Medical Image Analysis 67 (2021) 101830

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

A novel CERNNE approach for predicting Parkinson's Disease-associated genes and brain regions based on multimodal imaging genetics data

Xia-an Bi[a,b,*], Xi Hu[a,b], Yiming Xie[a,b], Hao Wu[a,b]

[a] Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, Hunan Normal University, Changsha 410081, China
[b] College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China

ARTICLE INFO

Article history:
Received 26 April 2020
Revised 24 July 2020
Accepted 1 September 2020
Available online 10 October 2020

Keywords:
Imaging genetics
Ensemble learning
Parkinson's disease
Abnormal brain region

ABSTRACT

The detection and pathogenic factors analysis of Parkinson's disease (PD) has a practical significance for its diagnosis and treatment. However, the traditional research paradigms are commonly based on single neural imaging data, which is easy to ignore the complementarity between multimodal imaging genetics data. The existing researches also pay little attention to the comprehensive framework of patient detection and pathogenic factors analysis for PD. Based on functional magnetic resonance imaging (fMRI) data and single nucleotide polymorphism (SNP) data, a novel brain disease multimodal data analysis model is proposed in this paper. Firstly, according to the complementarity between the two types of data, the classical correlation analysis method is used to construct the fusion feature of subjects. Secondly, based on the artificial neural network, the fusion feature analysis tool named clustering evolutionary random neural network ensemble (CERNNE) is designed. This method integrates multiple neural networks constructed randomly, and uses clustering evolution strategy to optimize the ensemble learner by adaptive selective integration, selecting the discriminative features for PD analysis and ensuring the generalization performance of the ensemble model. By combining with data fusion scheme, the CERNNE is applied to forming a multi-task analysis framework, recognizing PD patients and predicting PD-associated brain regions and genes. In the multimodal data experiment, the proposed framework shows better classification performance and pathogenic factors predicting ability, which provides a new perspective for the diagnosis of PD.

1. Introduction

As a typical neurodegenerative disease, Parkinson's Disease (PD) mainly occurs in people over 60 years (Benka Walln et al., 2015). Owing to the degeneration and apoptosis of dopaminergic neurons in the midbrain, most patients with PD show symptoms of resting tremor, motor retardation and myotonia (Lv et al., 2019). In addition, some patients are accompanied by sleep disorders, autonomic nervous dysfunction and cognitive impairment (Sveinbjornsdottir, 2016; Videnovic, 2017). With the aging of population, PD is proceeding a serious challenge to public health. At present, the comprehensive and clear understanding of its pathogenesis is yet not available. Therefore, the detection and pathogenic factors analysis of PD are urgent task in brain science.

In the traditional research paradigms, due to the complexity of clinical manifestations of PD, neuroimaging, gait and finger movement are all included in the mainstream approaches of detection and pathogenic factors analysis of PD (Liu et al., 2017; Nettersheim et al., 2019). Dissimilar gait analysis and finger movement analysis, the neuroimaging methods such as electroencephalograph (EEG), computed tomography (CT) and functional magnetic resonance imaging (fMRI) have realized direct measurement of brain function and structure. With the intuition and reliability, neuroimaging techniques play a more important role in exploring the pathogenic factors of PD (Li et al., 2017). Furthermore, compared with EEG and CT, fMRI has obvious advantages in temporal and spatial resolution, so it has a broader application prospect in the study of PD (Griffanti et al., 2016).

Many existing studies have attempted to apply fMRI alone to identifying PD patients and detecting lesions, but rarely consider the fusion of multiple modal data, especially the fusion of imaging data and genetic data to explore this problem. For example, Trojsi et al. (2017) applied covariance projection approach in combination with a bootstrapped permutation test to classify the PD patients. Borchert et al. (2016) used the seed-based correlation

* Corresponding author.
E-mail address: bixiaan@hnu.edu.cn (X.-a. Bi).

https://doi.org/10.1016/j.media.2020.101830
1361-8415/

analysis to detect the lesions of PD. However, with the deepening of the fMRI-based research on PD, researchers have noticed that it is difficult task only by analyzing the fMRI data if we want to have a complete understanding of PD (Lei et al., 2019; Wang et al., 2018; Liu et al., 2019). On the other hand, the latest studies have shown that the pathogenic mechanism of PD is closely related to genes. To date, at least 23 loci and 19 genes have taken part in the process of PD (Chang et al., 2017). Moreover, some significant potential associations between genes and brain regions may reveal their interactions, which arouses the interest of researchers (Bregman et al., 2017). Therefore, the fusion of fMRI data and gene data is a feasible way to explore the pathogenic factors of PD comprehensively, which is also the focus of this paper.

In the multimodal fusion researches of PD, machine learning technology has some unique advantages in the processing of high-dimensional data and small sample size compared with conventional statistical methods (Liu et al., 2016). In particular, Tang et al. (2019) observed the great potential of artificial neural network (ANN) in multimodal data analysis of PD. Furthermore, Bi et al. (2018) proposed an improved neural network method, which greatly improved the recognition accuracy of patients with brain diseases. Although the methods of ANN have shown quite satisfactory performance in the multimodal fusion study of PD, its potential has not been fully studied. Therefore, how to tap the potential of ANN in PD multimodal fusion researches is a objective of this study.

Furthermore, most of current multimodal data fusion analyses of PD using imaging and genetic data pay close attention to a single process, which is easy to ignore the study of the comprehensive analysis framework. Hao et al. (2020) emphasized the significance of fusion features construction in multimodal fusion analysis by consistent metric constraint feature selection method. Zeng et al. (2016) demonstrated the practicability of classification method in fusion research by a new deterministic learning technology. Gupta et al. (2018) highlighted the role of feature selection method in fusion research of PD. However, it must be noted that the design of a multimodal data fusion analysis framework for PD based on machine learning technology is more conducive to utilizing complementary information of different modal data and providing assistance for clinical diagnosis. This problem is a potential driving force of this study. Therefore, the design of a multimodal fusion framework for PD based on machine learning is another objective of our study.

In order to realize the above objectives, the fMRI and single nucleotide polymorphism (SNP) data are applied to performing multimodal fusion analysis of PD. Firstly, We test a variety of correlation analysis methods, and select the optimal method to extract the associations between genes and brain regions as the fusion features of multimodal data. Secondly, an improved neural network model of clustering evolutionary random neural network ensemble (CERNNE) is proposed. This method chooses the appropriate neural network for large-scale random integration, and introduces the idea of hierarchical clustering for adaptive dynamic optimization. Finally, with CERNNE as the core, we construct the multimodal fusion framework for PD to realize multi-task analysis of feature construction, patient recognition and pathogenic factors prediction. The framework is assessed by the real multimodal data of PD patients. The experiments suggest that the proposed comprehensive framework can effectively realize the recognition of PD patients and lesion detection, which is helpful to understand the pathogenesis of PD.

2. Materials and methods

In this section, we mainly introduce the multimodal data fusion analysis framework of PD proposed in this paper, which realizes the functions of feature construction, patient recognition and lesion detection. The framework consists of the following four parts including multimodal data preprocessing, fusion features construction, sample classification and PD-associated genes and brain regions prediction. The overall flow chart of the framework is shown in Fig. 1.

2.1. Multimodal dataset and preprocessing

In this paper, 104 samples are used to train and validate the framework, including 55 PD patients (male / female: 37 / 18; mean age: 66.9 $\pm$ 4.5 years) and 49 healthy controls (male / female: 24 / 25; mean age: 69.3 $\pm$ 5.3 years). Specifically, 69 samples from Parkinson's Progression Markers Initiative (PPMI) database including 55 PD patients and 14 healthy controls (HC). In order to expand the sample size to realize better framework training effect, we further obtain 35 HC from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, which are matched to the samples of PPMI database in the term of gender and age, and have fMRI and SNP data. Both Alzheimer's Disease (AD) and PD are neurodegenerative diseases, which makes PPMI and ADNI have great similarity in determining the standard of HC. All HC from ADNI and PPMI databases are free from other neurological diseases. In order to eliminate the potential impact of data mixing on the model, the scanning parameters of fMRI data we selected from the two cohorts are almost the same. Meanwhile, the SNP explored in the experiment are shared by both database cohorts. All data acquisition work has been approved by relevant institutions, and participants have signed written informed consent.

All fMRI data is preprocessed by the software of DPARSF. The specific process is as below. The data format is converted into NIFTI format and the first 10 fMRI functional volumes are discarded. Subsequently, the head motion and time slices of fMRI images are realigned, and the EPI template is used for image registration. Finally, using Gauss smoothing (full width at half maxima = 6mm), linear model and signal filtering (0.01-0.08HZ), the irrelevant variable interferences such as global signal and white matter signal are removed.

SNP is included in this study as a genetic material. The preprocessing of SNP is performed by PLINK software. Firstly, the sample recall rate is set to 95%. Next, the minimum allele frequency, genotyping rate and the threshold of Hardy-Weinberg test are set at 0.03, 0.95 and 1e-5, respectively. After the above quality control, we retain high quality SNP information for the next experiment.

