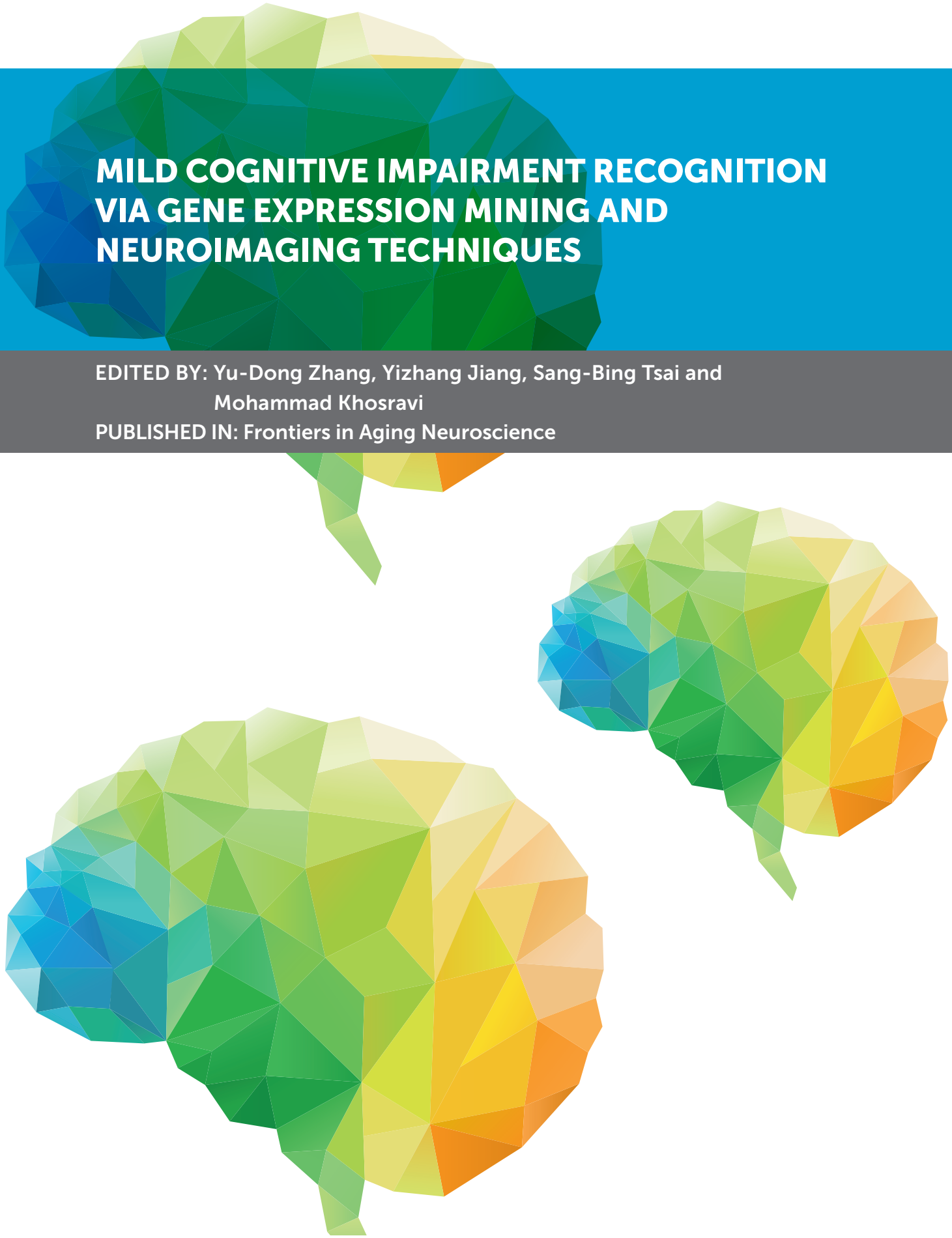
MILD COGNITIVE IMPAIRMENT RECOGNITION VIA GENE EXPRESSION MINING AND NEUROIMAGING TECHNIQUES

EDITED BY: Yu-Dong Zhang, Yizhang Jiang, Sang-Bing Tsai and Mohammad Khosravi

PUBLISHED IN: Frontiers in Aging Neuroscience



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-83250-840-4
DOI 10.3389/978-2-83250-840-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

MILD COGNITIVE IMPAIRMENT RECOGNITION VIA GENE EXPRESSION MINING AND NEUROIMAGING TECHNIQUES

Topic Editors:
Yu-Dong Zhang, University of Leicester, United Kingdom
Yizhang Jiang, Jiangnan University, China
Sang-Bing Tsai, Wuyi University, China
Mohammad Khosravi, Weifang University of Science and Technology, China

Citation: Zhang, Y.-D., Jiang, Y., Tsai, S.-B., Khosravi, M., eds. (2022). Mild Cognitive Impairment Recognition Via Gene Expression Mining and Neuroimaging Techniques. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-840-4

Table of Contents

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fnagi.2021.828214

Performance Analysis of Machine Learning and Deep Learning Architectures on Early Stroke Detection Using Carotid Artery Ultrasound Images

S. Latha[1], P. Muthu[2], Khin Wee Lai[3]*, Azira Khalil[4]* and Samiappan Dhanalakshmi[1]*

[1] Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Chennai, India, [2] Department of Biomedical Engineering, SRM Institute of Science and Technology, Chennai, India, [3] Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia, [4] Faculty of Science and Technology, Universiti Sains Islam Malaysia, Nilai, Malaysia

OPEN ACCESS

Edited by:
Sang-Bing Tsai,
Wuyi University, China

Reviewed by:
N. Venkateswaran,
SSN College of Engineering, India
Selvaraju Narayanasamy,
Indian Institute of Technology Guwahati, India

***Correspondence:**
P. Muthu
muthup@srmist.edu.in
Khin Wee Lai
lai.khinwee@um.edu.my
Azira Khalil
azira@usim.edu.my

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

Received: *03 December 2021*
Accepted: *28 December 2021*
Published: *27 January 2022*

Citation:
Latha S, Muthu P, Lai KW, Khalil A and Dhanalakshmi S (2022) Performance Analysis of Machine Learning and Deep Learning Architectures on Early Stroke Detection Using Carotid Artery Ultrasound Images. Front. Aging Neurosci. 13:828214. doi: 10.3389/fnagi.2021.828214

Atherosclerotic plaque deposit in the carotid artery is used as an early estimate to identify the presence of cardiovascular diseases. Ultrasound images of the carotid artery are used to provide the extent of stenosis by examining the intima-media thickness and plaque diameter. A total of 361 images were classified using machine learning and deep learning approaches to recognize whether the person is symptomatic or asymptomatic. CART decision tree, random forest, and logistic regression machine learning algorithms, convolutional neural network (CNN), Mobilenet, and Capsulenet deep learning algorithms were applied in 202 normal images and 159 images with carotid plaque. Random forest provided a competitive accuracy of 91.41% and Capsulenet transfer learning approach gave 96.7% accuracy in classifying the carotid artery ultrasound image database.

Keywords: carotid artery, ultrasound image, machine learning, deep learning, stroke

INTRODUCTION

Every year, in India, 26% of people die due to cardiovascular diseases, stroke because of artery stenosis is 75%, and heart attack is 42%. In the United States, one of the 19 deaths is due to stroke (Farah, 2018). Risk factors that may lead to stroke are physical inactivity, being obese, heavy drinking, use of illegal drugs, family history having a stroke and other cardiovascular diseases, cholesterol, high blood pressure, diabetes, and smoking. Other factors with increased stroke risk are race-, sex-, age-, and hormones-related problems.

Stroke is the third prominent reason for death in many developed countries (Benjamin et al., 2019). The common cause of stroke is the formation of atherosclerotic plaque in the carotid artery that can grow large enough to block blood flow leading to stenosis or rupture causing clots in the artery. Progressive intimal accumulation of protein, lipid, and cholesterol makes medium- and large-sized arteries, causing atherosclerosis. Atherosclerosis may be existing in body parts, such as infernal aorta, coronary artery, superficial femoral artery, and the common carotid artery

bifurcation region. Strain in the arterial wall causes variance in clinical, mechanical, and molecular levels in the artery. The plaque formation is compensated by artery enlargement with no changes in the lumen region, where blood flows.

The mapping of features to any one of the classes in a computer-assisted diagnostic system is called classification. Machine learning algorithms that are used for biomedical image classification are neural network, backpropagation, support vector machine (SVM), adaptive binary tree-based SVM, decision trees, such as linear regression, logistic regression, random forest, k-nearest neighbor (KNN), k-means, Boltzmann machine, mean shift clustering, Markov statistics nonparametric techniques, and fuzzy-based classification methods.

Stimulated by the function and structure of the brain, an artificial neural network (ANN) was developed. A subset of machine learning, called deep learning, performs classification tasks directly from the images. The accuracy of deep learning sometimes exceeds human performance. The model extracts all the necessary features by itself and performs the classification. Transfer learning is a kind of deep learning which uses the learnt knowledge from some other data and uses that for the application in hand. Some of the transfer learning algorithms are Alexnet, Mobilenet, Imagenet, Capsulenet, etc.

Carl Azzopardi et al. (2020) used a deep neural network (DNN) to delineate lumen-intima boundary (LIB) and media-adventitia boundary (MAB) with a fully automatic segmentation technique. For the network stochastic gradient descent optimization problem, a new objective function was formulated. The invariant intensity data input was given to the network with a bimodal synthesis of amplitude and phase congruency. The performance in MAB and LIB detection was 96.2 and 92.5%, respectively. The study was made with just 15 images in each stenosis category which is not a sufficient number for deep learning-based segmentation. Images from different sources were not considered for learning, missing generalizability (Azzopardi et al., 2020).

Roy-Cardinal et al. (2019) extracted noninvasive vascular ultrasound elastography (NIVE) and ultrasound features, such as homodyned-K (HK), Nakagami parametric maps, log-compressed images. The algorithm identified large lipid area, calcification, ruptured fibrous cap presence, differentiation of nonvulnerable and vulnerable plaques, and confirming symptomatic and asymptomatic patients using a random forest classifier. The study population was 91, and only 5 cases with fibrous caps were involved. A balanced dataset may give better classification performance. Based on elastography and B mode gray-level features, the AUC obtained was 0.90 (95% CI 0.80–0.92, $p < 0.001$). The area of calcification accuracy obtained was 0.95 (95% CI 0.94–0.96, $p < 0.001$), performed using the above features. Area under the curve variation for other tasks varied between 0.79 and 0.97 (Roy-Cardinal et al., 2019).

Loizou et al. (2017) studied the texture variability in the ultrasound video to identify the presence of vulnerable plaque. The videos were intensity normalized, denoised, IMT segmented, and texture feature learned to find systole and diastole states. The texture was visibly variable for diastolic and systolic states. More gray-scale average was recorded for systole compared to diastole. Plaque structures had variable textures in both the states. Systole and diastole features combined gave better results. Borders of type 1 plaque were not identified by this method. Acoustic shadowing was produced in type V plaque and was not recognizable. The state diagram was improper for 2% of cases (Loizou et al., 2017).

Lekadir et al. (2017) proposed a CNN classification model for the different plaque constituents. Lipid core, calcified tissues, and fibrous caps were detected with a correlation of 0.90 related to clinical results. Based on the patch batched technique, 56 images were converted into 90,000 patches for the process. SVM with predefined image features gave an accuracy of 78.5%. The testing time taken for classifying each image was 52 ± 13 ms, and changes in accuracy were reduced by 0.003 by changing the patches between 9×9, 11×11, 13×13, and 15×15 (Lekadir et al., 2017). Pazinato et al. (2016) used the features of neighboring pixels for carotid image classification. On a dataset with calcium, lipids, muscles, fibrous, and blood tissues texture, gradient, statistical, and local binary pattern (LBP) features were used. Pixel-based machine learning classification was carried out on the normalized image following multiscale description. The method was computationally complex and did not focus on any particular machine learning algorithm. The technique applied in ultrasound tissue engineering achieved a classification accuracy of 73%, and was statistically verified (Pazinato et al., 2016).

Gastounioti et al. (2015) explained the importance of kinematic features for plaque analysis for a computer-aided diagnosis (CAD). Fisher discriminant ratio-based feature selection and SVM-based classification were performed. Applying texture features gave 80% of accuracy and kinematic features recorded 88% of accuracy. The accuracy of this proposed CAD has still lots of scope for improvement. AUC, specificity, and sensitivity improved by 0.70, 0.83, and 0.67, respectively (Gastounioti et al., 2015). Vegas-Sánchez-Ferrero et al. (2014) defined a gamma mixture model (GMM) for the subsampled RF images, and their parameters are useful features to identify various plaque tissues. The method outperformed in terms of plaque echogenicity and characteristics. It achieved an accuracy of 95.16% for four-class classifications and 86.56% for three-class classification, which can still be improved (Vegas-Sánchez-Ferrero et al., 2014).

Saba et al. (2021) proposed a classification approach for carotid artery ultrasound images using four machine learning models, one deep learning model, and one transfer learning model. He used the scattering principle of the plaque, where the symptomatic ones are more scattered than the asymptomatic ones (Saba et al., 2021). He achieved stable results for the characterization and classification of the carotid artery ultrasound images.

Classification of the carotid artery images to identify the presence of plaque deposit is performed by machine learning algorithms, CART decision tree, random forest, and logistic regression. Convolutional neural network (CNN)-based deep learning classification and Mobilenet and Capsulenet transfer learning approaches are performed in the carotid artery image database. The performance of these classification methods is analyzed with the true values confirmed by three radiologists.

In this article, section 2 gives the methodology, section 3 describes the results and discussions, and section 4 concludes the article.

METHODOLOGY

This section defines the approach involved in the classification of the carotid artery ultrasound images. Feature extraction and selection are done to obtain the appropriate features. The selected features are given as input to the machine learning classification algorithms, CART decision tree, random forest, and logistic regression. The images are given as input to the CNN, transfer learning algorithms, Mobilenet and Capsulenet. The classification performance measures are used to identify the efficiency of the algorithms.

Figures 1A,B give the sample carotid artery ultrasound images with and without plaque deposit.

Database Creation

Ethical clearance is obtained from the SRM Medical College Hospital and Research Center, Kattankulathur, Tamil Nadu, India, to collect carotid artery ultrasound images. Database of the carotid artery ultrasound B mode images is collected from the Bharat Scans, Chennai and the SRM Medical College Hospital and Research Center, Kattankulathur, Chennai.

Feature Extraction

Machine learning involves high-dimensional data, where the analysis requires a considerable amount of data for learning and testing. The images obtained are denoised by curvelet decomposition to remove speckle and preserve useful edges. Feature reduction minimizes the effects of redundant variables by selecting feature subsets. Choosing the most significant features progresses the classification model performance and reduces over fitting.

Following preprocessing and segmentation of the images, 63 features are taken from the images in the database. A number of 33 texture features, 5 shape features, 10 histogram and correlogram features, and 15 morphology features are extracted from the images. Out of that, 22 most significant features are selected by principal component analysis (PCA) method (Parhizkar et al., 2021).

The most discriminant features from the extracted features are selected based on the following approach. Distance between two classes for every feature is computed as follows for mean m_1, m_2 and standard deviation σ_1, σ_2.

$$\text{distance} = \frac{|m_1 - m_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{1}$$

Features with more distance are those with more significance. From the 65 extracted features, 22 most significant features were selected for the classification task. PCA-based feature selection was performed in addition. The principal components are derived from the eigenvalues. A correlated feature set is converted into uncorrelated ones called principal components by an orthogonal transformation.

The selected features are texture, spatial structure, skewness, kurtosis, histogram, correlogram, histogram of oriented gradient (HOG), Gabor wavelet, angular 2nd moment, shape, sharpness, length irregularity, mean probability density function, gray-scale median, multiregion histogram, arterial wall ROI's randomness, absolute gradient, radian and angular sum of discrete Fourier transform for Fourier power spectrum, coarseness, convexity, connectivity, and plaque volume. The potential features are given as input to the machine learning classification algorithms.

Classification by Machine Learning Algorithms

Proper data preparation, automation and iterative learning, testing, scalability, and ensemble modeling are necessary for a classification algorithm. The classification of the carotid artery images database is performed with the machine learning algorithms, CART decision tree, logistic regression, and random forest algorithm.

Machine learning is to develop a mathematical model built by training the inputs. The inputs are the features selected from the ultrasound image dataset of the carotid artery. The learning experience is generalized so that it can give the correct output for the new image which is not in the database. The generalization of the model is improved by applying a validation set to the trained model. The resulting output and error are given as feedback to the input so that training of the model improves. After many iterations of tuning and training of the model, the trained model is used with new unseen test data to find the performance of the approach (Lundervold and Lundervold, 2019; Latha et al., 2020).

CART Decision Tree

The decision tree is a prediction-based machine learning model with parameters represented in the branches and target outputs represented in the form of leaves. Branch labels are represented by leaves and feature conjunctions that lead to the leaves are represented as branches. Target with continuous values is called regression trees. Classification and regression tree (CART) is a nonparametric decision tree algorithm (Seera and Lim, 2014). Information gain defines how to quantify the quality of the split. For attributes p and q, the information gain I is represented as

$$I(p, q) = -\frac{p}{p+q}\log_2\left(\frac{p}{p+q}\right) - \frac{q}{p+q}\log_2\left(\frac{q}{p+q}\right) \tag{2}$$

To create a tree from the available attributes, entropy is computed. It depends on how much variance the data has.

$$E(A) = \sum_{i=1}^{v} \frac{p_i + q_i}{p+q} I(p, q) \tag{3}$$

The training sets each attribute that is found from the gain. It is the variance between entropy and information gain.

$$\text{Gain} = I(p, q) - E(A) \tag{4}$$

Decision trees can identify the nonlinearity in the dataset and adapt accordingly. The data need not be standardized because a

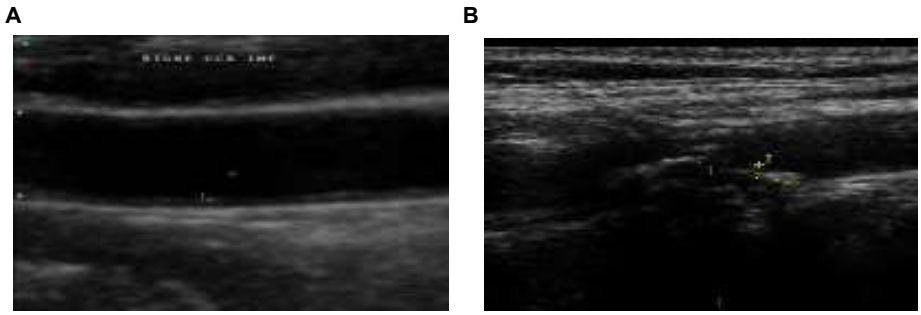


FIGURE 1 | (A) Sample image without plaque deposit **(B)** with plaque deposit.

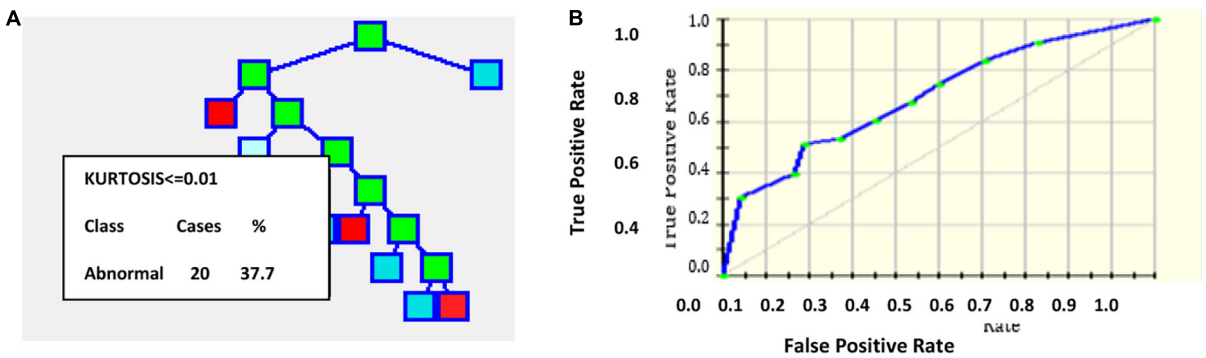


FIGURE 2 | (A) Tree formation for sample 53 images with kurtosis feature **(B)** ROC curve.

distance measure is not involved in the classification. Sigmoid activation is used to get the optimum classification result. The rules of CART and other decision trees are as follows:

1. Based on a variable's value, the splitting criteria for a node are formulated.
2. The stopping criteria are decided when to stop splitting a tree.
3. Final target variable at the end of each node is calculated.

An output of one implies the presence of plaque, and zero represents the absence of plaque in the image with a threshold of 0.5. **Figure 2** gives the results of applying the CART decision tree for the carotid artery ultrasound image database. Using the kurtosis feature, the tree formation for sample 53 images is projected in **Figure 2A**. Kurtosis ≤ 0.01 is separated and branches are formed from that node. **Figure 2B** is the ROC curve for which the AUC is 83.53%, which implies that CART is suitable for disease classification in the carotid artery.

Classification and regression tree is nonparametric and hence is independent on the distribution kind of the input data. The algorithm is not affected by the outliers in the input data. Without strictly following the stopping rule, the tree can be overgrown and can be pruned back to the optimal solution. Fit can be improved using a test set and validation sets. The input variable set can be selected by combining CART with other prediction methods. The drawbacks of CART include variance in the model when a small change in the database is made and imbalanced class data lead to underfit trees.

Logistic Regression

Binary logistics is more suitable for categorical targets with linear or nonlinear decision boundaries, with a threshold fixed. It applies the logistic or sigmoid function. For the curve's maximum value L, steepness parameter or growth rate k and x0 being the midpoint of x, the logistic function is given by

$$f = \frac{L}{1+e^{-k(x-x_0)}} \tag{5}$$

Assuming threshold 0.5, for probability 0.5, class = 1 is assigned. For probability < 0.5, class = 0 is assigned (Barui et al., 2018). The cost function J used is crossentropy since sigmoid activation is used.

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \text{cost}(h_\theta(x^i), (y^i)) \tag{6}$$

Where $\text{cost}(h\theta(x), y) = -\log(h\theta(x))$ for y = 1 and $\text{cost}(h(x), y) = -\log(1 - h(x))$ for y = 0. The natural log of odds called logit which transforms the line into the logistic curve is

$$\log\left(\frac{p(x)}{1-p(x)}\right) = \beta_0 + \beta_1(x) \tag{7}$$

The logistic regression coefficients are found by maximum likelihood estimation. Highly correlated inputs from the database are removed after calculating the pair-wise correlation of the features. It is done to prevent overfit because of multiple highly correlated inputs. The sparsity of the data is also reduced so that the likelihood estimation does not prevent target convergence (Zhang et al., 2018; Javeed et al., 2019; Zhang and Han, 2020). **Figures 3A,B** project the ROC curve and the number of trees with AUC 87.55%.

Random Forest

Random forest is an ensemble classification approach, protecting the structure from being affected by overfitting problems, introduced by Ho in 1995. The tree learners of the random forest follow bootstrap aggregation bagging. Without increasing, bias bootstrapping reduces the variance of the model. The trees are uncorrelated so the prediction of the average of many trees is not noise-sensitive. Bootstrapping gives different input sets for each training time. A forest is created randomly with root, internal, and terminal nodes. Algorithm efficiency improves for a bigger tree. Unlike other decision tree algorithms, random forest decides the root and other nodes randomly.

The classifier is efficient enough to handle missing values and is more suitable for categorical classification. Random forest is created first, and predictions are made from the created forest (Javeed et al., 2019; Wu et al., 2020). Sigmoid activation function is used. Using the random nodes, incorrect labeling can be identified using Gini impurity given by

$$I_G(n) = 1 - \sum_{i=1}^{j} (p_i)^2 \tag{8}$$

The algorithm for random forest creation is as follows.

1. From a total of m feature sets, K features are randomly selected $k < m$.
2. Find node from features after best split point.
3. From the best divided, segregate child node.
4. The above steps are repeated until l number of nodes is achieved.
5. Repeat the above steps for *n* times to achieve *n* nodes.

The prediction that forms the created random forest is done by the below procedure.

1. For each test feature, the rules of the model are applied to get the target.
2. For each predicted target, the votes are estimated.
3. The more voted target is considered the outcome.

Figure 4A projects the error rate which is least for nearly 85 number of trees, then increases, becomes constant, and the next drop is marked in nearly 920 trees. **Figure 4B** gives the ROC curve with AUC 90.63%.

Random forest combines individual tree's decisions and considers the maximum voted one, which makes it one of the best machine learning algorithms. Trees are modeled more diversely, thus implementing all possible models, and obtaining all possible outcomes improves model efficiency. Kernel-induced random forest (KIRF) is followed where trees are built till error no longer reduces. Out of bootstrap (OOB) samples are applied to get the error rate of the random forest by taking the mean of the error from all the bags using all the available features. The drawbacks of the random forest include model complexity, more time consuming than other decision trees, and less intuitive for large decision trees.

Deep Learning Algorithms

Deep learning, which is a class of ANN, extracts the semantic from the images directly, resulting in better classification performance. The deep learning model is built with multisource labeled data and provides more generalized results. The carotid artery ultrasound image classification is performed with a deep learning approach, CNN.

