AC
DDMFI	80.81
AS-Net	81.74
MSFN	88.65
Proposed	91.23

4. Conclusion

In this paper, a tristimania recognition method based on EEG brain network topology and CSP feature fusion is proposed. By calculating the PLV adjacencies matrix of multi-channel EEG signals, the brain network can be constructed, and the nonlinear EEG signals can be converted into more intuitive network parameters, which is not only beneficial to the further study of signals, but also helpful to the clinical diagnosis of tristimania. Considering that some redundant features increase the noise and training difficulty, this paper simplifies the

fusion features by combining the F-score method with the deep learning model RCNN. Experimental results show that the proposed method can improve the performance of tristimania identification. In addition, the results of PLV phase synchronization analysis show that the high-risk group and the low score control group display different frequency bands in different brain regions. In the future, the proposed method will be applied into the treatment of mental illness and will produce huge economic benefits.

Acknowledgments.

References

[1] Irene A Kretchy, Frances T Owusu-Daaku, Samuel A Danquah. (2014) Mental health in hypertension: assessing symptoms of anxiety, depression and stress on anti-hypertensive medication adherence. International Journal of Mental Health Systems 8(1):25.

[2] Pim Cuijpers, Annemieke van Straten, Gerhard Andersson. (2008) Internet-administered cognitive behavior therapy for health problems: a systematic review. Journal of Behavioral Medicine 31(2):169-177.

[3] Rupp, C., Gühne, D., Falke, C., et al. (2020) Comparing effects of detached mindfulness and cognitive restructuring in obsessive-compulsive disorder using ecological momentary assessment. Clinical Psychology & Psychotherapy 27(2).

[4] Masanori Higuchi, Yuko Suzuka. (2020) Properties of the coherence function between a sound envelope and magnetoencephalogram data in a selective listening study. Acoustical Science and Technology 41(1):201-203.

[5] Ma, J., Wang, X., Qiu, Q., et al. (2020) Changes in Empathy in Patients With Chronic Low Back Pain: A Structural–Functional Magnetic Resonance Imaging Study. Frontiers in Human Neuroscience 14.

[6] Yin, S., Bi, J. (2018) Medical Image Annotation Based on Deep Transfer Learning. 2018 IEEE International Congress on Cybermatics i-Things. Halifax, NS, Canada. DOI: 10.1109/Cybermatics_2018.2018.00042. 2019.6

[7] Parvinnia, E., Sabeti, M., Jahromi, M., et al. (2014) Classification of EEG Signals using adaptive weighted distance nearest neighbor algorithm. Journal of King Saud University-Computer and Information Sciences 26(1):1-6.

[8] Shi, Q., Yin, S., Wang, K., et al. (2021) Multichannel convolutional neural network-based fuzzy active contour model for medical image segmentation. Evolving Systems. https://doi.org/10.1007/s12530-021-09392-3

[9] Yin, S., Li, H., Liu, D., et al. (2020) Active Contour Modal Based on Density-oriented BIRCH Clustering Method for Medical Image Segmentation. Multimedia Tools and Applications 79: 31049-31068.

[10] Yin, X., Meng, M., She, Q, et al. (2021) Optimal channel-based sparse time-frequency blocks common spatial pattern feature extraction method for motor imagery classification. Mathematical Biosciences and Engineering 18(4): 4247-4263.

[11] Wang, F., Xu, Z., Zhang, W., et al. (2020) Motor imagery classification using geodesic filtering common spatial pattern and filter-bank feature weighted support vector machine. Review of Scientific Instruments 91(3):034106.

[12] Appel, S., Cohen, O., Chapman, J., et al. (2021) Spatial distribution of abnormal EEG activity in Creutzfeldt-Jakob disease. Neurophysiologie Clinique/Clinical Neurophysiology.

[13] Yin, S., Zhang, Y., Karim, S. (2019) Region search based on hybrid convolutional neural network in optical remote sensing images. International Journal of Distributed Sensor Networks 15(5).

[14] Gonuguntla, V., Wang, Y., Veluvolu, K. (2013) Phase synchrony in subject-specific reactive band of EEG for classification of motor imagery tasks. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 2784-2787, doi: 10.1109/EMBC.2013.6610118.

[15] Thomas, K., Robinson, N., and Vinod, A. (2019) Utilizing Subject-Specific Discriminative EEG Features for Classification of Motor Imagery Directions," 2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST), 2019, pp. 1-5, doi: 10.1109/ICAwST.2019.8923216.

[16] Hasan M, Khan M U, Mishra D. A Computationally Efficient Method for Hybrid EEG-fNIRS BCI Based on the Pearson Correlation[J]. BioMed Research International, 2020, 2020(3):1-13.

[17] Al R. Computer Aided Diagnosis of ASD based on EEG using RELIEFF and Supervised Learning Algorithm[J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(6):253-260.

[18] S. Dong, B. Kim and S. Lee, "EEG-Based Classification of Implicit Intention During Self-Relevant Sentence Reading," in IEEE Transactions on Cybernetics, vol. 46, no. 11, pp. 2535-2542, Nov. 2016, doi: 10.1109/TCYB.2015.2479240.

[19] Acharya U R, Oh S L, Hagiwara Y, et al. Automated EEG-based Screening of Depression Using Deep Convolutional Neural Network[J]. Computer Methods & Programs in Biomedicine, vol. 161, 2018, pp. 103-113.

