



# Diagnosis of Amnesic Mild Cognitive Impairment Using MGS-WBC and VGBN-LM Algorithms

Chunting Cai<sup>1</sup>, Jiangsheng Cao<sup>1</sup>, Chenhui Yang<sup>1\*</sup> and E. Chen<sup>2\*</sup>

<sup>1</sup> School of Informatics, Xiamen University, Xiamen, China, <sup>2</sup> Department of Neurology, Zhongshan Hospital Affiliated to Xiamen University, Xiamen, China

## OPEN ACCESS

### Edited by:

Yuanpeng Zhang,  
Nantong University, China

### Reviewed by:

Zhifan Gao,  
Sun Yat-sen University, China  
Cheng Wang,  
Huaqiao University, China

### \*Correspondence:

Chenhui Yang  
ych987@126.com  
E. Chen  
chenexm@163.com

### Specialty section:

This article was submitted to  
Alzheimer's Disease and Related  
Dementias,  
a section of the journal  
Frontiers in Aging Neuroscience

**Received:** 10 March 2022

**Accepted:** 20 April 2022

**Published:** 30 May 2022

### Citation:

Cai C, Cao J, Yang C and Chen E  
(2022) Diagnosis of Amnesic Mild  
Cognitive Impairment Using  
MGS-WBC and VGBN-LM  
Algorithms.  
Front. Aging Neurosci. 14:893250.  
doi: 10.3389/fnagi.2022.893250

Computer-aided diagnosis (CAD) has undergone rapid development with the advent of advanced neuroimaging and machine learning methods. Nevertheless, how to extract discriminative features from the limited and high-dimensional data is not ideal, especially for amnesic mild cognitive impairment (aMCI) data based on resting-state functional magnetic resonance imaging (rs-fMRI). Furthermore, a robust and reliable system for aMCI detection is conducive to timely detecting and screening subjects at a high risk of Alzheimer's disease (AD). In this scenario, we first develop the mask generation strategy based on within-class and between-class criterion (MGS-WBC), which primarily aims at reducing data redundancy and excavating multiscale features of the brain. Concurrently, vector generation for brain networks based on Laplacian matrix (VGBN-LM) is presented to obtain the global features of the functional network. Finally, all multiscale features are fused to further improve the diagnostic performance of aMCI. Typical classifiers for small data learning, such as naive Bayesian (NB), linear discriminant analysis (LDA), logistic regression (LR), and support vector machines (SVMs), are adopted to evaluate the diagnostic performance of aMCI. This study helps to reveal discriminative neuroimaging features, and outperforms the state-of-the-art methods, providing new insights for the intelligent construction of CAD system of aMCI.

**Keywords:** machine learning, aMCI, MGS-WBC, multi-scale features, VGBN-LM

## INTRODUCTION

Alzheimer's disease (AD), which occurs frequently in elderly individuals, is a chronic and irreversible neurodegenerative disease accompanied by brain impairments in memory, communication, and reasoning (Dadar et al., 2017). The new report indicates that more than 5.7 million individuals have been diagnosed with AD in the United States (Association, 2018). Unfortunately, as yet, there is no agreed medication or treatment protocol to cure and rehabilitate patients with AD (Xi et al., 2022). Amnesic mild cognitive impairment (aMCI), which is generally characterized by prominent deficits in memory, is widely considered as the early stage of AD due

to its greater risk of conversion to AD (Bian et al., 2014; Barone et al., 2016). Accordingly, a robust and reliable system for aMCI detection is conducive to timely detecting and screening the subjects with AD at high risk, thus providing an optimal treatment period for patients.

Resting-state functional magnetic resonance imaging (rs-fMRI), which reflects the neural functional activity of the brain by measuring spontaneous blood oxygen level dependency (BOLD) signal fluctuations in a non-invasive way, has been widely employed as one of the important techniques to diagnose brain-related diseases (Li et al., 2018; Li R. et al., 2020; Sundaram et al., 2020). However, rs-fMRI data are confronted with formidable challenges to the existing pattern classification methods due to its limited data samples and high dimensions, which is not favorable for aMCI detection (Kassani et al., 2020). In consequence, existing methods for observing alterations of spontaneous neural activity, such as regional homogeneity (ReHo), the amplitude of low-frequency fluctuation (ALFF), and functional connectivity analysis, are employed to explore the biological mechanisms of brain function and reduce the redundant information of brain (Harrison et al., 2019; Liao et al., 2019; Ting et al., 2020; Xiao et al., 2020). Furthermore, the aMCI can be diagnosed using rs-fMRI data processed by the use of abovementioned methods as reported in a few studies (Zhou et al., 2014; Kim et al., 2020; Li Y. et al., 2020). Consequently, the diagnosis of aMCI based on these data can be employed for timely prevention and screening of potential patients with AD.

In this study, the data sample after preprocessing is then conducted using ReHo calculation, ALFF calculation, and functional connectivity construction, and we obtain two three-dimensional whole-brain structures and the brain functional network. Then, we develop an MGS-WBC model, which mainly consists of the intraclass volatility detection method using variable coefficient and multiscale statistical thresholds, to excavate the multiscale features and reduce the redundancy information both in the whole-brain structures and functional network. It is remarkable that the traditional approaches only use the single local features of the functional network as the input of the classifier, while ignoring its global features. Therefore, we present a VGBN-LM model that extracts the global features and then concatenates local features obtained from the MGS-WBC model to further excavate features of the functional network. Then, we fuse the features that resulted from the MGS-WBC model and VGBN-LM model to further improve the diagnostic performance of aMCI. Concurrently, to evaluate the validity of obtained features and diagnostic performance of aMCI, typical machine learning classifiers applied to the limited number of data, such as naive Bayesian (NB), linear discriminant analysis (LDA), logistic regression (LR), and support vector machines (SVMs), are utilized. Noting that the leave-one-out cross-validation is employed to avert the overfitting problem in this study (Sangnawakij and Niwitpong, 2017). This study greatly improves the diagnostic performance of aMCI compared with the traditional methods, providing new insights for the construction of the computer-aided diagnosis (CAD) system for neurodegenerative diseases, especially for aMCI.

Herein, it is not ideal how to extract discriminative features from the limited and high-dimensional data, especially for aMCI data based on rs-fMRI. Furthermore, a robust and reliable system for aMCI detection is conducive to timely detecting and screening subjects at a high risk of AD. Hence, the novel aspect of this study is that we first develop an MGS-WBC model to extract discriminative features of whole-brain structures and functional networks using volatility detection and multiple significant thresholds. In the following section, we present a VGBN-LM model that excavates the global features of the functional network and fuses its local features, thus overcoming the low availability of the network feature. Finally, all features generated from the MGS-WBC model and VGBN-LM model are concatenated to further improve the diagnostic performance of aMCI.

The remainder of the study is structured as follows: We discuss the related studies on feature extraction and the diagnosis of aMCI in the “Related Work” Section. We present data source, data acquisition, and data preprocessing in the “Materials” Section. We describe the methodologies of aMCI’s diagnostic system in the “Methods” Section. We provide the experiment results and analysis in the subsequent section and give the conclusion in the final section.

## RELATED WORK

Advanced neuroimaging, machine learning techniques, and statistical algorithms provide an opportunity to understand how the brain works between healthy control (HC) and patients with brain diseases, and to analyze the pathological mechanism of the brain (Nadarajah and Kotz, 2006; Lei et al., 2019). Additionally, rs-fMRI, which is convenient and suitable for patients with cognitive impairment who feel difficult to complete tasks, is widely employed in the study of diseases related to the central nervous system, which can reflect the functional condition of the brain under the default state. Nevertheless, how to extract discriminative features from the limited and high-dimensional data is not ideal, especially for the aMCI data based on rs-fMRI.

To address this issue, numerous previous studies have focused on excavating biomarkers associated with aMCI disease from the perspective of neuroimaging. For example, compared with the HC group, abnormal patterns of diverse-club and rich-club organizations in the functional network are revealed using the two-sample *t*-test in the aMCI group, indicating that the overlapping nodes might be potential biomarkers in the diagnosis of aMCI (Xue et al., 2020). Also, a previous study has revealed that abnormal alterations of ALFF in patients with aMCI are found involved in brain regions, such as the right hippocampus, parahippocampal cortex, and left lateral temporal cortex, which may act as biomarkers of disease (Xi et al., 2013). The former investigation has also confirmed that disruptive patterns of ReHo are found using the two-sample *t*-test in the brain, which is conducive to better comprehending the neural substrates of aMCI and can be served as biomarkers (Zhen et al., 2018). It should be noted that all the studies mentioned above are conducted using statistical methods and existing medical background knowledge to analyze the disruptive patterns of brain

disease, ignoring the adoption of machine learning techniques to assess the validity and reliability of diagnostic results.

Currently, machine learning techniques have been actively adopted by researchers to design the CAD system of aMCI, which mainly aims at improving the classification performance and putting it into clinical practice (Zhang et al., 2015; Yang et al., 2021). To the limited number of data and information redundancy problems, feature selection is first implemented and then acted as the input of machine learning classifiers, which are suitable for small data learning, such as NB, LDA, LR, and SVM. A previous investigation based on the functional network classification of aMCI has elucidated that the obvious brain regions which resulted from the two-sample t-test are served as the input of SVM classifier, resulting in the ACC of 69% (Lee et al., 2015). Moreover, it has been reported that the identification ACC of aMCI is 75.35% when using significant regions of the ALFF as the input of SVM classifier (Yang et al., 2018). Up to now, the existing literature not only lacks quantity in the concerned study but also focuses on the single significant threshold method, ignoring the extraction of multiscale features in the brain.