Deep learning is a promising machine learning field that can unravel artificial intelligence problems efficiently. It uses a DNN where the solution depends on the database. Deep learning is superior in terms of nonlinearity, generalization, harmony, fault tolerance, parallelism, and learning. There are undisclosed neural network layers that perform the learning for the available data. Each layer holds a relationship with the next and the previous layers. Deep learning absorbs features and useful representations directly from the raw image bypassing the feature extraction step. This automatic learning of feature representation and learning both happen in the layers.

Due to complexity, the importance of the subject, carotid image analysis using machine learning is not efficient enough and needs a model learnt from a huge number of images. The analysis does not depend on the features extracted manually. The data may be patient-dependent and expert-dependent which may influence the outcomes. Deep learning extracts the hidden feature representations of the images and helps in efficient diagnosis. For example, deep learning algorithms are CNN, DNN, DBM, LSTM networks, and generative adversarial networks (GANs), each having their pros and cons which does not require any preprocessing of data. The extension of CNN called transfer learning algorithms, such as Alexnet, Leenet, Googlenet, and Resnet, has proved their efficiency in the testing phase to a huge extent in terms of complexity.

Deep learning stacks many neuron layers constructing a hierarchical feature representation. The layer count in the model is over 1,000 creating a gigantic model memorizing all features and thus makes more intelligent classification.

Deep learning executes feature engineering on its own by combining and correlating the necessary attributes of the image. Deep learning solves the classification problem end-to-end, which makes the model better than other machine learning approaches. There is a lot of scope of development of deep learning with emerging techniques, such as transfer learning. Other challenges of deep learning are interpretability, trust, data, regulations, and workflow integration.

Convolutional Neural Network

Convolutional neural network is a proven traditional deep learning network based on its translation invariance property

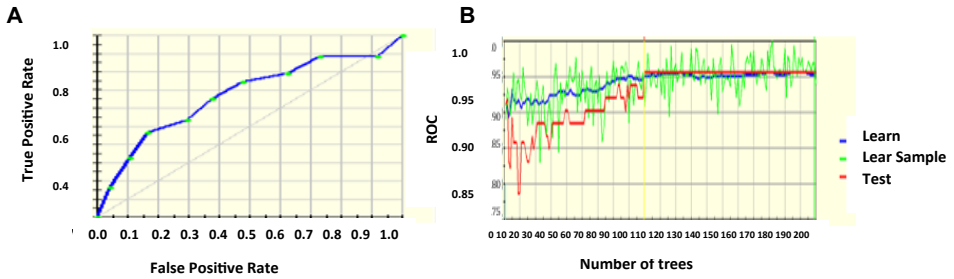


FIGURE 3 | (A) ROC curve **(B)** number of trees with respect to ROC.

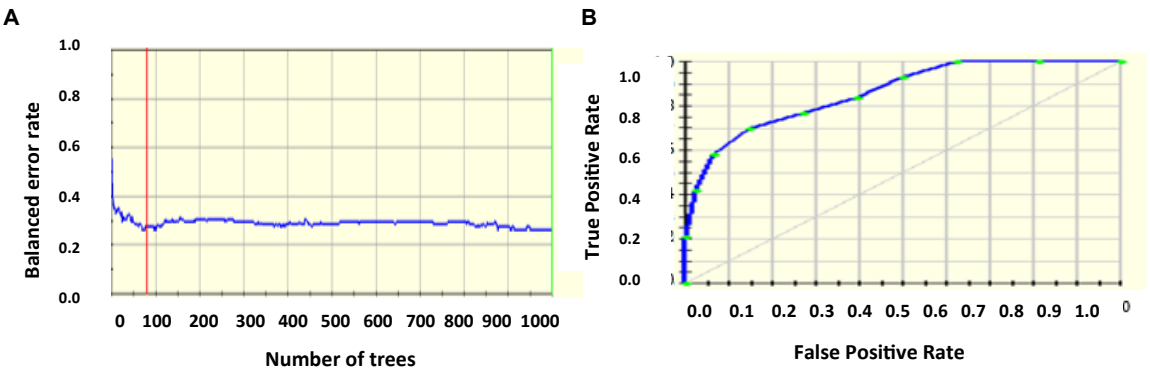


FIGURE 4 | (A) Error rate **(B)** ROC curve.

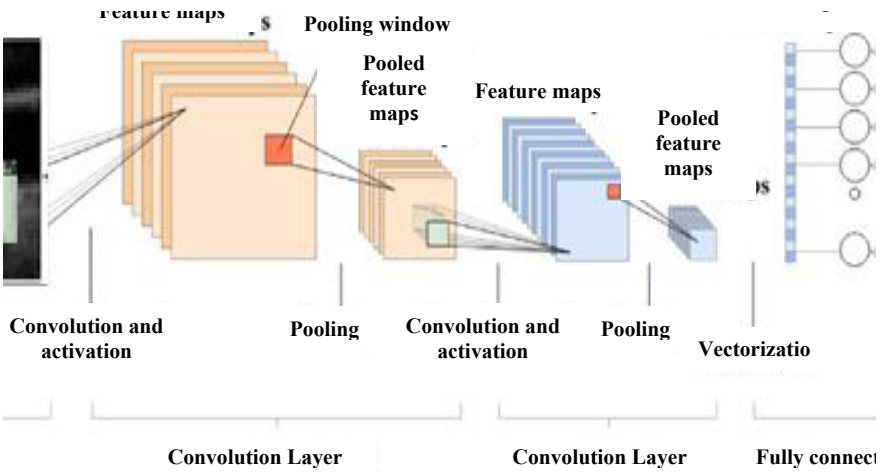


FIGURE 5 | CNN architecture.

and shared weights architecture. All nodes connected to all nodes in the other layers build a much complex system and may be inefficient. CNN uses the domain knowledge of the data preserving the spatial relationship, assembling complex patterns into small, simple patterns (Tajbakhsh et al., 2016).

Rectified linear unit (ReLU) activation function is used for CNN activation. In convolution layer activation, previous layer activations are convolved with parameterized filters of size 3×3. Learning the same weight reduces the complexity of weight calculation for each layer and node. The convolution layer outputs are polled in a pooling layer. For small grids, the polling layer provides single output by max-pooling or average pooling. Translational invariance is achieved after the pooling layer preventing a shift in activation maps because of the shift in the input. Increased stride length convolution leads to downsampled pooling reducing the model complexity.

Based on a stochastic sampling of the neural network, dropout regularization is performed. Different neurons are removed in different iterations leading to different outputs each time. Weights are updated each time to get more optimal results. Activation maps subtracted from the mean and divided by standard deviations for each training batch give batch normalized output (Lundervold and Lundervold, 2019). **Figure 5** gives CNN architecture. The image is directly fed as input to the model. The convolution layer extracts features, such as corners, edges, and colors from the input image. Deeper layers extract more deep features, such as plaque structure, kurtosis, texture of plaque, and nonplaque area. Dominant features from the restricted neighborhood are extracted in the pooling layer.

Max-pooling representation is used, which minimizes computational cost and provides translational in-variation to the internal representation. Alternate convolution and pooling layers are used to reduce the large feature space. Later, layers extract more disease-related features assisting the classification process and improve classification accuracy.

After the convolution and pooling, the data are converted into a column vector, suitable for multilevel fully connected architecture. It is followed by a feed-forward neural network and back-propagation architecture in successive training iterations. Dominant and low-level features are adequately identified and classification proceeds.

Transfer Learning Based on Mobile Network Architecture

A network pretrained on available images can be fine-tuned for the application to be performed. When the source and the target are nearly similar, transfer learning works best in terms of weight updating and optimization compared to random initializations.

Figure 6 gives Mobilenet architecture. The types of transfer learning are positive, negative, and neutral. Learning in a condition facilitating another condition is called positive transfer learning. Learning a task that makes learning another task harder is called negative learning. A learning which does not make a change in another learning is called neutral type of learning. A 1×1 convolution is associated with the depthwise convolution outputs in a pointwise convolution layer. In a single step, inputs and outputs are combined using a convolution filter. Using Mobilenet, computation and model size have drastically reduced. Transfer learning marks fast training, more accurate, and needs fewer data. The significant levels of transfer learning are

1. Full network adaptation—weights are updated from a pretrained network instead of arbitrary initialization and apprise them during the training phase (Wang et al., 2016).
2. Partial network adaptation—network parameters from the pretrained network are initialized and used as such for the first few layers and the last layers are updated for training (Zeng et al., 2017; Hesamian et al., 2019).
3. Zero adaptation—network parameters from a pretrained network are used and are not changed throughout.

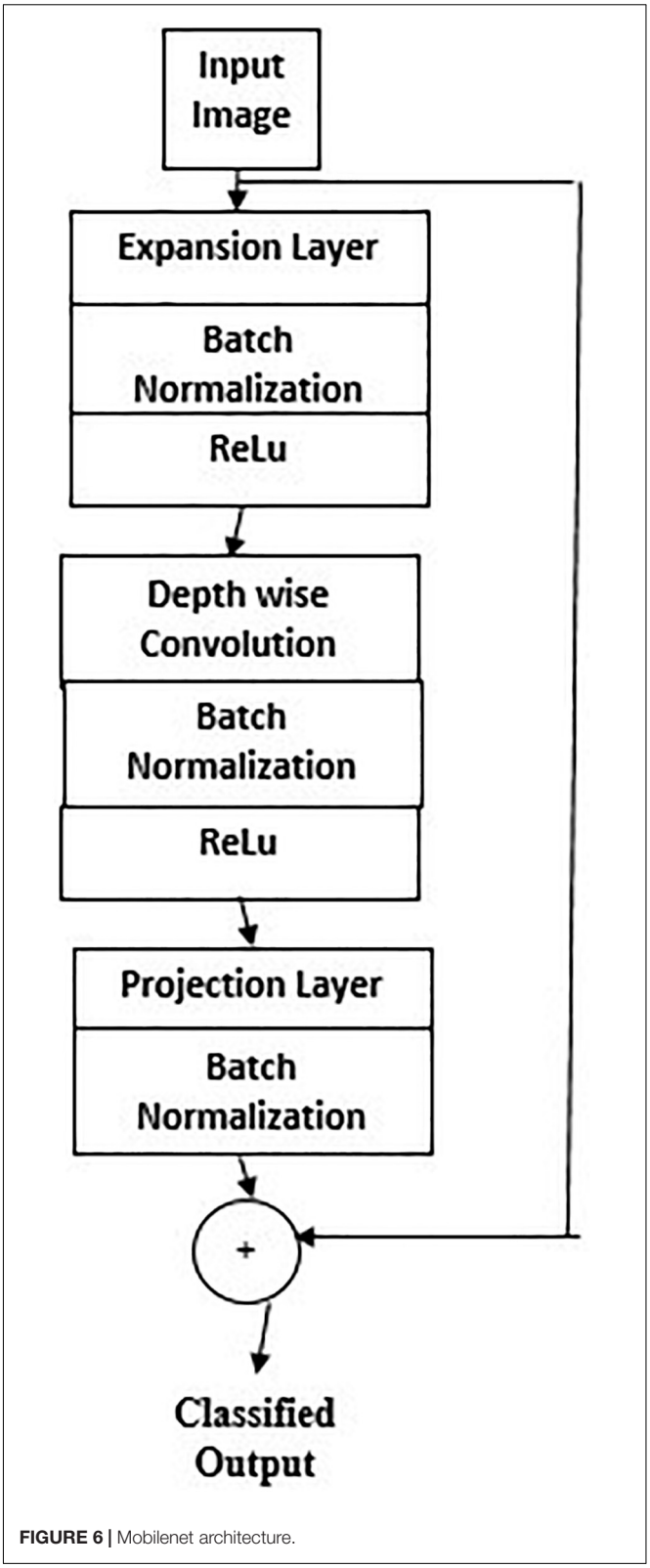


FIGURE 6 | Mobilenet architecture.

Zero adaptation may not be suitable for medical images trained with other organs or general images because they may not have similar properties of the carotid image. In using this carotid database for testing a pretrained network, since the available dataset is small than the training dataset, the following procedure is followed. Overfitting may be a concern because of the small

testing set (Akbarian et al., 2019; Latha et al., 2021). The extracted high-level features may not be similar to the target dataset. The key features of Mobilenet model compared with the CNN model are the following.

1. Most of the pretrained layers near the start of CNN are removed.
2. Instead, fully trained networks equal to the number of classes for the application are included.
3. The newly obtained weights are randomized and replaced instead of the removed network weights.
4. The network is trained to update the weights of the new fully connected layers.

Mobilenet is a family of mobile-first computer vision model for TensorFlow considering restricted data available and suited for embedded applications. The model is small, low latent, and low power designed by google researchers. A width multiplier parameter is introduced to overcome the resource-accuracy tradeoff. The resolution multiplier term reduces the layers' internal structure. ReLU activation function is used.

Figure 7 gives the transfer learning with mobile net architecture, which provides training accuracy 100% and validation accuracy 95%. Though the training performance is less than that of CNN, the validation performance has improved drastically on using mobile net architecture.

Capsulenet

Geoffrey Hinton proposed Capsulenet in 2017, which is a better representation of capsules than convolution. The neuron activities also have a viewpoint variance in addition. CNN requires augmentation and depends more on texture features, which led to these transfer learning approaches. CNN's max-pooling may lose valuable information because of

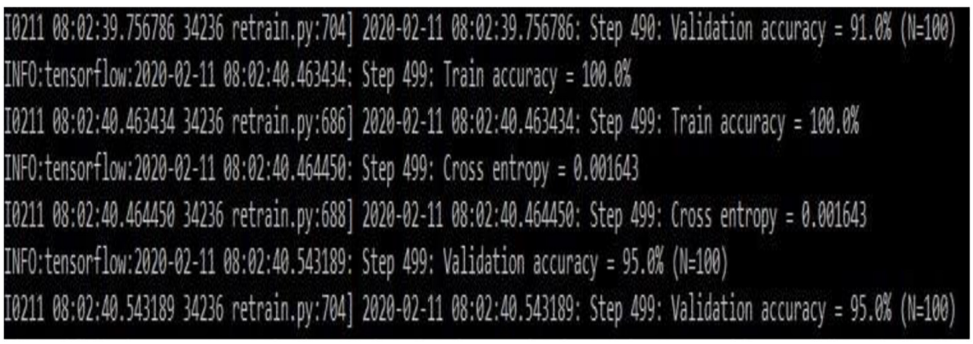


FIGURE 7 | Transfer learning based on the Mobilenet architecture (snapshot of the obtained results).

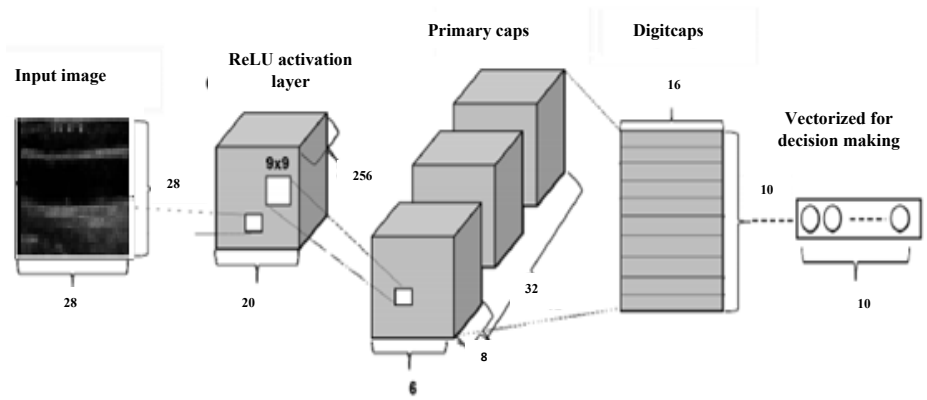


FIGURE 8 | Capsulenet architecture.

TABLE 1 | Confusion matrix of machine learning algorithms.

	CART Decision tree		Logistic regression		Random forest	
	Actual positive (1)	**Actual negative (0)**	**Actual positive (1)**	**Actual negative (0)**	**Actual positive (1)**	**Actual negative (0)**
Predicted positive (1)	123	34	120	27	132	23
Predicted negative (0)	23	181	14	200	8	198

```
Epoch 1/10
250/250 [==============================] - 356s 1s/step - loss: 0.2015 - accuracy: 0.9121 - val_loss: 3.5977 - val_accurac
y: 0.5500
Epoch 2/10
250/250 [==============================] - 264s 1s/step - loss: 0.0047 - accuracy: 0.9997 - val_loss: 5.0452 - val_accurac
y: 0.5500
Epoch 3/10
250/250 [==============================] - 241s 964ms/step - loss: 5.2712e-04 - accuracy: 1.0000 - val_loss: 5.3359 - val_a
ccuracy: 0.5500
Epoch 4/10
250/250 [==============================] - 247s 989ms/step - loss: 2.2240e-04 - accuracy: 1.0000 - val_loss: 5.9795 - val_a
ccuracy: 0.5500
Epoch 5/10
250/250 [==============================] - 242s 966ms/step - loss: 1.2134e-04 - accuracy: 1.0000 - val_loss: 6.4811 - val_a
ccuracy: 0.5500
Epoch 6/10
250/250 [==============================] - 242s 968ms/step - loss: 1.2433e-04 - accuracy: 1.0000 - val_loss: 7.0387 - val_a
ccuracy: 0.5500
Epoch 7/10
250/250 [==============================] - 239s 957ms/step - loss: 5.6958e-05 - accuracy: 1.0000 - val_loss: 7.0544 - val_a
ccuracy: 0.5500
Epoch 8/10
250/250 [==============================] - 251s 1s/step - loss: 4.7950e-05 - accuracy: 1.0000 - val_loss: 7.0233 - val_accu
racy: 0.5500
Epoch 9/10
250/250 [==============================] - 289s 1s/step - loss: 3.1103e-05 - accuracy: 1.0000 - val_loss: 7.2318 - val_accu
racy: 0.5500
Epoch 10/10
250/250 [==============================] - 341s 1s/step - loss: 2.1839e-05 - accuracy: 1.0000 - val_loss: 7.1682 - val_accu
racy: 0.5500
```

FIGURE 9 | CNN model applied to the carotid artery ultrasound image database (snapshot of the obtained results).

poor relationships between hierarchies of simple and complex objects. Capsulenet applies vector activation and outputs which encodes feature transformation information. ReLU activation function is used.

```
Layer (type)                 Output Shape              Param #   Connected to
==================================================================================================
input_1 (InputLayer)         (None, 200, 200, 1)       0
__________________________________________________________________________________________________
conv1 (Conv2D)               (None, 192, 192, 256      20992     input_1[0][0]
__________________________________________________________________________________________________
primarycap_conv2d (Conv2D)   (None, 92, 92, 256)       5308672   conv1[0][0]
__________________________________________________________________________________________________
primarycap_reshape (Reshape) (None, 270848, 8)         0         primarycap_conv
2d[0][0]
__________________________________________________________________________________________________
primarycap_squash (Lambda)   (None, 270848, 8)         0         primarycap_resh
ape[0][0]
__________________________________________________________________________________________________
digitcaps (CapsuleLayer)     (None, 1, 16)             34668544  primarycap_squa
sh[0][0]
__________________________________________________________________________________________________
input_2 (InputLayer)         (None, 1)                 0
__________________________________________________________________________________________________
mask_1 (Mask)                (None, 16)                0         digitcaps[0][0]
                                                                 input_2[0][0]
__________________________________________________________________________________________________
capsnet (Length)             (None, 1)                 0         digitcaps[0][0]
__________________________________________________________________________________________________
decoder (Sequential)         (None, 200, 200, 1)       41534016  mask_1[0][0]
==================================================================================================
Total params: 81,532,224
Trainable params: 81,532,224
```

FIGURE 10 | Capsulenet implementation for the carotid artery database images (snapshot of the obtained results).

Figure 8 gives Capsulenet architecture with ReLU activation. Capsules are convolutions with block nonlinearity and routing. The iterations are slow but require few parameters than CNN. Inside the knowledge representations, Capsulenet builds a better model hierarchy. Capsule structures are added to the CNN model, and the outputs are reused to get more stable higher representations. Max-pooling is used instead of dynamic routing and hence achieves translation invariance. It improves the ability of the network to detect an object even wherever it lies in the image.

RESULTS AND DISCUSSION

Choice of performance measures to evaluate the machine learning algorithms gives hope for its practical use. An unsuitable incorrect measure will mislead to wrong results and a flawed model which is not suitable for the application. The available data are imbalanced, and thus, analyzing more number of metrics assists in proper model selection. It involves comparing the proposed model with an existing model or predicting the class label for a given image set.

Performance Metrics

The classification of a carotid artery ultrasound image as symptomatic or asymptomatic is a binary classification problem. The performance depends on the count of correctly classified samples to their class (true positive (TP)), not belonging to the class, correctly classified as (true negative (TN)), samples misclassified to that class (false positive (FP)), and those that are

TABLE 2 | Performance comparison of carotid artery image classification using machine learning approaches.

Algorithm	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)	F score (%)	AUC (%)
CART Decision Tree	84.21	88.72	78.34	84.25	81.19	83.53
Logistic Regression	88.64	93.46	81.63	89.55	85.41	87.55
Random Forest	91.41	96.11	85.16	94.29	89.49	90.63

misrecognized as belonging to that category (false negative (FN)) (Sokolova and Lapalme, 2009). The overall effectiveness of the model is given by

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{9}$$

The labels class agreement with positive labels in the algorithm is given by

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{10}$$

Positive label identification efficiency is expressed by recall or sensitivity. The relevant data points are identified using. F score measures the relationship between the positive labeled data and that given in the classifier. Specificity explains how effective the model identifies a negative label. FPR is the false alarm probability and TPR is the recall parameter. The model's ability to identify false classification is derived from the area under the ROC curve (AUC). An AUC rate 1 is expected for an ideal classification model. These measures signify the classification model performance.

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{12}$$

$$\text{F score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{13}$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{14}$$

$$\text{AUC} = \left(\frac{1}{2}\right)\left(\frac{\text{TP}}{\text{TP} + \text{FN}} + \frac{\text{TN}}{\text{TN} + \text{FP}}\right) \tag{15}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{16}$$

ReLu activation function is used in the classification models.

TABLE 3 | Performance comparison of carotid artery image classification by deep learning approaches.

Algorithm	Accuracy (%)
CNN	55
Mobilenet	95
Capsulenet Transfer Learning	96.7

Machine Learning

Table 1 gives the confusion matrix of the machine learning algorithms applied in the dataset containing 361 images, out of which 159 are abnormal and 202 are those without any disease indications.

The CART model gives an accuracy of 84.21%, specificity 88.72%, sensitivity 78.34%, and precision of 84.25%. The results prove that the model is useful in identifying the negative cases better than the positive ones. Logistic regression records an accuracy of 88.64% for the carotid database. The obtained specificity is 93.46%, sensitivity is 81.63%, and precision is 89.55%. More number of features added to the logistic regression model will increase the variance in the odds and may lead to overfitting. This reduces the generalization of the model fit. Based on the chi-square test, Hosmer–Lemeshow goodness-of-fit measure can improve model performance. The algorithm that assumes the data is noise-free. Outliers from the training data must be removed to prevent misclassification. Random forest gives an accuracy of 91.41%, specificity 96.11%, sensitivity 85.16%, and precision of 94.29%. The above results prove that random forest is a more accurate classifier than logistic regression and CART decision tree for classifying the carotid artery ultrasound images.

Deep Learning

Convolutional neural network model is applied on ultrasound image database for the classification of the images as with and without plaque deposit. The model achieved training accuracy of 100% and validation accuracy of 55% as given in **Figure 9**. **Figure 10** gives the result of the capsulenet implementation in the database.

Convolutional neural network requires a wide number of data for training the model. Because of the limited number of data, the validation performance is nearly half, though the training is efficient. To overcome this, transfer learning was introduced to perform a deep learning architecture with limited training dataset.

Capsules group neurons and thus require fewer parameters between layers. Pose matrix in Capsulenet defines the rotation and translation of an object, which represents its change in viewpoint. It makes the model better generalized to new viewpoints. The spatial relationship between part of the image and the whole is learnt which makes the image identification simple. It is a viewpoint-dependent neural activity which does not require image normalization and can also identify multiply transformed images (Samiappan and Chakrapani, 2016; Arun et al., 2019; del Mar Vila et al., 2020; Samiappan et al., 2020). Underfitting problem was seen in the classification problem by CNN, which has led to poor performance and generalization.

The carotid artery ultrasound image dataset is small and was not sufficient for a deep learning-based classification.

Initially, 300 training images and 61 validation images were used. Data augmentation methods, such as rotation, flipping, and translation were done to improve the classification accuracy.

Table 2 gives the performance of the three machine learning techniques applied for the carotid artery ultrasound image database. Random forest gives computationally faster and improved performance results compared to CART and logistic regression. Since the dataset was small (361 images), machine learning algorithms were not computationally complex, lags in accuracy of identification of the disease. Capsules group neurons and thus require fewer parameters between layers. Pose matrix captures rotated and translated versions as linear transformations, and so, Capsulenet is better generalized to new viewpoints. The spatial relationship between part of image and the whole is learnt, which makes the image identification simple. Capsulenet achieves accuracy of 96.7%, which is the highest for the carotid artery database images.