2.2. Multimodal fusion features

In computer and data science, the biological interactions between brain regions and genes can usually be detected by coding correlation of brain region and SNP sequences, which is also one of accepted methods at present. Researchers have tried and achieved some satisfactory results to verify the feasibility of this method (Hao et al., 2017; Du et al., 2020). Our research is based on this concept for in-depth exploration and tries to use a more practical and generalizable method.

The preprocessed fMRI image is segmented into 90 regions of interest (ROIs) by anatomical automatic labeling (AAL) template, and the functional time series of each ROI is extracted, its length is fl. Subsequently, for SNPs through quality control, we query their reference SNP (rs) number in National Coalition Building Institute (NCBI) and Ensemble website to determine the genes they belong to. Then we group the SNPs according to their corresponding genes and arrange the SNP groups according to SNP position in gene. We select groups with SNP count more than sl as the candidate genes and recode the genes discretely. Specifically, the four bases A, T, C and G in SNP are recoded to 1, 2, 3 and 4 respectively according

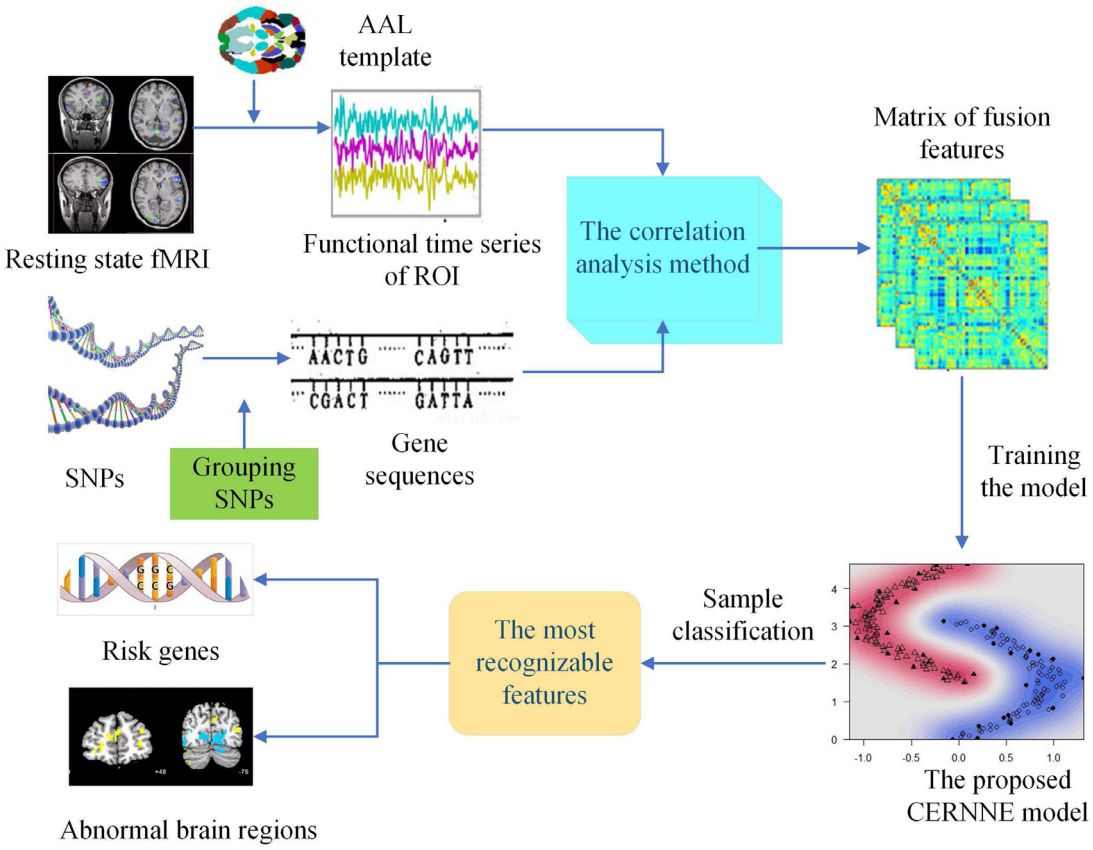


Fig. 1. Overview of the proposed method.

to the PLINK (1.07) Documentation (Purcell, 2012), thus the digital sequence of gene group is obtained. Ultimately, the time series length of ROI is usually larger than the gene sequence length. Therefore, the time series length of ROI is intercepted to make it equal to the gene sequence length. Pearson correlation analysis is used to calculate the associations between ROIs and genes as fusion features (Eq. 1).

$$Pea_{r,s} = \frac{l\sum w_r s_s - \sum w_r \sum s_s}{\sqrt{l\sum {w_r}^2 - (\sum w_r)^2}\sqrt{l\sum {s_s}^2 - (\sum s_s)^2}} \tag{1}$$

where w_r represents the functional time series of a ROI, s_s is the gene sequence, and l represents the length of each ROI or gene.

In addition, in order to verify the robustness of Pearson correlation analysis in multimodal data fusion, we use canonical correlation analysis (CCA, Eq. 2) and correlation distance (CD, Eq. 3) as the comparison methods, and the calculation equations are as follows.

$$CCA_{r,s} = \frac{\beta^T \sum_{12} \gamma}{\sqrt{\beta^T \sum_{11} \beta}\sqrt{\gamma^T \sum_{22} \gamma}} \tag{2}$$

$$CD_{r,s} = 1 - \frac{l\sum w_r s_s - \sum w_r \sum s_s}{\sqrt{l\sum {w_r}^2 - (\sum w_r)^2}\sqrt{l\sum {s_s}^2 - (\sum s_s)^2}} \tag{3}$$

where β represents the weights of ROIs and γ represents the weights of genes. $\sum_{11}$ represents a covariance matrix of ROIs and $\sum_{22}$ represents a covariance matrix of genes. $\sum_{12}$ represents a covariance matrix of ROIs and genes.

2.3. Clustering evolutionary random neural network ensemble

As an important branch of machine learning technology, the ANN still has considerable potential in high-dimensional data processing. A novel improved method of neural network is proposed in this paper. We introduce the idea of evolutionary learning to design an ensemble learning model based on the neural network. In this model, a suitable neural network is selected as a single computing node for large-scale random integration, and clustering evolution strategy is used to enhance the variousness of the ANN base learners in the model through dynamic adaptive selection. The implementation process is as follows.

In the initial stage of building the model, all sample data are marked as $\{x_i, y_i\}^N$, where $x_i = \{f_1, f_2, f_3, \ldots, f_m\}$ denotes the feature set composed of m fusion features of the $sample_i$, y_i denotes the category label of $sample_i$, "$+1$" denotes the normal person, and "-1" denotes the patient with PD. We randomly extract 30% of all samples as test set marked as $\{x_i, y_i\}^t_{test}$ at the first. Then we randomly choose the left samples at a ratio of 7 : 3 each time, as training set $\{x_i, y_i\}^r_{train}$ and corresponding validation set $\{x_i, y_i\}^s_{validate}$ of a neural network base classifier in the ensemble learning model. That is to say, the training set of each base classifier is different, which is helpful to improve the diversity of base classifiers.

After determining the data partitioning strategy, the training set $\{x_i, y_i\}^r_{train}$, which is randomly extracted each time, is used as a training sample of the neural network base classifier. At the same time, the input features of base classifier are randomly selected from all fusion features, and the number of input features is defined as follows:

$$s = fix\sqrt{m} \tag{4}$$

where m equals to the quantity of all fusion features. Through the above methods, we construct a single neural network base classifier, and use the corresponding validation set $\{x_i, y_i\}^s_{validate}$ to evaluate the classification performance of the base classifier. Furthermore, in order to construct an initial ensemble learner, the process

of constructing the base classifier is repeated z times, and once the training of the base classifier is completed, the base classifier does not need to be retrained in the subsequent clustering evolution process.

The initial ensemble learner is clustering evolved to improve the robustness and accuracy. The setting of clustering index is the precondition of clustering evolution. We take the similarity between neural networks as the clustering criterion which is the disagreement measure (DM). Suppose there are two neural network base classifiers ANN_U and ANN_V. We count the classification results of the two classifiers. Among them, z_{11} equals the quantity of validation samples correctly recognized by ANN_U and ANN_V at the same time. On the contrary, z_{00} represents the quantity of validation samples misidentified by both ANN_U and ANN_V. Similarly, z_{10} equals the quantity of validation samples that can only be recognized by ANN_U, and z_{01} equals the number of validation samples that can only be recognized by ANN_V. As a result, the DM_{UV} between ANN_U and ANN_V is defined as:

$$DM_{UV} = \frac{z_{01} + z_{10}}{z_{01} + z_{10} + z_{00} + z_{11}}. \tag{5}$$

The value of DM_{UV} is negatively correlated with the similarity. A smaller value indicates that there is a higher similarity between ANN_U and ANN_V, and it is more likely to be viewed as the same cluster.