In this study, we first develop an MGS-WBC model using validation detection and multiple significant thresholds methods to extract multiscale features both in the whole-brain structures and functional network. Concurrently, we present the VGBN-LM model to extract the global features of the functional network and fuse its local features. Finally, we fuse the extracted features generated from the MGS-WBC model and VGBN-LM model to further improve the diagnostic performance of aMCI. Compared with the traditional method, the results elucidate that the proposed methods provide the accuracy (ACC) of 89.55, 91.04, 92.54, and 94.03% with the same input data in four typical classifiers, including NB, LDA, LR, and SVM, and maximum improvements are 1.49, 1.49, 1.50, and 4.48%, respectively. Moreover, the area under the curve (AUC) on the four sequence classifiers mentioned above are 95.63, 92.69, 97.59, and 97.33%, increasing by 2.49, 1.96, 2.67, and 2.23% compared with the traditional method. It can be inferred that our aMCI diagnostic system is more suitable and reliable than the traditional method. In addition, this study can help reveal impressive and discriminative neuroimaging features, providing new insights for the construction of the CAD system of aMCI.

## MATERIALS

### Description of the Data Source

In this study, the raw data samples are composed of two parts, including 39 patients with aMCI and 38 HCs. These raw data samples are obtained from the second stage of Alzheimer's Disease Neuroimaging Initiative (ADNI) database, which helps to explore neurological biomarkers of cognitive disease and assist in the diagnosis and treatment of the disease<sup>1</sup>.

It deserves to be further mentioned that not all data samples are conformed to work demands. One data sample has been precluded due to undue head movement (cumulative translation

or rotation > 2 mm or 2°) and poor quality of image registration (Wang et al., 2018). Concurrently, we exclude 6 patients with aMCI and 4 HCs for inconsistent voxel size or dimension of the brain. To sum up, we end up with 33 patients with aMCI and 34 HCs in this study (Wang et al., 2020).

### Data Acquisition and Preprocessing

All the data samples we focus on in this study are based on rs-fMRI obtained using a clinical 3.0-Tesla (T) scanner. Noting that each raw data sample based on rs-fMRI is composed of brains at 140 points in time. To maintain the reliability of the raw data samples, the subjects are instructed to lay flat, to close their eyes without thinking, and to keep their head in position throughout the imaging processing. We perform the scanner parameters as follows (Li W. et al., 2020): Echo time (TE) = 30 ms, repetition time (TR) = 3,000 ms, voxel size = 3.31 mm × 3.31 mm × 3.31 mm, flip angle (FA) = 80°. More detailed information about scanner parameters based on rs-fMRI data can be inquired on the ADNI's website.

We perform raw data sample preprocessing using Resting-State fMRI Data Analysis Toolkit plus (RESTplus)<sup>2</sup>, which is based on math software MATLAB2012a<sup>3</sup> and Statistical Parametric Mapping software (SPM12)<sup>4</sup>. The raw data preprocessing steps we adopt are depicted as follows: first, owing to the machine and human factors, the imaging signal acquired at the beginning of scanning may exhibit unstable signal ingredient factors. Thus, the first 5 time points of each data sample after imaging are discarded in this study. Also, we calibrate the brains of the remaining time points to eliminate the effects induced by discrepancies between various brains or slices. We further normalize the brains of all data samples using the echo-plane imaging (EPI) template from the original coordinate system to Montreal Neurological Institute (MNI) coordinate system due to the shape of each subject's brain being inconsistent. Specifically, we smooth out the noise that exists in the brain using a Gaussian kernel of 6-mm full width at half maximum (FWHM) for functional network and ALFF analysis (Yang et al., 2021). After that, we remove the variables that affect the dependent variables, including nuisance cerebrospinal fluid signal, global mean signal, white matter signal, and 6 head motion parameters. Finally, the signals with important physiology meaning are obtained through a bandpass filter of 0.01–0.08 Hz (Vicente et al., 2018). Based on the steps mentioned above, the data samples after preprocessing are utilized in the following section.

## METHODS

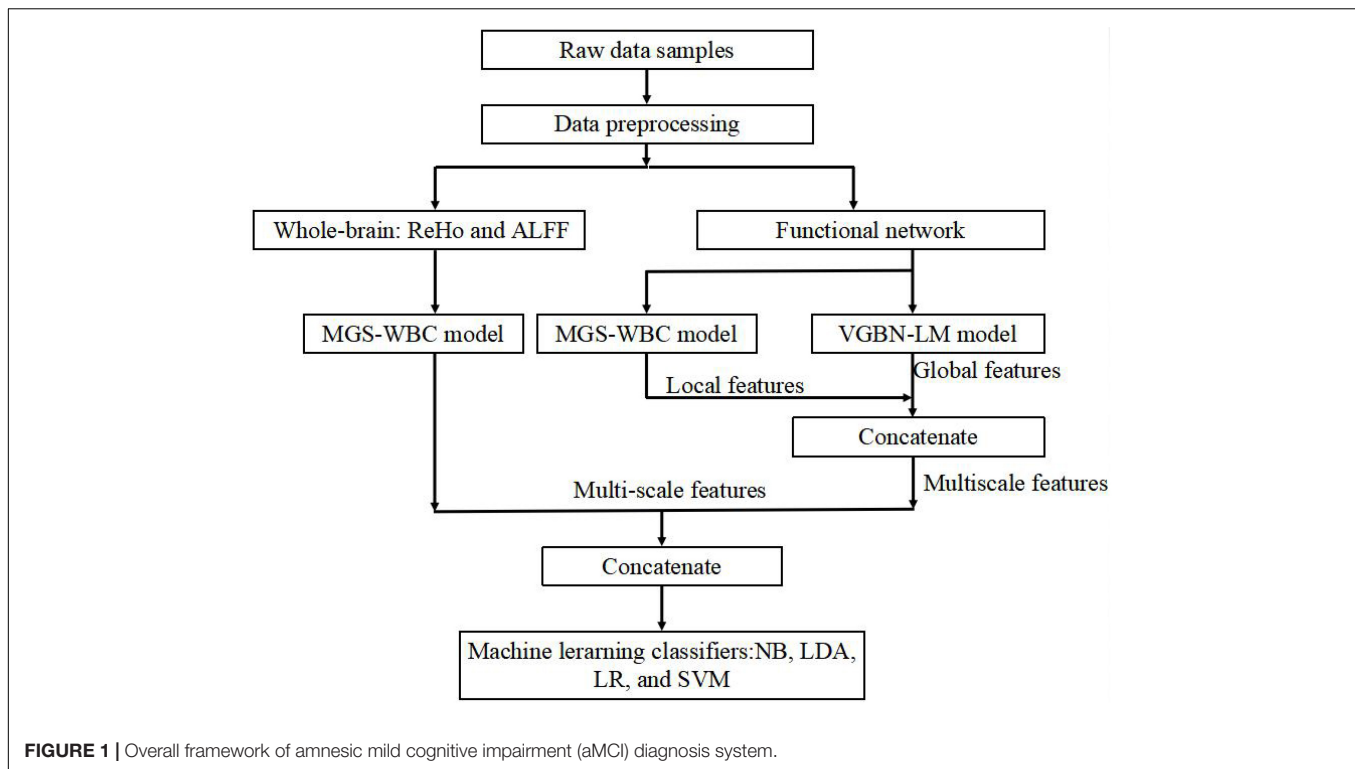
### Overview of the Amnesic Mild Cognitive Impairment Diagnosis System

In this study, we develop a diagnosis system for aMCI. At first, the raw data samples based on the rs-fMRI are pre-processed as depicted in the "Related Work" Section. In the next step,

<sup>2</sup><http://restfmri.net/forum/RESTplusV1.2>

<sup>3</sup><http://www.mathworks.com/products/matlab/>

<sup>4</sup><https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>



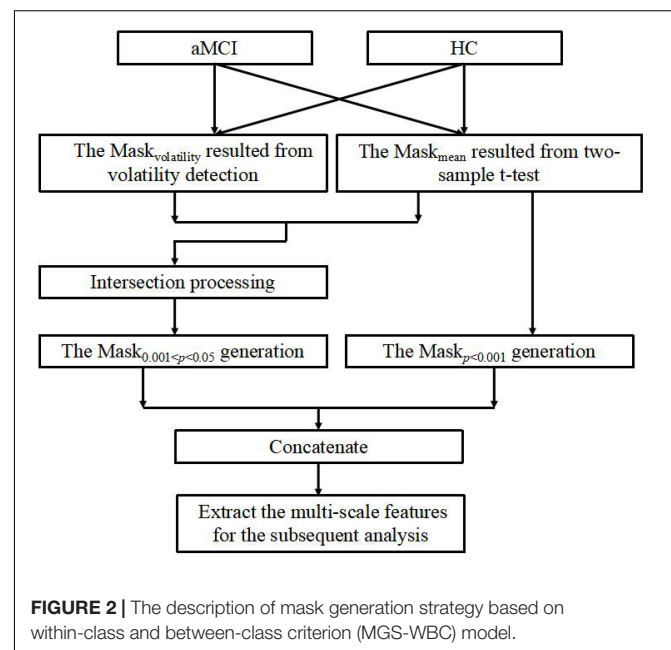
the data samples after preprocessing are calculated using typical approaches used to observe alterations of neuronal activity, such as ReHo, ALFF, and functional connectivity, and we obtain two whole-brain structures and one functional network for each data sample. Then, the calculated data samples are acted as the input of MGS-WBC model and VGBN-LM model we propose to extract biological features of whole-brain structures and functional networks. To evaluate the validity of obtained features, NB, LDA, LR, and SVM are utilized in this study. Furthermore, the leave-one-out cross-validation is carried out to avert the overfitting problem in this study.

In the results, the features generated by the MGS-WBC model and VGBN-LM model yield better performance than the traditional method in the aMCI detection process. The framework of our aMCI diagnosis system is illustrated in **Figure 1**.