The images in the database were flipped to both plane axis rotated to $\pi/4$ axis. **Table 3** gives the performance of the three deep learning techniques applied in the carotid artery image database.

Proposed Capsulenet with max-pooling gives 12.91, 8.33, 5.47, 43.12, and 1.75% improvement in accuracy compared with a CART decision tree, logistic regression, random forest, CNN, and Mobilenet classification algorithms, respectively. Negative transfer is the interference of the previous knowledge in the new learning. It has not affected the classification performance of the carotid artery ultrasound images. It is proved with improved performance measures.

It is proved that deep learning approaches give improved accuracy of 95.7% for Capsulenet compared to other machine learning and deep learning algorithms reported in the literature.

CONCLUSION

A number of 361 images were processed to form a database with the help of radiologists. Extracted features from the database images are applied to the machine learning algorithms CART decision tree, random forest, logistic regression, CNN model, Mobilenet, and Capsulenet transfer learning algorithms for classifying the images as normal or abnormal. Machine learning algorithms were able to perform with an accuracy of 84.21, 88.64, and 91.41%, respectively, for CART, logistic regression, and random forest. Proposed Capsulenet transfer learning approach eliminates the need for large amount of training data. Proposed Capsulenet with max-pooling gives 12.91, 8.33, 5.47, 43.12, and 1.75% improvement in accuracy compared with CART decision tree, logistic regression, random forest, CNN, and Mobilenet classification algorithms, respectively.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

Ethical clearances were obtained from SRM Medical College Hospital and Research Center, India. Ethics Clearance Number: 1736/IEC/2019. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SL performed conceptualization, methodology, design, data collection, data visualization, formal analysis, reviewing, and editing. PM carried out conceptualization, methodology, design, investigation, data collection, data analysis, and writing original draft preparation. KL done conceptualization, methodology, data collection, data visualization, formal analysis, reviewing, and editing. AK involved in data curation, critical analysis, writing, reviewing, and editing. SD contributed in formal analysis, reviewing, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by The Institution of Engineers India, RDDR2016064.

ACKNOWLEDGMENTS

The authors are grateful to the SRM Institute of Science and Technology, the University of Malaya, and the Universiti Sains Islam Malaysia for supporting this research.

REFERENCES

Akbarian, S., Delfi, G., Zhu, K., Yadollahi, A., and Taati, B. (2019). Automated non-contact detection of head and body positions during sleep. *IEEE Access* 7, 72826–72834. doi: 10.1109/ACCESS.2019.2920025

Arun, P. V., Buddhiraju, K. M., and Porwal, A. (2019). Capsulenet-based spatial–spectral classifier for hyperspectral images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 12, 1849–1865. doi: 10.1109/JSTARS.2019.2913097

Azzopardi, C., Camilleri, K. P., and Hicks, Y. (2020). Bimodal automated carotid ultrasound segmentation using geometrically constrained deep neural networks. *IEEE J. Biomed. Health Inform.* 24, 1004–1015. doi: 10.1109/JBHI.2020.2965088

Barui, S., Latha, S., Samiappan, D., and Muthu, P. (2018). SVM pixel classification on colour image segmentation. *J. Phys. Conf. Ser.* 1000:012110. doi: 10.1088/1742-6596/1000/1/012110

Benjamin, J., Muntner, P., Alonso, A., Bittencourt, M. S., Callaway, C. W., Carson, A. P., et al. (2019). Heart disease and stroke statistics- 2019 update a report from

the American Heart Association. *Circulation* 139, e56–e528. doi: 10.1161/CIR.0000000000000659

del Mar Vila, M., Remeseiro, B., Grau, M., Elosua, R., Betriu, A., Fernandez-Giraldez, E., et al. (2020). Semantic segmentation with DenseNets for carotid artery ultrasound plaque segmentation and CIMT estimation. *Artif. Intell. Med.* 103:101784. doi: 10.1016/j.artmed.2019.101784

Farah, A. (2018). *Nutritional and Health Effects of Coffee*. Rio de Janeiro: Federal University of Rio de Janeiro. doi: 10.19103/AS.2017.0022.14

Gastounioti, A., Makrodimitris, S., Golemati, S., Kadoglou, N. P. E., Liapis, C. D., and Nikita, K. S. (2015). A novel computerized tool to stratify risk in carotid atherosclerosis using kinematic features of the arterial wall. *IEEE J. Biomed. Health Inform.* 19, 1137–1145.

Hesamian, M. H., Jia, W., He, X., and Kennedy, P. (2019). Deep learning techniques for medical image segmentation: achievements and challenges. *J. Digit. Imaging* 32, 582–596. doi: 10.1007/s10278-019-00227-x

Javeed, A., Zhou, S., Yongjian, L., Qasim, I., Noor, A., and Nour, R. (2019). An intelligent learning system based on random search algorithm and optimized random forest model for improved heart disease detection. *IEEE Access* 7, 180235–180243. doi: 10.1109/ACCESS.2019.2952107

Latha, S., Samiappan, D., and Kumar, R. (2020). Carotid artery ultrasound image analysis: a review of the literature. *Proc. Inst. Mech. Eng. H J. Eng. Med.* 234, 417–443. doi: 10.1177/0954411919900720

Latha, S., Samiappan, D., Muthu, P., and Kumar, R. (2021). Fully automated integrated segmentation of carotid artery ultrasound images using DBSCAN and affinity propagation. *J. Med. Biol. Eng.* 41, 260–271. doi: 10.1007/s40846-020-00586-9

Lekadir, K., Galimzianova, A., Betriu, A., Del Mar Vila, M., Igual, L., Rubin, D. L., et al. (2017). A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. *IEEE J. Biomed. Health Inform.* 21, 48–55. doi: 10.1109/JBHI.2016.2631401

Loizou, C. P., Pattichis, C. S., Pantziaris, M., Kyriacou, E., and Nicolaides, A. (2017). Texture feature variability in ultrasound video of the atherosclerotic carotid plaque. *IEEE J. Transl. Eng. Health Med.* 5:1800509. doi: 10.1109/JTEHM.2017.2728662

Lundervold, A. S., and Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Deep Learn. Med. Phys.* 29, 102–127. doi: 10.1016/j.zemedi.2018.11.002

Parhizkar, T., Rafieipour, E., and Parhizkar, A. (2021). Evaluation and improvement of energy consumption prediction models using principal component analysis based feature reduction. *J. Clean. Prod.* 279:123866. doi: 10.1016/j.jclepro.2020.123866

Pazinato, D. V., Stein, B. V., de Almeida, W. R., Werneck, R. O., Mendes Júnior, P. R., Penatti, O. A., et al. (2016). Pixel-level tissue classification for ultrasound images. *IEEE J. Biomed. Health Inform.* 20, 256–267. doi: 10.1109/JBHI.2014.2386796

Roy-Cardinal, M., Destrempes, F., Soulez, G., and Cloutier, G. (2019). Assessment of carotid artery plaque components with machine learning classification using homodyned-K parametric maps and elastograms. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 66, 493–504. doi: 10.1109/TUFFC.2018.2851846

Saba, L., Sanagala, S. S., Gupta, S. K., Koppula, V. K., Laird, J. R., Viswanathan, V., et al. (2021). A multicenter study on carotid ultrasound plaque tissue characterization and classification using six deep artificial intelligence models: a stroke application. *IEEE Trans. Instr. Meas.* 70, 1–12. doi: 10.1109/TIM.2021.3052577

Samiappan, D., and Chakrapani, V. (2016). Classification of carotid artery abnormalities in ultrasound images using an artificial neural classifier. *Int. Arab J. Inf. Technol.* 13, 756–762.

Samiappan, D., Latha, S., Rao, T. R., Verma, D., and Sriharsha, C. S. A. (2020). Enhancing machine learning aptitude using significant cluster identification for augmented image refining. *Int. J. Pattern Recognit. Artif. Intell.* 34:2051009. doi: 10.1142/S021800142051009X

Seera, M., and Lim, C. P. (2014). Online motor fault detection and diagnosis using a hybrid FMM-CART model. *IEEE Trans. Neural Netw. Learn. Syst.* 25, 806–812. doi: 10.1109/TNNLS.2013.2280280

Sokolova, M., and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* 45, 427–437. doi: 10.1016/j.ipm.2009.03.002

Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Todd Hurst, R., Kendall, C. B., Gotway, M. B., et al. (2016). Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans. Med. Imaging* 35, 1299–1312. doi: 10.1109/TMI.2016.2535302

Vegas-Sánchez-Ferrero, G., Seabra, J., Rodriguez-Leor, O., Serrano-Vida, A., Aja-Fernández, S., Palencia, C., et al. (2014). Gamma mixture classifier for plaque detection in intravascular ultrasonic images. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 61, 44–61. doi: 10.1109/TUFFC.2014.6689775

Wang, J., MacKenzie, J. D., Ramachandran, R., and Chen, D. Z. (2016). "A deep learning approach for semantic segmentation in histology tissue images," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, eds S. Ourselin, L. Joskowicz, M. Sabuncu, G. Unal, and W. Wells (Cham: Springer), 176–184. doi: 10.1007/978-3-319-46723-8_21

Wu, J., Chen, P., Lin, C., Chen, S., and Shung, K. K. (2020). Breast benign and malignant tumors rapidly screening by ARFI-VTI elastography and random decision forests based classifier. *IEEE Access* 8, 54019–54034. doi: 10.1109/ACCESS.2020.2980292

Zeng, G., Yang, X., Li, J., Yu, L., Heng, P. A., and Zheng, G. (2017). "3D U-net with multi-level deep supervision: fully automatic segmentation of proximal femur in 3D MR images," in *Proceedings of the International Workshop on Machine Learning in Medical Imaging*, eds Q. Wang, Y. Shi, H. I. Suk, and K. Suzuki (Cham: Springer), 274–282. doi: 10.1007/978-3-319-67389-9_32

Zhang, C., Yao, L., Song, S., Wen, X., Zhao, X., and Long, Z. (2018). Euler elastica regularized logistic regression for whole-brain decoding of fMRI data. *IEEE Trans. Biomed. Eng.* 65, 1639–1653. doi: 10.1109/TBME.2017.2756665

Zhang, Z., and Han, Y. (2020). Detection of ovarian tumors in obstetric ultrasound imaging using logistic regression classifier with an advanced machine learning approach. *IEEE Access* 8, 44999–45008. doi: 10.1109/ACCESS.2020.2977962

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.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 17 February 2022
doi: 10.3389/fnagi.2022.848511

Multi-Frequent Band Collaborative EEG Emotion Classification Method Based on Optimal Projection and Shared Dictionary Learning

*Jiaqun Zhu†, Zongxuan Shen† and Tongguang Ni**

School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, China

Affective computing is concerned with simulating people's psychological cognitive processes, of which emotion classification is an important part. Electroencephalogram (EEG), as an electrophysiological indicator capable of recording brain activity, is portable and non-invasive. It has emerged as an essential measurement method in the study of emotion classification. EEG signals are typically split into different frequency bands based on rhythmic characteristics. Most of machine learning methods combine multiple frequency band features into a single feature vector. This strategy is incapable of utilizing the complementary and consistent information of each frequency band effectively. It does not always achieve the satisfactory results. To obtain the sparse and consistent representation of the multi-frequency band EEG signals for emotion classification, this paper propose a multi-frequent band collaborative classification method based on optimal projection and shared dictionary learning (called MBCC). The joint learning model of dictionary learning and subspace learning is introduced in this method. MBCC maps multi-frequent band data into the subspaces of the same dimension using projection matrices, which are composed of a common shared component and a band-specific component. This projection method can not only make full use of the relevant information across multiple frequency bands, but it can also maintain consistency across each frequency band. Based on dictionary learning, the subspace learns the correlation between frequency bands using Fisher criterion and principal component analysis (PCA)-like regularization term, resulting in a strong discriminative model. The objective function of MBCC is solved by an iterative optimization algorithm. Experiment results on public datasets SEED and DEAP verify the effectiveness of the proposed method.

Keywords: cognitive computing, EEG-based emotion classification, multi-frequency band EEG signals, subspace learning, dictionary learning

OPEN ACCESS

Edited by:
Yuanpeng Zhang,
Nantong University, China

Reviewed by:
Lu Chen,
Nanjing Institute of Technology (NJIT), China
Runmin Liu,
Wuhan Sports University, China

***Correspondence:**
Tongguang Ni
ntg@cczu.edu.cn

† These authors have contributed equally to this work

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

Received: *04 January 2022*
Accepted: *27 January 2022*
Published: *17 February 2022*

Citation:
Zhu J, Shen Z and Ni T (2022) Multi-Frequent Band Collaborative EEG Emotion Classification Method Based on Optimal Projection and Shared Dictionary Learning. Front. Aging Neurosci. 14:848511. doi: 10.3389/fnagi.2022.848511

INTRODUCTION

Affective computing focuses on how to actively learn, reason, and perceive the surrounding world, as well as realize a certain level of brain-inspired cognitive intelligence by simulating people's psychological cognitive processes (Aranha et al., 2019; Samsonovich, 2020). Researchers in psychology and neurobiology investigate the changes and relationships in the human physiological

systems that occur during various emotional states and activities (Li et al., 2020). More and more evidences show that with the progress of neuroscience research, there is a connection between human emotional activity and the activity of specific areas of the brain, especially the cerebral cortex and central nervous system. For example, the amygdale is associated with emotions like fear and anxiety in the limbic system of the brain. Anger can activate the left frontal lobe of the brain (Davis and Whalen, 2001). Researchers have also studied the relationship between certain diseases and emotional activities, such as cancer, cardiovascular disease, and depression (Zhao et al., 2018). Wirkner et al. (2017) and Guil et al. (2020) studied the impact of emotional activity on the progression of breast cancer patients. Nurillaeva and Abdumalikova (2021) studied the pathways of communication between the heart and the brain, as well as the relationship between heart rhythm and cognitive and emotional functions. According to the study of Gianaros et al. (2014), there is a link between affective regulation and cardiovascular disease. The author discussed how intense emotional activity and the immune system interact, and how these close interactions affect the treatment of rheumatic cardiovascular disease. Tennant and McLean (2001) associated mood disorders such as anxiety, depression, and anger with coronary heart disease. Authors classified mood disorders as an important risk factor for coronary heart disease, and concluded that mood disorders are frequently associated with coronary heart disease events. Klatzkin et al. (2021) studied the food intake of emotional dieters during various emotional and stress responses. Researchers are also interested in the impact of emotional activities in the business field. According to research on the effect of emotion on commercial advertising, advertisements with emotional expression and influence are easier for consumers to remember, and publicity images with emotional color can influence consumers' access behavior (Shareef et al., 2018). It is clear that research on human emotional activities is important not only in the study and understanding of humanity, but also in medical health and commercial activities. As a result, the study of human emotions, including emotional activity intervention, can be regarded as scientific and practical.

Electroencephalogram activities are closely related to people's psychological attention consciousness and emotional experience. An emotional EEG signal is a physiological electrical signal collected by the human brain in a specific emotional state. EEG signals, as a window into brain thinking activities, cognitive processes, and mental states, are an important technical means for studying brain function and its neural mechanism. Wearable devices placed on the top of the head collect emotional EEG signals. The acquisition electrode's placement position is typically determined using the international standard 10–20 and other systems. Researchers in the field of artificial intelligence study the relationship between emotional activities caused by internal and external stimuli and the content of stimuli. Machine learning technology in artificial intelligence is widely used in EEG signals-based emotion classification. For example, Liu et al. (2020) developed a multi-level features guided capsule network to describe the internal relationship of multiple EEG signal channels. The advantage of this model is that different levels of feature mapping are integrated during the process of forming the primary capsule, which can improve feature representation ability. Zhong et al. (2020) proposed a regularized graph neural network to mine both local and global relationships between various EEG channels. This method can alleviate the problem of time dependence in emotional process. Ni et al. (2021) developed a domain adaptation sparse representation classification model to alleviate the problem of insufficient training data in the new scene. This method employed the discriminative knowledge of historical data or related data to aid in establishing the classification model of the current scene.

According to intra-band correlation with a distinct psychological state, the EEG signals can be split into five frequency bands. Different frequency band EEG signals reflect the different states of brain state. **Table 1** briefly describes the information of five frequent bands of EEG signals (Gu et al., 2021a; Shen et al., 2021). Many scholars have studied EEG signals in different frequency bands. Mohammadi et al. (2017) used wavelet transform to decompose EEG signals into five sub-band signals, then extracted entropy and energy features from each sub-band signal and sent them to support vector machine and *k*-nearest neighbor, respectively. Li and Lu (2009) proposed a frequency band search method to find the best frequency band for emotion classification. According to their findings, the gamma frequency band is appropriate for EEG-based emotion classification using still images as stimuli. Zheng and Lu (2015) built a Multi-frequent band emotion recognition classifier using deep neural networks. This study had shown that the beta and gamma bands contained more discriminative information for emotion classification. Li et al. (2018) used the hierarchical deep learning model to train numerous classifiers on EEG signals. They verified that high-frequency bands played the most important role in emotion classification. Yang et al. (2018) developed a 3D representation of signal segment to extract representative features on bands. They integrated multiple frequency bands and used the constructed 3D signal cube as model input. Li et al. (2019) developed a sparse linear regression model using the technologies of graph regularization and sparse regularization. The authors compared the effects of different frequency band signals in emotion recognition on various EEG datasets.

Because there are internal relationships and differences between different frequency bands, a new learning method is required to make full use of the information in multi-frequency band data. Despite extensive research on the use of different frequency bands of EEG signals for emotion recognition, one traditional strategy is to directly concatenate features from

TABLE 1 | The basic information of five frequent bands of EEG signals.

Patterns	Frequency	Brain state
Delta (δ)	1–3 Hz	Slowest "sleep waves"
Theta (θ)	4–7 Hz	Light meditation and sleeping
Alpha (α)	8–13 Hz	Closing the eyes, relaxation
Beta (β)	14–30 Hz	Waking consciousness and reasoning waves
Gamma (γ)	30–100 Hz	Sensory and high-level information processing

multi-frequent bands in high dimensional space and consider this single feature vector as the model's input. Obviously, this strategy does not account for the complementarity and consistency of the data in each frequency band.

Our previous work named as optimized projection and Fisher discriminative dictionary learning (OPFDDL) (Gu et al., 2021a) extracted multi-frequent band EEG features in the optimal sparse representation subspace, and adopted the Fisher discrimination criterion to build a discriminative classifier. This method did not directly concatenate the features of each frequency band, but regarded each band signal as an independent feature. It incorporated the band-correlation knowledge into a dictionary learning model by learning independent projection matrices for each frequency band signal. Inspired of this work, we further use multi-frequent band shared information to exploit the intrinsic knowledge of EEG signals and achieve correlation modeling of multiple band data. Thus, in this study we propose a multi-frequent band collaborative EEG emotion classification method based on optimal projection and shared dictionary learning (MBCC). We construct a projection matrix for each frequency band. The projection matrix is composed of a common shared matrix (called shared component) and a frequency band-specific matrix (called specific component). The shared matrix well reflects the relationship between frequency bands. The EEG signal of each frequency band is projected to the subspace through the projection matrix, and the dictionary shared by each frequency band is learned in the subspace. The corresponding sparse representation is then obtained from the new data features using dictionary learning. According to Fisher's criterion, the MBCC method ensures that the coding reconstruction errors of the same class are as small as possible, while the coding reconstruction errors of different classes are as large as possible. Considering the information available in the original domain should not be lost in the projection space, we provide a regularization term similar to principal component analysis (PCA) that can retain discriminative knowledge to improve the discrimination ability of the model. An efficient alternating iterative optimization algorithm is designed to solve the proposed model. The experiment yielded good classification results on the public EEG emotion datasets SEED (Zheng and Lu, 2015) and DEAP (Koelstra et al., 2011).

The advantages of MBCC are as follows: (1) An effective discriminative dictionary is trained using the dictionary learning model framework by capturing common shared feature information from multi-frequency band data. The first correlation between data in multiple frequency bands is represented by the common shared dictionary. It creates a link between data from different frequency bands in order to obtain a new feature representation of EEG data. (2) Take into account the complementarity and difference of frequency band data, the projection matrix of each frequency band has the common shared and independent components. The common shared component reflects the second correlation between multiple frequency bands and can keep each frequency band consistent. (3) To assess the model's discriminative ability, the Fisher criterion based on coding error is introduced in the projection space. Furthermore, the PCA-like regularization term based on the common shared projection component contributes to obtain more discriminative sparse coding.

BACKGROUND

Let $\mathbf{Z} = [\mathbf{z}_1, ..., \mathbf{z}_n] \in \mathbf{R}^{d\times n}$ be a set of d-dimensional n training signals. The traditional dictionary learning is to learn a dictionary matrix to sparsely represent the EEG signals Z. The problem of dictionary learning (Jiang et al., 2013; Gu et al., 2021b) is,

$$\begin{aligned} &\min_{\mathrm{D,A}} ||\mathbf{Z} - \mathbf{DA}||_F^2 + \lambda||\mathbf{A}||_1, \\ &s.t.\ \forall i,\ ||\mathbf{d}_i||_0 = 1, \end{aligned} \tag{1}$$

where $D = [\mathbf{d}_1, ..., \mathbf{d}_k] \in \mathbf{R}^{d\times K}$ is the learned dictionary, K is the dictionary size. $\mathbf{A} = [\mathbf{a}_1, ..., \mathbf{a}_n] \in \mathbf{R}^{K\times n}$ is the sparse coding coefficient matrix. The first term in Eq. (1) is to minimize the reconstruction errors of **Z**. The second term is the sparsity constraints.

In our previous work OPFDDL method (Gu et al., 2021a), $\mathbf{Z}^m = [\mathbf{z}_1^m, ..., \mathbf{z}_{n_m}^m] \in \mathbf{R}^{d\times n_m}$is the class m frequent band signal set, $m = 1, 2, ..., M$, $n = \sum_m n_m$. By introducing the frequent band specific projection matrix $\mathbf{G}^m \in \mathbf{R}^{d\times p}$, each training signal $\mathbf{z}_j^m$ is projected into a low-dimensional space, as $\mathbf{G}^m\mathbf{z}_j^m$. Suppose $\mathbf{S}_w^m$ and $\mathbf{S}_b^m$ are within-class and between-class reconstruction errors of the m-th frequent band signals, respectively. $\boldsymbol{S}_w^m$ and $\boldsymbol{S}_b^m$ are defined as,

$$\begin{aligned} \mathbf{S}_w^m &= Tr(\textstyle\sum_{j=1}^{n_m}((\mathbf{G}^m)^T\mathbf{z}_j^m - (\mathbf{G}^m)^T\mathbf{D}\delta(\mathbf{a}_j^m)) \times ((\mathbf{G}^m)^T\mathbf{z}_j^m \\ &\quad -\mathbf{G}^{mT}\mathbf{D}\delta(\mathbf{a}_j^m))^T) \\ &= Tr(\mathbf{G}^{mT}\mathbf{W}_w^m\mathbf{G}^m) \end{aligned} \tag{2}$$

where $\mathbf{W}_w^m = \sum_j^{n_m}(\mathbf{z}_j^m - \mathbf{D}\delta(\mathbf{a}_j^m)) \times (\mathbf{z}_j^m - \mathbf{D}\delta(\mathbf{a}_j^m))^T$. The function $\delta(\mathbf{a}_j^m)$ returns the coding coefficients consistent with the class of $\mathbf{z}_j^m$.

$$\begin{aligned} \mathbf{S}_b^m &= Tr(\textstyle\sum_j^{n_m}((\mathbf{G}^m)^T\mathbf{z}_j^m - (\mathbf{G}^m)^T\mathbf{D}\xi(\mathbf{a}_j^m)) \times ((\mathbf{G}^m)^T\mathbf{x}_j^m \\ &\quad -\mathbf{G}^{mT}\mathbf{D}\xi(\mathbf{a}_j^m))^T) \\ &= Tr((\mathbf{G}^m)^T\mathbf{W}_b^m\mathbf{G}^m) \end{aligned} \tag{3}$$

where $\mathbf{W}_b^m = \sum_j^{n_m}(\mathbf{z}_j^m - \mathbf{D}\xi(\mathbf{a}_j^m)) \times (\mathbf{z}_j^m - \mathbf{D}\xi(\mathbf{a}_j^m))^T$. The function $\xi(\mathbf{a}_j^m)$ returns the coding coefficients not consistent with the class of $\mathbf{z}_j^m$.