Based on the above similarity index, we calculate the similarity between neural network classifiers in the initial ensemble learning model, and construct the following similarity matrix $Matrix_{(s)}$:

$$Matrix_{(s)} = \begin{bmatrix} DM_{(1,1)} & \dots & DM_{(1,z)} \\ \dots & \dots & \dots \\ DM_{(z,1)} & \dots & DM_{(z,z)} \end{bmatrix} \tag{6}$$

where $DM_{(z,1)}$ indicates the disagreement measure between ANN_z and ANN_1. This paper applies the linkage hierarchical clustering algorithm to clustering the base classifiers with high similarity. The specific process is as follows. In the initial stage of clustering evolution, each base classifier is regarded as an initial cluster. Next, we calculate the disagreement measures of different base classifiers as the clustering criteria. If there is a maximum similarity between two base classifiers (the value of DM is minimum), the two base classifiers will be defined as a new cluster in the clustering evolution. Then the base classifier with the best performance in a cluster is retained for the next clustering evolution. The above process is a process of clustering evolution. For training the model, this process is iterated multiple times, and the termination condition is that the evolution times reaches the preset threshold. In the process of clustering evolutions, the number of base classifiers is decreased. In order to control the iteration process, the step size of iteration is set to cl, which makes us gradually find out the set of high similarity base classifiers. The quantity of neural network base classifiers retained in the final ensemble model is as follows:

$$ANN_{ensemble} = z - w \times cl \tag{7}$$

where w denotes the evolution times corresponding to the peak performance of the ensemble learner and z denotes the number of initial base classifiers. All the above procedure of the CERNNE is summarized in Algorithm 1.

2.4. Classification mechanism

The neural network ensemble model evolved by multiple clustering evolutions is regarded as the final model of classification. We design a voting method as the final decision strategy of the ensemble learner. Considering that each base classifier in the final CERNNE has satisfactory classification performance, each classifier is given equal voting right. When a test sample x is input into the model, each learner in the model will give a classification result, forming a classification result set:

$$result = \{f_1(x), f_2(x), \dots, f_k(x)\} \tag{8}$$

where $f_k(x)$ indicates the classification result of the $k-th$ base classifier. We count all the classification results by the Eq. (9):

$$Label_A = \sum_{i=1}^{k} I(f_i(x) = A) \tag{9}$$

where $I(*)$ is the indicator function. If the test sample x is predicted to be belong to category A by the $i-th$ base classifier, the value of $I(f_i(x) = A)$ is 1, otherwise the value is 0. Then, the label with the largest value is the final category of the unclassified sample. The calculation equation is defined as:

$$L = argmax(Label). \tag{10}$$

2.5. Parkinson's Disease-associated genes and brain regions

After several clustering evolutions of neural network ensemble, the base classifiers with the greatest diversity and the best classification performance in ensemble learning model are retained. It is well known that input features have a great influence on the classification ability of the base classifier, which means that the input features in the retained base classifier have important potential information. Therefore, we design a multi-stage analysis scheme to look for the most recognizable features.

In the first stage, due to multiple clustering evolutions, different fusion features in the base classifier contribute to classification to a certain extent. The features occur in multiple base classifiers repeatedly, which means that it may have a significant positive impact on classification. Therefore, all features in the retained base classifier are counted. And we use frequency as the criterion to select e high-frequency features, which will be used to further search for the PD-associated genes and brain regions.

In the second stage, we have completed the preliminary screening of features through the frequency in the process of extracting high-frequency features, however, the correlations between partly high-frequency features with relatively low frequency and disease may not be stable. Therefore, we use the backward sequence search algorithm to test all the e high-frequency features in descending frequency order, and determine the most recognizable fusion features. The selected e high-frequency features are divided into several feature subsets, and the partitioning strategy is as follows. The first g features of high-frequency features are taken as feature $subset1$. Then, with h as step size, the quantity of features in the feature subset is gradually increased in descending order of frequency until all high-frequency features are contained within the feature subset. Finally, different feature subsets are used as training features of the traditional neural network ensemble. The classification performances of different subsets will be tested. These features with strong recognition ability can better reflect the difference between PD and HC, which also means that the brain regions and genes fused by these features may have more strong correlations with PD and deserve further analysis.

In the third stage, the most recognizable fusion features are analyzed separately. On the one hand, the fusion features represent the correlations between brain regions and genes, which has outstanding explicability. On the other hand, these features have strong recognition ability, which shows that they have obvious distinction between patients and normal people. The brain regions contained in these features are more likely to have functional or structural lesions, and the genes contained in these features are more likely to have abnormal expressions. Therefore, the brain regions and genes contained in the fusion features are extracted as a single component, and their frequencies appearing in the most

Algorithm 1 CERNNE learning process.

Input: experimental data set {X,Y}
Output: The clustering evolutionary random neural network ensemble
Initialize {X,Y}, cl, w.
{X,Y} is experimental data set.
z is the number of initial ANN.
cl is the step size of clustering evolution.
w is the clustering evolution times.
Partitioned the {X,Y} into $\{X,Y\}_{train\_1}$,$\{X,Y\}_{validate\_1}$,⋯,
$\{X,Y\}_{train\_n}$, $\{X,Y\}_{validate\_n}$ and $\{X, Y\}_{test}$
i=1
for k = 1 to z :
  select $\{X,Y\}_{train\_k}$
  Randomly select a subset of features as $\{Features\}_{train\_k}$
  $\{X,Y\}_{train\_k}$ and $\{Features\}_{train\_k}$ → ANN_k
  $\{X, Y\}_{validate\_k}$ → test the classification accuracy of ANN_k
end for
ANN Ensemble = Ensemble of $\{ANN_1 \cdots\cdots ANNz\}$
Do
  Calculate a similarity measure between ANNs according to Eq. (5)
  Clustering the ANNs in the ANN ensemble into several clusters
  Retain the ANNs with the highest accuracy in each cluster→ remove the inefficient ANNs
  z = z -w*cl, w=w+1
  $\{ANN\ Ensemble\}_{new}$ = Ensemble of the retained ANNs
  Calculate the classification ability of $\{ANN\ Ensemble\}_{new}$
Until the evolution times reach the preset threshold
  CERNNE = {ANN Ensemble} with the highest classification ability

recognizable features are counted separately. The high-frequency components (i.e. specific brain regions or genes) imply that they may be more closely related to the disease.

2.6. Parameters optimization

The type of base classifier, the number of cluster evolutions and the number of initial ANN base classifiers are free parameters in the CERNNE model. Optimizing these parameters can effectively promote the overall performance of the model.

Firstly, the types of neural networks are various and have their own characteristics. To find the best type of neural network as the base classifier, we preliminarily test the integrated performance of back propagation neural network (BP NN), probabilistic neural network (PNN), Elman neural network (Elman NN), learning vector quantization neutral network (LVQ NN) and competitive neural network (Competitive NN). Specifically, we use different types of neural networks as base classifiers, and set the number of base classifiers to K to construct a initial random neural network ensemble. Then, for five different types of neural network ensemble, the linkage hierarchical clustering evolution strategy is adopted to find the performance peaks. We choose the base classifier type according to the pre-determined standard. We mainly consider the performance peak of the ANN ensemble learner. Specifically, the neural network ensemble with the highest performance peak shows that this type of neural network is more suitable as the base classifier of ensemble learning model.

After determining the type of base classifier, the number of initial base classifiers and the number of clustering evolutions also have an impact on the construction speed and classification performance of ensemble learning model. In this paper, a grid search strategy is designed to find the optimal parameter combination. Assuming that the optimal ANN base classifiers number is between the interval of $[c, d]$ by repeated experiments, we use different ANN base classifiers number in the interval to construct ensemble learners. Next, clustering evolution is used to make each neural network ensemble reach its peak performance, and the number of clustering evolutions corresponding to different ensemble learners' peak is counted. When building the optimal CERNNE model, we choose the minimum number of clustering evolutions to reduce the time cost of building the ensemble learner as much as possible. At the same time, the number of base classifiers is also selected a less value, which also helps to reduce time complexity.