### Mask Generation Strategy Based on Within-Class and Between-Class Criterion

In this section, we develop an MGS-WBC model to reduce the redundancy both in the whole-brain structures and functional networks and to further excavate the corresponding features of multiscale significant thresholds. Remarkably, the connectivity of the functional network is constructed using the Pearson correlation coefficient (PCC), and the inputs of PCC come from the time series of the corresponding brain regions. Also, to better extract the time series of the brain, we employ the general template of anatomical automatic labeling (AAL) to segment

the brain into 90 regions (Ju et al., 2019). The structure of the MGS-WBC model shown in **Figure 2** mainly consists of within-class volatility detection using variable coefficient and multi-scale significant thresholds. Since the mask size using a single significant threshold of  $0.001 < p < 0.05$  is much larger than  $p < 0.001$ , the traditional method only uses the mask with  $p < 0.001$  for feature extraction and ignores the effective



utilization of features. Besides, for the limited amount of data with a high dimension, how to better extract discriminative features effectively corresponding to the significant threshold of  $0.001 < p < 0.05$  is no clear. First, intraclass volatility detection mask of the whole-brain structures and  $90 \times 90$  functional network, which is defined as  $\text{Mask}_{\text{volatility}}$ , are generated using variable coefficient within the group to extract the smallest values of the first 5% (Xu et al., 2017), and the variable coefficient (VC) is defined as follows:

$$VC = \frac{\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}}{\bar{x}} \times 100\% \quad (1)$$

where  $\bar{x}$  denotes the average value of the selected pixel or selected brain region,  $n$  denotes the number of subjects within the group, and  $x_j$  denotes the value of the selected pixel or selected brain region corresponding to the  $i$ th subject. To obtain the features under multiple significant thresholds, we use  $\text{Mask}_{\text{volatility}}$  and a two-sample  $t$ -test to generate the mask of  $0.001 < p < 0.05$ , defined as  $\text{Mask}_{0.001 < p < 0.05}$ , which makes a large difference between the averages of the classes and the fluctuations within the classes are stable. Next, we further use the traditional method that extracts the features obtained from  $p < 0.001$  to generate a mask, defined as  $\text{Mask}_p < 0.001$ . And finally,  $\text{Mask}_{0.001 < p < 0.05}$  and  $\text{Mask}_p < 0.001$  are concatenated as the input mask of whole-brain structures or functional networks to extract the final features. It is worth mentioning that the whole-brain structures employed in this study are calculated using the ALFF method and ReHo method, respectively, which are commonly employed to analyze neural activity in the brain and reduce the dimensions of brain data.

## Vector Generation for Brain Networks Based on Laplacian Matrix

In this section, we present a VGBN-LM model for extracting the global features of the functional network. Laplacian Eigenmaps (LE) is a non-linear dimensionality reduction method, which constructs the relationship between samples from the local point of view, and can reflect the local relationship of samples to a certain extent. That is, the points related to each other are expected to be as close as possible to each other in the space after dimensionality reduction. Accordingly, we mainly use LE's conception to extract global features of the functional network. Here is a brief introduction to the LE algorithm (Belkin and Niyogi, 2001):

Step 1: The construction of the graph is defined as  $G(V, W)$ : Given  $n$  samples  $x_1, x_2, \dots, x_n$  in  $R^d$ , the edge connection  $w_{ij}$  between sample  $i$  and sample  $j$  is established using  $b$  nearest neighbors, and the  $w_{ij}$  can usually be constructed using Gaussian kernel function shown in (2). Notably,  $w_{ij}$  represents symmetric,  $D$  represents the node degree matrix shown in (3), and  $x_n$  represents the eigenvector corresponding to the sample  $n$ .

$$w_{ij} = e^{-\frac{\|x_i - x_j\|^2}{t}} \quad (2)$$

$$D_{ii} = \sum_j w_{ij} \quad (3)$$

Step 2: To maintain the identical geometric property as possible after reducing the dimension of the functional network, it can be solved as the minimization problem of  $y^T L y$ . That is, if the connection value  $w_{ij}$  is larger, the values of  $y_i$  and  $y_j$  will be closer, as shown in (4). Thus, this reduces now to (5).

$$\begin{aligned} y^T L y &= y^T (D - W) y \\ &= y^T D y - y^T W y \\ &= \sum_{i=1}^n d_i y_i^2 - \sum_{i,j=1}^n y_i y_j w_{ij} \\ &= \frac{1}{2} (\sum_{i=1}^n d_i y_i^2 - 2 \sum_{i,j=1}^n y_i y_j w_{ij} + \sum_{j=1}^n d_j y_j^2) \\ &= \frac{1}{2} \sum_{i,j=1}^n w_{ij} (y_i - y_j)^2 \geq 0 \end{aligned} \quad (4)$$

$$Y_{opt} = \arg \min_{Y^T D Y = I} \text{tr}(Y^T L Y) \quad (5)$$

Step 3: Finally, after using the Lagrange multiplier method, the optimal result can be obtained by (6).

$$L y = \lambda D y \quad (6)$$

However, the goal of the LE algorithm is to reduce the dimension of the sample features, and the dimension of the reduced features depends on the number of minimum non-zero eigenvalues, which does not conform to the requirements of extracting the global features from the functional network in this study. Moreover, it is unable to transform the functional network from two dimensions to one dimension that maintains the global features. Thus, considering the limited and high-dimensional data, the VGBN-LM model based on the Laplacian matrix is proposed to reduce the dimension of  $90 \times 90$  functional network and obtain the global features of functional network. Concurrently, we expect that the functional network after dimension reduction operation can maintain and reflect the identical geometric property of functional network. That is, we represent the global features of  $90 \times 90$  functional network in terms of reduced dimension vector.

First, we construct the minimization problem based on the Laplacian matrix. More specifically, let

$$\begin{aligned} \min \quad & y^T L y = \frac{1}{2} \sum_{i,j} w_{ij} (y_i - y_j)^2 \\ \text{s.t.} \quad & y^T D y = \sigma \end{aligned} \quad (7)$$

where  $\sigma$  represents the special real number and its value will be introduced later. We then use Lagrange to acquire optimal result, which can be formulated as follows:

$$G(y, \lambda) = y^T L y + \lambda (\sigma - y^T D y) \quad (8)$$

Then, we take the derivative of  $y$  and  $\lambda$  of  $G(y, \lambda)$ , respectively.

$$\begin{aligned} \frac{\partial G(y, \lambda)}{\partial y} &= \frac{\partial (y^T L y + \lambda (\sigma - y^T D y))}{y} \\ &= \frac{\partial (y^T (L - \lambda D) y + \lambda \sigma)}{y} \\ &= 0 \\ \Rightarrow L y &= \lambda D y \end{aligned} \quad (9)$$

$$\frac{\partial G(y, \lambda)}{\lambda} = \frac{\partial (y^T Ly + \lambda(\sigma - y^T Dy))}{\lambda} \Rightarrow y^T Dy = \sigma \tag{10}$$

Thus, the optimal result turns to determine the solution of (11).

$$\begin{cases} Ly = \lambda Dy \\ y^T Dy = \sigma \end{cases} \tag{11}$$

Nevertheless, the LE algorithm employs constraint conditions of  $y^T Dy = 1$  to obtain the eigenvectors corresponding to all non-zero eigenvalues, ignoring that all the optimal results may not satisfy the original constraint condition. Moreover, the generated eigenvectors of  $90 \times 90$  functional network cannot present its global features due to the excessive redundant features and limited amount of data.

Next, we deform  $Ly = \lambda Dy$  to obtain the optimal result, and its expression could be written as:

$$\begin{aligned} Ly &= \lambda Dy \\ \Leftrightarrow (D - W)y &= \lambda Dy \\ \Leftrightarrow Dy - \lambda Dy &= Wy \\ \Leftrightarrow (1 - \lambda)Dy &= Wy \\ \Leftrightarrow y^T (1 - \lambda)Dy &= y^T Wy \\ \Leftrightarrow (1 - \lambda)y^T Dy &= y^T Wy \end{aligned} \tag{12}$$

If  $\lambda \neq 1$ , we obtain

$$\begin{aligned} (1 - \lambda)y^T Dy &= y^T Wy \\ \Rightarrow y^T Dy &= \frac{y^T Wy}{1 - \lambda} \end{aligned} \tag{13}$$

Then, if  $\lambda = 1$ , it can be inferred as below:

$$\begin{aligned} (1 - \lambda)y^T Dy &= y^T Wy \\ \Leftrightarrow Ly &= \lambda Dy \\ \xrightarrow{\lambda=1} Ly &= Dy \\ \Leftrightarrow Wy &= 0 \end{aligned} \tag{14}$$

Therefore, the optimal result of (11) can be further deformed as:

$$\begin{cases} y^T Dy = \sigma \\ Ly = \lambda Dy \end{cases} \Rightarrow \begin{cases} \begin{cases} y^T Dy = \frac{y^T Wy}{1 - \lambda} \\ y^T Dy = \sigma \end{cases}, & \text{if } \lambda \neq 1 \\ \begin{cases} Dy = Ly \Leftrightarrow Wy = 0 \\ y^T Dy = \sigma \end{cases} & \text{if } \lambda = 1 \end{cases} \tag{15}$$