According to the classification rule of Fisher criterion (Peng et al., 2020; Zhang et al., 2021), the OPFDDL method proposes the discriminative model on M frequent bands in the projection space,

$$\begin{aligned} &\min_{\mathrm{G}^m,\mathrm{D}} \frac{\sum_m Tr(\mathbf{G}^{mT}\mathbf{W}_w^m\mathbf{G}^m)}{\sum_m Tr(\mathbf{G}^{mT}\mathbf{W}_b^m\mathbf{G}^m)}, \\ &s.t.\ (\mathbf{G}^m)^T(\mathbf{G}^m) = \mathbf{I},\ m = 1, 2, ..., M \end{aligned} \tag{4}$$

Then the matrices $\tilde{\mathbf{G}}$, $\tilde{\mathbf{W}}_w$, and $\tilde{\mathbf{W}}_b$ are defined as $\tilde{\mathbf{G}} = [\mathbf{G}^1, \mathbf{G}^2, ..., \mathbf{G}^M]$, $\tilde{\mathbf{W}}_w = \begin{bmatrix} \mathbf{W}_w^1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{W}_w^M \end{bmatrix}$, and

$\tilde{W}_b = \begin{bmatrix} \mathbf{W}_b^1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{W}_b^M \end{bmatrix}$. With these definitions, the objective function of OPFDDL has the following form,

$$\begin{aligned} &\min_{\tilde{\mathbf{G}},\mathbf{D},\lambda} \lambda^2 Tr(\tilde{\mathbf{G}}^T \tilde{\mathbf{W}}_w \tilde{\mathbf{G}}) - \lambda Tr(\tilde{\mathbf{G}}^T \tilde{\mathbf{W}}_b \tilde{\mathbf{G}}), \\ &s.t.\ \tilde{\mathbf{G}}^T \tilde{\mathbf{G}} = \mathbf{I}, \end{aligned} \tag{5}$$

where λ is the adaptive weight parameter.

The training procedure of OPFDDL is given by **Algorithm 1**.

ALGORITHM 1 | The OPFDDL algorithm.

Repeat

1. Calculate the coding coefficients by solving the following problem:

$$\min_{\mathrm{A}} ||\tilde{\mathbf{G}}^T \mathbf{Z} - \mathbf{DA}||_F^2 + \lambda ||\mathbf{A}||_1 \tag{6}$$

2. Calculate the projection matrix by solving the following problem:

$$(\lambda^2 \tilde{\mathbf{W}}_w - \lambda \tilde{\mathbf{W}}_b)\tilde{\mathbf{G}} = \gamma \tilde{\mathbf{G}} \tag{7}$$

3. Calculate the dictionary **D** by:

$$\mathbf{D} = \mathbf{D} - \frac{\lambda_D}{n} \sum_k \frac{\partial L(\tilde{\mathbf{Z}}_k)}{\partial \mathbf{D}}, k = 1, \ldots, d \tag{8}$$

$$\frac{\partial L(\tilde{\mathbf{Z}}_k)}{\partial \mathbf{D}} = 2\tilde{\mathbf{G}}\tilde{\mathbf{G}}^T \mathbf{D}(\mathbf{\Lambda}\mathbf{\Lambda}^T + \mathbf{H}\mathbf{H}^T) - 2\tilde{\mathbf{G}}\tilde{\mathbf{G}}^T \tilde{\mathbf{Z}}_k(\mathbf{\Lambda}^T + \mathbf{H}^T) \tag{9}$$

$$\mathbf{\Lambda} = [\delta^1, \delta^2, ..., \delta^M], \mathbf{H} = [\zeta^1, \zeta^2, ..., \zeta^M] \tag{10}$$

4. Calculate the adaptive weightλ by:

$$\lambda = \frac{tr(\tilde{\mathbf{G}}^T \tilde{\mathbf{W}}_b \tilde{\mathbf{G}})}{2tr(\tilde{\mathbf{G}}^T \tilde{\mathbf{W}}_w \tilde{\mathbf{G}})} \tag{11}$$

Until convergence

MULTI-FREQUENT BAND COLLABORATIVE EEG EMOTION CLASSIFICATION METHOD BASED ON OPTIMAL PROJECTION AND SHARED DICTIONARY LEARNING

Objective Function of MBCC

The OPFDDL method can be regarded as the baseline algorithm of MBCC. The primary distinction between the MBCC method and OPFDDL is that, although OPFDDL also employs a projection matrix to project each frequency band to the subspace, the correlation between projection matrices is weak. The common shared component defined in MBCC is a key part of multi-frequent band collaborative learning. In addition, according to the consistency principle, the PCA-like regularization term in the shared potential space further captures the discriminative information contained among multiple frequency bands. Thus, the MBCC method can balance discriminative knowledge and multi-frequent band correlation in the projection space.

We look for a projection matrix in the MBCC method to project the data from d-dimensional space to p-dimensional space. This study assumes that the projection matrix $\boldsymbol{G}^m \in \boldsymbol{R}^{d\times p}$ for each frequency band has two parts: the shared component $\boldsymbol{G}^0 \in \boldsymbol{R}^{d\times p}$, which is a common shared matrix that reflects the correlation between different frequency bands, and the band specific component $\tilde{\boldsymbol{G}}^m \in \boldsymbol{R}^{d\times p}$, which is the projection matrix for each frequency band. The matrix is equal to the sum of the shared component and the band specific component,

$$\mathbf{G}^m = (1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m, \tag{12}$$

where σ ∈ [0, 1]is the balance parameter. When σ = 1, the projection matrix $\mathbf{G}^m$ is degenerated into the band specific matrix $\tilde{\mathbf{G}}^m$, which is equivalent to the projection matrix in the OPFDDL method. When σ = 0, the model only learns the common shared matrix.

The projection of the signal in each frequency band is represented as,

$$(\mathbf{G}^m)^T \mathbf{z}_j^m = ((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)^T \mathbf{z}_j^m. \tag{13}$$

The within-class reconstruction error of the m-th frequent band in the projected space can be represented as

$$\begin{aligned} J_w^m &= Tr(\textstyle\sum_j^{n_m} \sum_k^p [(1-\sigma)\mathbf{G}^0(:,k)^T(\mathbf{z}_j^m - \mathbf{D}\delta(\mathbf{a}_j^m)) + \\ &\sigma\tilde{\mathbf{G}}^m(:,k)^T(\mathbf{z}_j^m - \mathbf{D}\delta(\mathbf{a}_j^m))]^2) \\ &= Tr\left(((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)^T \mathbf{W}_w^m ((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)\right). \end{aligned} \tag{14}$$

The between-class reconstruction error of the m-th frequent band in the projected subspace can be represented as

$$\begin{aligned} J_b^m &= Tr(\textstyle\sum_j^{n_m} \sum_k^p [(1-\sigma)\mathbf{G}^0(:,k)^T(\mathbf{z}_j^m - \mathbf{D}\xi(\mathbf{a}_j^m)) + \\ &\sigma\tilde{\mathbf{G}}^m(:,k)^T(\mathbf{z}_j^m - \mathbf{D}\xi(\mathbf{a}_j^m))]^2) \\ &= Tr\left(((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)^T \mathbf{W}_b^m ((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)\right). \end{aligned} \tag{15}$$

Thus, the Fisher criterion of all frequent bands is written as,

$$\min_{\mathbf{D},\mathrm{G}^0,\tilde{\mathbf{G}}^m} \frac{\sum_m Tr\left(((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)^T \mathbf{W}_w^m ((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)\right)}{\sum_m Tr\left(((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)^T \mathbf{W}_b^m ((1-\sigma)\mathbf{G}^0 + \sigma\tilde{\mathbf{G}}^m)\right)}. \tag{16}$$

Because different frequent band data describe the same object, there must be an internal connection between them. The model maximizes the commonality of multiple frequent band data in the shared projection space using the consistency principle. When projecting the data from multiple bands to the optimal subspace, we need to preserve the discriminative information available in the original space. To solve this problem, we use a PCA-like regularization term as follows,

$$J(\mathbf{G}^0) = \min_{\mathrm{G}^0} \sum_m ||\mathbf{Z}^m - \mathbf{G}^0(\mathbf{G}^0)^T \mathbf{Z}^m||_F^2. \tag{17}$$

Ignoring the constant terms in $J(\mathbf{G}^0)$, Eq. (17) can be represented as,

$$\begin{aligned} J(\mathrm{G}^0) &= -\min_{\mathrm{G}^0} \textstyle\sum_m Tr\left(((\mathbf{G}^0)^T\mathbf{Z}^m)((\mathbf{G}^0)^T\mathbf{Z}^m)^T\right) \\ &= -\min_{\mathrm{G}^0} \textstyle\sum_m Tr\left((\mathbf{G}^0)^T\mathbf{Z}^m(\mathbf{Z}^m)^T\mathbf{G}^0\right). \end{aligned} \tag{18}$$

Let $\mathbf{\Theta}^m = \mathbf{Z}^m(\mathbf{Z}^m)^T$, Eq. (18) can be written as,

$$J(\mathbf{G}^0) = -\min_{\mathrm{G}^0} \sum_m Tr\left((\mathbf{G}^0)^T\mathbf{\Theta}^m\mathbf{G}^0\right). \tag{19}$$

Combined the Fisher criterion and PCA-like regularization term, the objective function of MBCC is,

$$\begin{aligned} &\min_{\mathrm{G}^0,\tilde{\mathrm{G}}^m,\mathrm{D}} \frac{\begin{array}{c}\sum_m Tr(((1-\sigma)\mathbf{G}^0+\sigma\tilde{\mathbf{G}}^m)^T\mathbf{W}_w^m((1-\sigma)\mathbf{G}^0 \\ +\sigma\tilde{\mathbf{G}}^m)) - \alpha\sum_m Tr((\mathbf{G}^0)^T\mathbf{\Theta}^m\mathbf{G}^0)\end{array}}{\begin{array}{c}\sum_m Tr(((1-\sigma)\mathbf{G}^0+\sigma\tilde{\mathbf{G}}^m)^T \\ \mathbf{W}_b^m((1-\sigma)\mathbf{G}^0+\sigma\tilde{\mathbf{G}}^m))\end{array}}, \\ &s.t.\ \forall m,\ \left((1-\sigma)\mathbf{G}^0+\sigma\tilde{\mathbf{G}}^m\right)^T\left((1-\sigma)\mathbf{G}^0+\sigma\tilde{\mathbf{G}}^m\right)=\mathbf{I}. \end{aligned} \tag{20}$$

The projection matrix is orthogonal and it will result in an efficient procedure for optimization. We can see that the dictionary learned in the MBCC method have the stronger discriminative ability.

Define $\mathbf{G} = [\mathbf{G}^0; \mathbf{G}^1; ..., \mathbf{G}^M] \in \mathbf{R}^{(M+1)d\times p}$, $\mathbf{\Omega}^m = [(1-\sigma)\mathbf{I}_d, \sigma\mathbf{I}_d, ..., \sigma\mathbf{I}_d] \in \mathbf{R}^{d\times(M+1)d}$, $\mathbf{\Delta}^m = [\mathbf{I}_d, \mathbf{0}_{d\times d}, ..., \mathbf{0}_{d\times d}] \in \mathbf{R}^{d\times(M+1)d}$, $\mathbf{\Lambda} = \sum_m^M(\mathbf{\Omega}^m)^T\mathbf{W}_w^m\mathbf{\Omega}^m$, $\mathbf{\Theta} = \sum_m^M(\mathbf{\Delta}^m)^T\mathbf{\Theta}^m\mathbf{\Delta}^m$, $\mathbf{H} = \sum_m^M(\mathbf{\Omega}^m)^T\mathbf{W}_b^m\mathbf{\Omega}^m$, Eq. (20) is equivalent to,

$$\begin{aligned} &\min_{\mathrm{G,D}} \frac{Tr\left(\mathbf{G}^T\mathbf{\Lambda}\mathbf{G}\right) - \alpha Tr\left(\mathbf{G}^T\mathbf{\Theta}\mathbf{G}\right)}{Tr\left(\mathbf{G}^T\mathbf{H}\mathbf{G}\right)}, \\ &s.t.\ \mathbf{G}^T\mathbf{G} = \mathbf{I}. \end{aligned} \tag{21}$$

By combining the two terms on the numerator, we can get,

$$\begin{aligned} &\min_{\mathrm{G,D}} \frac{Tr\left(\mathbf{G}^T(\mathbf{\Lambda}-\alpha\mathbf{\Theta})\mathbf{G}\right)}{Tr\left(\mathbf{G}^T\mathbf{H}\mathbf{G}\right)}, \\ &s.t.\ \mathbf{G}^T\mathbf{G} = \mathbf{I}. \end{aligned} \tag{22}$$

Optimization

It is not easy to directly solve the variables **G** and **D** in the objective function. Therefore, we will take the alternative iterative optimization scheme to decompose the original problem into two sub-optimization problems.

Update **G**. For the given dictionary **D**, there must be a minimum ρ, which satisfies the following formulation,

$$\frac{Tr\left(\mathbf{G}^T(\mathbf{\Lambda}-\alpha\mathbf{\Theta})\mathbf{G}\right)}{Tr\left(\mathbf{G}^T\mathbf{H}\mathbf{G}\right)} \geq \rho, \tag{23}$$

We have $F(\rho) = \min_{\mathrm{G}} Tr\left(\mathbf{G}^T(\mathbf{\Lambda}-\alpha\mathbf{\Theta})\mathbf{G}\right) - \rho Tr\left(\mathbf{G}^T\mathbf{H}\mathbf{G}\right)$.

As a result, we can define the function of ρ by,

$$Tr\left(\mathbf{G}^T(\mathbf{\Lambda}-\alpha\mathbf{\Theta})\mathbf{G}\right) - \rho Tr\left(\mathbf{G}^T\mathbf{H}\mathbf{G}\right) \geq 0, \tag{24}$$

According to Zhang et al. (2017), (1) $F(\rho)$ is a decreasing function of ρ. (2) $F(\rho) = 0$ if $\rho = \rho^*$. In addition, the minimum ρ always exists.

Then ρ can be updated by,

$$\begin{aligned} &\rho = \rho + \lambda_\rho\frac{F(\rho)}{F'(\rho)}, \\ &F'(\rho) = -Tr(\mathbf{G}^T\mathbf{H}\mathbf{G}), \end{aligned} \tag{25}$$

where λ_ρ is the learning rate.

With the fixed ρ and **D**, the objective function of **G** is,

$$\begin{aligned} &\min_{\mathrm{G}} Tr\left(\mathbf{G}^T(\mathbf{\Lambda}-\alpha\mathbf{\Theta}-\rho\mathbf{H})\mathbf{G}\right), \\ &s.t.\ \mathbf{G}^T\mathbf{G} = \mathbf{I}, \end{aligned} \tag{26}$$

The optimization of **G** can be solved by the following eigenvalue decomposition,

$$(\mathbf{\Lambda}-\alpha\mathbf{\Theta}-\rho\mathbf{H})\mathbf{G} = \gamma\mathbf{G}. \tag{27}$$

The columns of the matrix **G** are the eigenvectors with respect to the first p minimum eigenvalues of Eq. (27).

Update **D**. With the fixed **G**, the objective function of **D** is,

$$\min_{\mathrm{D}} \frac{Tr(\mathbf{G}^T\mathbf{\Lambda}\mathbf{G})}{Tr(\mathbf{G}^T\mathbf{H}\mathbf{G})}, \tag{28}$$

Let $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, ..., \mathbf{D}_C]$ be the learned dictionary, and $\mathbf{D}_j$ is the j-th class sub-dictionary. The Eq. (28) can be re-written as,

$$J(\mathbf{D}_j) = \min_{\mathrm{D}_j} \sum_{j=1}^{C} \frac{Tr(\mathbf{G}^T\mathbf{\Lambda}_j\mathbf{G})}{Tr(\mathbf{G}^T\mathbf{H}\mathbf{G})}, \tag{29}$$

where $\mathbf{\Lambda}_j = \sum_{m=1}\sum_{j\neq s}^{c}(\mathbf{Z}_j^m - \mathbf{D}_s\mathbf{\Gamma}_{j,s}^m) \times (\mathbf{Z}_j^m - \mathbf{D}_s\mathbf{\Gamma}_{j,s}^m)^T$, $\mathbf{H} = \sum_{m=1}\sum_{j=1}^{c}(\mathbf{Z}_j^m - \mathbf{D}_j\mathbf{\Gamma}_{j,j}^m) \times (\mathbf{Z}_j^m - \mathbf{D}_j\mathbf{\Gamma}_{j,j}^m)^T$. $\mathbf{\Gamma}_{j,s}^m$ and $\mathbf{\Gamma}_{j,j}^m$ are the coding coefficient matrices corresponding to classes s and j of the m-th frequent band, respectively, where $s \neq j$.

$\mathbf{D}_j$ can be updated by gradient descent method, in which $\mathbf{D}_j$ is computed as,

$$\begin{aligned} &\mathbf{D}_j = \mathbf{D}_j + \eta\partial J(\mathbf{D}_j), \\ &\partial J(\mathbf{D}_j) = \frac{\partial J(\mathbf{D}_j)}{\partial \mathbf{\Lambda}_j}\frac{\partial \mathbf{\Lambda}_j}{\partial \mathbf{D}_j} + \frac{\partial J(\mathbf{D}_j)}{\partial \mathbf{H}}\frac{\partial \mathbf{H}}{\partial \mathbf{D}_j}. \end{aligned} \tag{30}$$

There is no connection between $\mathbf{\Lambda}_j$ and $\mathbf{D}_j$, i.e., $\frac{\partial \mathbf{\Lambda}_j}{\partial \mathbf{D}_j} = 0$. Therefore, we only need compute $\frac{\partial J(\mathbf{D}_j)}{\partial \mathbf{H}}\frac{\partial \mathbf{H}}{\partial \mathbf{D}_j}$.

$$\frac{\partial J(\mathbf{D}_j)}{\partial \mathbf{H}} = \frac{-Tr(\mathbf{G}^T\mathbf{\Lambda}_j\mathbf{G})(\mathbf{G})^T\mathbf{G}}{\left(Tr(\mathbf{G}^T\mathbf{H}\mathbf{G})\right)^2}, \tag{31}$$

$$\frac{\partial \mathbf{H}}{\partial \mathbf{D}_j} = (\mathbf{\Gamma}_{j,j}^m)^T(\mathbf{D}_j\mathbf{\Gamma}_{j,j}^m - \mathbf{Z}_j^m). \tag{32}$$

Update **A**. With the fixed **D** and **G**, the sparse coding coefficient matrix **A** can be computed as,

$$\min_{\mathrm{A}} ||\mathbf{G}^T\mathbf{Z} - \mathbf{D}\mathbf{A}||_F^2 + \lambda||\mathbf{A}||_1, \tag{33}$$

Since $\mathbf{A}$ is differentiable, it can be obtained by,

$$\mathbf{A} = (\mathbf{D}^T\mathbf{D} + \lambda\mathbf{I})^{-1}(\mathbf{D}^T\mathbf{G}^T\mathbf{Z}). \quad (34)$$

The alternating optimization procedure of MBCC is summarized in **Algorithm 2**.

ALGORITHM 2 | The MBCC algorithm.

Repeat

1. Calculate the coding coefficient matrix **A** by Eq. (34)
2. Calculate the projection matrix **G** by Eq. (27)
3. Calculate the dictionary **D** by Eq. (30)

Until convergence

Testing

For the testing procedure, each frequency band feature of the signal $\mathbf{z}$ is represented as $\mathbf{z}^m$. With the obtained$\{\mathbf{G}^m, \mathbf{D}\}$by Algorithm 2, its label $l(z^m)$ on the m-th frequency band can be computed by the following optimization problem,

$$l(\mathbf{z}^m) = \min_{\mathbf{z}^m} ||(\mathbf{Q}^m)^T\mathbf{z}^m - \mathbf{D}_j(\mathbf{D}_j^T\mathbf{D}_j)^{-1}\mathbf{D}_j^T\mathbf{z}^m||_2. \quad (35)$$

Then the majority voting strategy is used to determine the class label of signal $\mathbf{z}$,

$$y = \arg\max_m l(\mathbf{z}^m). \quad (36)$$

EXPERIMENT

Datasets and Experimental Settings

Two EEG emotion recognition data sets used in the experiment, SEED and DEAP datasets, which are described as follows. The SEED dataset is an emotional EEG dataset collected and provided by Shanghai Jiao Tong University's BCMI Laboratory. The dataset is completed by requiring participants to wear EEG acquisition equipment and recording the emotional EEG signals produced by watching three different types of movie clips. Sixty-two channel electrodes are used in the SEED dataset. The dataset was compiled from 15 participants. With a total of 15 clips, the films are classified as positive, negative, or neutral in terms of their emotional impact. There are five clips of each type, and each movie clip lasts about 4 min. To ensure the experiment's validity and accuracy, the playback sequence of the 15 videos is random, with no repeated clips. Every participant repeated the experiment three times. A few days were set aside in the middle of each experiment to allow participants to adjust their emotions so that they had a consistent emotional response to the same movie clip. In the experiment, EEG signals are divided into 5-s segments

TABLE 2 | The accuracy (standard deviations) of all methods on SEED dataset in session 1.

Methods	$\beta+\gamma$	$\alpha+\beta+\gamma$	$\theta+\alpha+\beta+\gamma$	$\delta+\theta+\alpha+\beta+\gamma$
SVM	77.87	77.96	79.07	81.02
	(8.69)	(9.42)	(9.96)	(8.26)
LC-KSVD	78.39	80.10	81.65	83.50
	(9.81)	(8.13)	(9.19)	(8.38)
MvCVM	81.03	81.89	83.01	83.84
	(9.84)	(8.72)	(9.67)	(9.93)
GLSRM	81.62	82.84	83.33	84.17
	(9.66)	(8.95)	(8.11)	(9.39)
MVU	81.65	82.85	83.08	84.27
	(9.93)	(9.50)	(8.67)	(9.04)
OPFDDL	81.18	83.67	84.81	86.52
	(8.51)	(8.18)	(8.76)	(8.59)
MBCC	**81.85**	**84.61**	**86.07**	**87.91**
	(7.98)	(8.69)	(8.82)	(8.26)

The best performance of each comparison is emphasized by the bold font.

TABLE 3 | The accuracy (standard deviations) of all methods on SEED dataset in session 2.

Methods	$\beta+\gamma$	$\alpha+\beta+\gamma$	$\theta+\alpha+\beta+\gamma$	$\delta+\theta+\alpha+\beta+\gamma$
SVM	77.25	78.04	79.67	80.92
	(8.53)	(7.81)	(7.96)	(9.92)
LC-KSVD	78.31	79.63	80.88	82.37
	(7.74)	(7.23)	(9.65)	(9.62)
MvCVM	80.53	82.55	83.37	83.94
	(8.28)	(8.33)	(8.24)	(9.54)
GLSRM	80.78	82.66	82.70	83.94
	(7.09)	(8.75)	(8.56)	(9.11)
MVU	81.76	82.51	82.84	84.24
	(8.88)	(8.89)	(9.57)	(9.82)
OPFDDL	81.30	83.22	84.76	86.21
	(8.75)	(8.78)	(9.90)	(9.43)
MBCC	**81.90**	**84.24**	**86.14**	**87.87**
	(8.52)	(7.99)	(8.85)	(8.07)

The best performance of each comparison is emphasized by the bold font.

TABLE 4 | The accuracy (standard deviations) of all methods on SEED dataset in session 3.

Methods	$\beta+\gamma$	$\alpha+\beta+\gamma$	$\theta+\alpha+\beta+\gamma$	$\delta+\theta+\alpha+\beta+\gamma$
SVM	77.19	78.24	79.31	80.78
	(9.23)	(9.65)	(9.41)	(9.50)
LC-KSVD	77.61	79.76	80.12	81.92
	(8.12)	(7.09)	(7.32)	(9.87)
MvCVM	79.87	82.14	83.18	83.53
	(8.67)	(8.12)	(8.02)	(9.13)
GLSRM	80.45	81.30	82.43	83.28
	(9.09)	(9.59)	(9.34)	(8.84)
MVU	80.84	81.32	83.00	83.94
	(9.27)	(8.13)	(9.72)	(9.28)
OPFDDL	81.05	83.18	84.61	86.43
	(7.84)	(9.67)	(9.08)	(9.86)
MBCC	**81.81**	**84.24**	**85.63**	**87.74**
	(8.04)	(8.98)	(8.22)	(8.90)

The best performance of each comparison is emphasized by the bold font.

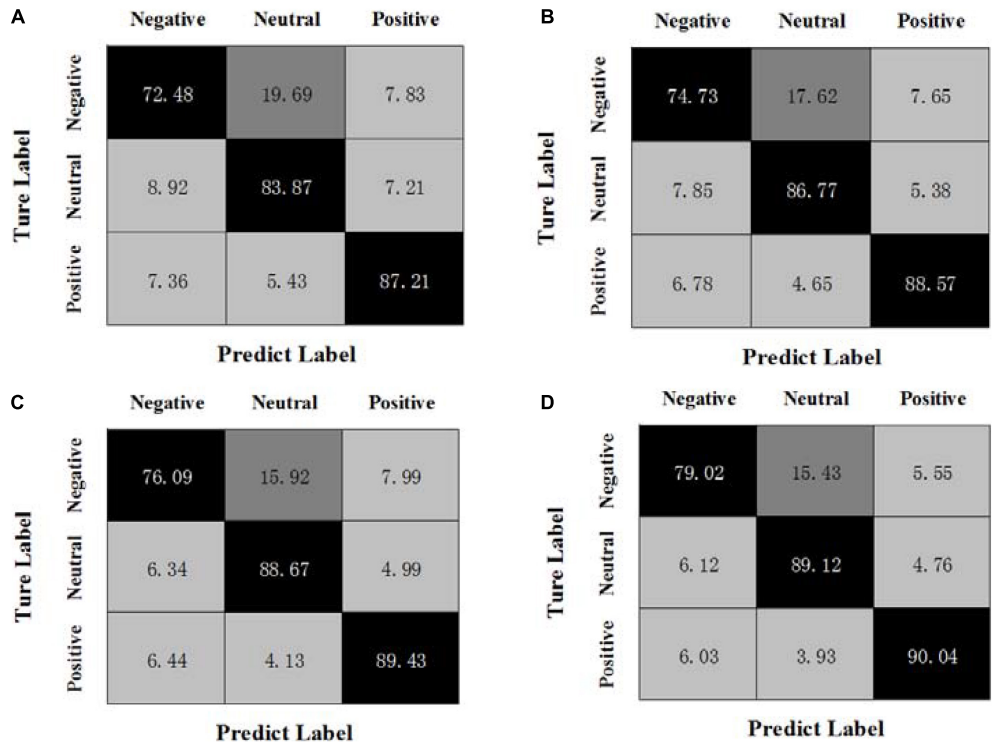


FIGURE 1 | Confusion matrices of MBCC on the SEED dataset, **(A)** $\beta + \gamma$, **(B)** $\alpha + \beta + \gamma$, **(C)** $\theta + \alpha + \beta + \gamma$, **(D)** $\delta + \theta + \alpha + \beta + \gamma$.