3. Results

3.1. Constructing fusion features

For each preprocessed sample, the functional time series of 90 brain regions were obtained from fMRI data, and 23595 SNPs were extracted from genetic data. In order to ensure the effectiveness of fusion feature construction, we controlled the length of ROI and SNP sequences so that all sequences were converted to the same length. In detail, we retained 45 SNP groups with more than 40 SNPs, and discretized the first 40 SNPs of each SNP group to obtain a gene digital sequence with a length of 80. Meanwhile, the functional time series of brain regions were also adjusted to 80. The above length threshold we selected was determined by repeated experiments. Because the relatively long sequence could ensure the robustness of correlation analysis, and provide better differentiated fusion features. According to the multimodal fusion feature construction scheme, the Pearson correlation analysis of brain re-

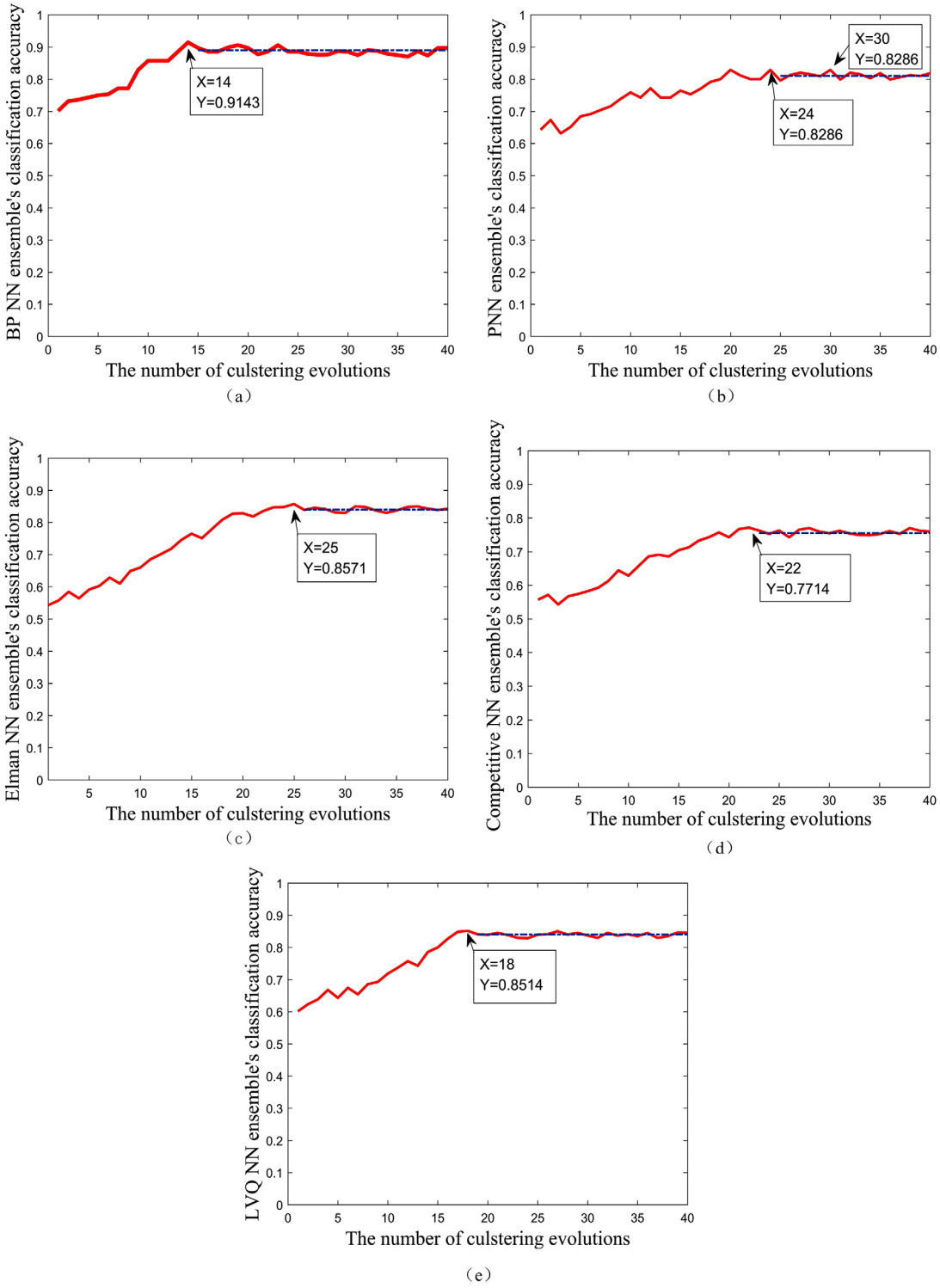


Fig. 2. The ensemble performance of different base classifiers. We compare the performances of different classifiers, including BP NN, PNN, Elman NN, Competitive NN and LVQ NN.

gions and genes was carried out. Consequently, each sample was abstracted into a set of 4050-dimensional fusion features, which were used as the basis for subsequent analysis.

3.2. Selecting the type of base classifier

We used various types of neural networks as base classifiers to construct different neural network ensembles, and compared the performance of the ensemble learners to determine the most suitable base classifier type. The candidate base classifier types included BP NN, PNN, Elman NN, Competitive NN and LVQ NN. Specifically, the number of hidden layers of BP NN, Elman NN and LVQ NN was set to 5, and the parameters of other networks were set to the default parameters of MATLAB platform. In addition, the number of iterative training for all neural network base classifiers was 300. For any type neural network ensemble, we set the

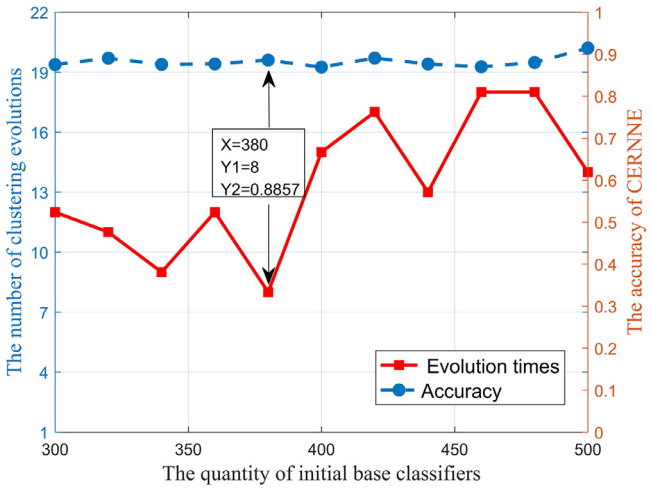


Fig. 3. The optimal quantity of base classifiers. In fact, we search for the optimal combination of parameters in a wider range, and determine the optimal range of parameters, and display the parameter search results in this range in detail.

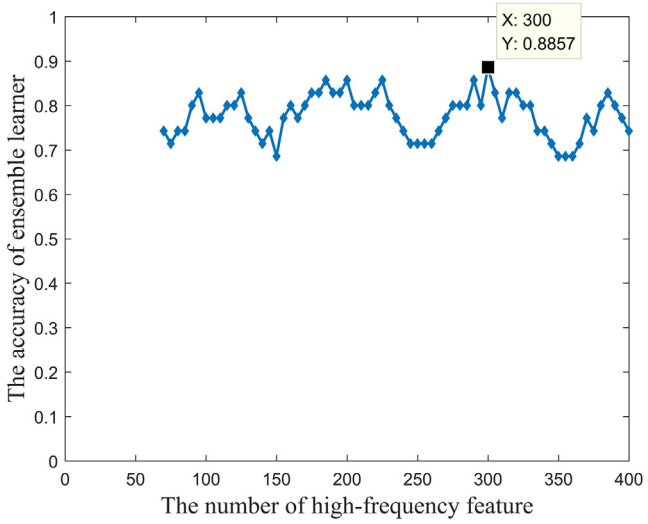


Fig. 4. The most recognizable features. The first 300 fusion features have the highest classification performance, reaching 88.57%.

quantity of training samples and random input features of its base classifier to 50 and 64 respectively, and the quantity of initial base classifiers in the ensemble learner was set to 500. Subsequently, five different types of neural network ensemble models were clustering evolved 40 times with 10 as step size, and their performance curves in all the clustering evolutions were depicted in Fig. 2.

In Fig. 2, with the clustering evolution times increased, the classification performance of the ensemble learner rose steadily, and started to converge after reaching the peak value. Even though the accuracy fluctuated in a small range, the overall performance tended to be stable. According to the pre-determined basis classifier selection criteria, we found that the BP NN ensemble had the highest classification performance and the fastest convergence speed among all types of neural network ensemble, which showed that it has the potential to be the best base classifier. In addition, it is noteworthy that there are two peaks in the evolution process of random PNN ensemble, which may indicate that it is difficult for them to achieve stable performance in the evolution process. Based on the above analysis, it is concluded that BP NN has better potential in clustering evolution, so it is regarded as the optimal type of base classifier to build the final CERNNE model proposed in this paper.

3.3. Training the optimal CERNNE model

We employed BP NN to construct the optimal clustering evolutionary random neural network ensemble. In the experiment, we found that the different initial base classifier numbers required the different optimal clustering evolution times, which affected the performance of CERNNE model. However, we only tested the case when the number of initial base classifiers equals 500 (Fig. 2(a)). Therefore, we further searched for the optimal quantity of initial base classifiers to improve the construction efficiency and performance. In the experiment, we analyzed the evolution process of ensemble learners with different initial base classifier numbers and preliminarily determined the search range of the optimal initial base classifier number in the interval [300,500] through a large range of parameter search. At the same time, the relationship between the initial base classifier numbers and the optimal clustering evolution times was also displayed in Fig. 3. We found that when the number of base classifiers was enough, the clustering evolutions could make the ensemble learners with different initial base classifiers number reach the peak performances, and the performance peaks were relatively close. But we noticed that when the number of BP NN base classifiers was 380, the number of cluster evolutions was the least. Therefore, considering the cost of clustering evolutions, we finally used 380 as the initial base classifiers number. This meant that when the performance peaks were at the approximate level (86.5%-91%), the CERNNE was constructed by using 380 as the quantity of initial base classifiers. It took only 8 clustering evolutions to achieve the performance peak and had the fastest construction efficiency.