Significantly, the solution of (15) is reached by use of the backward induction. Let  $(\lambda_{\xi_i}, y_{\xi_i})$  represents the solutions of  $\begin{cases} y^T Dy = \frac{y^T Wy}{1 - \lambda} \\ y^T Dy = \sigma \end{cases} \Leftrightarrow Ly = \lambda Dy$  if  $\lambda \neq 1$ , then we can deduce that  $\sigma_{\xi_i} = y_{\xi_i}^T Dy_{\xi_i}$ . Based on this, we can obtain the following

form:

$$\begin{aligned} \arg \min_{y^T Dy = \sigma} y^T Ly &\xrightarrow{L\lambda_{\xi_i} = \lambda_{\xi_i} Dy_{\xi_i}, y = y_{\xi_i}} \min_{\lambda_{\xi_i}, y_{\xi_i}} \lambda_{\xi_i} y_{\xi_i}^T Dy_{\xi_i} \\ &\xrightarrow{y_{\xi_i}^T Dy_{\xi_i} = \frac{y_{\xi_i}^T Wy_{\xi_i}}{1 - \lambda_{\xi_i}}} \min_{\lambda_{\xi_i}} \lambda_{\xi_i} \frac{y_{\xi_i}^T Wy_{\xi_i}}{1 - \lambda_{\xi_i}} \\ &= \min(\lambda_{\xi_1} \frac{y_{\xi_1}^T Wy_{\xi_1}}{1 - \lambda_{\xi_1}}, \lambda_{\xi_2} \frac{y_{\xi_2}^T Wy_{\xi_2}}{1 - \lambda_{\xi_2}}, \dots, \lambda_{\xi_i} \frac{y_{\xi_i}^T Wy_{\xi_i}}{1 - \lambda_{\xi_i}}) \\ &\quad , i \in (1, 2, \dots, d) \end{aligned} \tag{16}$$

where  $d$  represents the number of solutions in  $Ly = \lambda Dy$ . Concurrently, we assume that  $Dy = Ly$  have the solutions, defined as  $y_{\beta_j}$ , and then we make  $\sigma_{\beta_j} = y_{\beta_j}^T Dy_{\beta_j}$ . From this, it can

be deduced that the equation  $\begin{cases} Dy = Ly \Leftrightarrow Wy = 0 \\ y^T Dy = \sigma \end{cases}$  if  $\lambda = 1$  have the solution. Therefore, we can deduce the expression as follows:

$$\begin{aligned} \arg \min_{y^T Dy = \sigma} y^T Ly &\xrightarrow{y = y_{\beta_j}} \min_{y_{\beta_j}} y_{\beta_j}^T Ly_{\beta_j} \\ &= \min(y_{\beta_1}^T Ly_{\beta_1}, y_{\beta_2}^T Ly_{\beta_2}, \dots, y_{\beta_j}^T Ly_{\beta_j}), j \in (1, 2, \dots, h) \end{aligned} \tag{17}$$

where  $h$  represents the number of solutions in  $Dy = Ly$ . To sum up, the optimal result of (7) can be written as follows:

$$\arg \min_{y^T Dy = \sigma} y^T Ly = \min(\min_{\lambda_{\xi_i}} \lambda_{\xi_i} \frac{y_{\xi_i}^T Wy_{\xi_i}}{1 - \lambda_{\xi_i}}, \min_{y_{\beta_j}} y_{\beta_j}^T Dy_{\beta_j}) \tag{18}$$

$, i \in (1, 2, \dots, d)$  and  $j \in (1, 2, \dots, h)$

Remarkably, the value of  $\sigma$  is obtained from the optimal solution in (18). Here, the procedure of VGBN-LM model can be summarized as follows:

Step 1: Given  $n$  brain regions extracted using AAL template, we use the absolute value of PCC shown in (19) to construct the edge connection  $w_{ij}$  between brain region  $i$  and brain region  $j$  instead of Gaussian kernel function and  $b$  nearest neighbors method.

$$w_{ij} = \left| \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \right| \tag{19}$$

where  $x_i$  and  $y_i$  denote the two signals from two different brain regions,  $\bar{x}$  denotes the average value corresponding to  $x_i$ ,  $\bar{y}$  denotes the average value corresponding to  $y_i$ , and  $n$  denotes the total number of brain regions in AAL template.

Step 2: We present (7–19) to reduce the dimension of the matrix  $w_{ij}$  from two dimensions to one dimension. Meanwhile, we maintain and reflect the global structure information in the lower dimension space to a certain degree. That is, the closer the connections between brain regions before dimension reduction, the nearer the corresponding values after dimension reduction. In this way, our generated vector  $y_{opt}$  resulted from (18) contains global features of functional network.

Step 3: Due to the limited amount of data, the two-sample  $t$ -test method is employed to select the global features with obvious differences ( $p < 0.05$ ) in the generated vector  $y_{opt}$ , and then, we concatenate the local features obtained from MGS-WBC model for the following analysis.

## Multiscale Feature Fusion

We concatenate all the generated features obtained from MGS-WBC model and VGBN-LM model to further improve the diagnostic performance of aMCI. That is, if the dimensions of two input features  $x_{input}$  and  $y_{input}$  are  $p$  and  $q$ , then the dimension of the output feature  $z$  is equal to  $p + q$ .

## Classification Using Naive Bayesian, Linear Discriminant Analysis, Logistic Regression, and Support Vector Machine

As can be seen from the above data, we have obtained multiscale features of whole-brain structures and functional networks using the MGS-WBC model and VGBN-LM mode. Considering the limited amount of data, the selection of an appropriate classifier adopted in this study is crucial to evaluate the validation of obtained features and improve the diagnostic performance of aMCI. Fortunately, the SVM classifier and the LR classifier exhibit effective diagnostic performance in terms of mild cognitive impairment (MCI), and it has been exploited by the majority of researchers (Ciulli et al., 2016; Khatun et al., 2019). In addition, typical classifiers applied to small data learning, such as NB classifier and LDA classifier, are also employed to better assess the generalization ability of features. The following is a brief introduction to the classifiers mentioned above.

The NB classifier mainly uses the probability of known data to determine the classification of the unknown data (Liu et al., 2016):

(1) Each data sample  $x$  is composed of  $m$ -dimensional features, denoted by  $a_1, a_2, a_3, \dots, a_m$ .

(2) We calculate the conditional probability for each category under each data sample. The category set is denoted as  $C = \{l_1, l_2, \dots, l_k\}$ , hence, we calculate  $p(l_1|x), p(l_2|x), \dots, p(l_k|x)$  separately.

$$p(l_k|x) = \frac{p(l_k)p(x|l_k)}{p(x)} \quad (20)$$

where  $p(l_k|x)$  can be deduced by (20) and  $k$  is the class number of  $x$ .

(3) If  $p(l_f|x) = \max(p(l_1|x), p(l_2|x), \dots, p(l_k|x))$ , then  $l_f$  is the category corresponding to  $x$ , where  $f \in \{1, 2, \dots, k\}$ .

The main idea of the LDA classifier is to ensure that the intraclass variance of each class is small, and the mean difference between classes is large in the space after the projection (Ji and Ye, 2008). Given input data set  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ ,  $y \in \{0, 1\}$ , meanwhile,  $N_j$  ( $j = 0, 1$ ) denotes the number of data samples corresponding to the category  $j$ ,  $x$  denotes the input and  $X_j$  ( $j = 0, 1$ ) denotes the set of data samples corresponding to the category  $j$ . The mean vector of the data sample corresponding to class  $j$  can be expressed as follows:

$$u_j = \frac{1}{N_j} \sum_{x \in X_j} x, \quad j = 0, 1 \quad (21)$$

The covariance matrix of the data sample corresponding to class  $j$  is denoted as  $\sum_j$ .

$$\sum_j = \sum_{x \in X_j} (x - u_j)(x - u_j)^T, \quad j = 0, 1 \quad (22)$$

Then, the divergence matrix within class is defined as  $S_w$ .

$$S_w = \sum_{x \in X_0} (x - u_0)(x - u_0)^T + \sum_{x \in X_1} (x - u_1)(x - u_1)^T \quad (23)$$

where  $u$  denotes mean value.

Simultaneously, the divergence matrix between classes can be described as  $S_b$ .

$$S_b = (u_0 - u_1)(u_0 - u_1)^T \quad (24)$$

In this way, the optimization objective is rewritten as follows:

$$\arg \max J(w) = \frac{w^T S_b w}{w^T S_w w} \quad (25)$$

According to (21–25), it can be inferred as follows:

$$w = S_w^{-1} (u_0 - u_1) \quad (26)$$

That is, the optimal projection direction  $w$  can be determined by calculating the mean and variance of the original two types of samples, and then, the classification is conducted on this basis.

The LR classifier is a probabilistic statistical classification model, which uses a probability score as the predicted value of the dependent variable to evaluate the mutual relation between the dependent variable and the independent variable. Specifically, we utilize the L2 regularization considering the overfitting problem in this study (Chen et al., 2011). The loss function with L2 regularization can be described as follows:

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m [-y^{(i)} \log(h_\theta(x^{(i)}))] - \frac{1}{m} \sum_{i=1}^m [(1 - y^{(i)}) \log(1 - h_\theta(x^{(i)}))] + \frac{\eta}{2m} \sum_{j=1}^n \theta_j^2 \quad (27)$$

where  $h_\theta(x)$  is defined as:

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}} \quad (28)$$

Then, we take the derivative concerning (27).

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^m (h_\theta(x_i) - y_i) x_{ij} + \frac{\eta}{m} \theta_j \quad (29)$$

where  $m$  represents the number of data samples, and  $\eta$  represents the regularization coefficient. Finally, the parameters  $\theta_j$  corresponding to the minimum loss function are obtained through continuous iteration.

The SVM classifier is to obtain the optimal separating hyperplane in the feature space to maximize the interval between positive and negative data samples on the training set. It is also worth noting that the kernel function adopted in this study is the Gaussian Radial Basis Function kernel (Zhou et al., 2014). The

SVM classifier mainly works around the following optimization problems (Cristovao et al., 2022).

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \psi_i \\ \text{s.t.} \quad & y_i [w^T x_i + b] \geq 1 - \psi_i, \quad \psi_i \geq 0 \end{aligned} \quad (30)$$

The Gaussian kernel is described as follows:

$$K(x, x_i) = \exp\left(-\frac{\|x-x_i\|^2}{\sigma^2}\right) \quad (31)$$

where  $C$  represents the penalty coefficient,  $x_i$  represents the data sample,  $y_i$  represents class corresponding to  $i$ th data sample,  $w$  and  $b$  represent the optimal parameters of the model,  $N$  represents the number of data, and  $\psi_i$  represents the relaxation variable that corresponds to the  $i$ th data sample.

In this study, the leave-one-out cross-validation is employed to avert the overfitting problem of classifiers and overcome the limited number of data, which means that one data sample is employed as the testing set and the remaining samples are employed as the training set. Besides, the parameter adjustment process of the classifier follows the optimal principle.

## Evaluation Criteria

In this study, we obtain features using the MGS-WBC model and VGBN-LM model. To evaluate the validity of obtained features, the most widely used measurements for binary classification problems, such as ACC, F1-Score, and AUC, are adopted based on the confusion matrix (Kam et al., 2020).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (32)$$

$$Precision = \frac{TP}{TP + FP} \quad (33)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (34)$$

$$F1 - Score = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \quad (35)$$

True positive (TP), false negative (FN), false positive (FP), and true negative (TN) are defined to calculate the evaluation indicators of classifiers, as is shown in **Figure 3**.