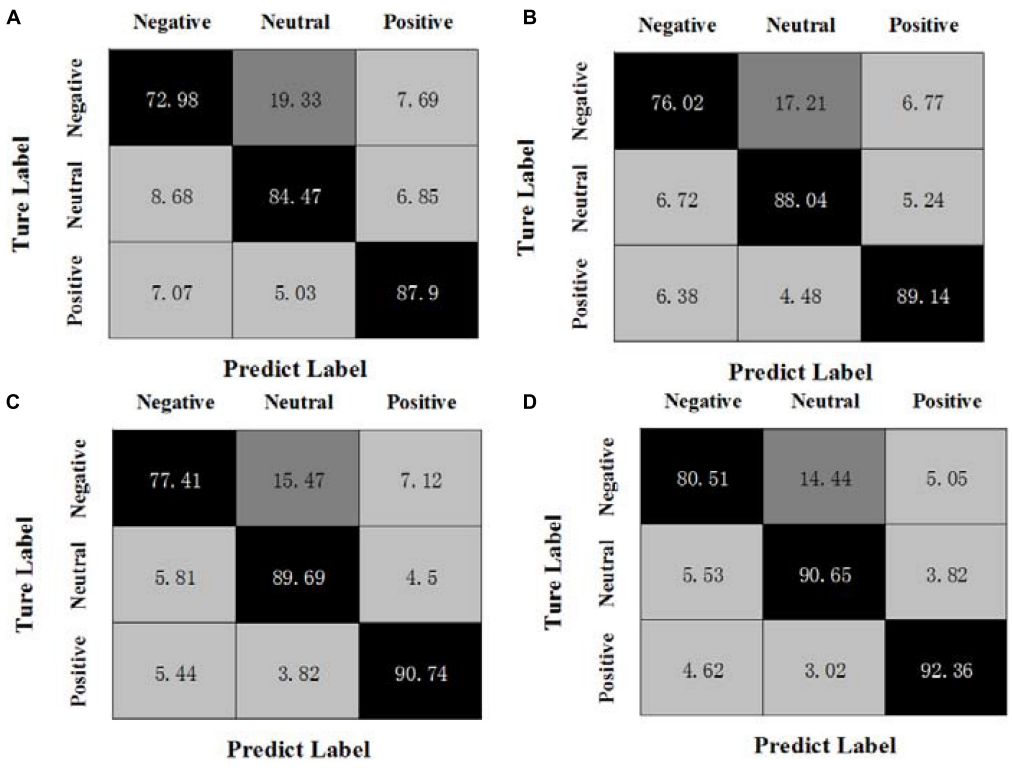


FIGURE 2 | Confusion matrices of OPFDDL on the SEED dataset, **(A)** $\beta + \gamma$, **(B)** $\alpha + \beta + \gamma$, **(C)** $\theta + \alpha + \beta + \gamma$, **(D)** $\delta + \theta + \alpha + \beta + \gamma$.

and features are extracted every 0.5 s. Thus, the sequence length of each segment is 19.

The DEAP dataset is another open database for emotion recognition research that uses EEG and peripheral physiological signals. The dataset recorded the EEG data and 13 peripheral physiological signals of 32 participants using music videos as stimulus materials. The DEAP dataset employs 40 music videos, each of which is 1 min long, as stimulus materials. These music videos are labeled and screened using the general three-dimensional model of valence, arousal, and dominance.

To illustrate the effectiveness of the MBCC method, the comparison methods in the experiment are: SVM (Cortes and Vapnik, 1995), LC-KSVD (Jiang et al., 2013), multi-view CVM (MvCVM) (Huang et al., 2016), global and local structural risk minimization (GLSRM) (Zhu et al., 2016), multi-view learning with universum (MVU) (Wang et al., 2014), and OPFDDL (Gu et al., 2021a). In detail, the Gaussian kernel is used in MvCVM. The kernel parameter and the weight parameter are searched in the grid {1/64, 1/32, …, 64} and {1, 10^1, …, 10^3}, respectively. The weights and offsets in GLSRM are searched in the grid{0.1, 0.2, …, 1}, and its regularization parameters are searched in the grid {1, 10^1, …, 10^3}. In MVU, the learning rate is 0.99, and the relaxation of views is 10^{-6}. In OPFDDL and MBCC, the number of atoms in each class is selected in {5, 10, …, 35}. The dimension of matrix **G** is set to be 90% of the dimension of the EEG signal features. The parameter α is searched in the grid{0.1, 0.2, …, 1}. The parameter σ is set as $\sigma = 1 - \alpha$. The regularization parameter in Eq. (2) was set as 0.01. All methods are implemented in MATLAB.

TABLE 5 | The accuracy (standard deviations) of all methods on the DEAP dataset in valence.

Methods	$\beta+\gamma$	$\alpha+\beta+\gamma$	$\theta+\alpha+\beta+\gamma$
SVM	62.21	63.04	63.55
	(8.30)	(8.09)	(8.76)
KSVD	62.63	63.51	63.94
	(8.59)	(8.28)	(8.43)
MvCVM	63.87	64.20	64.65
	(8.57)	(8.42)	(9.07)
GLSRM	64.15	66.26	66.84
	(9.60)	(8.71)	(9.35)
MVU	64.14	66.18	66.79
	(9.02)	(9.29)	(9.13)
OPFDDL	66.04	68.42	69.08
	(8.74)	(8.40)	(8.95)
MBCC	**66.64**	**68.85**	**69.97**
	(8.65)	(8.20)	(8.46)

The best performance of each comparison is emphasized by the bold font.

TABLE 6 | The accuracy (standard deviations) of all methods on the DEAP dataset in arousal.

Methods	$\beta+\gamma$	$\alpha+\beta+\gamma$	$\theta+\alpha+\beta+\gamma$
SVM	64.77	65.37	65.85
	(10.85)	(11.67)	(10.94)
KSVD	65.07	66.07	66.20
	(10.46)	(11.09)	(11.86)
MvCVM	66.31	66.90	67.19
	(11.01)	(11.48)	(11.24)
GLSRM	66.49	69.05	69.49
	(10.33)	(10.27)	(10.48)
MVU	66.38	69.10	69.27
	(10.79)	(10.75)	(11.12)
OPFDDL	68.46	70.35	70.59
	(10.56)	(10.87)	(10.06)
MBCC	**69.14**	**70.96**	**71.55**
	(10.39)	(10.88)	(10.70)

The best performance of each comparison is emphasized by the bold font.

Experiments on the SEED Dataset

The commonly used power spectral density (PSD) features (Jenke et al., 2014) are adopted in δ, θ, α, β, and γ frequent bands. We obtain 62 dimensional features on each band. We divided the EEG signal data corresponding to the 15 movie clips collected and used 12 clips as training data and the remaining three clips as test data. In both the training and test sets, the proportion of three classes of EEG signals is the same. After the final preprocessing, the samples of three different classes of EEG signals in the training and test sets are balanced. The SEED dataset is divided into three sessions (sessions 1–3) according to the time interval of signal acquisition. The classification results of all methods in three sessions are shown in **Tables 2–4**. We can see that the MBCC method performs the best in terms of accuracy in all three sessions. In **Table 2**, the accuracies of the MBCC method are 0.67, 0.94, 1.26, and 1.39% better than the second best method OPFDDL in multi-frequent bands $\beta+\gamma$,$\alpha+\beta+\gamma$,$\theta+\alpha+\beta+\gamma$,$\delta+\theta+\alpha+\beta+\gamma$. The results in **Tables 3**, **4** are similar to those in **Table 2**. Compared with the OPFDDL method, the proposed MBCC has the ability to take into account the complementarity and consistency between frequency bands while maintaining the PCA constraints of the data structure in the projection space, which is conducive to improving classification performance. Thus, the dictionary learned in the projection space by MBCC has good discriminative performance. The SVM and LC-KSVD methods merge all frequency band data into a vector for learning, and they cannot effectively find the internal connection between each frequency band. For joint learning of multiple perspectives, MvCVM, GLSRM, and MVU treat each frequency band as a learning view. Obviously, the MBCC method obtains a more discriminative model based on dictionary learning and subspace learning.

By calculating the average results of all experiments on three sessions, **Figures 1**, **2** show the confusion matrices of MBCC and OPFDDL on the SEED dataset, respectively. The real label is represented by the ordinate of the confusion matrix, while the predicted label is represented by the abscissa.

It can be seen from **Figures 1**, **2** that (1) the classification results of positive emotional EEG signals are relatively good on the SEED dataset, while the classification results of negative emotions are relatively poor. Positive emotion is easier to identify than negative and neutral ones. This shows that positive emotions

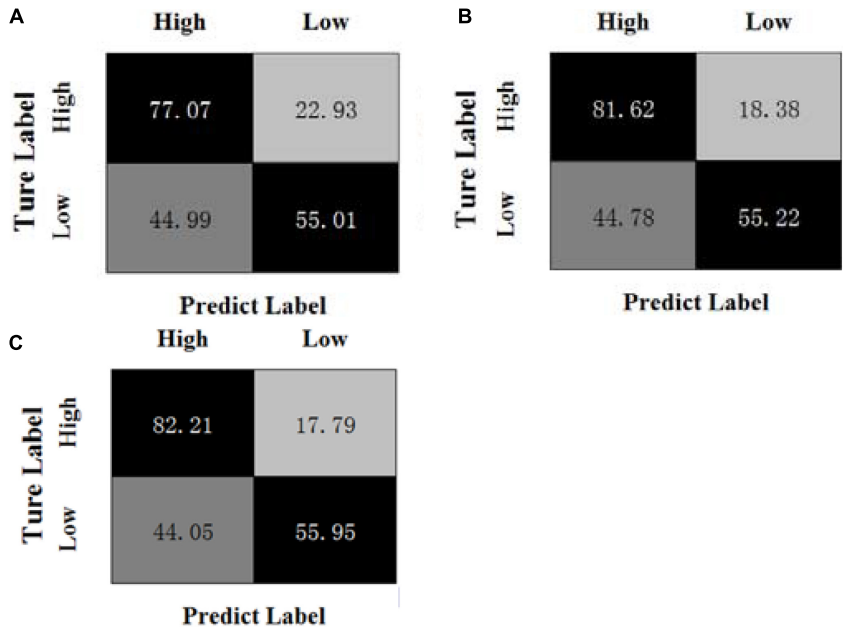


FIGURE 3 | Confusion matrices of OPFDDL on the DEAP dataset in valence, **(A)** $\beta + \gamma$, **(B)** $\alpha + \beta + \gamma$, **(C)** $\theta + \alpha + \beta + \gamma$.

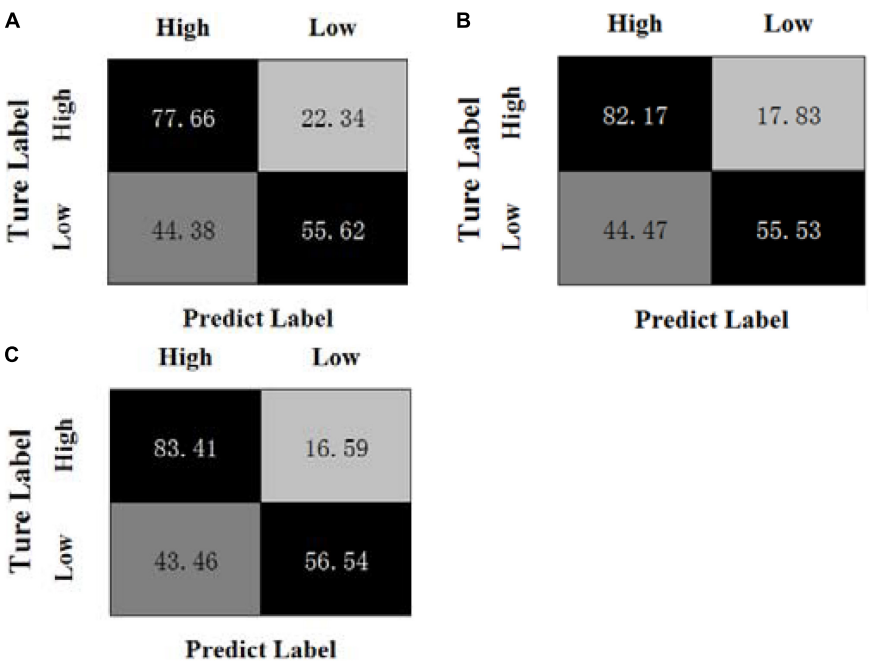


FIGURE 4 | Confusion matrices of MBCC on the DEAP dataset in valence, **(A)** $\beta + \gamma$, **(B)** $\alpha + \beta + \gamma$, **(C)** $\theta + \alpha + \beta + \gamma$.

can cause similar brain feedback between frequency bands more than neutral and negative emotions. (2) The data of different frequency bands are projected into subspaces, and the common shared component of the projection matrix represents the correlation between frequency bands. Obviously, the OPFDDL method does not have this characteristic. (3) In addition, the MBCC method use the PCA-like regularization term based on shared projection matrix to make full use of the discriminative information of EEG data. Thus, the MBCC method achieves better classification accuracy on the SEED dataset.

Experiments on the DEAP Dataset

In the DEAP dataset, music video stimulation is a three-dimensional emotion model based on valence, arousal, and

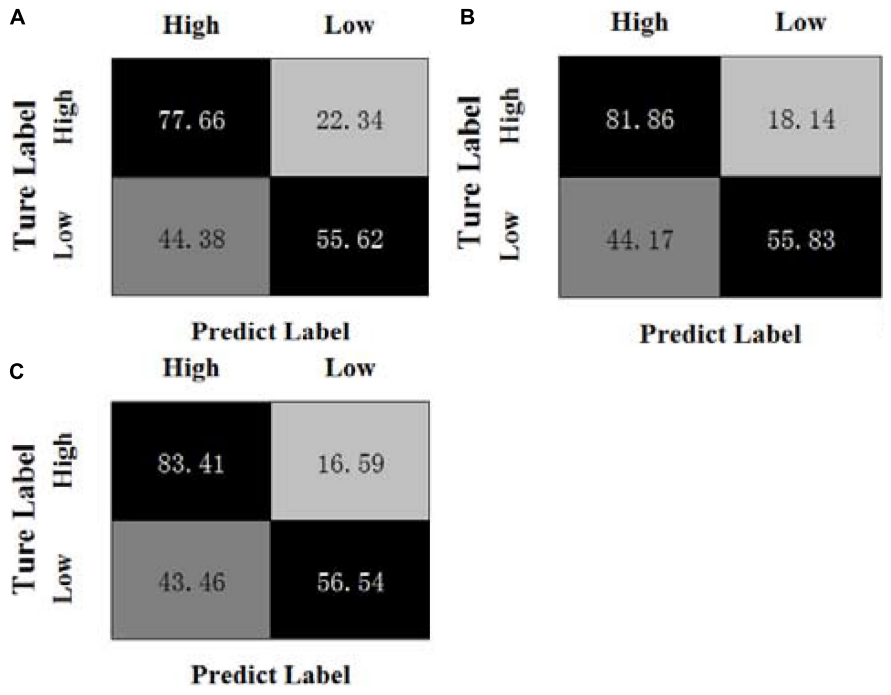


FIGURE 5 | Confusion matrices of OPFDDL on the DEAP dataset in arousal, **(A)** $\beta + \gamma$, **(B)** $\alpha + \beta + \gamma$, **(C)** $\theta + \alpha + \beta + \gamma$.

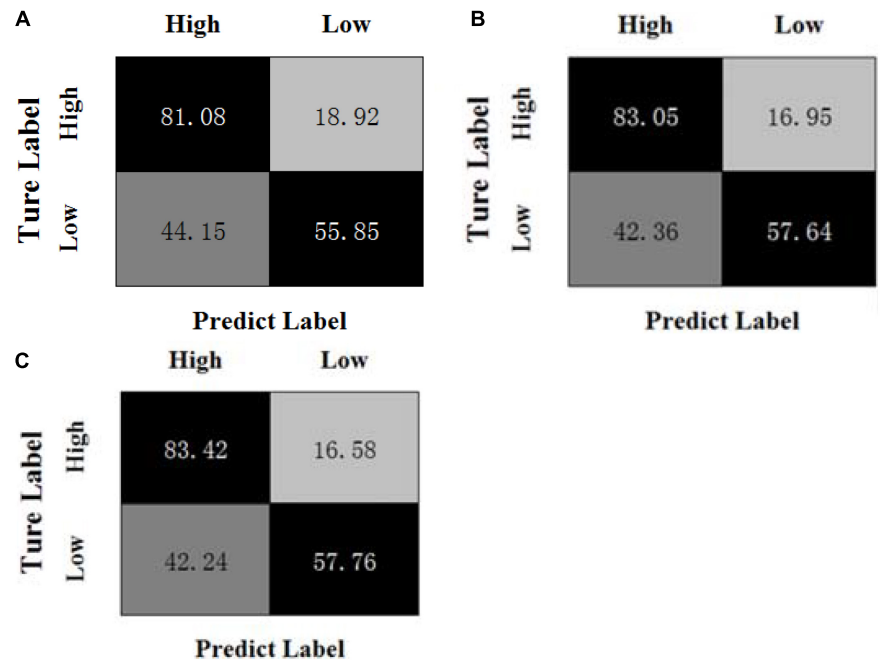


FIGURE 6 | Confusion matrices of MBCC on the DEAP dataset in arousal, **(A)** $\beta + \gamma$, **(B)** $\alpha + \beta + \gamma$, **(C)** $\theta + \alpha + \beta + \gamma$.

dominance. The valence and arousal of emotion are classified in this subsection. The binary valence-oriented classification refers to the classification of emotions according to high valence and low valence. Also, the binary arousal-oriented classification refers to the classification of emotions according to high arousal and low arousal. The classification threshold is set to 5, the participant's score $\in$ [1, 5]for valence is low valence, and score $\in$ (5, 9] is high valence. Similarly, the participant's score $\in$ [1, 5]for arousal is low arousal, and score $\in$ (5, 9] is high arousal. The EEG signals are segmented by a 4-s time window with an overlap 2 s

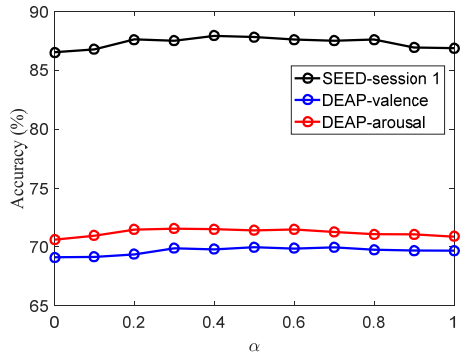


FIGURE 7 | Classification accuracy of MBCC vs. different α on the SEED and DEAP datasets.

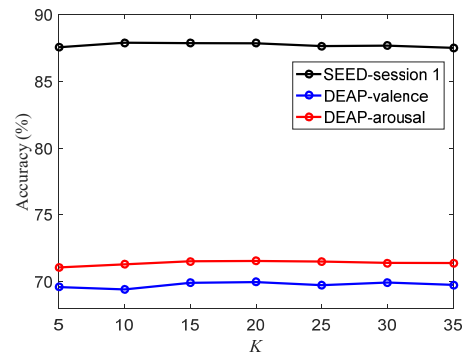


FIGURE 8 | Classification accuracy of MBCC vs. different *K* on the SEED and DEAP datasets.

for each frequency band. Similar to the feature extraction strategy in subsection "Experiments on the SEED Dataset," PSD features are used in the DEAP dataset. Following (Shen et al., 2021), four frequency bands (θ, α, β, γ) are used in the experiment.

Tables 5, **6** compare the average recognition results on valence and arousal on the DEAP dataset, respectively. We can see that (1) all methods have achieved the better classification accuracy for the arousal than the valence on the DEAP dataset. The reason may be that arousal, as an indicator of physiological arousal, reflects the degree of activation of neurophysiological activities, which can be directly reflected in changes in physiological signals. The valence-oriented classification is a more complex task involving mental state, and PSD features may not fully reflect valence's state. (2) Compared with the benchmark methods SVM and LC-KSVD, the MBCC method has achieved much better results. Compared with GLSRM, MVU, and OPFDDL methods, the classification performance of the MBCC method has further improved. The MBCC method has the accuracy rate of 69.97% for the valence-oriented classification, and 71.55% for the arousal-oriented classification using four frequency bands. The classification accuracies of the MBCC method are increased by 0.89 and 0.96%, respectively, when compared to the second best method. This is due to that the multi-frequent band data maintains the consistency between feature similarity and semantic similarity in the learned subspace and can learn a more discriminative dictionary shared by frequency bands.

Figures 3, **4** show the confusion matrices of the OPFDDL method and the MBCC method in valence, respectively. **Figures 5**, **6** show the confusion matrices of the OPFDDL method and the MBCC method in arousal, respectively. Compared with OPFDDL, MBCC has obvious advantages in valence-oriented and arousal-oriented classifications. When different band data describe the same object, there must be an internal connection between each band data. According to the consistency principle, the MBCC method maximizes the commonness of multiple frequent bands in the shared projection space. Furthermore, the Fisher criterion and PCA-like regularization term aids in learning more discriminative sparse representation and maintaining data structure.

Parameter Analysis

The parameter involved in the objective function of the MBCC method is α, and its impact on MBCC's performance is analyzed here. The set value range specifies how to conduct experiments on the SEED session 1 and DEAP dataset, respectively. **Figure 7** depicts the accuracy values at various values of α. The figure shows that MBCC achieves the highest accuracy value when taking 0.4, 0.5, and 0.6 on the SEED session 1, DEAP in valence, and DEAP in arousal, respectively.

The atomic number K of the dictionary also directly determines the performance of the MBCC method. **Figure 8** shows the accuracy values under different K values. We can see that when K reaches 15 and 20 on the SEED session 1 and DEAP dataset, respectively, the accuracy rate tends to stabilize. It indicates that the learned dictionary well represents the data characteristics of the EEG data. Also it shows that the MBCC method can be well applied to the SEED and DEAP datasets using a small size of dictionary.

CONCLUSION

According to the consistent complementarity of multi-frequent band EEG signals and the internal correlation of data itself, this study proposes multi-frequency band collaborative EEG emotion classification method based on the idea of dictionary learning and subspace learning. Using a projection matrix, this method maps different frequency band data to the subspaces of the same dimension. Unlike most existing projection strategies, the projection matrix we designed is divided into two parts, a common shared component and a band-specific component. This strategy can fully use the relevance of different frequency bands and their shared semantics. In the subspace, the MBCC method learns the common shared dictionary between the frequency bands, which can represent the correlation and discrimination of the EEG data. Simultaneously, the incorporation of Fisher criterion and PCA-like regularization term into the subspace *via* dictionary learning makes the learned model more discriminative.

However, the time computation of MBCC is relatively high. It may be not suitable for real-time predicting emotional states

in applications of human-computer interaction. This is the problem we need to solve in the next stage. Furthermore, the work that can be studied further in the future includes: (1) Human emotions are susceptible to external influences. For example, the emotions of the subjects may change while watching a film. The overall emotions of watching the film may be consistent, but the emotions may be inconsistent with expectations at times. As a result, the collected EEG signals are mixed with abnormal samples. In practice, selecting the appropriate abnormal sample processing method is important. The use of the correct processing method can improve the accuracy of emotional EEG signal recognition. (2) EEG signals have the characteristics of randomness. That is, for the same individual subjects, EEG signals are different even in the same emotional state at different times. How to improve the robustness of emotion classification model in multiple domains still needs further research. In the future, we will continue to design and improve our method to be suitable in across time and individuals scenes.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The SEED and DEAP datasets analyzed for this study can be found in the following links, respectively: http://bcmi.sjtu.edu.cn/~seed/seed.html, http://www.eecs.qmul.ac.uk/mmv/datasets/deap/.

AUTHOR CONTRIBUTIONS

JZ and TN conceived and designed the proposed model. ZS performed the experiment. All authors drafted and approved the manuscript.

FUNDING

This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grants Bk202101333.