3.4. Extracting fusion features with recognition ability

After many tests on the ensemble performance of the optimal CERNNE, the average classification accuracy was 88.6%. The results showed that clustering evolution improves the diversity of base classifiers and the validity of input features. According to the multi-stage analysis scheme in the section of methods, in the first stage, we counted the frequency of input features in the base classifiers which were in the optimal CERNNE model, and extracted 400 high-frequency features for the next stage of analysis. The above operation reduced the searching range of the most recognizable feature to 400 dimensions.

Then, in order to extract the most recognizable part of high-frequency features, the backward sequence search algorithm was used to test all possible combinations of high-frequency features in descending order of frequency. The 400 high-frequency features were divided into several subsets of fusion features. Specifically, the first feature subset consisted of the first 70 high-frequency features. Then we increased the number of features in the subset gradually according to the frequency descending with 5 as the step size until the subset contains all 400 high-frequency features. These feature subsets were inputted the traditional random BP NN ensemble to test their classification ability. The results were shown in Fig. 4. There were several local peaks in the curve, but the peak point was $(300, 0.8857)$. In addition, the trend of the curve near this point changed obviously, and the curve showed a downward trend after the number of features exceeded 300, which indicated that the recognition ability of the last 100 fusion features was weak and may interfere with classification. Therefore, we regarded the first 300 dimensional high-frequency features as the most recognizable features, and visualized the first 20 fusion features in Fig. 5.

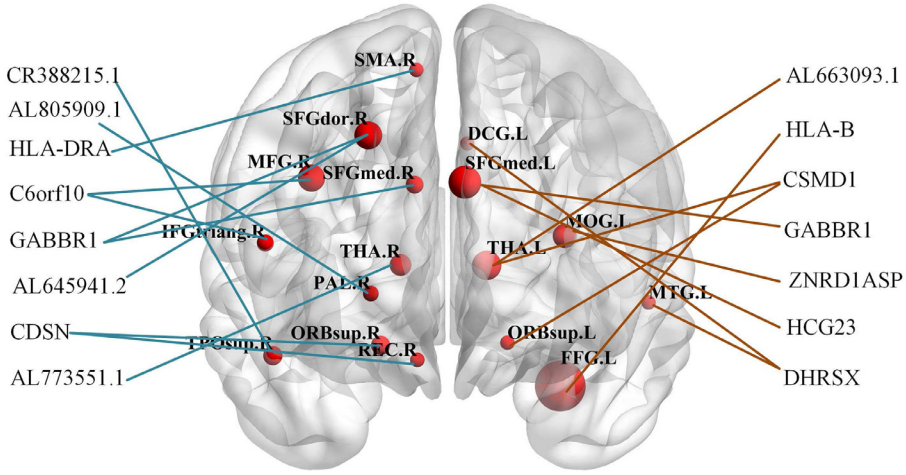


Fig. 5. The top 20 features with the strongest recognition ability. The node represents a single brain region or gene.

Table 1
The quantitative comparison of optimal fusion feature on multimodal data.

Method	Optimal Fusion Feature	Accuracy of SVM	Overlap with GERNNE
Pearson + CERNNE	300	0.904	
Pearson + RF	670	0.785	202 (p = 2.764806e-27)
Pearson + RSVME	260	0.619	94 (p = 1.642645e-6)
Pearson + t-test	499	0.714	147 (p = 2.764806e-32)
CCA + t-test	412	0.785	192 (p = 3.817671e-03)
CD + t-test	447	0.738	159 (p = 4.08077e-12)

The t-test is two-sample t-test.
The p value is acquired by the hypergeometric test.

3.5. Method comparison and verification

Although satisfactory results have been achieved in extracting the most recognizable features based on the proposed model. The multimodal fusion as an emerging field, the validity of fusion features still needs to be verified. In this study, a variety of fusion feature construction and analysis methods were combined into a new model and compared with the model proposed in this paper. The fusion feature construction methods included canonical correlation analysis and correlation distance, feature analysis methods included random support vector machine ensemble (RSVME), random forest (RF) and two-sample t-test. Among them, the construction parameters of RF and RSVME, such as the number of base classifiers and the number of input features, were the same as the proposed CERNNE model. By analyzing the quantity and classification performance of the most recognizable feature subsets detected by other models, we verified the rationality of the most recognizable features extracted by our model. The results were displayed in Table 1.

From the comparison results, in the number of the most recognizable features, Pearson + RF was the largest, while the proposed model was the lesser. In addition, we found an interesting result. If the most recognizable features extracted by different methods were used as the training features of support vector machine (SVM), the classification performance of feature extracted by our model was the best, which showed that the features extracted by proposed model were more effective. In order to further verify this hypothesis, we analyzed the overlapping parts of the most recognizable features extracted by our method and the most recognizable features extracted by other methods. The results showed that the larger the overlapping parts were, the better the performance of the method was. At the same time, the results of hypergeometric distribution test also proved that these overlaps were not randomly generated. Based on the above analysis, we found that the features extracted by other methods may contain more interference information, while our method improved this situation.

In order to further prove the validity of the fusion features and clustering evolutions, CERNNE was compared with other models including single classifier method and statistical method under multimodal fusion feature and unimodal feature respectively. The experimental results were shown in Fig. 6. Under the classical statistical method, the classification performance of multimodal fusion feature was better than that of unimodal feature, which showed the benefit of data fusion to classification. After the optimization of the model proposed in this paper, the classification ability was obviously improved compared with the other methods, especially the single classifier method. It also proved the rationality and advantages of clustering evolution strategy and ensemble learning in this paper.

3.6. Predicting abnormal brain regions and risk genes

We isolated genes and brain regions from the most recognizable features extracted by our model, and analyzed the frequencies of genes and brain regions respectively. The higher the frequencies of brain regions and genes were, the greater the difference between patients and normal people was. Therefore, they may be more likely to have a potential association with the disease. We showed the information of high-frequency brain regions and genes in Fig. 7 and Fig. 8. The high-frequency brain regions and genes included Thalamus, Caudate nucleus, Amygdala, Angular gyrus, GABBR1, CSMD1, C6orf10, HLA-DRA and BTNL2. We noticed that most of the high frequency brain regions and genes had influence on human emotion, memory, cognition and behavior, and many literatures revealed that they were directly or indirectly related to PD from the aspects of pathology, neurocognition and genetics (Gilat et al., 2018; Bot et al., 2018; Saeed, 2018).

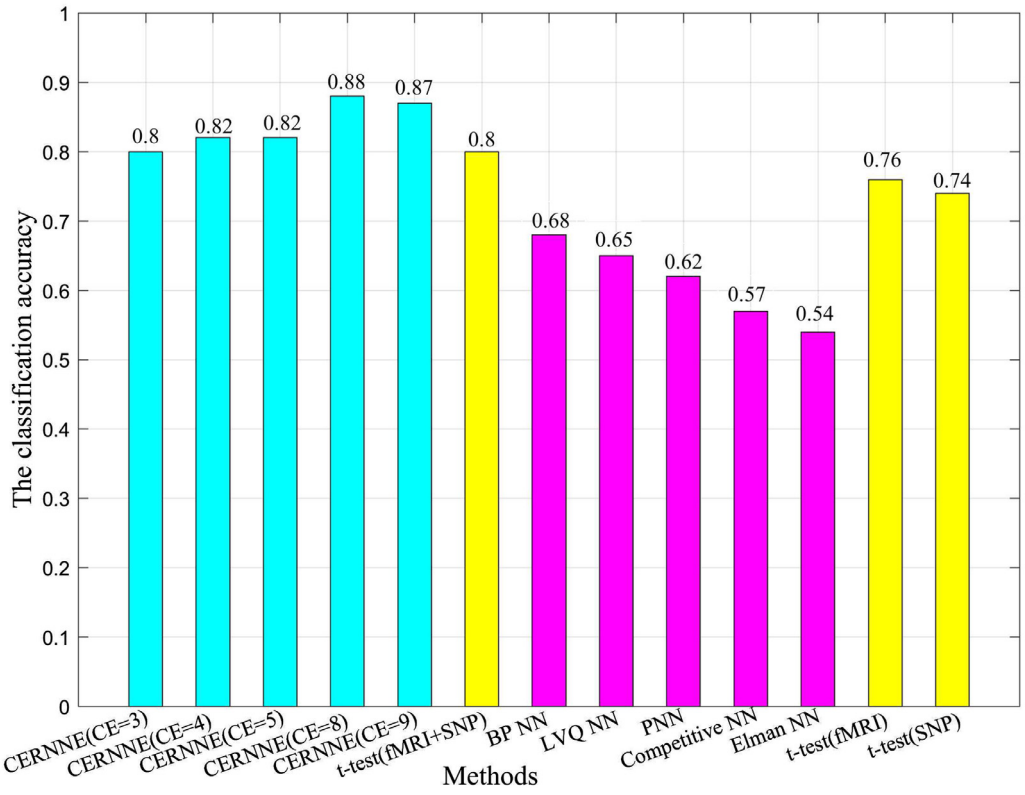


Fig. 6. The comparison results of classical single modal features with the multimodal features. The CE denotes the number of clustering evolutions, where the performance of CERNNE is tested when CE is 3, 4, 5, 8 and 9, respectively.