		Predicted Class	
		y=1, aMCI	y=0, HC
True Class	y=1, aMCI	True Positive (TP)	False Negative (FN)
	y=0, HC	False Positive (FP)	True Negative (TN)

**FIGURE 3** | Confusion matrix used to measure the binary classification problems.

## EXPERIMENT RESULTS AND ANALYSIS

In this section, compared with the traditional method, we first analyze the diagnostic performance of aMCI by using the whole-brain structural features obtained from the MGS-WBC model as the input of four classifiers. Then, we analyze the diagnostic performance of aMCI by using the global features obtained from VGBN-LM model and local features obtained from MGS-WBC model as the input of four classifiers compared with the traditional method. In the end, we fuse all features, including whole-brain structural features, local features, and global features, to further improve the diagnostic performance of aMCI. It is notable that the order of the above four classifiers is as follows: NB, LDA, LR, and SVM.

### The Performance Analysis of Whole-Brain Structural Features

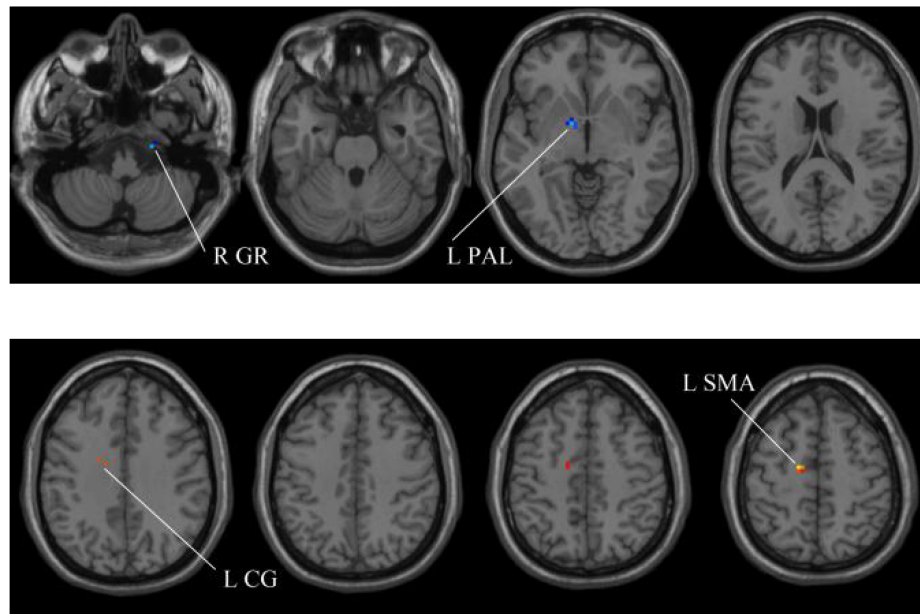
In our experiments, the MGS-WBC model is adopted in two whole-brain structures, including ReHo and ALFF, to reduce the data redundancy and extract multiscale features. For the whole-brain structure calculated by ReHo, as shown in **Table 1** and **Figures 4, 5**, we obtain 10 obvious regions of the brain using MGS-WBC model, including the right gyrus rectus (GR), left pallidum (PAL), left cingulate gyrus (CG), left supplementary motor area (SMA), left inferior cerebellum (IC), right IC, left middle temporal gyrus (MTG), left middle frontal gyrus (MFG), right MFG, and left middle occipital gyrus (MOG), while only six obvious brain regions, namely, left IC, right IC, left MTG, left MFG, right MFG, and left MOG, are found using a single significant threshold (two-sample  $t$ -test,  $p < 0.001$ ). Significantly, each significant region is composed of the activated voxels in the brain. Then, the features extracted from 10 significant regions and six significant regions are separately employed as the input of classifiers, including NB, LDA, LR, and SVM, to evaluate the aMCI's diagnostic

**TABLE 1** | The extracted clusters using MGS-WBC model after ReHo calculation.

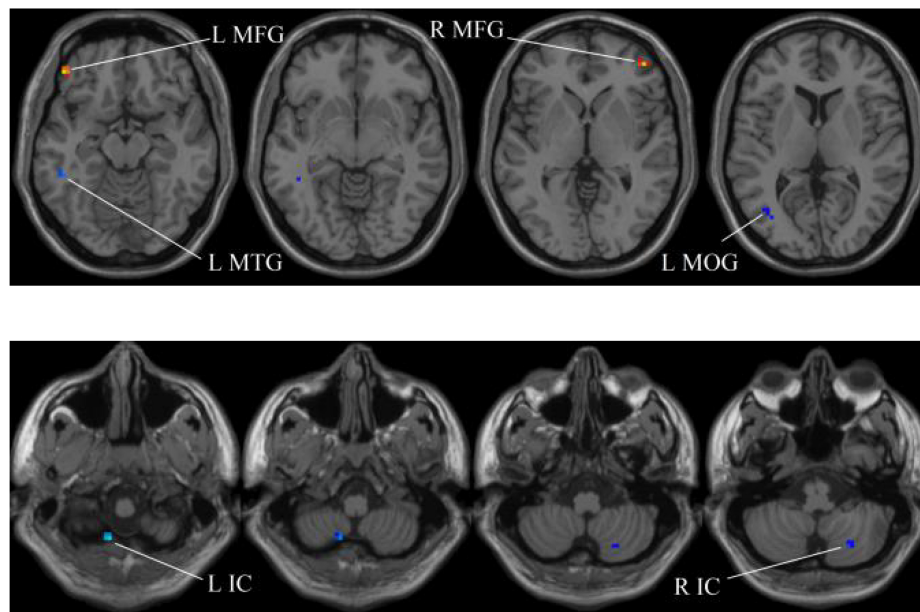
Region	Peak/MNI			t-score	Cluster size
	x	y	z		
R GR	21	-18	-45	-2.9046	5
L PAL	-12	0	-3	-2.6854	5
L CG	-15	-6	36	2.3872	5
L SMA	-15	-9	54	3.3572	9
L IC	-12	-63	-60	-4.2059	8
R IC	24	-69	-48	-3.8093	24
L MTG	-51	-39	-12	-4.0141	11
L MFG	-48	42	-15	4.2324	7
R MFG	45	48	0	4.1363	7
L MOG	-45	-66	3	-3.7574	10

The  $x$ ,  $y$ , and  $z$  coordinates are the primary peak locations in the MNI space. Cluster size  $\geq 5$  voxels in two-sample  $t$ -test. L, left; R, right; GR, gyrus rectus; PAL, pallidum; CG, cingulate gyrus; SMA, supplementary motor area; IC, inferior cerebellum; MTG, middle temporal gyrus; MFG, middle frontal gyrus; MOG, middle occipital gyrus.





**FIGURE 4** | Compared with the HC group, the ReHo in aMCI group exhibits prominent differences based on MGS-WBC<sup>2</sup>. L, left; R, right; GR, gyrus rectus; PAL, pallidum; CG, cingulate gyrus; SMA, supplementary motor area.



**FIGURE 5** | Compared with the HC group, the ReHo in aMCI group exhibits prominent differences based on SSW<sup>1</sup>. L, left; R, right; IC, inferior cerebellum; MTG, middle temporal gyrus; MFG, middle frontal gyrus; MOG, middle occipital gyrus.

performance. **Table 2** presents the diagnostic performance of aMCI using different feature extraction methods of the whole-brain structure (ReHo). The results of MGS-WBC<sup>2</sup> all exceed 66.67% in the four classifiers. Besides, compared with the single significant threshold in the whole-brain structures (SSW<sup>1</sup>) (Zhang et al., 2015; Yang et al., 2018), the results of MGS-WBC<sup>3</sup>

provide the ACC of 88.06, 89.55, 91.04, and 91.04% in the four classifiers, improving by 4.48, 2.98, 10.44, and 4.47%, respectively. Concurrently, the AUC are 94.12, 96.08, 96.88, and 95.72% in four classifiers, and the increases are 2.24, 1.96, 5.53, and 1.51% compared with SSW<sup>1</sup>. For F1-Score, we provide about 87.88, 89.55, 91.18, and 90.91% in four classifiers,

**TABLE 2** | Performance metrics of different classifiers using MGS-WBC model after ReHo calculation.

Classifier	Method	ACC	F1-Score	AUC
NB	SSW <sup>1</sup>	0.8358	0.8358	0.9189
	MGS-WBC <sup>2</sup>	0.6866	0.6667	0.7460
	MGS-WBC <sup>3</sup>	0.8806	0.8788	0.9412
LDA	SSW <sup>1</sup>	0.8657	0.8732	0.9412
	MGS-WBC <sup>2</sup>	0.7164	0.6885	0.7799
	MGS-WBC <sup>3</sup>	0.8955	0.8955	0.9608
LR	SSW <sup>1</sup>	0.8060	0.8169	0.9135
	MGS-WBC <sup>2</sup>	0.7164	0.7077	0.7861
	MGS-WBC <sup>3</sup>	0.9104	0.9118	0.9688
SVM	SSW <sup>1</sup>	0.8657	0.8696	0.9421
	MGS-WBC <sup>2</sup>	0.7612	0.7419	0.7709
	MGS-WBC <sup>3</sup>	0.9104	0.9091	0.9572

SSW<sup>1</sup> refers to the single significant threshold (two-sample *t*-test,  $p < 0.001$ ) in the whole-brain structures (Zhang et al., 2015; Yang et al., 2018). MGS-WBC<sup>2</sup> refers to the combination of volatility detection and significant threshold ( $0.001 < p < 0.05$ ) in the MGS-WBC model. MGS-WBC<sup>3</sup> refers to the fused features, that is, the final output of MGS-WBC model.

**TABLE 3** | The extracted clusters using MGS-WBC model after ALFF calculation.

Region	Peak/MNI			t-score	Cluster size
	x	y	z		
VER	27	0	33	2.6341	21
VER	27	-21	45	-2.7783	5

The x, y, and z coordinates are the primary peak locations in the MNI space. Cluster size  $\geq 5$  voxels in two-sample *t*-test. VER, vermis.

and maximum improvements are 4.30, 2.23, 9.49, and 3.95% compared with SSW<sup>1</sup>.