REFERENCES

Aranha, R. V., Corrêa, C. G., and Nunes, F. (2019). Adapting software with affective computing: a systematic review. *IEEE Trans. Affect. Comp.* 12, 883–899. doi: 10.1109/TAFFC.2019.2902379

Cortes, C., and Vapnik, V. (1995). Support vector machine. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018

Davis, M., and Whalen, P. (2001). The amygdala: vigilance and emotion. *Mol. Psychiatry* 6, 13–34. doi: 10.1038/sj.mp.4000812

Gianaros, P. J., Marsland, A. L., Kuan, C. H., Schirda, B. L., Jennings, J. R., Sheu, L. K., et al. (2014). An inflammatory pathway links atherosclerotic cardiovascular disease risk to neural activity evoked by the cognitive regulation of emotion. *Biol. Psychiatry* 75, 738–745. doi: 10.1016/j.biopsych.2013.10.012

Gu, X., Fan, Y. Q., Zhou, J., and Zhu, J. Q. (2021a). Optimized projection and fisher discriminative dictionary learning for eeg emotion recognition. *Front. Psychol.* 12:705528. doi: 10.3389/fpsyg.2021.705528

Gu, X., Zhang, C., and Ni, T. (2021b). A hierarchical discriminative sparse representation classifier for EEG signal detection. *IEEE/ACM Trans. Comp. Biol. Bioinform.* 18, 1679–1687. doi: 10.1109/TCBB.2020.3006699

Guil, R., Ruiz-González, P., Merchán-Clavellino, A., Morales-Sánchez, L., Zayas, A., and Gomez-Molinero, R. (2020). Breast cancer and resilience: the controversial role of perceived emotional intelligence. *Front. Psychol.* 11:595713. doi: 10.3389/fpsyg.2020.595713

Huang, C., Chung, F., and Wang, S. (2016). Multi-view L2-SVM and its multi-view core vector machine. *Neural Networks* 75, 110–125. doi: 10.1016/j.neunet.2015.12.004

Jenke, R., Peer, A., and Buss, M. (2014). Feature extraction and selection for emotion recognition from EEG. *IEEE Trans. Affect. Comput.* 5, 327–339. doi: 10.1109/TAFFC.2014.2339834

Jiang, Z., Lin, Z., and Davis, L. (2013). Label consistent K-SVD: learning a discriminative dictionary for recognition. *IEEE Trans. Patt. Anal. Mach. Intell.* 35, 2651–2664. doi: 10.1109/TPAMI.2013.88

Klatzkin, R. R., Nolan, L. J., and Kissileff, H. R. (2021). Self-reported emotional eaters consume more food under stress if they experience heightened stress reactivity and emotional relief from stress upon eating. *Physiol. Behav.* 243:113638. doi: 10.1016/j.physbeh.2021.113638

Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., et al. (2011). Deap: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15

Li, J., Zhang, Z., and He, H. (2018). Hierarchical convolutional neural networks for EEG-based emotion recognition. *Cogn. Comp.* 10, 368–380. doi: 10.1007/s12559-017-9533-x

Li, M., and Lu, B. L. (2009). "Emotion classification based on gamma-band EEG". *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN. 1223–1226. doi: 10.1109/IEMBS.2009.5334139

Li, W., Zhang, Z., and Song, A. (2020). Physiological-signal-based emotion recognition: an odyssey from methodology to philosophy. *Measurement* 172:108747.

Li, Y., Zheng, W., Cui, Z., Zong, Y., and Ge, S. (2019). EEG emotion recognition based on graph regularized sparse linear regression. *Neural Proc. Lett.* 49, 555–571. doi: 10.1007/s11063-018-9829-1

Liu, Y., Ding, Y., Li, C., Cheng, J., Cheng, J., Song, R. C., et al. (2020). Multi-channel EEG-based emotion recognition *via* a multi-level features guided capsule network. *Comp. Biol. Med.* 123:103927. doi: 10.1016/j.compbiomed.2020.103927

Mohammadi, Z., Frounchi, J., and Amiri, M. (2017). Wavelet-based emotion recognition system using EEG signal. *Neural Comp. Appl.* 28, 1985–1990. doi: 10.1007/s00521-015-2149-8

Ni, T. G., Ni, Y. Y., Xue, J., and Wang, S. H. (2021). A domain adaptation sparse representation classifier for cross-domain Electroencephalogram-based emotion classification. *Front. Psychol.* 12:721266. doi: 10.3389/fpsyg.2021.721266

Nurillaeva, N. M., and Abdumalikova, F. B. (2021). Predictive importance of psycho-emotional syndrome of patients with coronary heart disease in the violation of platelet hemostatic system. *Atherosclerosis* 331:E204. doi: 10.1016/j.atherosclerosis.2021.06.625

Peng, Y., Liu, S., Wang, X., and Wu, X. (2020). Joint local constraint and fisher discrimination based dictionary learning for image classification. *Neurocomputing* 398, 505–519. doi: 10.1016/j.neucom.2019.05.103

Samsonovich, A. V. (2020). Socially emotional brain-inspired cognitive architecture framework for artificial intelligence. *Cogn. Syst. Res.* 60, 57–76. doi: 10.1016/j.cogsys.2019.12.002

Shareef, M. A., Mukerji, B., Alryalat, M., Wright, A., and Dwivedi, Y. K. (2018). Advertisements on Facebook: identifying the persuasive elements in the development of positive attitudes in consumers. *J. Retail. Cons. Serv.* 43, 258–268. doi: 10.1016/j.jretconser.2018.04.006

Shen, F., Peng, Y., Kong, W., and Dai, G. (2021). Multi-scale frequency bands ensemble learning for EEG-based emotion recognition. *Sensors* 21:1262. doi: 10.3390/s21041262

Tennant, C., and McLean, L. (2001). The impact of emotions on coronary heart disease risk. *J. Cardiovas. Risk* 8, 175–183. doi: 10.1177/174182670100800309

Wang, Z., Zhu, Y., Liu, W., Chen, Z., and Gao, D. (2014). Multi-view learning with Universum. *Knowledge-Based Syst.* 70, 376–391. doi: 10.1016/j.knosys.2014.07.019

Wirkner, J., Weymar, M., Lw, A., Hamm, C., Anne-Marie, S., Kirschbaum, C., et al. (2017). Cognitive functioning and emotion processing in breast cancer survivors and controls: an ERP pilot study. *Psychophysiology* 54, 1209–1222. doi: 10.1111/psyp.12874

Yang, Y., Wu, Q., Fu, Y., and Chen, X. (2018). "Continuous Convolutional Neural Network with 3D Input for EEG-Based Emotion Recognition," in *International Conference on Neural Information Processing*, Siem Reap, 433–443. doi: 10.1007/978-3-030-04239-4_39

Zhang, G., Sun, H., Porikli, F., Liu, Y., and Sun, Q. (2017). Optimal couple projections for domain adaptive sparse representation-based classification. *IEEE Trans. Image Process*. 26, 5922–5935. doi: 10.1109/TIP.2017.2745684

Zhang, G., Yang, J., Zheng, Y., Luo, Z., and Zhang, J. (2021). Optimal discriminative feature and dictionary learning for image set classification. *Inform. Sci.* 547, 498–513. doi: 10.1016/j.ins.2020.08.066

Zhao, G., Zhang, Y., and Yan, G. (2018). Frontal EEG asymmetry and middle line power difference in discrete emotions. *Front. Behav. Neurosci.* 11:225. doi: 10.3389/fnbeh.2018.00225

Zheng, W. L., and Lu, B. L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Mental Dev* 7, 162–175. doi: 10.1109/tamd.2015.2431497

Zhong, P., Wang, D., and Miao, C. (2020). EEG-Based emotion recognition using regularized graph neural networks. *IEEE Trans. Affect. Comp. Early Access* 2020:2994159. doi: 10.1109/TAFFC.2020.2994159

Zhu, C., Wang, Z., and Gao, D. (2016). New design goal of a classifier: global and local structural risk minimization. *Knowledge-Based Syst.* 100, 25–49. doi: 10.1016/j.knosys.2016.02.002

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 10 May 2022
doi: 10.3389/fnagi.2022.888577

Alzheimer's Disease Analysis Algorithm Based on No-threshold Recurrence Plot Convolution Network

Xuemei Li[1], *Tao Zhou*[2]* *and Shi Qiu*[3]*

[1] School of Mechanical and Electrical Engineering, Chengdu University of Technology, Chengdu, China, [2] School of Computer Science and Engineering, North Minzu University, Yinchuan, China, [3] Key Laboratory of Spectral Imaging Technology CAS, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China

OPEN ACCESS

Edited by:
Yuanpeng Zhang,
Nantong University, China

Reviewed by:
Qing Tian,
Nanjing University of Information Science and Technology, China
Zhiyu Jiang,
Northwestern Polytechnical University, China
Kalyana C. Veluvolu,
Kyungpook National University, South Korea
Rupesh Kumar Chikara,
University of Texas at Arlington, United States

***Correspondence:**
Tao Zhou
zhoutaonxmu@126.com
Shi Qiu
qiushi@opt.ac.cn

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

Received: 03 March 2022
Accepted: 07 April 2022
Published: 10 May 2022

Citation:
Li X, Zhou T and Qiu S (2022) Alzheimer's Disease Analysis Algorithm Based on No-threshold Recurrence Plot Convolution Network. Front. Aging Neurosci. 14:888577. doi: 10.3389/fnagi.2022.888577

Alzheimer's disease is a neurological disorder characterized by progressive cognitive dysfunction and behavioral impairment that occurs in old. Early diagnosis and treatment of Alzheimer's disease is great significance. Electroencephalography (EEG) signals can be used to detect Alzheimer's disease due to its non-invasive advantage. To solve the problem of insufficient analysis by single-channel EEG signal, we analyze the relationship between multiple channels and build PLV framework. To solve the problem of insufficient representation of 1D signal, a threshold-free recursive plot convolution network was constructed to realize 2D representation. To solve the problem of insufficient EEG signal characterization, a fusion algorithm of clinical features and imaging features was proposed to detect Alzheimer's disease. Experimental results show that the algorithm has good performance and robustness.

Keywords: Alzheimer's disease, EEG, PLV, recursive graph, no-threshold

INTRODUCTION

Alzheimer's disease is a degenerative disease of the central nervous system, mainly manifested as progressive memory impairment, cognitive dysfunction, personality change and language impairment, and other neuropsychiatric symptoms, which seriously affect social, career, and life functions. Alzheimer's disease is a common disease in the elderly, and its prevalence and incidence are extremely high. According to statistics, the incidence of Alzheimer's disease is 5%, the disease is the most common type of dementia in the elderly, accounting for 50–70% of Alzheimer's disease, common in people over 65 years old. It is of great significance to study it.

Alzheimer's Disease

Alzheimer's disease occurred in elderly and senile prophase, characterized by progressive cognitive dysfunction and behavioral impairment of nervous system diseases, main show is memory disorders, aphasia, disuse, agnosia, visual spatial ability damage, abstract thinking and calculation ability damage, personality and behavior change, and so on, can be improved by drug treatment, and the disease is not cured. The etiology and pathogenesis of Alzheimer's disease are extremely complex, and may be related to genetic factors, brain pathological changes and other factors. Generally, Alzheimer's disease tends to occur in people over 65 years old. Mental stimulation, trauma, neurological diseases and other factors can induce Alzheimer's disease. The main pathological changes were amyloid precursor protein gene on chromosome 21, *PSEN1* gene on

chromosome 14, and *PSEN2* gene mutation on chromosome 1. The brain was reduced in size and weight, and the typical histopathological changes were neuroinflammatory plaques, neurofibrillary tangles, and neuron loss (Yoon et al., 2022). Alzheimer's disease is usually silent onset, pre-dementia, and dementia stage symptoms are different, but generally manifested as memory impairment, speech loss or emotional apathy, crying and laughing impermanent, severe patients can be complicated with lung infection, urinary tract infection and pressure ulcers, and other diseases. Early diagnosis and early treatment is of great significance.

Method

The current examination methods mainly include: neuropsychological test, hematological examination, neuroimaging examination, Electroencephalography (EEG), cerebrospinal fluid testing, genetic testing. Due to the convenience of EEG collection, it has a good detection effect for early Alzheimer's disease to become the main research direction. To this end, we used EEG for the study. Morabito et al. (2012) constructs the model analysis of Alzheimer's disease EEG from the perspective of energy entropy. Anh et al. (2012) used support vector machine (SVM) to cluster EEG. Falk et al. (2012) analyzed the disease by the variability in EEG amplitude. Hulbert and Adeli (2013) combine EEG and imaging information to make a diagnosis of the disease. Morabito et al. (2013) proposed the EEG enhancement algorithm to highlight the area where the lesion signal is located. Zhao and He (2014) used a deep learning network for disease diagnosis. Cassani et al. (2014) extracted useful information from the EEG to conduct the research on Alzheimer's disease. Bhat et al. (2015) combined the clinical neural data and EEG to conduct the study. Al-Jumeily et al. (2015) was diagnosed by EEG analysis. Al-Nuaimi et al. (2016) analyzed EEG from the perspective of amplitude to diagnose early Alzheimer's disease. Yu et al. (2016) analyzed EEG, the signal transmission process. Kulkarni and Bairagi (2017) used SVM to extract the significant features of the EEG signal. Deng et al. (2017) constructed a multiscale model from an entropy perspective to analyze the complex EEG. Chikara et al. (2018) proposed monetary reward and punishment to response inhibition modulate activation and synchronization within the inhibitory brain network. Houmani et al. (2018) built multiple networks to implement disease analysis. Kim and Kim (2018) analyzed the correlation between the signals and extracted the features. Yang et al. (2018) studied the multi-channel data of EEG and proposed parallel revolutionary recurrent neural network to realize Alzheimer's disease recognition. Chen et al. (2020) constructed a model from the perspective of classification to realize signal analysis. Yu et al. (2019) introduced the fuzzy learning theory to analyze the EEG signals. Maturana-Candelas et al. (2019) constructed a multiscale model to extract EEG features. Chikara and Ko (2019) used hierarchical model to neural activities classification of human inhibitory control, which achieved good results. Rossini et al. (2020) proposed markers for early Alzheimer's disease diagnosis, demonstrating the validity of the EEG analysis. Qiu et al. (2020) analyzed the EEG transmission process. Oltu et al. (2021) proposed a novel Alzheimer's disease detection algorithm based on EEG. Li et al. (2021) analyzed the correlation between multiple channels to diagnose the disease. Puri et al. (2022) proposed the Kolmogorov Complexity diagnosis of Alzheimer's disease. Ding et al. (2022) proposed the Alzheimer's disease automatic detection system based on EEG.

In conclusion, the diagnosis of Alzheimer's disease based on EEG has achieved some results. However, there are still the following problems in computer processing: (1) The correlation between different channels is not studied. (2) The EEG signal is not well visualized and difficult to analyze. (3) With limited characteristics and insufficient characterization.

In view of the difficult problem of analysis of Alzheimer's disease, we use computer to assist diagnosis. (1) Analyze the corresponding relationship between different channels at the same time and build PLV network structure. (2) One-dimensional EEG signals are converted into two-dimensional recurrence plot to achieve visual analysis of features. (3) Combining clinical features with EEG signals features to realize diagnosis of Alzheimer's disease.

ALGORITHM FRAMEWORK

Through the analysis of EEG signals, we constructed a new Alzheimer's disease analysis algorithm, and the block diagram is shown in **Figure 1**. The model is constructed from the perspective of cognition, and the EEG signal analysis model based on Phase Locking Value (PLV) is proposed to simulate the EEG transmission process. From the correlation of EEG time series, the EEG signal analysis algorithm based on recurrence plot is proposed to convert one-dimensional information into two-dimensional information for intuitive analysis. From the perspective of feature correlation, multi-source features are extracted in order to build a model, and finally realize the fusion of decision sets and Alzheimer's disease analysis.

EEG Signal Analysis Based on PLV

Research shows that the cognitive process of human brain designs the activities of various brain regions and the information dissemination and interaction between different functional regions (Sarma and Barma, 2022). From the perspective of computer, this process can be regarded as building a network between relevant brain regions to reflect the relationship between mutual transmissions and processing. Since EEG signal has phase synchronization relationship, we use PLV to measure EEG phase synchronization relationship:

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

$$\Delta\varphi\left(t\right) = \varphi_x\left(j\Delta t\right) - \varphi_y\left(j\Delta t\right) \tag{2}$$

Where, $\phi_x(t)$ and $\phi_y(t)$ represent the instantaneous phase of $x(t)$ and $y(t)$, respectively, $\Delta\phi(t)$ represents the phase difference, Δt represents the period of application. Clustering coefficient can measure the degree of brain function separation, and the

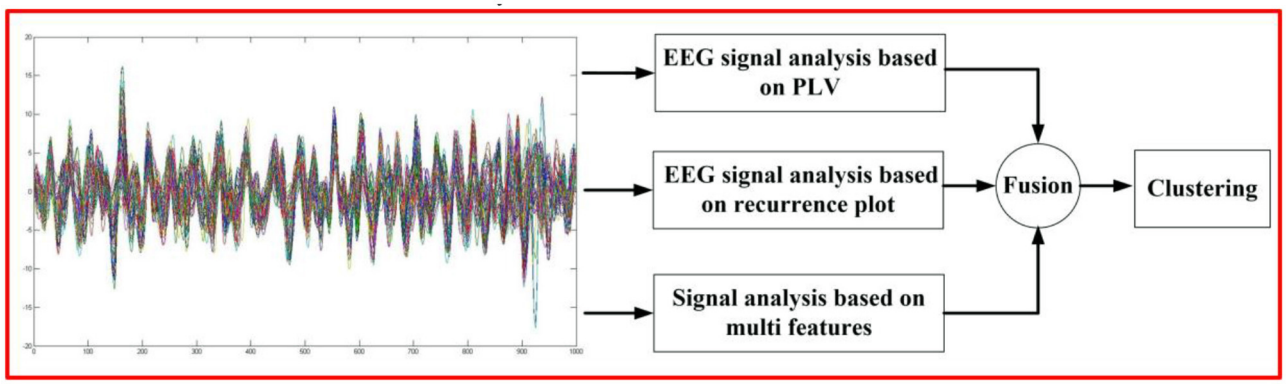


FIGURE 1 | The proposed algorithm flow chart.

proportion of the number of connections and the maximum number of connections between a node and adjacent nodes.

The clustering coefficient of node i is defined as:

$$C_i = \frac{\sum_{k\neq i}\sum_{l\neq i,l\neq k} c_{ik}c_{il}c_{kl}}{\sum_{k\neq i}\sum_{l\neq i,l\neq k} c_{ik}c_{il}} \tag{3}$$

where c_{ij} is the weight between nodes i and j of the adjacency matrix. The characteristic path length L represents the minimum number of edges of two nodes connected in the network.

The weighted network is expressed as:

$$L = \frac{N(N-1)}{\sum_{i=1}^{N}\sum_{j\neq i}^{N}\left(1/L_{ij}\right)} \tag{4}$$

where N represents the number of weighted nodes and L_{ij} represents the number of edges of the shortest path of nodes i and j.

$$G = \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}^{N} L_{ij}^{-1} \tag{5}$$

Local subnet efficiency is

$$L_{ei} = \frac{1}{N_{G_i}\left(N_{G_i}-1\right)}\sum_{i,k\in G_i} L_{j,k}^{-1} \tag{6}$$

where N_{G_i} is the number of nodes of the subgraph G_i. The centrality of the network is introduced for measurement:

$$b_i = \sum_{m\neq i,\neq n}\frac{\sigma_{mn}(i)}{\sigma_{mn}} \tag{7}$$

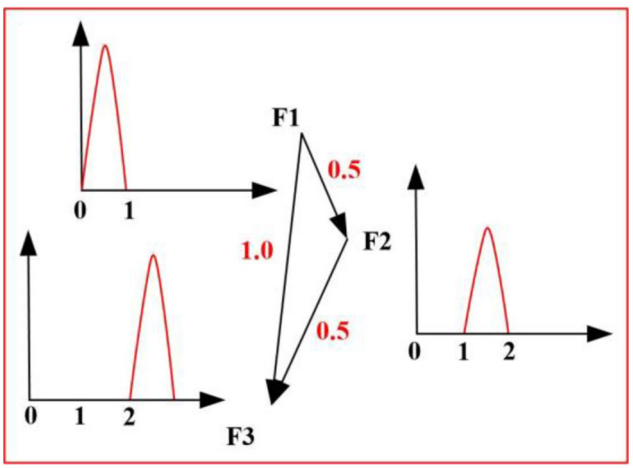


FIGURE 2 | The calculation process.

Where, $\sigma_{mn}(i)$ represents the number of shortest paths from node m to node n, which goes through i. σ_{mn} represents the shortest path length from m to n. As shown in **Figure 2**, the signal starts F1 and ends F3 through two branches. We take i =2 and $b_2 = \frac{0.5+0.5}{0.5+0.5+1.0} = \frac{1}{2}$ to achieve the centrality measure.

Under the condition of network establishment, it is necessary to extract features from the signal as input. Common Space Pattern (CSP) is used to extract airspace information. It is an efficient spatial filtering algorithm whose goal is to create an optimal common spatial filter (Kumar et al., 2017). We use CSP to extract features. CSP obtains the most distinguishing feature vector by diagonalizing the task covariance matrix. The specific process is as follows:

Given two types of data samples X_1 and X_2, the corresponding covariance matrix is

$$R_i = \frac{X_i X_i^T}{trace\left(X_i X_i^T\right)} \tag{8}$$

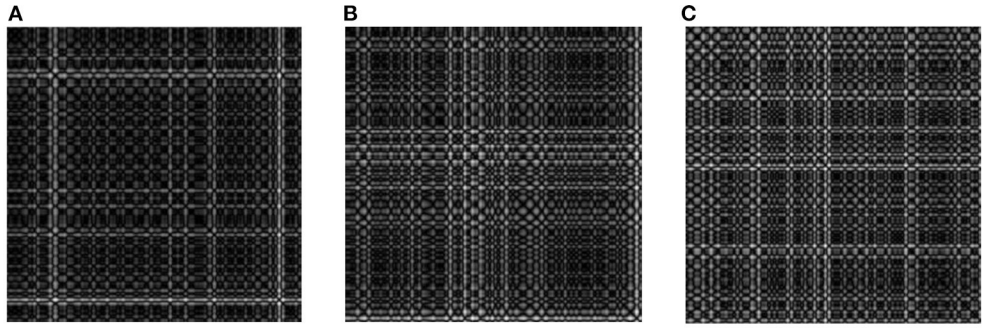


FIGURE 3 | Alzheimer's disease recurrence plot. **(A)** Calm. **(B)** Morbidity. **(C)** Transitional period.

The corresponding mixed space covariance matrix is

$$R_c = \bar{R}_1 + \bar{R}_2 \tag{9}$$

Where, $\bar{R}_1$ and$\bar{R}_2$ represent the average covariance matrix of two types of tasks.

Principal component analysis is applied to decompose eigenvalues of R_c:

$$R_c = U_c \Lambda_c U_c^T \tag{10}$$

Where, U_c represents eigenvector matrix and Λ_c represents eigenvalue. The corresponding whitening matrix is

$$P = \frac{U_c^T}{\sqrt{\Lambda_c}} \tag{11}$$

The spatial filter P is constructed to meet the following conditions:

$$\begin{cases} S_1 = PR_1P^T = B\Lambda_1 B^T \\ S_1 = PR_2P^T = B\Lambda_2 B^T \\ \Lambda_1 + \Lambda_2 = I \end{cases} \tag{12}$$

Calculate the projection matrix, and whiten the transformation of the eigenvector corresponding to the maximum eigenvalue in EEG and max (Λ_1, Λ_2) to achieve the best classification. To do so, a projection matrix is built:

$$W = \left(B^T P\right)^T \tag{13}$$

EEG data characteristics are obtained:

$$Z_{M\times N} = W_{M\times M} * X_{M\times N} \tag{14}$$

Select the maximum values of $2m$ row from $Z_{M\times N}$ as feature input, which is input into the constructed PLV network to realize feature classification.

EEG Signal Analysis Algorithm Based on Recurrence Plot

Recurrence plot can be used to measure the correlation of time series. Its core idea is to map the trajectory of moving state to the plane, which can realize visualization as shown in **Figure 3**. The set of time series is marked as X, and the corresponding recursion diagram is:

$$\begin{cases} R_{ij} = \varphi\left(\varepsilon - r_{ij}\right) i, j \in \{1, 2, ..., N - (m-1)\,\tau\} \\ r_{ij} = \left\| X\left(i\right) - X\left(j\right) \right\| \\ \varphi = \begin{cases} 1\ x \geq 0 \\ 0\ other \end{cases} \end{cases} \tag{15}$$

According to the recursive state of two times, i and j represent the horizontal and vertical coordinates of the image, and the matrix R composed of 0 and 1 is obtained.

Although the recurrence plot can intuitively express the time series, it increases the threshold ϕ. The richer nonlinear dynamic characteristics are lost and the characterization is incomplete. Thus, we improve it as follows to retain its characteristics to the greatest extent:

$$ER_{ij} = \left|\varepsilon - r_{ij}\right| i, j \in \{1, 2, ..., N - (m-1)\,\tau\} \tag{16}$$

Convolutional neural network (CNN) network has shown unique advantages in target segmentation and recognition, and has the invariance of translation, scaling and tilt of network structure. It is usually composed of input layer, convolution layer, pooling layer, full connection layer, and output layer.