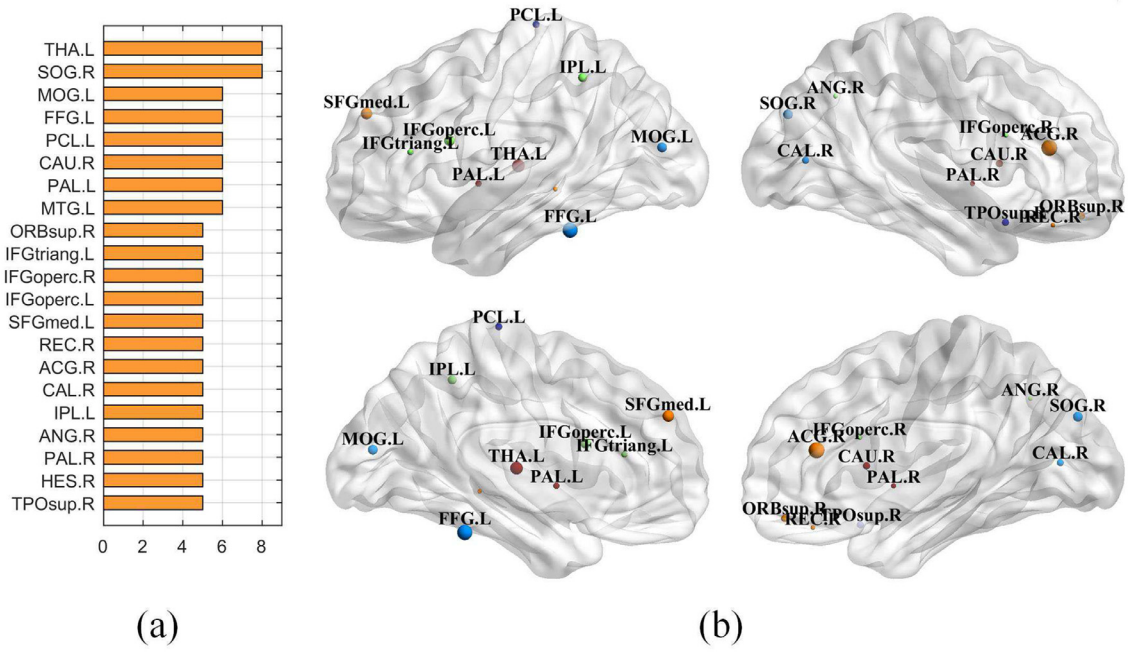


Fig. 7. The locations and frequencies of pathogenic brain regions. The figure 7(a) denotes the frequency information of abnormal brain areas. The figure 7(b) denotes the location information of abnormal brain regions.

4. Discussion

4.1. Consistency with the literature

We verified the validity of our results from some conventional studies. For example, Garg et al. (2015) noticed thalamic morphological abnormalities in PD patients. Dirkx et al. (2017) reported that the changes of cerebello-thalamo-cortical circuit were responsible for resting tremor in PD patients. Abnormal dopaminergic denervation in the caudate nucleus led researchers to believe that this region was the key area for the onset of PD (Bohnen et al., 2015). The pathological analysis also provided evidence to confirm

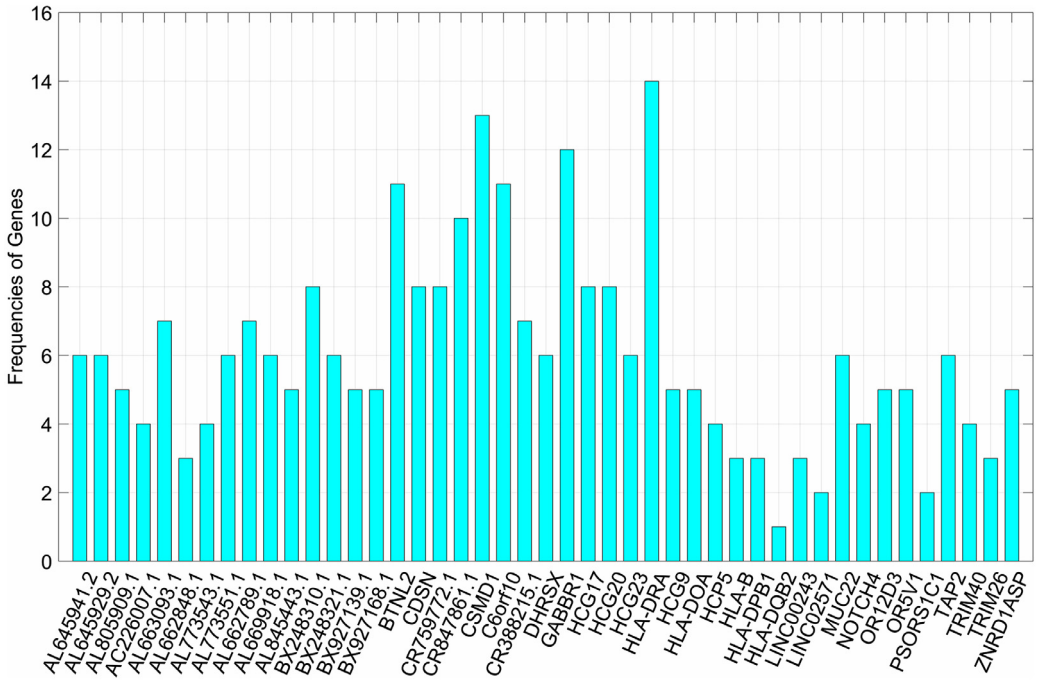


Fig. 8. The frequencies of genes. The total number of candidate genes examined in this study was 45, and the high-frequency genes were more likely to be risk genes.

that caudate denervation could lead to excessive daytime sleepiness in PD patients (Politis, 2018). van Mierlo et al. (2015) found a significant association between amygdala atrophy and depression in PD patients by voxe-based morphometry (VBM) test. Subsequently, many researchers also reported that amygdala was involved in apathy, anxiety and other pathologies of PD (Thobois et al., 2017). Garcia-Esparcia et al. (2018) identified abnormal mitochondrial activity in the angular gyrus of PD patients by metabolic tests.

In the risk genes predicted by our method, the potential association between some genes and PD also has been confirmed by many genomics studies. The whole exon sequencing showed that the mutation of CSMD1 gene resulted in abnormal complement activation, which was an important single gene mutation factor in PD (Patel, 2017; Ruiz-Mart-nez et al., 2017; Shahmohammadibeni et al., 2016). The results of multiple genome-wide association analysis and meta-analysis also showed that HLA-DRA gene was involved in the pathogenesis of PD and was an important risk gene (Alexis, 2012; Jamshidi et al., 2014). Recently, Gusev et al. (2019) further found that the multipoint chromatin interaction between HLA-DRB gene and BTNL2 gene was an important cause of many nervous system diseases including PD through the analysis of chromosome conformation capture in human brain cells. In addition, some literatures also reported that GABBR1, C6orf10 and other genes were more or less related to PD (Consortium, 2011; Patel, 2017; Perrone-Bizzozero, 2019).

4.2. Method performance analysis

The results of our study reveal that altered fusion features of patients with PD could be detected automatically by CERNNE. The proposed method not only provides an efficient and reliable approach to identify PD patients, but also explores the multimodal pathogenic factors of PD.

Distinguishing the patients accurately is a long-term challenge in traditional PD studies, which attracts the attentions of many researchers. In the classical single classifier methods, Zhang et al. (2016) tried to apply SVM with linear kernel to speech detection of PD patients, and the results showed that the recognition accuracy was only 68.5%. Faris et al. (2016) extracted features from speech data, and combined multi-source optimization algorithm with BP feedforward neural network to achieve 77.31% recognition accuracy for PD patients. Adeli et al. (2016) improved the accuracy to 73.6% through joint feature sample selection (JFSC) and robust linear discriminant analysis (RLDA) based on fMRI data. On the other hand, the conventional ensemble learning model is also applied in the classification of PD. According to the characteristics of resting fMRI data, Galdi et al. (2018) proposed a method of combining consensus-based feature extraction with random forest, and achieved nearly 80% recognition accuracy. Drotr et al. (2016) also tested a variety of ensemble learning methods in the experiment. Among them, the ensemble AdaBoost classifier achieved 78.9% accuracy, 79.2% specificity and 82.4% sensitivity. With the rise of deep learning methods, Um et al. (2017) used the popular convolutional neural metwork (CNN) method to analyze the sensor data of wearable devices. The classification accuracy is between 77% and 80%, which indicates the potential of CNN method in the detection of PD. Prince et al. (2019) suggested that CNN-ensemble can be used to analyze multimodal data of PD patients including gait, voice and memory data, and the classification accuracy of CNN-ensemble was stabilized at 82%.

We found that the accuracy of CERNNE model was better than that of most above methods, and even not inferior to the current popular deep learning technology. Our classification performance advantages are mainly due to the following reasons. Compared with single classifier method, multi-learner ensemble can effectively avoid performance fluctuations and increase the robustness of classification performance (Chen et al., 2019). Different from the traditional random integration of multiple base classifiers, the introduction of clustering evolution eliminates redundant or invalid features and improves the diversity of different base classifiers in the ensemble model, which ensures the performance of the ensemble learning model proposed in this paper. Finally, the information complementary advantage of multimodal data and appropriate parameter optimization can also help to promote the classification performance.