Next, for the whole-brain structure after ALFF computation, as shown in **Table 3** and **Figure 6**, only two significant regions of the brain corresponding to different locations are found, whereas no significant region exists using SSW<sup>1</sup>. As shown in **Table 4**, the results of MGS-WBC<sup>2</sup> or MGS-WBC<sup>3</sup> provide the ACC of 65.67, 73.13, 71.64, and 67.17% in four classifiers. In addition, we get the AUC of 65.67, 77.72, 77.18, and 75.22% in four classifiers. For F1-Score, we obtain 65.67, 72.73, 69.84, and 60.71% in four classifiers.

**TABLE 4** | Performance metrics of different classifiers using MGS-WBC model after ALFF calculation.

Classifier	Method	ACC	F1-Score	AUC
NB		0.6567	0.6567	0.6567
LDA	MGS-WBC <sup>2</sup>	0.7313	0.7273	0.7772
LR	/MGS-WBC <sup>3</sup>	0.7164	0.6984	0.7718
SVM		0.6716	0.6071	0.7522

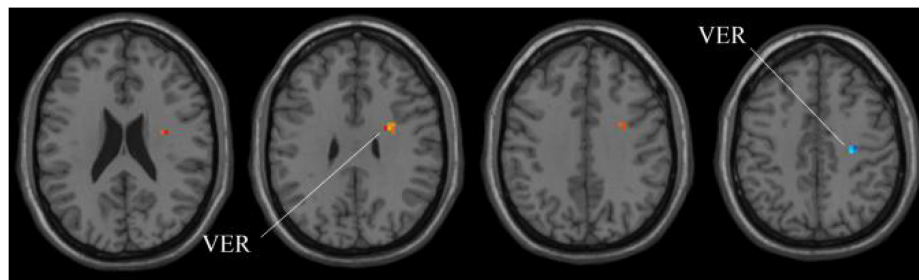
MGS-WBC<sup>2</sup> refers to the combination of volatility detection and significant threshold ( $0.001 < p < 0.05$ ) in MGS-WBC model. MGS-WBC<sup>3</sup> refers to the fused features, that is, the final output of MGS-WBC model.

Our goal is to make the average difference between the classes larger and the variance within the classes more stable. Consequently, we develop an MGS-WBC model using multiple significant thresholds and validation detection to generate the fused features of whole-brain structures. Through ablation experiments, it can be inferred that the features of whole-brain structures extracted using MGS-WBC can significantly improve the diagnostic performance of aMCI in typical classifiers compared with the SSW<sup>1</sup>. It indicates that the MGS-WBC model we present is valid, and the model can also effectively remove redundant information from the brain.

## The Performance Analysis of Functional Network Features

In our experiments, to fully excavate the features of the functional network, we develop an MGS-WBC model to extract local features. Meanwhile, we present a VGBN-LM model to extract global features of the functional network and fuse its local features to improve the diagnostic performance of aMCI.

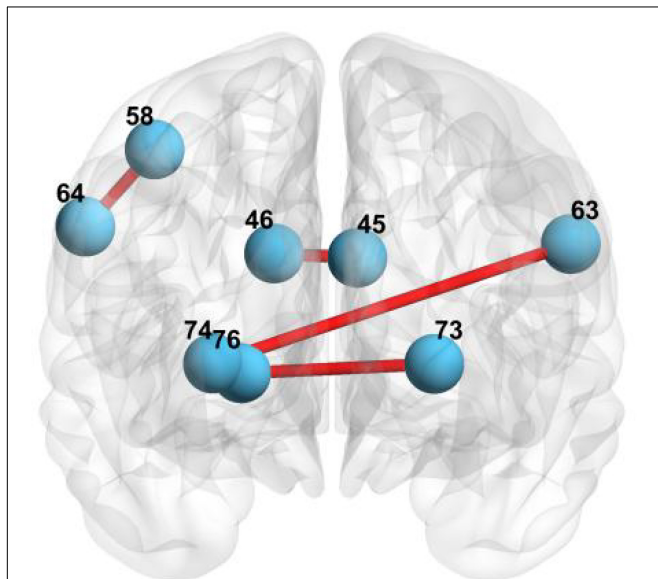
First, we select the local features of the functional network using MGS-WBC model to improve the diagnosis performance of aMCI. Our findings elucidate that three pairs of connected brain regions with notable differences are found using MGS-WBC<sup>2</sup>, including (45, 46), (73, 76), and (74, 76), as can be seen from **Table 5**. Also, two pairs of connected brain regions, such as (64, 58) and (63, 74), are found using SSW<sup>1</sup>. The brain regions involved in the local feature selection are shown in **Figure 7**, where the light blue ball denotes the selected brain regions, and the red line indicates that the two selected brains are connected. Notably, the number in **Figure 7** and **Table 6** denotes the brain regions segmented using the AAL template.

**FIGURE 6** | Compared with HC group, the ALFF in aMCI group exhibits prominent differences based on volatility detection in the MGS-WBC model. VER, Vermis.

**TABLE 5** | Performance metrics of local features using MGS-WBC model.

Classifier	Method	ACC	F1-Score	AUC
NB	SSF <sup>1</sup>	0.7313	0.7188	0.7496
	MGS-WBC <sup>2</sup>	0.5672	0.5538	0.6346
	MGS-WBC <sup>3</sup>	0.7463	0.7463	0.8048
LDA	SSF <sup>1</sup>	0.7313	0.7188	0.7531
	MGS-WBC <sup>2</sup>	0.6269	0.6032	0.6854
	MGS-WBC <sup>3</sup>	0.7164	0.7164	0.8066
LR	SSF <sup>1</sup>	0.6567	0.5490	0.7308
	MGS-WBC <sup>2</sup>	0.6119	0.5938	0.6961
	MGS-WBC <sup>3</sup>	0.7015	0.6667	0.7995
SVM	SSF <sup>1</sup>	0.7612	0.7576	0.7647
	MGS-WBC <sup>2</sup>	0.6418	0.5862	0.6943
	MGS-WBC <sup>3</sup>	0.7761	0.7826	0.8324

SSF<sup>1</sup> refers to the single significant threshold ( $p < 0.001$ ) in the functional network (Zhang et al., 2015; Yang et al., 2018). MGS-WBC<sup>2</sup> refers to the combination of volatility detection and significant threshold ( $0.001 < p < 0.05$ ) in MGS-WBC model. MGS-WBC<sup>3</sup> refers to the fused features, that is, the final output of MGS-WBC model.



**FIGURE 7** | The brain regions involved in local feature selection: 45 represents the left cuneus; 46 represents right cuneus; 73 represents left putamen; 76 represents right pallidum; 74 represents right putamen; 64 represents right supramarginal gyrus; 58 represents postcentral gyrus; and 63 represents left supramarginal gyrus.

The diagnosis performance of aMCI using local features resulted from the MGS-WBC model is sorted in **Table 5** compared with SSF<sup>1</sup>. When the features generated from MGS-WBC<sup>2</sup> are used as the input of the four classifiers, including NB, LDA, LR, and SVM, the classification metrics ranged from 55.38 to 69.61%. Meanwhile, the results elucidate that except for LDA, each of NB, LR, and SVM classifiers makes large improvements using MGS-WBC<sup>3</sup> compared with SSF<sup>1</sup>. Although the ACC and F1-Score of LDA decreased by 1.49 and 0.24%, respectively, the AUC value increased by 5.35%.

**TABLE 6** | Local feature selection of functional network using MGS-WBC model.

Connected regions	Group	Volatility detection	p-value
(45, 46)	aMCI	0.2611	0.0157
	HC	0.3063	0.0462
(73, 76)	aMCI	0.6176	0.0335
	HC	0.5038	0.0335
(74, 76)	aMCI	0.5172	0.0133
	HC	0.3487	0.0133
(64, 58)	aMCI and HC	–	0.0006
(63, 74)	aMCI and HC	–	0.0006

45 represents the left cuneus; 46 represents right cuneus; 73 represents left putamen; 76 represents right pallidum; 74 represents right putamen; 64 represents right supramarginal gyrus; 58 represents postcentral gyrus; and 63 represents left supramarginal gyrus.

**TABLE 7** | Global features analysis using VGBN-LM model.

Classifier	Method	ACC	F1-Score	AUC
NB	VGBN-LM <sup>1</sup>	0.4478	0.3934	0.5116
	VGBN-LM	0.6418	0.6757	0.6996
LDA	VGBN-LM <sup>1</sup>	0.5672	0.5246	0.3137
	VGBN-LM	0.6269	0.6377	0.6185
LR	VGBN-LM <sup>1</sup>	0.5224	0.5000	0.5330
	VGBN-LM	0.6567	0.6567	0.6979
SVM	VGBN-LM <sup>1</sup>	0.4627	0.4375	0.3966
	VGBN-LM	0.6119	0.6286	0.6729

VGBN-LM<sup>1</sup> refers to VGBN-LM model while without using two-sample  $t$ -test ( $p < 0.05$ ). VGBN-LM refers to the method we propose in the "Vector Generation for Brain Networks Based on Laplacian Matrix (VGBN-LM)" Section.

We further develop a VGBN-LM model to extract global features of the functional network, including left MFG ( $p = 0.035$ , two-sample  $t$ -test), left hippocampus ( $p = 0.030$ , two-sample  $t$ -test), right hippocampus ( $p = 0.038$ , two-sample  $t$ -test), left amygdala ( $p = 0.009$ , two-sample  $t$ -test), and right amygdala ( $p = 0.013$ , two-sample  $t$ -test), and the global features, are then used as the input of classifiers to evaluate the validity of the extracted features. As summarized in **Table 7**, the results reveal that the ACC, F1-Score, and AUC in the four classifiers can achieve more than 61% using VGBN-LM. Moreover, the results of VGBN-LM can provide the ACC of 64.18, 62.69, 65.67, and 61.19% in four classifiers, and maximum improvements are 19.40, 5.97, 13.43, and 14.92% compared with VGBN-LM<sup>1</sup>. Also, compared to VGBN-LM<sup>1</sup>, the AUCs of the four classifiers are 69.96, 61.85, 69.79, and 67.29%, and the growth rates are 18.80, 30.48, 16.49, and 27.63%. For F1-Score, we provide about 64.18, 62.69, 65.67, and 61.19% in four classifiers, increasing by 28.23, 11.31, 15.67, and 19.11% compared with VGBN-LM<sup>1</sup>. Furthermore, we reduce the number of features per data sample from 4,005 ( $\frac{8100-90}{2}$ ) to 5 using VGBN-LM.