With the increase of network layers, the network has nonlinear fitting ability and improves the performance of the model. But it will also be accompanied by the phenomenon of gradient disappearance. In order to solve this problem, we introduce the residual block to construct the relationship between input and output through fitting the residual mapping of multi-layer networks is shown in **Figure 4**. The problem of difficult convergence of the deep-seated network can be solved according to certain overlapping rules. The structure is shown in **Table 1**.

Based on the above introduction, PLV was used to analyze the correlation between signals and calculate the probability PE of signal attributes. To obtain the probability RE of signal attributes,

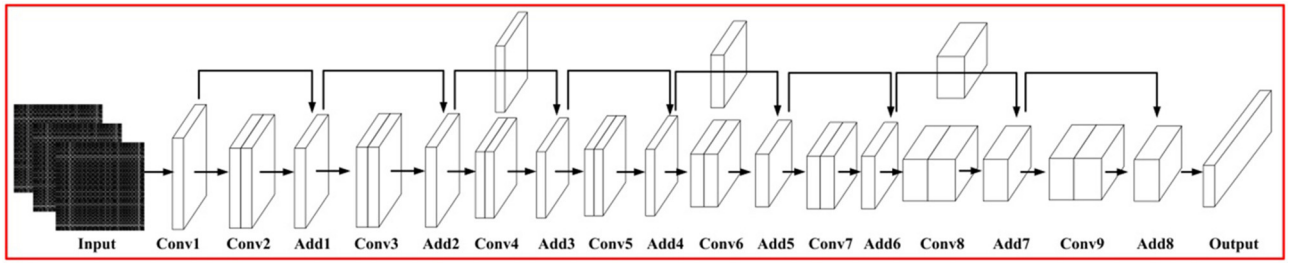


FIGURE 4 | Network structure.

TABLE 1 | Network parameters.

The network layer	Parameters
Conv 1	(7, 7, 64); D = 2
Conv 2	(3, 3, 64) × 2; Maxpooling; D = 2
Conv 3	(3, 3, 64) × 2; Maxpooling; D = 2
Conv 4	(3, 3, 128) × 2
Conv 5	(3, 3, 128) × 2
Conv 6	(3, 3, 256) × 2
Conv 7	(3, 3, 256) × 2
Conv 8	(3, 3, 512) × 2
Conv 9	(3, 3, 128) × 2

a network based on non-threshold recursive plot was built from the time correlation of EEG signals. We collected the age, gender, basic diseases (hypertension, hyperlipidemia, diabetes), eye movement test, etc., and selected the patients with statistically significant characteristics using $p < 0.05$. Age, diabetes, and eye movement tests were significant by screening.

EXPERIMENT AND RESULT ANALYSIS

There are two data, (1) http://adni.loni.usc.edu/; (2) Clinical data collected by the hospital. The frequency of signal acquisition is 8–30 Hz, 62 channels of data. With the consent of the patients, 100 patients with Alzheimer's disease at different stages were collected including 48 women and 52 men aged 55–80 years. The EEG collected was divided into calm, morbidity, and transitional period according to professional physicians and clinical manifestations. Total 1,000 points of data were collected in each period. Construct data sets and conduct experiments.

Introduction of Experimental Parameters and Evaluation Indexes

We analyzed the characteristics of EEG signals and sampled the data. For each EEG signal accord to the principle of average sampling, we obtained 1,000 data points, and formed the recursive plot data of 1,000 × 1,000 data. Then, subsequent experiments were conducted on this basis. In order to ensure the consistency of the experiment, we preprocessed the EEG signal data. Through data analysis, to ensure the consistency of the experiment, EEG signal data were preprocessed and representative Fp1, Fp2, F3, and F4 were normalized.

Accuracy A is used to measure the performance of different algorithms:

$$A = \frac{TP + TN}{TP + FP + TN + FN} \tag{17}$$

Where, TP is the positive sample with correct model classification, FP is the negative sample with wrong model classification, TN is the negative sample with correct model classification, and FN is the positive sample with wrong model classification.

Performance of PLV Algorithm

We build the brain network graph $G = (V, E)$ and using EEG click as network nodes. The graph side shows the channel relationship. The PLV can be used as a synchronicity measure to represent the connection strength in a weighted network analysis. The results for Alzheimer's disease are shown in **Figure 5**, with a low degree of connection in Fp1. The connection degree between Fp2 and F3 and F4 is high, so the study is carried out based on this.

EEG Signal Analysis of Recurrence Plot

We explored the recurrence plot by selecting EEG during periods of calm, transition and onset as shown in **Figure 6**. From the analysis of EEG signal, during the calm period, EEG does not fluctuate much, and the signal in the lower right corner of the recursion graph is strong. In transition, the EEG considerably began from smooth band, in the middle of the recursive plot chart presents signal is stronger. During the onset of the disease, the EEG amplitude was further enlarged, but it was not obvious on the EEG alone and could not be distinguished effectively, and the signal intensity around the recursion diagram was strong. Based on this, the three can be distinguished. Subsequent fusion of PLV and clinical features can further improve the detection effect.

The ROC curve corresponding to our algorithm is compared with the mainstream algorithm, as shown in **Figure 7**. SVM algorithm (Anh et al., 2012) constructed the classifier and realized the classification of Alzheimer's disease. Parallel revolutionary Cyclic Neural Network (PCRNN) (Yang et al., 2018) established depth model and analyzed signal composition. Libsvm classifier (Chen et al., 2020) constructs the model from the perspective

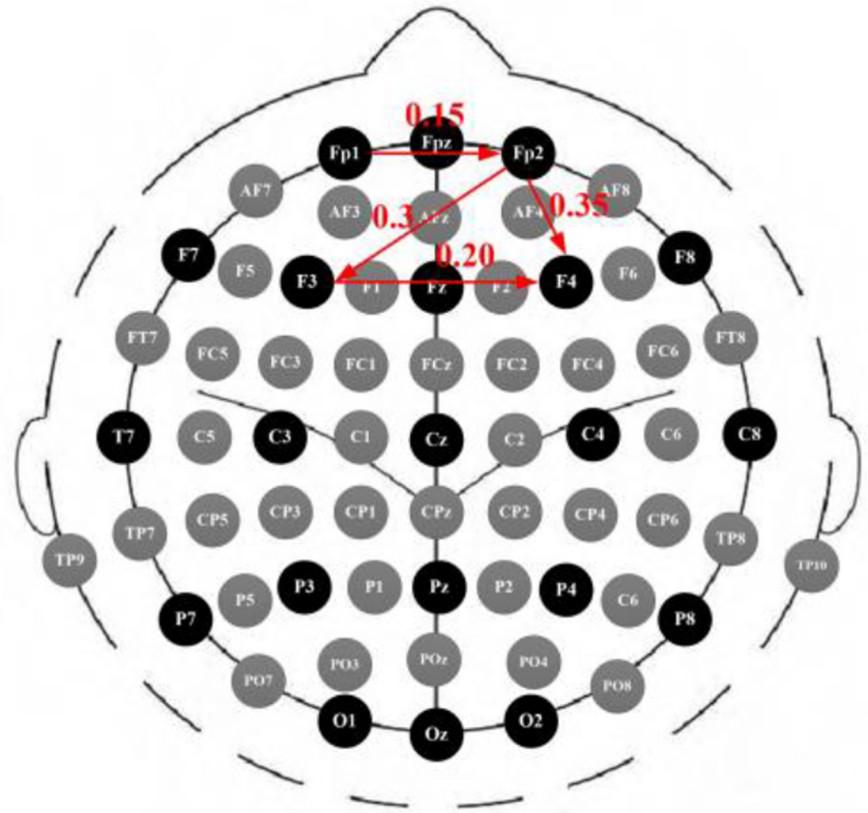


FIGURE 5 | EEG PLV.

of classification to realize signal analysis. DTW can realize the measurement of time series. The idea of DTW is to extend and shorten two time series to represent signal similarity with the shortest distance. However, EEG signals have a strong correlation, and the number of points collected by the EEG signal is too large, which will lead to information loss and error information introduction through DTW extension and shortening. Result in poor effect. **Figure 7** quiet period for acquisition of data, due to the quiet period EEG signals is relatively stable, our algorithm and comparison algorithm can better on the test. **Figure 7** shows the data collected in the transitional period. The EEG gradually fluctuates from a relatively stable signal. However, due to the limited amplitude of fluctuation, the detection effect of the algorithm decreases compared with that in the calm period. **Figure 7** shows the data collected during the onset of the disease, and the EEG fluctuates greatly, which can be detected by changing the amplitude. Overall, all algorithms performed best for quiet Alzheimer's disease, followed by morbidity and transitional Alzheimer's disease. In addition, the algorithm establishes a model from the perspective of EEG, carries out processing, recurrence plot and auxiliary features of EEG Alzheimer's disease, and realizes EEG Alzheimer's disease analysis, with high performance.

We added comparative experiments, and the PR algorithm proposed transformed 1D features into 2D PR without threshold, which achieved certain results in the diagnosis of Alzheimer's disease. On this basis, PLV was adopted to analyze the correlation between different channels at the same time, and the detection effect was further improved. Finally, we simulate the process of physician diagnosis, and fuse the clinical features into the model to achieve the best effect.

CONCLUSIONS AND DISCUSSIONS

Alzheimer's disease is a central nervous system variable disease, although there is no effective treatment method, but it has a positive effect on its early diagnosis and early treatment. Studies show that EEG has non-invasive and easy acquisition characteristics, which has proved to be an effective means to detect Alzheimer's disease, for which we propose a new Alzheimer's disease analysis algorithm.

Early AI algorithms conducted the analysis only from a single signal perspective, ignoring the response relationship between different channels at the same time, resulting in the limited representational ability of the established model. With the improvement of medical and information acquisition ability, scholars focus their attention to the signal transmission process to

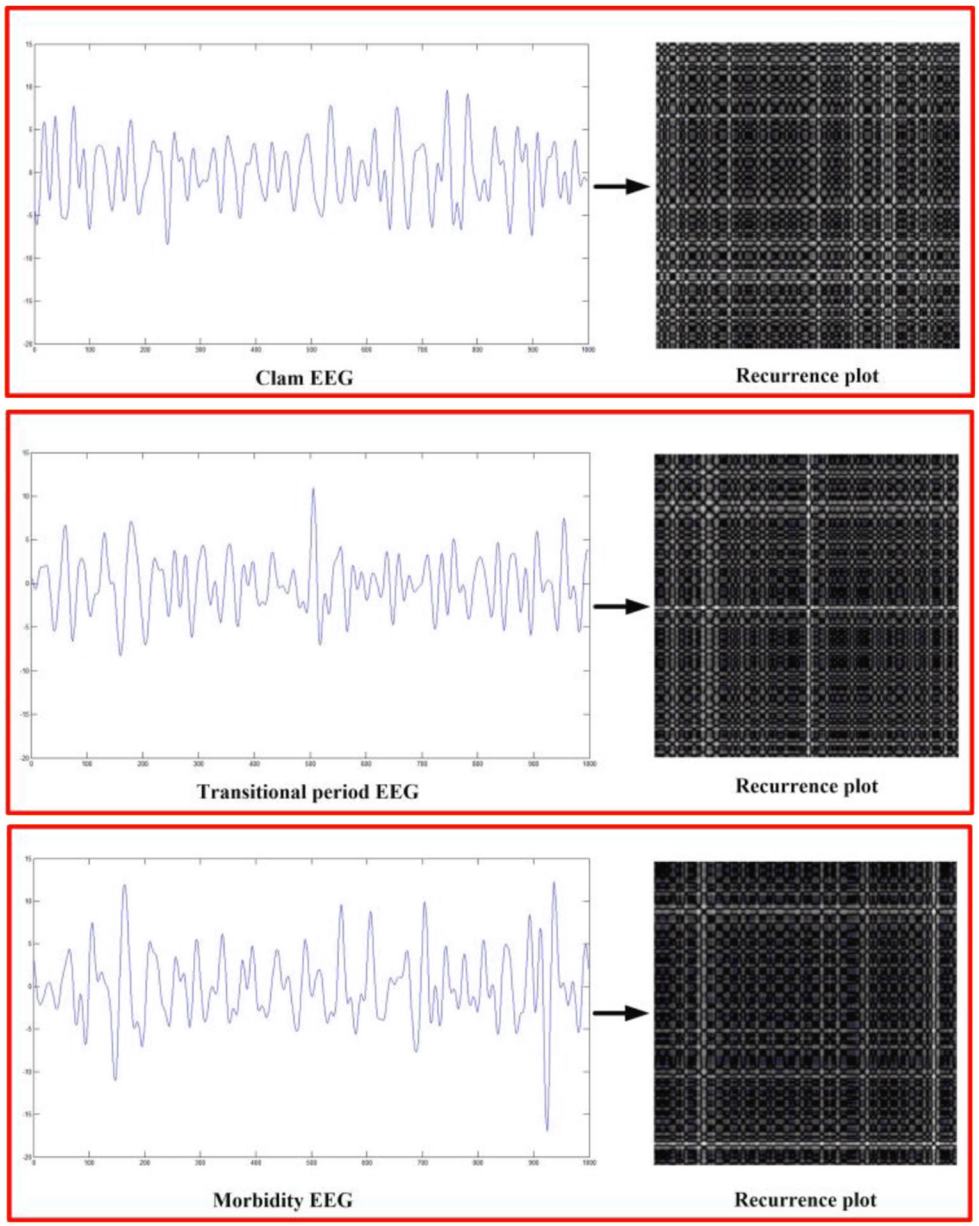


FIGURE 6 | Recurrence plot.

build a model, which enhances the model representation ability. After deeply studying the EEG transmission process, the PLV model is constructed to simulate the EEG transmission process to obtain the Alzheimer's disease transmission characteristics.

EEG can be regarded as a time series signal, and the traditional algorithm only builds the model from the 1-dimension perspective to carry out the study of similarity measures. Due to the complexity of the signal, a unified 1-dimensional model cannot be constructed. In this paper, 1-dimensional EEG is transformed into 2D recurrence plot to measure signal similarity in an intuitive way, and construct a threshold-free mechanism to quantify similarity. On this basis, a deep-learning network is constructed to simulate the cognitive process of physicians and obtain Alzheimer's disease signal characteristics.

A large number of clinical data show that Alzheimer's disease is very closely related to clinical characterization, and modeling from the signaling perspective alone has certain limitations. Clinical data collected from patients show that people with hypertension and diabetes have a high probability and rapid progression of Alzheimer's disease.

In this paper, based on EEG signals, signal transmission, signal similarity, and clinical characterization are combined to achieve the detection of Alzheimer's disease. Experiments

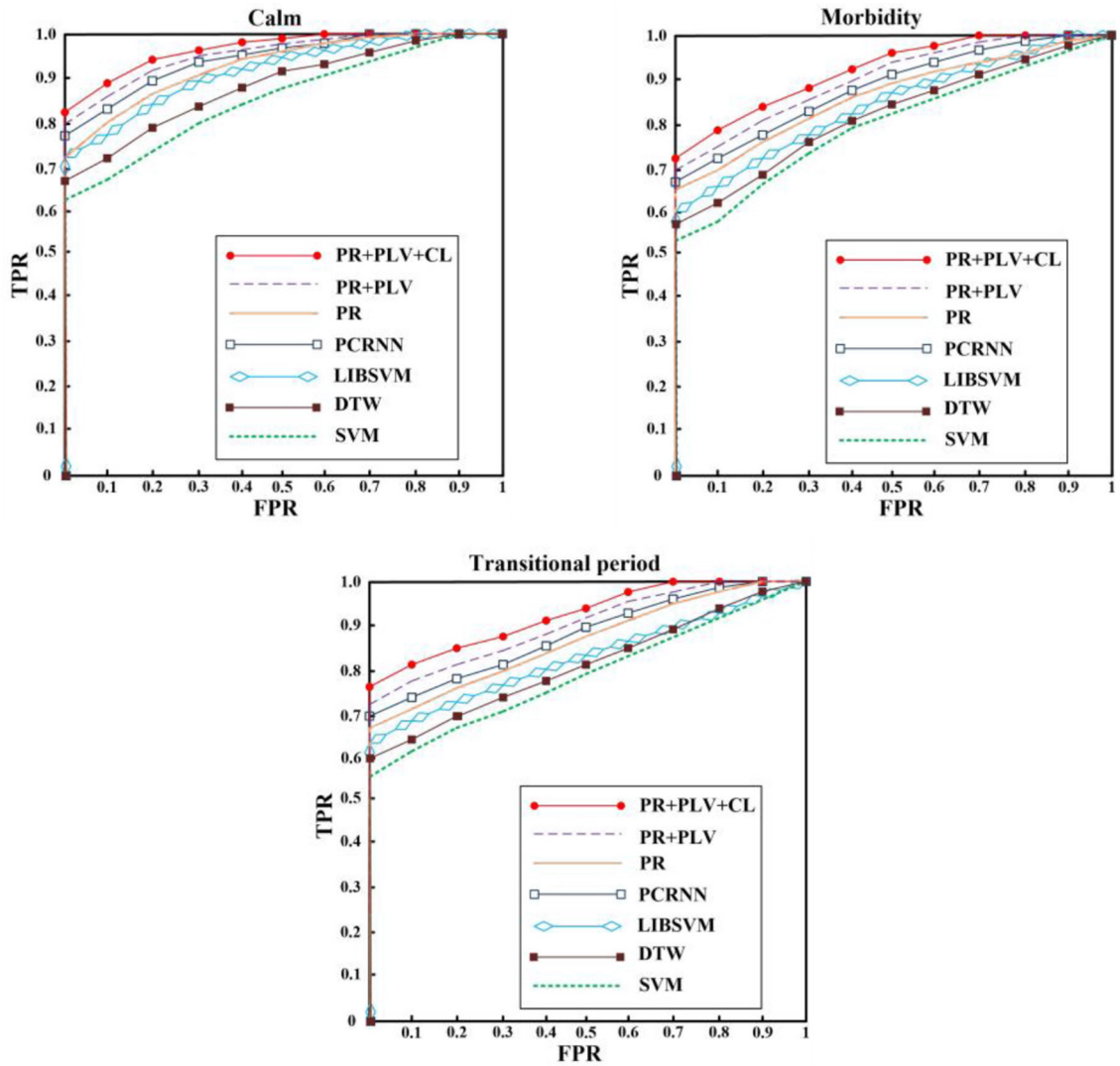


FIGURE 7 | Alzheimer's disease ROC curve.

show that the algorithm has strong robustness and detection rate. Subsequently, we will continue to collect data to expand the data set and carry out annotation and feature mining of typical data to assist doctors in accurate diagnosis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

All authors made significant contributions to the manuscript.

FUNDING

This work is supported by Natural Science Foundation of China under Grant No. 62062003, Key Research and Development Project of Ningxia (Special projects for talents) under Grant No. 2020BEB04022, and North Minzu University Research Project of Talent Introduction under Grant No. 2020KYQD08. Science and Technology Rising Star of Shaanxi Youth under Grant No. 2021KJXX-61.

ACKNOWLEDGMENTS

The authors are very grateful to the editor and peer reviewers and all those people who helped us a lot during the writing of this article.

REFERENCES

Al-Jumeily, D., Iram, S., Vialatte, F. B., Fergus, P., and Hussain, A. (2015). A novel method of early diagnosis of Alzheimer's disease based on EEG signals. *Sci. World J.* 2015, 931387. doi: 10.1155/2015/931387

Al-Nuaimi, A. H., Jammeh, E., Sun, L., and Ifeachor, E. (2016). "Changes in the EEG amplitude as a biomarker for early detection of Alzheimer's disease." *in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Orlando, FL: IEEE). 993–996. doi: 10.1109/EMBC.2016.7590869

Anh, V. H., Van, M. N., Ha, B. B., and Quyet, T. H. (2012). "A real-time model based support vector machine for emotion recognition through EEG." *in 2012 International Conference on Control, Automation and Information Sciences (ICCAIS)* (Saigon, Vietnam: IEEE). 191–196. doi: 10.1109/ICCAIS.2012.6466585

Bhat, S., Acharya, U. R., Dadmehr, N., and Adeli, H. (2015). Clinical neurophysiological and automated EEG-based diagnosis of the alzheimer's disease. *Eur. Neurol.* 74, 202–210. doi: 10.1159/000441447

Cassani, R., Falk, T. H., Fraga, F. J., Kanda, P. A., and Anghinah, R. (2014). The effects of automated artifact removal algorithms on electroencephalography-based Alzheimer's disease diagnosis. *Front. Aging Neurosci.* 6, 55. doi: 10.3389/fnagi.2014.00055

Chen, T., Ju, S., Ren, F., Fan, M., and Gu, Y. (2020). EEG emotion recognition model based on the LIBSVM classifier. *Measurement*. 164, 108047. doi: 10.1016/j.measurement.2020.108047

Chikara, R. K., Chang, E. C., Lu, Y. C., Lin, D. S., Lin, C. T., and Ko, L. W. (2018). Monetary reward and punishment to response inhibition modulate activation and synchronization within the inhibitory brain network. *Front. Hum. Neurosci.* 12, 27. doi: 10.3389/fnhum.2018.00027

Chikara, R. K., and Ko, L. W. (2019). Neural activities classification of human inhibitory control using hierarchical model. *Sensors.* 19, 3791. doi: 10.3390/s19173791

Deng, B., Cai, L., Li, S., Wang, R., Yu, H., Chen, Y., et al. (2017). Multivariate multi-scale weighted permutation entropy analysis of EEG complexity for Alzheimer's disease. *Cogn. Neurodyn.* 11, 217–231. doi: 10.1007/s11571-016-9418-9

Ding, Y., Chu, Y., Liu, M., Ling, Z., Wang, S., Li, X., et al. (2022). Fully automated discrimination of Alzheimer's disease using resting-state electroencephalography signals. *Quant. Imaging. Med. Surg.* 12, 1063. doi: 10.21037/qims-21-430

Falk, T. H., Fraga, F. J., Trambaiolli, L., and Anghinah, R. (2012). EEG amplitude modulation analysis for semi-automated diagnosis of Alzheimer's disease. *EURASIP J. Adv. Signal Process.* 2012, 1–9. doi: 10.1186/1687-6180-2012-192

Houmani, N., Vialatte, F., Gallego-Jutgl,à, E., Dreyfus, G., Nguyen-Michel, V. H., Mariani, J., et al. (2018). Diagnosis of Alzheimer's disease with Electroencephalography in a differential framework. *PloS ONE.* 13, e0193607. doi: 10.1371/journal.pone.0193607

Hulbert, S., and Adeli, H. (2013). EEG/MEG-and imaging-based diagnosis of Alzheimer's disease. *Rev Neurosci.* 24, 563–576. doi: 10.1515/revneuro-2013-0042

Kim, D., and Kim, K. (2018). "Detection of early stage Alzheimer's disease using EEG relative power with deep neural network." *in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Honolulu, HI: IEEE). 352–355. doi: 10.1109/EMBC.2018.8512231

Kulkarni, N. N., and Bairagi, V. K. (2017). Extracting salient features for EEG-based diagnosis of Alzheimer's disease using support vector machine classifier. *IETE J. Res.* 63, 11–22. doi: 10.1080/03772063.2016.1241164

Kumar, S., Mamun, K., and Sharma, A. (2017). CSP-TSM: optimizing the performance of Riemannian tangent space mapping using common spatial pattern for MI-BCI. *Comput. Biol. Med.* 91, 231–242. doi: 10.1016/j.compbiomed.2017.10.025

Li, K., Wang, J., Li, S., Yu, H., Zhu, L., Liu, J., et al. (2021). Feature extraction and identification of Alzheimer's disease based on latent factor of multi-channel EEG. *IEEE Trans. Neural Syst. Rehabilitation Eng.* 29, 1557–1567. doi: 10.1109/TNSRE.2021.3101240

Maturana-Candelas, A., Gómez, C., Poza, J., Pinto, N., and Hornero, R. (2019). EEG characterization of the Alzheimer's disease continuum by means of multiscale entropies. *Entropy*. 21, 544. doi: 10.3390/e21060544

Morabito, F. C., Labate, D., Bramanti, A., La Foresta, F., Morabito, G., Palamara, I., et al. (2013). Enhanced compressibility of eeg signal in alzheimer's disease patients. *IEEE Sens. J.* 13, 3255–3262. doi: 10.1109/JSEN.2013.2263794

Morabito, F. C., Labate, D., La Foresta, F., Bramanti, A., Morabito, G., and Palamara, I. (2012). Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG. *Entropy*. 14, 1186–1202. doi: 10.3390/e14071186

Oltu, B., Akşahin, M. F., and Kibaroglu, S. (2021). A novel electroencephalography based approach for Alzheimer's disease and mild cognitive impairment detection. *Biomed Signal Process Control.* 63, 102223. doi: 10.1016/j.bspc.2020.102223