4.3. Limitations and future directions

Although our method has achieved satisfactory results in sample classification and lesion detection, it must be noted that the method still has some limitations. Even though mixed use of data can contribute to model generalization, considering the limitations of the experimental data in quantity and acquisition methods, we will obtain more homologous data in the future to conduct more in-depth research and verification on the disease-associated factors detected in this paper. This study mainly focuses on the fusion analysis of fMRI and gene data by sequence coding correlation. However, the proposed method also has considerable potential for the fusion of other multimodal data (Chen et al., 2018). In future studies, we will try to use more complex and effective coding correlation methods to encode different sequences, and incorporate other modal data such as protein data and gait data into the fusion research of PD (Wang et al., 2018; Wei et al., 2019). Finally, for a few atypical pathogenic factors we found, due to the lack of relevant existing research, we will collect more data, design new algorithms, and further cooperate with clinicians to explain its role and rationality in PD.

5. Conclusions

This paper attempts to design a practical multimodal fusion analysis framework. The main contributions include three parts. The interactions between genes and brain regions are detected by correlation analysis, and the fusion features with more recognition ability are constructed according to the advantages of multimodal information complementarity. In this paper, a new ensemble learner is proposed to analyze the fusion features, and the recognition accuracy of 88.57% in patients with PD is achieved. The pathogenic factors of PD at the level of gene and brain function image are detected by searching for features that are meaningful for classification. These works provide a new perspective for the diagnosis and lesions analysis of PD.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Xia-an Bi: Conceptualization, Methodology, Supervision. **Xi Hu:** Data curation, Writing - original draft. **Yiming Xie:** Software, Validation. **Hao Wu:** Writing - review & editing.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 62072173), the Natural Science Foundation of Hunan Province, China (Grant No.2020JJ4432), the Degree and Postgraduate Education Reform Project of Hunan Province (Grant No.2019JGYB091), the Hunan Provincial Science and Technology Project Foundation (Grant No. 2018TP1018).

References

Adeli, E., Shi, F., An, L., Wee, C.-Y., Wu, G., Wang, T., Shen, D., 2016. Joint feature-sample selection and robust diagnosis of Parkinson's disease from MRI data. Neuroimage 141, 206–219.

Alexis, E., 2012. Association between parkinson's disease and the HLA-DRB1 locus. Movement Disorders 27 (9), 1104–1110.

Benka Walln, M., Franzn, E., Nero, H., Hagstrmer, M., 2015. Levels and patterns of physical activity and sedentary behavior in elderly people with mild to moderate parkinson disease. Phys Ther 95 (8), 1135–1141.

Bi, X.-a., Liu, Y., Jiang, Q., Shu, Q., Sun, Q., Dai, J., 2018. The diagnosis of autism spectrum disorder based on the random neural network cluster. Front Hum Neurosci 12 (1), 257–266.

Bohnen, N.I., Albin, R.L., Mller, M.L.T.M., Petrou, M., Kotagal, V., Koeppe, R.A., Scott, P.J.H., Frey, K.A., 2015. Frequency of cholinergic and caudate nucleus dopaminergic deficits across the predemented cognitive spectrum of parkinson disease and evidence of interaction effectsdopaminergic deficits in parkinson diseasedopaminergic deficits in parkinson disease. JAMA Neurol 72 (2), 194–200.

Borchert, R.J., Rittman, T., Passamonti, L., Ye, Z., Sami, S., Jones, S.P., Nombela, C., Vzquez Rodr-guez, P., Vatansever, D., Rae, C.L., Hughes, L.E., Robbins, T.W., Rowe, J.B., 2016. Atomoxetine enhances connectivity of prefrontal networks in Parkinsons disease. Neuropsychopharmacology 41, 2171.

Bot, M., van den Munckhof, P., Schmand, B.A., de Bie, R.M.A., Schuurman, P.R., 2018. Electrode penetration of the caudate nucleus in deep brain stimulation surgery for Parkinson's disease. Stereotact Funct Neurosurg 96 (4), 223–230.

Bregman, N., Thaler, A., Mirelman, A., Helmich, R.C., Gurevich, T., Orr-Urtreger, A., Marder, K., Bressman, S., Bloem, B.R., Giladi, N., 2017. A cognitive fMRI study in non-manifesting LRRK2 and GBA carriers. Brain Structure and Function 222 (3), 1207–1218.

Chang, D., Nalls, M.A., Hallgr-msdttir, I.B., Hunkapiller, J., van der Brug, M., Cai, F., International Parkinson's Disease Genomics, C., andMe Research, T., Kerchner, G.A., Ayalon, G., Bingol, B., Sheng, M., Hinds, D., Behrens, T.W., Singleton, A.B., Bhangale, T.R., Graham, R.R., 2017. A meta-analysis of genome-wide association studies identifies 17 new parkinson's disease risk loci. Nat. Genet. 49, 1511–1516.

Chen, X., Wang, L., Qu, J., Guan, N.-N., Li, J.-Q., 2018. Predicting mirna"disease association based on inductive matrix completion. Bioinformatics 34 (24), 4256–4265.

Chen, X., Zhu, C.-C., Yin, J., 2019. Ensemble of decision tree reveals potential mirna-disease associations. PLoS Comput. Biol. 15 (7), e1007209.

Consortium, I.P.D.G., 2011. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. The Lancet 377 (9766), 641–649.

Dirkx, M.F., den Ouden, H.E.M., Aarts, E., Timmer, M.H.M., Bloem, B.R., Toni, I., Helmich, R.C., 2017. Dopamine controls Parkinsons tremor by inhibiting the cerebellar thalamus. Brain 140 (3), 721–734.

Drotr, P., Mekyska, J., Rektorov, I., Masarov, L., Smkal, Z., Faundez-Zanuy, M., 2016. Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. Artif Intell Med 67, 39–46.

Du, L., Liu, K., Yao, X., Risacher, S.L., Han, J., Saykin, A.J., Guo, L., Shen, L., 2020. Detecting genetic associations with brain imaging phenotypes in alzheimers disease via a novel structured SCCA approach. Med Image Anal 61, 101656.

Faris, H., Aljarah, I., Mirjalili, S., 2016. Training feedforward neural networks using multi-verse optimizer for binary classification problems. Applied Intelligence 45 (2), 322–332.

Galdi, P., Fratello, M., Trojsi, F., Russo, A., Tedeschi, G., Tagliaferri, R., Esposito, F., 2018. Consensus-based feature extraction in rs-fmri data analysis. Soft comput 22 (11), 3785–3795.

Garcia-Esparcia, P., Koneti, A., Rodr-guez-Oroz, M.C., Gago, B., del Rio, J.A., Ferrer, I., 2018. Mitochondrial activity in the frontal cortex area 8 and angular gyrus in Parkinson's disease and Parkinson's disease with dementia. Brain Pathology 28 (1), 43–57.

Garg, A., Appel-Cresswell, S., Popuri, K., McKeown, M.J., Beg, M.F., 2015. Morphological alterations in the caudate, putamen, pallidum, and thalamus in parkinson's disease. Front Neurosci 9 (1), 101–111.

Gilat, M., Ehgoetz Martens, K.A., Miranda-Dom-nguez, O., Arpan, I., Shine, J.M., Mancini, M., Fair, D.A., Lewis, S.J.G., Horak, F.B., 2018. Dysfunctional limbic circuitry underlying freezing of gait in parkinson's disease. Neuroscience 374, 119–132.

Griffanti, L., Rolinski, M., Szewczyk-Krolikowski, K., Menke, R.A., Filippini, N., Zamboni, G., Jenkinson, M., Hu, M.T.M., Mackay, C.E., 2016. Challenges in the reproducibility of clinical studies with resting state FMRI: an example in early parkinson's disease. Neuroimage 124 (Prat A), 704–713.

Gupta, D., Julka, A., Jain, S., Aggarwal, T., Khanna, A., Arunkumar, N., de Albuquerque, V.H.C., 2018. Optimized cuttlefish algorithm for diagnosis of parkinsons disease. Cogn Syst Res 52, 36–48.

Gusev, F.E., Reshetov, D.A., Mitchell, A.C., Andreeva, T.V., Dincer, A., Grigorenko, A.P., Fedonin, G., Halene, T., Aliseychik, M., Filippova, E., Weng, Z., Akbarian, S., Rogaev, E.I., 2019. Chromatin profiling of cortical neurons identifies individual epigenetic signatures in schizophrenia. Transl Psychiatry 9 (1), 256.

Hao, X., Bao, Y., Guo, Y., Yu, M., Zhang, D., Risacher, S.L., Saykin, A.J., Yao, X., Shen, L., 2020. Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease. Med Image Anal 60, 101625.