[20] Li X, Zhang X, Zhu J, et al. Depression recognition using machine learning methods with different feature generation strategies[J]. Artificial Intelligence in Medicine, 2019, 99.

[21] X. Zhang et al., "EEG-based Depression Detection Using Convolutional Neural Network with Demographic Attention Mechanism," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 128-133, doi: 10.1109/EMBC44109.2020.9175956.

[22] J. Wang, H. Li, S. Yin and Y. Sun, "Research on Improved Pedestrian Detection Algorithm Based on Convolutional Neural Network," 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2019, pp. 254-258.

[23] Xiaoyu Zhang, Xiaowei Wang, and Shoulin Yin. Multi-modal Data Transfer Learning-based LSTM Method for Speech Emotion Recognition. International Journal of Electronics and Information Engineering. Vol. 13, No. 2, 2021, pp. 54-65.

[24] Nakagome S, Luu T P, He Y, et al. An empirical comparison of neural networks and machine learning algorithms for EEG gait decoding[J]. Scientific Reports, 2020, 10(1).

[25] Liang W, Pei H, Cai Q, et al. Scalp EEG epileptogenic zone recognition and localization based on Long-term Recurrent Convolutional Network[J]. Neurocomputing, 2019.

[26] Borghini, G., Aricò, P., Graziani, I. et al. Quantitative Assessment of the Training Improvement in a Motor-Cognitive Task by Using EEG, ECG and EOG Signals. Brain Topogr 29, 149–161 (2016). https://doi.org/10.1007/s10548-015-0425-7

[27] Yixin Wang, Xiangru Zhu, Lijun Yang. EEG depression recognition based on feature fusion of common spatial pattern and brain connectivity. Computer Engineering and Applications, 2021. doi: 10.3778/j.issn.1002-8331.2104-0329

[28] Jeffrey R Wozniak, Bryon A Mueller, Christopher J Bell, Ryan L Muetzel, Kelvin O Lim. Global Functional Connectivity Abnormalities in Children with Fetal Alcohol Spectrum Disorders[J]. Alcoholism Clinical and Experimental Research, 2012, 37(5):748-756.

[29] Yang F P, Chao A S, Lin S H, et al. Functional human brain connectivity during labor and its alteration under epidural analgesia[J]. Brain Imaging and Behavior, 2020, 14(6):2647-2658.

[30] Duan X, Gou M, Liu N, et al. High-Capacity Image Steganography Based on Improved Xception[J]. Sensors, 2020, 20(24):7253.

[31] Alhagry S, Aly A, Reda A. Emotion Recognition based on EEG using LSTM Recurrent Neural Network[J].

International Journal of Advanced Computer Science and Applications, 2017, 8(10).

[32] Xu Y, Sun M. Convolution Neural Network Image Defogging Based on Multi-Feature Fusion[J]. Laser & Optoelectronics Progress, 55(3) 2018.

[33] Dan Zheng, Hang Li and Shoulin Yin. Action Recognition Based on the Modified Two-stream CNN. International Journal of Mathematical Sciences and Computing (IJMSC), vol. 6, no.6, pp.15-23, 2020. doi: 10.5815/IJMSC.2020.06.03

[34] Dubey S R, Chakraborty S. Average biased ReLU based CNN descriptor for improved face retrieval[J]. Multimedia Tools and Applications, 2021:1-26.

[35] Suraci S S, Oliveira L, Klein I, et al. Monte Carlo-Based Covariance Matrix of Residuals and Critical Values in Minimum L1-Norm[J]. Mathematical Problems in Engineering, 2021, 2021(1):1-9.

[36] A. -R. Mohammadi-Nejad, G. -A. Hossein-Zadeh and H. Soltanian-Zadeh, "Structured and Sparse Canonical Correlation Analysis as a Brain-Wide Multi-Modal Data Fusion Approach," in IEEE Transactions on Medical Imaging, vol. 36, no. 7, pp. 1438-1448, July 2017, doi: 10.1109/TMI.2017.2681966.

[37] Tomokawa S, Asakura T, Keosada N, et al. Reliability and Validity of the Center for Epidemiologic Studies Depression (CES-D) Scale for Adolescents in Lao PDR[J]. Psychology, 2020, 11(11):1646-1660.

[38] Abma I L, Butje B, Klooster P, et al. Measurement properties of the Dutch–Flemish patient-reported outcomes measurement information system (PROMIS) physical function item bank and instruments: a systematic review[J]. Health and Quality of Life Outcomes, 2021, 19(1).

[39] Miseon S, Chang-Hwan I, Yong-Wook K, et al. Altered cortical functional network in major depressive disorder: A resting-state electroencephalogram study[J]. NeuroImage: Clinical, 2018, 19: 1000-1007.

[40] Fingelkurts A A, Fingelkurts A A, Rytsl H, et al. Composition of brain oscillations in ongoing EEG during major depression disorder[J]. Neuroscience Research, 2006, 56(2):133-144.

[41] Chen, S., Xua, K., Yao, X., et al. (2021) Psychophysiological data-driven multi-feature information fusion and recognition of miner fatigue in high-altitude and cold areas. Computers in Biology and Medicine, 33.

[42] Guo M, H Lan, Yang C, et al. (2021) AS-Net: Fast Photoacoustic Reconstruction with Multi-feature Fusion from Sparse Data. arXiv:2101.08934

[43] Y. Pei, B. Sun and S. Li. (2021) Multifeature Selective Fusion Network for Real-Time Driving Scene Parsing. IEEE Transactions on Instrumentation and Measurement, 70: 1-12.