Then, we fuse the multiscale features of functional network, including local features and global features, to further improve the diagnostic performance of aMCI. The local features of functional network are obtained using SSF<sup>1</sup> and MGS-WBC<sup>2</sup>. As summarized in **Table 8**, using MGS-VGBN<sup>1</sup>, each of NB, LDA,

**TABLE 8** | Performance metrics of local and global features of functional network.

Classifier	Method	ACC	F1-Score	AUC
NB	SSF <sup>1</sup>	0.7313	0.7188	0.7496
	MGS-VGBN <sup>1</sup>	0.7313	0.7353	0.7772
	MGS-VGBN <sup>2</sup>	0.7612	0.7647	0.8164
LDA	SSF <sup>1</sup>	0.7313	0.7188	0.7531
	MGS-VGBN <sup>1</sup>	0.7463	0.7536	0.7763
	MGS-VGBN <sup>2</sup>	0.7761	0.7761	0.8119
LR	SSF <sup>1</sup>	0.6567	0.5490	0.7308
	MGS-VGBN <sup>1</sup>	0.7164	0.6545	0.7879
	MGS-VGBN <sup>2</sup>	0.7761	0.7692	0.8342
SVM	SSF <sup>1</sup>	0.7612	0.7576	0.7647
	MGS-VGBN <sup>1</sup>	0.7761	0.7761	0.8057
	MGS-VGBN <sup>2</sup>	0.7910	0.7879	0.8520

SSF<sup>1</sup> refers to the single significant threshold ( $p < 0.001$ ) in the functional network (Zhang et al., 2015; Yang et al., 2018). MGS-VGBN<sup>1</sup> refers to local features generated from volatility detection of MGS-WBC model and global features generated from VGBN-LM model in functional network. MGS-VGBN<sup>2</sup> refers to the fused features of functional network.

LR, and SVM classifiers has a large value in ACC, F1-Score, and AUC, and its values ranged from 65.45 to 80.57%. Specifically, the results of MGS-VGBN<sup>2</sup> provide the ACC of 76.12, 77.61, 77.61, and 79.10%, and the growth rates are 2.99, 4.48, 11.94, and 2.98% compared with SSF<sup>1</sup>. In addition, we obtain the AUC of 81.64, 81.19, 83.42, and 85.20% in four classifiers, and the improvements are 6.68, 5.88, 10.34, and 8.73% compared with SSF<sup>1</sup>. For F1-Score, we provide about 76.47, 77.61, 76.92, and 78.79% in four classifiers, and the improvements are 4.59, 5.73, 22.02, and 3.03% compared with SSF<sup>1</sup>.

Regarding limited and high-dimensional data, we present a VGBN-LM model that excavates the global features of the functional network and fuses local features obtained by the MGS-WBC model. By ablation experiments, multiscale features of functional network we propose can greatly improve the diagnostic performance of aMCI compared to SSF<sup>1</sup>, indicating that the global features we propose can play a certain role in the intelligent diagnosis of aMCI. Besides, the features extracted by this study can accurately locate the significant regions in the brain, which is convenient for doctors to conduct further studies.

## The Performance Analysis of Fused Features

In our experiments, we mainly fuse the multiscale features, including whole-brain structural features and features of functional network, to improve the diagnostic performance of aMCI.

As summarized in **Table 9**, compared with the single significant threshold ( $p < 0.001$ ) in the whole-brain structures and function network (SSWF<sup>1</sup>) (Zhang et al., 2015; Yang et al., 2018), the ACC of NB, LDA, LR, and SVM classifiers are about 89.55, 91.04, 92.54, and 94.03%, increasing by 1.49, 1.49, 1.50, and 4.48%, respectively. Besides, this study provides the AUC of 95.63, 92.69, 97.59, and 97.33% in the four classifiers, increasing by 2.49, 1.96, 2.67, and 2.23%, respectively. For F1-Score, this

**TABLE 9** | The fused features analysis.

Classifier	Method	ACC	F1-Score	AUC
NB	SSWF <sup>1</sup>	0.8806	0.8788	0.9314
	FUSE	0.8955	0.8955	0.9563
LDA	SSWF <sup>1</sup>	0.8955	0.8986	0.9465
	FUSE	0.9104	0.9091	0.9269
LR	SSWF <sup>1</sup>	0.9104	0.9091	0.9492
	FUSE	0.9254	0.9254	0.9759
SVM	SSWF <sup>1</sup>	0.8955	0.8986	0.9510
	FUSE	0.9403	0.9412	0.9733

SSWF<sup>1</sup> refers to the single significant threshold ( $p < 0.001$ ) in the whole-brain structures and function network (Zhang et al., 2015; Yang et al., 2018). FUSE refers to the fused features obtained from MGS-WBC and VGBN-LM models in the whole-brain structures and functional network.

study provides about 76.47, 77.61, 76.92, and 78.79% in the four classifiers, increasing by 1.67, 1.05, 1.64, and 4.26%, respectively.

Over the years, the extraction of biomarkers of aMCI based on rs-fMRI data has been reported by numerous studies (Yang et al., 2018, 2021), yet very few studies about multiscale feature extraction of whole-brain structures and functional network due to limited data, and how to use machine learning methods to verify the effectiveness of the features. Furthermore, a robust and reliable system for aMCI detection is conducive to timely detecting and screening patients at a high risk of AD. Based on this, we extract multiscale features according to the characteristics of the brain and develop the framework of the diagnostic system. Our proposed models, including MGS-WBC model and VGBN-LM model, outperform the traditional approaches (SSWF<sup>1</sup>) (Zhang et al., 2015; Yang et al., 2018) in this study. It turns out that the fused features obtained from the MGS-WBC model and VGBN-LM model are more important than the traditional single-scale features in terms of the diagnosis of aMCI. Furthermore, our feature selection methods will make the diagnosis of aMCI more accurate and reliable, providing novel insights for the extraction of discriminative neuroimaging features.

## CONCLUSION

In this study, we present a system for the diagnosis of aMCI. We first develop an MGS-WBC model to extract discriminative features of whole-brain structures and functional networks. Then, we propose a VGBN-LM model that excavates the global features of the functional network and fuses its local features, thus overcoming the low availability of the functional network features. Finally, we fuse all the features generated from the MGS-WBC model and VGBN-LM model to further improve the diagnostic performance of aMCI. The results demonstrate that this study outperforms the traditional method. In conclusion, the proposed feature extraction methods can be utilized to detect other similar neurological diseases of the brain, providing new insights for the intelligent construction of the CAD system. The future study contains applying our approaches to other brain diseases and verifying the robustness of the system. Besides, more

data will be collected so that deep learning models can be used for accurate classification.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <http://adni.loni.usc.edu/>.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. Ethical review and approval was not required for the animal study because publicly available datasets were analyzed in this study.

## REFERENCES

- Association, A. S. (2018). 2018 Alzheimer's disease facts and figures. *Alzheimer's Dement.* 14, 367–429. doi: 10.21926/obm.geriatri.1904079
- Barone, E., Domenico, F. Di, Cassano, T., Arena, A., Tramutola, A., Lavecchia, M., et al. (2016). Impairment of biliverdin reductase-A promotes brain insulin resistance in Alzheimer disease: a new paradigm. *Free Radic. Biol. Med.* 91, 127–142. doi: 10.1016/j.freeradbiomed.2015.12.012
- Belkin, M., and Niyogi, P. (2001). "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, (Columbia: MIT Press), 585–591. doi: 10.1162/0899766041732396
- Bian, Z., Li, Q., Wang, L., Lu, C., Yin, S., and Li, X. (2014). Relative power and coherence of EEG series are related to amnesic mild cognitive impairment in diabetes. *Front. Aging Neurosci.* 6:11. doi: 10.3389/fnagi.2014.0011
- Chen, C. C. M., Schwender, H., Keith, J., Nunkesser, R., Mengersen, K., and Macrossan, P. (2011). Methods for Identifying SNP Interactions: a Review on Variations of Logic Regression, Random Forest and Bayesian Logistic Regression. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 8, 1580–1591. doi: 10.1109/TCBB.2011.46
- Ciulli, S., Citi, L., Salvadori, E., Valenti, R., Poggesi, A., Inzitari, D., et al. (2016). Prediction of Impaired Performance in Trail Making Test in MCI Patients With Small Vessel Disease Using DTI Data. *IEEE J. Biomed. Health Inform.* 20, 1026–1033. doi: 10.1109/JBHI.2016.2537808
- Cristovao, F., Cascianelli, S., Canakoglu, A., Carman, M., Nanni, L., Pinoli, P., et al. (2022). Investigating Deep Learning Based Breast Cancer Subtyping Using Pan-Cancer and Multi-Omic Data. *IEEE/ACM Transac. Comput. Biol. Bioinf.* 19, 121–134. doi: 10.1109/TCBB.2020.3042309
- Dadar, M., Pascoal, T., Manitsirikul, S., Misquitta, K., Fonov, V., Tartaglia, M., et al. (2017). Validation of a Regression Technique for Segmentation of White Matter Hyperintensities in Alzheimer's Disease. *IEEE Transac. Med. Imag.* 36, 1758–1768. doi: 10.1109/TMI.2017.2693978
- Harrison, T., Maass, A., Adams, J., Du, R., Baker, S., and Jagust, W. (2019). Tau deposition is associated with functional isolation of the hippocampus in aging. *Nat. Comm.* 10:4900. doi: 10.1038/s41467-019-12921-z
- Ji, S., and Ye, J. (2008). Generalized linear discriminant analysis: a unified framework and efficient model selection. *IEEE Transact. Neur. Netw.* 19, 1768–1782. doi: 10.1109/TNN.2008.2002078
- Ju, R., Hu, C., Zhou, P., and Li, Q. (2019). Early Diagnosis of Alzheimer's Disease Based on Resting-State Brain Networks and Deep Learning. *IEEE/ACM Trans. Comp. Biol. Bioinform.* 16, 244–257. doi: 10.1109/TCBB.2017.2776910

This data can be found here: <http://adni.loni.usc.edu/>. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

CC was responsible for writing the manuscript and doing the experiments. CY and EC instructed the experiments. JC was responsible for drawing pictures in the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the Fujian Province Innovation Strategy Research Program (No. 2020R01020196) and the YongTai Institute of Artificial Intelligence.