Hao, X., Li, C., Yan, J., Yao, X., Risacher, S.L., Saykin, A.J., Shen, L., Zhang, D., for the Alzheimers Disease Neuroimaging, I., 2017. Identification of associations between genotypes and longitudinal phenotypes via temporally-constrained group sparse canonical correlation analysis. Bioinformatics 33 (14), i341–i349.

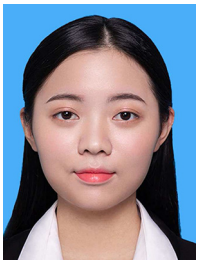
Jamshidi, J., Movafagh, A., Emamalizadeh, B., Bidoki, A.Z., Darvish, H., 2014. Hla–dra is associated with parkinson's disease in iranian population. Int. J. Immunogenet. 41 (6), 508–511.

Lei, D., Kefei, L., Lei, Z., Xiaohui, Y., L, R.S., Lei, G., J, S.A., Li, S., Initiative, A.D.N., 2019. Identifying progressive imaging genetic patterns via multi-task sparse canonical correlation analysis: a longitudinal study of the adni cohort. Bioinformatics 14 (14), i474–i483.

Li, C., Huang, B., Zhang, R., Ma, Q., Yang, W., Wang, L., Wang, L., Xu, Q., Feng, J., Liu, L., Zhang, Y., Huang, R., 2017. Impaired topological architecture of brain

structural networks in idiopathic Parkinsons disease: a DTI study. Brain Imaging Behav 11 (1), 113–128.
Liu, B., Zhang, X., Cui, Y., Qin, W., Tao, Y., Li, J., Yu, C., Jiang, T., 2016. Polygenic risk for schizophrenia influences cortical gyrification in 2 independent general populations. Schizophr Bull 43 (3), 673–680.
Liu, J.Z., Erlich, Y., Pickrell, J.K., 2017. Casecontrol association mapping by proxy using family history of disease. Nat. Genet. 49, 325.
Liu, S., Li, A., Liu, Y., Li, J., Wang, M., Sun, Y., Qin, W., Yu, C., Jiang, T., Liu, B., 2019. MIR137 polygenic risk is associated with schizophrenia and affects functional connectivity of the dorsolateral prefrontal cortex. Psychol Med 1–9.
Lv, D.-j., Li, L.-x., Chen, J., Wei, S.-Z., Wang, F., Hu, H., Xie, A.-M., Liu, C.-F., 2019. Sleep deprivation caused a memory defects and emotional changes in a rotenone-based zebrafish model of Parkinsons disease. Behav. Brain Res. 372, 112031.
van Mierlo, T.J., Chung, C., Foncke, E.M., Berendse, H.W., van den Heuvel, O.A., 2015. Depressive symptoms in Parkinson's disease are related to decreased hippocampus and amygdala volume. Movement Disorders 30 (2), 245–252.
Nettersheim, F.S., Loehrer, P.A., Weber, I., Jung, F., Dembek, T.A., Pelzer, E.A., Dafsari, H.S., Huber, C.A., Tittgemeyer, M., Timmermann, L., 2019. Dopamine substitution alters effective connectivity of cortical prefrontal, premotor, and motor regions during complex bimanual finger movements in parkinson's disease. Neuroimage 190, 118–132.
Patel, M., 2017. Csmd1 gene mutations can lead to familial parkinson disease. Nature Reviews Neurology 13 (1), 641–652.
Perrone-Bizzozero, N., 2019. Neuropsychiatric implications of rna-binding proteins hud and ksrp revealed by genome-wide identification of their targets. Eur. Neuropsychopharmacol. 29 (Supplement 3), S721.
Politis, M., 2018. Excessive daytime sleepiness may be associated with caudate denervation in parkinson disease. J. Neurol. Sci. 387, 220–227.
Prince, J., Andreotti, F., Vos, M.D., 2019. Multi-source ensemble learning for the remote prediction of parkinson's disease in the presence of source-wise missing data. IEEE Trans. Biomed. Eng. 66 (5), 1402–1411.
Purcell, S., 2012. PLINK (1.07). Documentation, pp. 32–33.
Ruiz-Mart-nez, J., Azcona, L.J., Bergareche, A., Mart-Mass, J.F., Paisn-Ruiz, C., 2017. Whole-exome sequencing associates novel csmd1 gene mutations with familial parkinson disease. Neurology. Genetics 3 (5), e177.
Saeed, M., 2018. Genomic convergence of locus-based gwas meta-analysis identifies axin1 as a novel parkinsons gene. Immunogenetics 70 (9), 563–570.
Shahmohammadibeni, N., Rahimi-Aliabadi, S., Jamshidi, J., Emamalizadeh, B., Shahmohammadibeni, H.A., Zare Bidoki, A., Akhavan-Niaki, H., Eftekhari, H., Abdollahi, S., Shekari Khaniani, M., Shahmohammadibeni, M., Fazeli, A., Motallebi, M., Taghavi, S., Ahmadifard, A., Shafiei Zarneh, A.E., Andarva, M., Dadkhah, T., Khademi, E., Alehabib, E., Rahimi, M., Tafakhori, A., Atakhorrami, M., Darvish, H., 2016. The analysis of association between snca, huseyo and csmd1 gene variants and parkinsons disease in iranian population. Neurological Sciences 37 (5), 731–736.
Sveinbjornsdottir, S., 2016. The clinical symptoms of parkinson's disease. J. Neurochem. 139 (S1), 318–324.
Tang, J., Yang, B., Adams, M.P., Shenkov, N.N., Klyuzhin, I.S., Fotouhi, S., Davoodi-Bojd, E., Lu, L., Soltanian-Zadeh, H., Sossi, V., Rahmim, A., 2019. Artificial neural network based prediction of outcome in parkinsons disease patients using datscan spect imaging features. Molecular Imaging and Biology 21 (1), 1–9.
Thobois, S., Prange, S., Sgambato-Faure, V., Tremblay, L., Broussolle, E., 2017. Imaging the etiology of apathy, anxiety, and depression in parkinsons disease: implication for treatment. Curr Neurol Neurosci Rep 17 (10), 76.
Trojsi, F., Di Nardo, F., Santangelo, G., Siciliano, M., Femiano, C., Passaniti, C., Caiazzo, G., Fratello, M., Cirillo, M., Monsurr, M.R., Esposito, F., Tedeschi, G., 2017. Resting state fMRI correlates of theory of mind impairment in amyotrophic lateral sclerosis. Cortex 97, 1–16.
Um, T.T., Pfister, F.M.J., Pichler, D., Endo, S., Lang, M., Hirche, S., Fietzek, U., Kuli, D., 2017. Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks. Proceedings of the 19th ACM International Conference on Multimodal Interaction 216–220.
Videnovic, A., 2017. Management of sleep disorders in Parkinson's disease and multiple system atrophy. Movement Disorders 32 (5), 659–668.
Wang, M., Hao, X., Huang, J., Shao, W., Zhang, D., 2018. Discovering network phenotype between genetic risk factors and disease status via diagnosis-aligned multi-modality regression method in alzheimers disease. Bioinformatics 35 (11), 1948–1957.
Wei, L., Su, R., Luan, S., Liao, Z., Manavalan, B., Zou, Q., Shi, X., 2019. Iterative feature representations improve n4-methylcytosine site prediction. Bioinformatics 35 (23), 4930–4937.
Zeng, W., Liu, F., Wang, Q., Wang, Y., Ma, L., Zhang, Y., 2016. Parkinson'S disease classification using gait analysis via deterministic learning. Neurosci. Lett. 633, 268–278.
Zhang, H.-H., Yang, L., Liu, Y., Wang, P., Yin, J., Li, Y., Qiu, M., Zhu, X., Yan, F., 2016. Classification of parkinsons disease utilizing multi-edit nearest-neighbor and ensemble learning algorithms with speech samples. Biomed Eng Online 15 (1), 122–130.

Xia-an Bi received the Ph.D. degree in computer applications from Hunan University in 2012. He is currently an associate professor in the Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, and College of Information Science and Engineering in Hunan Normal University. His current research interests include machine learning, brain science, artificial intelligence. He is the author or coauthor of several technical papers including several ESI Highly Cited Papers according to the most recent Clarivate Analytics ESI report and also a very active reviewer for many international journals and conferences. He is a member of IEEE, MICCAI and CCF.



Xi Hu received the B.E. degree in Information Engineering from Hunan Institute of Science and Technology, Yueyang, China, in 2019. She is currently pursuing the M.S. degree in College of Information Science and Engineering, Hunan Normal University, Changsha, China. The major of her is computer science and technology. Her main research fields include data mining, brain science and artificial intelligence.

Yiming Xie received the B.E. degree in Network Engineering from Chengdu Technological University, Chengdu, China, in 2019. He is currently pursuing the M.S. degree in College of Information Science and Engineering, Hunan Normal University, Changsha, China. The major of him is computer science and technology. His main research fields include data mining, brain science and artificial intelligence.

Hao Wu received the B.E. degree in Computer Science and Technology from Jinggangshan University, Ji'n, China, in 2019. He is currently pursuing the M.S. degree in College of Information Science and Engineering, Hunan Normal University, Changsha, China. The major of him is computer technology. His main research fields include data mining, brain science and artificial intelligence.