- Kam, T., Zhang, H., Jiao, Z., and Shen, D. (2020). Deep Learning of Static and Dynamic Brain Functional Networks for Early MCI Detection. *IEEE Transact. Med. Imag.* 39, 478–487. doi: 10.1109/TMI.2019.2928790
- Kassani, P., Gossman, A., and Wang, Y. (2020). Multimodal Sparse Classifier for Adolescent Brain Age Prediction. *IEEE J. Biomed. Health Inform.* 24, 336–344. doi: 10.1109/JBHI.2019.2925710
- Khatun, S., Morshed, B., and Bidelman, G. (2019). A Single-Channel EEG-Based Approach to Detect Mild Cognitive Impairment via Speech-Evoked Brain Responses. *IEEE Transac. Neur. Syst. Rehabil. Eng.* 27, 1063–1070. doi: 10.1109/TNSRE.2019.2911970
- Kim, H.-G., Park, S., Rhee, Y. H., Lee, M. K., Ryu, C.-W., Lee, Y. S., et al. (2020). Evaluation and Prediction of Early Alzheimer's Disease Using a Machine Learning-based Optimized Combination-Feature Set on Gray Matter Volume and Quantitative Susceptibility Mapping. *Curr. Alzheimer Res.* 17, 428–437. doi: 10.2174/1567205017666200624204427
- Lee, J., Kim, Y., Jeong, Y., Na, D., Kim, J., Lee, K., et al. (2015). Inference of brain pathway activities for Alzheimer's disease classification. *BMC Med. Inform. Dec. Making* S1. doi: 10.1186/1472-6947-15-S1-S1
- Lei, B., Yang, P., Zhuo, Y., Zhou, F., Ni, D., Chen, S., et al. (2019). Neuroimaging Retrieval via Adaptive Ensemble Manifold Learning for Brain Disease Diagnosis. *IEEE J. Biomed. Health Inform.* 23, 1661–1673. doi: 10.1109/JBHI.2018.2872581
- Li, R., Rui, G., Zhao, C., Wang, C., Fang, F., and Zhang, Y. (2020). Functional Network Alterations in Patients With Amnesic Mild Cognitive Impairment Characterized Using Functional Near-Infrared Spectroscopy. *IEEE Transact. Neural Syst. Rehabil. Eng.* 28, 123–132. doi: 10.1109/TNSRE.2019.2956464
- Li, W., Zhang, L., Qiao, L., and Shen, D. (2020). Toward a Better Estimation of Functional Brain Network for Mild Cognitive Impairment Identification: a Transfer Learning View. *IEEE J. Biomed. Health Inform.* 24, 1160–1168. doi: 10.1109/JBHI.2019.2934230
- Li, X., Zhu, Z., Zhao, W., Yu, S., Wen, D., Xie, Y., et al. (2018). Decreased resting-state brain signal complexity in patients with mild cognitive impairment and Alzheimer's disease: a multi-scale entropy analysis. *Biomed. Optics Exp.* 9:1916. doi: 10.1364/BOE.9.001916
- Li, Y., Liu, J., Tang, Z., and Lei, B. (2020). Deep Spatial-Temporal Feature Fusion From Adaptive Dynamic Functional Connectivity for MCI Identification. *IEEE Transac. Med. Imag.* 39, 2818–2830. doi: 10.1109/TMI.2020.2976825
- Liao, W., Li, J., Ji, G., Wu, G., Long, Z., Xu, Q., et al. (2019). Endless Fluctuations: Temporal Dynamics of the Amplitude of Low Frequency Fluctuations. *IEEE Transac. Med. Imag.* 38, 2523–2532. doi: 10.1109/TMI.2019.2904555
- Liu, X., Lu, R., Ma, J., Chen, L., and Qin, B. (2016). Privacy-Preserving Patient-Centric Clinical Decision Support System on Naïve Bayesian Classification. *IEEE J. Biomed. Health Inform.* 20, 655–668. doi: 10.1109/JBHI.2015.2407157

- Nadarajah, S., and Kotz, S. (2006). Statistical Distribution of the Measure of Coherence. *IEEE Transac. Biomed. Eng.* 53, 2409–2410. doi: 10.1109/TBME.2006.883625
- Sangnawakij, P., and Niwitpong, S.-A. (2017). Confidence intervals for coefficients of variation in two-parameter exponential distributions. *Comm. Stat.* 46, 6618–6630. doi: 10.1006/tpbi.2000.1504
- Sundaram, P., Luessi, M., Bianciardi, M., Stufflebeam, S., Hamalainen, M., and Solo, V. (2020). Individual Resting-State Brain Networks Enabled by Massive Multivariate Conditional Mutual Information. *IEEE Trans. Med. Imag.* 39, 1957–1966. doi: 10.1109/TMI.2019.2962517
- Ting, C., Ombao, H., Salleh, S., and Latif, A. Z. A. (2020). Multi-Scale Factor Analysis of High-Dimensional Functional Connectivity in Brain Networks. *IEEE Trans. Netw. Sci. Eng.* 7, 449–465. doi: 10.1109/tNSE.2018.2869862
- Vicente, T., Hoai, M., and Samaras, D. (2018). Leave-One-Out Kernel Optimization for Shadow Detection and Removal. *IEEE Trans. Patt. Analy. Mach. Intell.* 40, 682–695. doi: 10.1109/TPAMI.2017.2691703
- Wang, M., Lian, C., Yao, D., Zhang, D., Liu, M., and Shen, D. (2020). Spatial-Temporal Dependency Modeling and Network Hub Detection for Functional MRI Analysis via Convolutional-Recurrent Network. *IEEE Trans. Bio-Med. Eng.* 67, 2241–2252. doi: 10.1109/TBME.2019.2957921
- Wang, Z., Zheng, Y., Zhu, D., Bozoki, A., and Li, T. (2018). Classification of Alzheimer's Disease, Mild Cognitive Impairment and Normal Control Subjects Using Resting-State fMRI Based Network Connectivity Analysis. *IEEE J. Transl. Eng. Health Med.* 6:1801009. doi: 10.1109/JTEHM.2018.2874887
- Xi, Q., Zhao, X., Wang, P., Guo, Q., and He, Y. (2013). Abnormal intrinsic brain activity in amnesic mild cognitive impairment revealed by amplitude of low-frequency fluctuation: a resting-state functional magnetic resonance imaging study. *Chin. Med. J.* 126, 2912–2917.
- Xi, Y., Chen, Y., Jin, Y., Han, G., Song, M., Song, T., et al. (2022). Versatile nanomaterials for Alzheimer's disease: pathogenesis inspired disease-modifying therapy. *J. control. Rel.* 345, 38–61. doi: 10.1016/j.jconrel.2022.02.034
- Xiao, L., Wang, J., Kassani, P., Zhang, Y., Bai, Y., Stephen, J., et al. (2020). Multi-Hypergraph Learning-Based Brain Functional Connectivity Analysis in fMRI Data. *IEEE Transac. Med. Imag.* 39, 1746–1758. doi: 10.1109/TMI.2019.2957097
- Xu, B., Li, Y., and Kim, Y. (2017). Classification of Finger Movements Based on Reflection Coefficient Variations of a Body-Worn Electrically Small Antenna. *IEEE Anten. Wirel. Prop. Lett.* 16, 1812–1815.
- Xue, C., Sun, H., Hu, G., Qi, W., Yue, Y., Rao, J., et al. (2020). Disrupted Patterns of Rich-Club and Diverse-Club Organizations in Subjective Cognitive Decline and Amnesic Mild Cognitive Impairment. *Front. Neurosci.* 14:575652. doi: 10.3389/fnins.2020.575652
- Yang, L., Yan, Y., Wang, Y., Hu, X., Lu, J., Chan, P., et al. (2018). Gradual Disturbances of the Amplitude of Low-Frequency Fluctuations (ALFF) and Fractional ALFF in Alzheimer Spectrum. *Front. Neurosci.* 12:975. doi: 10.3389/fnins.2018.00975
- Yang, P., Zhou, F., Ni, D., Xu, Y., Chen, S., Wang, T., et al. (2021). Fused Sparse Network Learning for Longitudinal Analysis of Mild Cognitive Impairment. *IEEE Transac. Cybernet.* 51, 233–246. doi: 10.1109/TCYB.2019.2940526
- Zhang, X., Hu, B., Ma, X., and Xu, L. (2015). Resting-state whole-brain functional connectivity networks for MCI classification using L2-regularized logistic regression. *IEEE Transact. Nanobiosci.* 14, 237–247. doi: 10.1109/TNB.2015.2403274
- Zhen, D., Xia, W., Yi, Z., Zhao, P., Zhong, J., Shi, H., et al. (2018). Alterations of brain local functional connectivity in amnesic mild cognitive impairment. *Transl. Neurodegen.* 7:26.
- Zhou, Q., Goryawala, M., Cabrerizo, M., Wang, J., Barker, W., Loewenstein, D., et al. (2014). An optimal decisional space for the classification of Alzheimer's disease and mild cognitive impairment. *IEEE Transact. Bio-Med. Eng.* 61, 2245–2253. doi: 10.1109/TBME.2014.2310709

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Cai, Cao, Yang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.