Puri, D., Nalbalwar, S., Nandgaonkar, A., and Wagh, A. (2022). "EEG-Based Diagnosis of Alzheimer's Disease Using Kolmogorov Complexity." *in Applied Information Processing Systems.* Springer, Singapore. 157–165. doi: 10.1007/978-981-16-2008-9_15

Qiu, S., Li, J., Cong, M., Wu, C., Qin, Y., and Liang, T. (2020). Detection of solitary pulmonary nodules based on brain-computer interface. *Comput. Math. Methods Med.* 2020, 4930972. doi: 10.1155/2020/4930972

Rossini, P. M., Di Iorio, R., Vecchio, F., Anfossi, M., Babiloni, C., Bozzali, M., et al. (2020). Early diagnosis of Alzheimer's disease: the role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts. *Clin Neurophysiol.* 131, 1287–1310. doi: 10.1016/j.clinph.2020.03.003

Sarma, P., and Barma, S. (2022). Emotion Recognition by Discriminating EEG Segments With High Affective Content From Automatically Selected Relevant Channels. *IEEE Trans. Instrum. Meas.* 71, 4000812. doi: 10.1109/TIM.2022.3147876

Yang, Y., Wu, Q., Qiu, M., Wang, Y., and Chen, X. (2018). "Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network." *in 2018 International Joint Conference on Neural Networks (IJCNN)* (Rio de Janeiro: IEEE). 1–7. doi: 10.1109/IJCNN.2018.8489331

Yoon, S., Kim, S. E., Ko, Y., Jeong, G. H., Lee, K. H., Lee, J., et al. (2022). Differential expression of MicroRNAs in Alzheimer's disease: a systematic review and meta-analysis. *Mol. Psychiatr.* 1–9. doi: 10.1038/s41380-022-01534-6

Yu, H., Lei, X., Song, Z., Liu, C., and Wang, J. (2019). Supervised network-based fuzzy learning of EEG signals for Alzheimer's disease identification. *IEEE Trans Fuzzy Syst.* 28, 60–71. doi: 10.1109/TFUZZ.2019.2903753

Yu, M., Gouw, A. A., Hillebrand, A., Tijms, B. M., Stam, C. J., van Straaten, E. C., et al. (2016). Different functional connectivity and network topology in behavioral variant of frontotemporal dementia and Alzheimer's disease: an EEG study. *Neurobiol. Aging.* 42, 150–162. doi: 10.1016/j.neurobiolaging.2016.03.018

Zhao, Y., and He, L. (2014). "Deep learning in the EEG diagnosis of Alzheimer's disease," *in Asian Conference on Computer Vision.* Springer, Cham. 340–353. doi: 10.1007/978-3-319-16628-5_25

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 13 May 2022
doi: 10.3389/fnagi.2022.912283

3D FRN-ResNet: An Automated Major Depressive Disorder Structural Magnetic Resonance Imaging Data Identification Framework

Jialin Hong[1†], Yueqi Huang[2†], Jianming Ye[3†], Jianqing Wang[1], Xiaomei Xu[1], Yan Wu[1], Yi Li[1], Jialu Zhao[1], Ruipeng Li[4], Junlong Kang[5*] and Xiaobo Lai[1,6*]*

[1] School of Medical Technology and Information Engineering, Zhejiang Chinese Medical University, Hangzhou, China, [2] Department of Psychiatry, Hangzhou Seventh People's Hospital, Hangzhou, China, [3] First Affiliated Hospital, Gannan Medical University, Ganzhou, China, [4] Hangzhou Third People's Hospital, Hangzhou, China, [5] Zhongshan Hospital, Xiamen University, Xiamen, China, [6] Department of Nephrology Surgery, Hangzhou Hospital of Traditional Chinese Medicine, Hangzhou, China

OPEN ACCESS

Edited by:
Yizhang Jiang,
Jiangnan University, China

Reviewed by:
T. Niu,
Georgia Institute of Technology, United States
Linfei Huang,
China Pharmaceutical University, China
Gayathri Rajagopal,
Sri Venkateswara College of Engineering, India

***Correspondence:**
Ruipeng Li
705213095@qq.com
Junlong Kang
kjl295@163.com
Xiaobo Lai
dmia_lab@zcmu.edu.cn

[†] These authors have contributed equally to this work and share first authorship

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

Received: 04 April 2022
Accepted: 19 April 2022
Published: 13 May 2022

Citation:
Hong J, Huang Y, Ye J, Wang J, Xu X, Wu Y, Li Y, Zhao J, Li R, Kang J and Lai X (2022) 3D FRN-ResNet: An Automated Major Depressive Disorder Structural Magnetic Resonance Imaging Data Identification Framework. Front. Aging Neurosci. 14:912283. doi: 10.3389/fnagi.2022.912283

Major Depressive Disorder (MDD) is the most prevalent psychiatric disorder, seriously affecting people's quality of life. Manually identifying MDD from structural magnetic resonance imaging (sMRI) images is laborious and time-consuming due to the lack of clear physiological indicators. With the development of deep learning, many automated identification methods have been developed, but most of them stay in 2D images, resulting in poor performance. In addition, the heterogeneity of MDD also results in slightly different changes reflected in patients' brain imaging, which constitutes a barrier to the study of MDD identification based on brain sMRI images. We propose an automated MDD identification framework in sMRI data (3D FRN-ResNet) to comprehensively address these challenges, which uses 3D-ResNet to extract features and reconstruct them based on feature maps. Notably, the 3D FRN-ResNet fully exploits the interlayer structure information in 3D sMRI data and preserves most of the spatial details as well as the location information when converting the extracted features into vectors. Furthermore, our model solves the feature map reconstruction problem in closed form to produce a straightforward and efficient classifier and dramatically improves model performance. We evaluate our framework on a private brain sMRI dataset of MDD patients. Experimental results show that the proposed model exhibits promising performance and outperforms the typical other methods, achieving the accuracy, recall, precision, and $F1$ values of 0.86776, 0.84237, 0.85333, and 0.84781, respectively.

Keywords: major depressive disorder, deep learning, feature graph reconstruction network, structural magnetic resonance imaging, automated identification

INTRODUCTION

Major Depressive Disorder (MDD), one of the most common diseases associated with suicidal behavior, has become increasingly prevalent in recent years and is expected to be the largest contributor to the world's disease burden by 2030 (GBD 2017 Disease and Injury Incidence and Prevalence Collaborators, 2018). People with MDD are at higher risk for obesity, cardiovascular

disease, stroke, diabetes, cognitive impairment, cancer, and Alzheimer's disease. Approximately 8% of men and 15% of women suffer from depressive disorders during their lifetime, and nearly 15% of them choose to commit suicide (Gold et al., 2015). Therefore, it is crucial to diagnose MDD early and provide timely treatment.

Currently, the clinical diagnosis of MDD is mainly based on the relevant criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM), combined with the patient's interview and the subjective judgment of the clinician (Sakai and Yamada, 2019). The rapid development of medical imaging technology has provided more possibilities for pathological and identification studies of psychiatric disorders. Common medical imaging available includes Computerized Tomography (CT), Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI). Compared with other types of medical images, brain structural MRI (sMRI) images can describe changes in brain tissue volume or structure and reflect changes in neural activity in the brain. Therefore, sMRI is widely used to detect and treat psychiatric disorders. On the other hand, Segall et al. (2009) have found that sMRI of the brain can generate reliable and accurate brain volume estimates, making it practical to study the classification of depression based on brain sMRI images. However, due to the lack of clear physiological indicators, images of MDD patients cannot visually present abnormalities or lesions. Therefore, automated MDD identification is urgently needed in clinical practice.

Under the deep learning method, it is not easy to obtain many training samples, and the heterogeneity of MDD is substantial. Furthermore, most current deep learning networks rarely involve 3D data. How to apply deep learning framework to the identification task of MDD sMRI data has become a research hotspot and challenge. So far, many outstanding studies have been presented, such as Seal et al. (2021) proposed a deep learning-based convolutional neural network named DeprNet to classify Electroencephalogram (EEG) data from MDD patients and normal subjects. Baek and Chung (2020) proposed a contextual Deep Neural Network (DNN) model using multiple regression to efficiently detect depression risk in MDD patients. However, the methods above use only a 2D deep convolutional neural network, which cannot obtain the image's shallow and deep semantic features. It also easily leads to overfitting, which seriously affects the accuracy and robustness of the system and requires a considerable computational cost.

Previous methods rarely use sMRI data to identify MDD automatically and lack of MDD sMRI dataset, motivating us to start this study. Moreover, the primary purpose of this paper is to improve the automated identification accuracy of MDD effectively to help clinicians make a medical diagnosis. Therefore, we propose and develop an automated MDD sMRI data identification framework (3D FRN-ResNet), which introduces the Feature Map Reconstruction Network (FRN) based on the ResNet model. Its network structure is shown in **Figure 1**. Compared with other methods, our novel framework can preserve the granular information and details of the feature maps without overfitting the model. The contributions of our study are: (1) A feature map reconstruction network is proposed. (2) Building a 3D residual connectivity network to learn more deep features of sMRI images. (3) Preserving more texture details in sMRI images of MDD patients. (4) To get better identification results.

The remainder of this paper is organized as follows. After reviewing the state-of-the-art in the field of traditional machine learning-based methods, deep learning-based methods, and mental illness detection methods in Section "Related Works." Then, we explain our approach for solving the problem of MDD identification with sMRI data in Section "Materials and Methods." Then, we describe MDD sMRI dataset and the evaluation metrics, also the experimental details in Section "Experiments." Finally, the results and the discussions are described in Sections "Results and Discussion," followed by the conclusion in Section "Conclusion."

RELATED WORKS

Traditional Machine Learning

In recent years, machine learning techniques have been widely used to mine medical images as computer-aided diagnostic methods. Multivariate pattern analysis (MVPA), a data-driven machine learning method, has been used in diagnostic classification studies of mental disorders at the individual level (Bachmann et al., 2017). Researchers have classified feature selection algorithms into Filter-style feature selection algorithms and Wrapper-style feature selection algorithms based on the different feature evaluation strategies (Lazli et al., 2019). In the Filter feature selection model, Mwangi et al. (2012) used the *T*-test algorithm to implement feature selection and classification on a multicenter MDD dataset. Moreover, in the Wrapper model, Guyon et al. (2002) proposed a support vector machine-based recursive feature elimination (RFE-SVM) algorithm for gene sequence feature selection. This algorithm has been widely used in machine learning tasks for medical image analysis, such as Hidalgo-Muñoz et al. (2014) used the RFE-SVM algorithm to classify structural image features of Alzheimer's disease, which outperformed the *T*-test feature selection algorithm.

However, the Filter model usually has low computational intensity but poor classification accuracy; the Wrapper model has high classification accuracy but runs slowly, which is challenging to apply to datasets with many features. Therefore, researchers combined the advantages of both and proposed a combined Filter and Wrapper feature selection method to improve the classification accuracy while reducing the computational time overhead. Among them, Ding and Fu (2018) used the feature selection method combining the Filter model and Wrapper model to conduct experiments on several different types of datasets. The experimental results showed that the hybrid algorithm has high computational efficiency and classification accuracy (Ding and Fu, 2018). However, the drawback of the above methods is that they usually require manual feature design and redundant feature removal to extract useful distinguishable features.

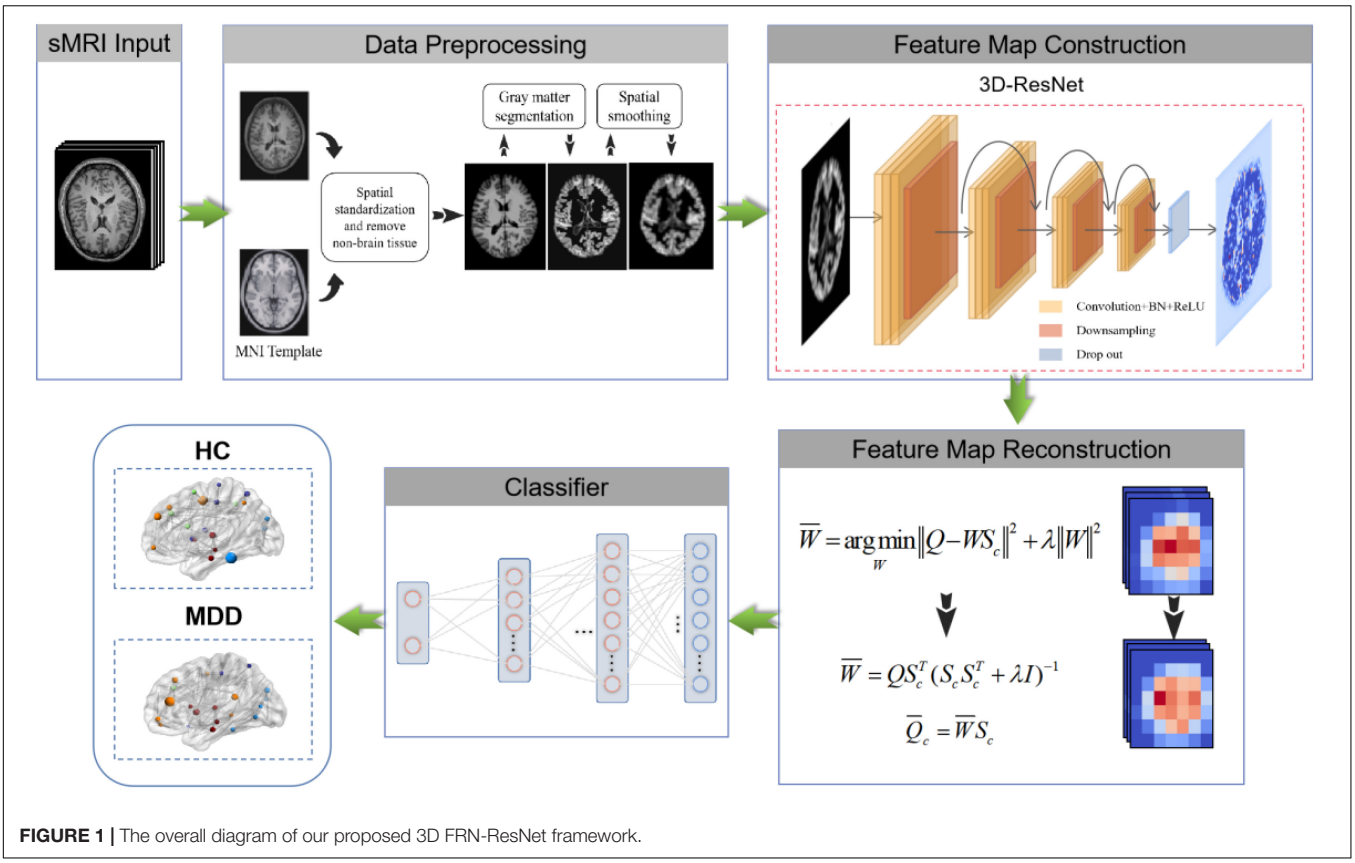


FIGURE 1 | The overall diagram of our proposed 3D FRN-ResNet framework.

Deep Learning

Deep learning techniques have led to remarkable progress in machine learning methods and promising results in medical image classification applications. Chen et al. (2021) proposed a cyclic Convolutional Neural Network (CNN) framework that can take full advantage of multi-scale and multi-location contexts in a single-layer convolution (LeCun et al., 1989). Cyclic CNNs can be easily plugged into many existing CNN pipelines, such as the ResNet family (He et al., 2016), resulting in highly low-cost performance gains (Chen et al., 2021). Liang and Wang (2022) proposed a novel model which uses involution and convolution (I-CNet) to improve the accuracy of image classification tasks by extracting feature representations on the channel and spatial domains. Wang et al. (2021) proposed a semi-supervised generative adversarial network (CCS-GAN) for image classification. It employs a new cluster consistency loss to constrain its classifier to maintain local discriminative consistency in each unlabeled image cluster. At the same time, an enhanced feature matching approach is used to encourage its generator to generate adversarial images from low-density regions of the true distribution, thus enhancing the discriminative ability of the classifier during adversarial training. The model achieves a competitive performance in semi-supervised image classification tasks (Wang et al., 2021). For fine-grained image classification, it has been a challenge to quickly and efficiently focus on the subtle discriminative details that make subclasses different from each other. Zhang et al. (2021) proposed a new multi-scale erasure and confusion method (MSEC) to address the challenge of fine-grained image classification.

Furthermore, Dai et al. (2021) proposed a model named TransMed for multimodal medical image classification in terms of medical image. TransMed combines the advantages of CNN and transformer to efficiently extract low-level features of images and establish long-range dependencies between modalities. The method has great potential to be applied to many medical image analysis tasks. Karthikeyan et al. (2020) used three pre-trained models-VGG16 (Simonyan and Zisserman, 2014), VGG19 (Simonyan and Zisserman, 2014), RESNET101 (He et al., 2016), on a dataset of X-ray images from patients with common bacterial pneumonia, COVID-19 patients, and healthy individuals to investigate migration learning methods. The proposed method obtained the best results (Karthikeyan et al., 2020). Talaat et al. (2020) proposed an improved hybrid image classification method that uses CNN for feature extraction and a swarm-based feature selection algorithm to select relevant features.

Mental Illness Detection

There are numerous mental illness detection algorithms, most of which are based on improvements to the basic deep learning framework. Payan and Montana (2015) used sparse autoencoder and 3D convolutional neural networks based on the Alzheimer's Disease Neuroimaging Initiative (ADNI) datasets to build algorithms that could predict patients' disease status, outperforming the latest research findings at the time. Similarly,

Farooq et al. (2017) applied deep convolutional neural networks such as Goolenet and ResNet on the ADNI dataset to learn discriminative features, achieving the purpose of classifying Alzheimer's disease (AD), mild cognitive impairment (MCI), advanced mild cognitive impairment (LMCI), and healthy individuals. Moreover, the prediction accuracy of the proposed technique was significantly improved compared (Farooq et al., 2017). Li and Liu (2018) applied the deep dense network (DenseNet) to the ADNI dataset. The original sMRI images did not need to be standardized preprocessing and directly extracted and classified features. The experimental results proved the effectiveness of the proposed method. Yang et al. (2022) proposed a spatial similarity-based perceptual learning and fusion deep polynomial network model to learn further robust information to detect obsessive-compulsive disorder (OCD); the model achieved promising performance in the rs-fMRI dataset of OCD patients. Ulloa et al. (2015) proposed a classification architecture using synthetic sMRI scans to scale up the sample size efficiently. A simulator that can capture statistical properties from observed data using independent component analysis (ICA) and random variable sampling methods was also designed to generate synthetic samples. Afterward, the DNN was specially trained on continuously generated synthetic data, and it significantly improved the generalization ability in classifying Schizophrenia patients and healthy individuals (Ulloa et al., 2015). Eslami et al. (2019) devised a data augmentation strategy to generate the synthetic dataset required to train the ASD-DiagNet model. The model consists of an auto-encoder and single-layer perceptron to improve the quality of extracted features and improve the detection efficiency of autism spectrum disorder.

Our Work

Although various deep learning frameworks have been proposed and significant progress has been made in the classification of brain tumor images. There are still challenges, such as insufficient sample size for training (Wertheimer et al., 2021), overfitting or underfitting due to the increased dimensionality of images (from 2D to 3D), and excessive consumption of computational resources (Pathak et al., 2019). In addition, the use of deep learning feature representation has weakened the interpretability of the features and is not conducive to the pathological analysis and understanding of the learned features (Zadeh Shirazi et al., 2020). These challenges limit the application of deep learning in medical images, so more innovative deep learning models are needed to achieve better results in medical images.

We propose a 3D FRN-ResNet framework for MDD sMRI images identification, which uses 3D-ResNet as the base framework. The conventional ResNet network incorporates pooling operations to extract global features, discarding a large amount of local detail information and thus reducing the resolution of the data. Specifically, during sMRI image processing of the brain, changes in neural activity in abnormally active (or inactive) brain regions are difficult to capture, but these small changes may be necessary for MDD. To solve this problem, we introduce the FRN method so that the granularity information and details of the feature map can be retained without overfitting the model. Its network structure is shown in **Figure 1**. It achieves this by framing class membership as a problem in reconstructing the feature map. Given a set of images belonging to a single class, we generate the associated feature maps and collect the component feature vectors across locations and images into a single pool of support features. For each query image, we attempt to reconstruct each location in the feature map as a weighted sum of the support features with a negative mean squared reconstruction error as the class score. Images from the same class should be easier to reconstruct because their feature maps contain similar embeddings, while images from different classes are more complex and produce larger reconstruction errors. By evaluating the reconstruction of the complete feature map, FRN preserves the spatial details of the appearance. Additionally, by allowing this reconstruction to use feature vectors from any location in the support image, FRN explicitly discards the annoying location information. An auxiliary loss function is also introduced, which encourages orthogonality between features of different classes to focus on feature differences.

We evaluate the performance of the proposed model on a constructed sMRI dataset of MDD patients and compare it with other methods. The results show that our model has good performance in automated MDD sMRI data identification. (1) A novel identification network structure based on feature map reconstruction is proposed in this paper. (2) Feature extraction followed by feature map reconstruction of sMRI images retains more fine spatial details and dramatically improves the identification performance. (3) Classification-assisted loss functions are developed to distinguish between different features classes.

MATERIALS AND METHODS

Our goal is to identify MDD using sMRI images automatically. In order to obtain good identification performance, a robust network structure is usually required. Therefore, we propose the 3D FRN-ResNet model for automated MDD sMRI data identification, consisting of a feature extraction network and a feature map reconstruction network. The network structure of 3D FRN-ResNet is shown in **Figure 1**. This section describes the preprocessing process, the network structure of 3D FRN-ResNet, and the loss function used in detail.

Data Preprocessing

The sMRI data preprocessing work is implemented using the MATLAB-based SPM12 toolkit (Ashburner et al., 2021). The main contents of preprocessing include AC-PC calibration, non-brain tissue removal, gray matter segmentation, spatial standardization, and spatial smoothing. The size of sMRI data for each subject after processing is $121 \times 145 \times 122$ voxels.

Anterior Commissure-Posterior Commissure Calibration

The calibration procedure focuses on the anterior commissure (AC) and posterior commissure (PC) calibration. We use MATLAB software to perform AC-PC calibration, resampling the images in the standard $256 \times 256 \times 256$ mode, and then the

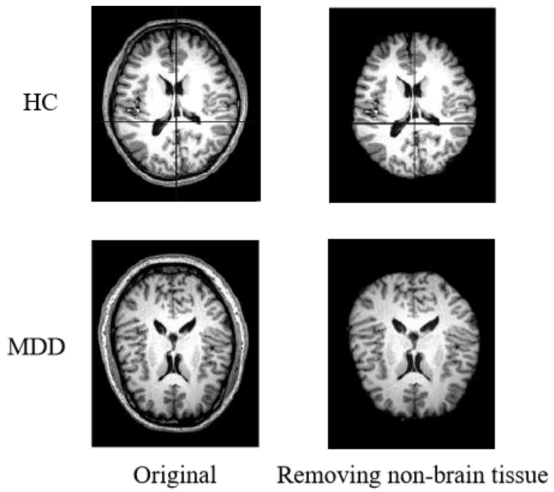


FIGURE 2 | Results of removing non-brain tissue.

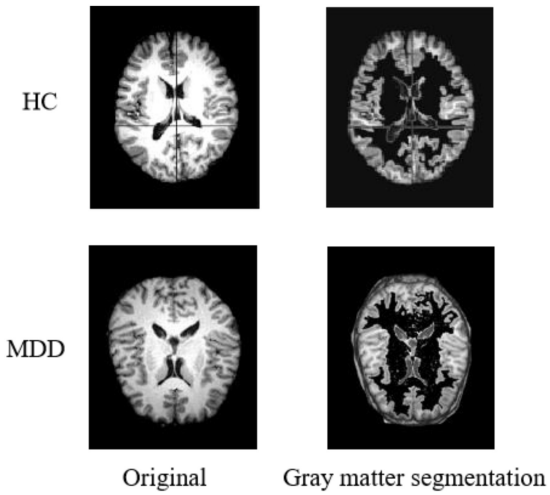


FIGURE 3 | Results of gray matter segmentation.

N3 algorithm is used to correct for non-uniform tissue intensity. We also perform skull stripping and cerebellar resection after correcting the images by AC-PC correction.

Non-brain Tissue Removal

The original images of sMRI contain some non-brain structures, such as skulls. In order to avoid increasing the computational workload and to avoid subsequent image preprocessing, which may affect the experimental results. Non-brain structures such as skulls need be removed from the images during the image preprocessing operation. **Figure 2** shows the comparison of a sample before and after removing non-brain tissue.

Gray Matter Segmentation

During sMRI image processing, sometimes only the state of specific regions is focused on, which requires tissue extraction from the target area according to the brain's anatomy. In the preprocessing process, we segment the sMRI into three different images by brain gray matter, white matter, and cerebrospinal fluid structures. Considering the critical influence of the gray matter region on the diagnosis of MDD (Arnone et al., 2013), only the gray matter part is used for the experiments in this paper. **Figure 3** shows the result of gray matter segmentation.

Spatial Standardization

Standardization is the alignment of the images from the previous preprocessing process to the standard brain template space Montreal Neurological Institute (MNI) to unify the coordinate space of all images. The algorithms used for standardization are non-rigid body alignment algorithms, including affine and non-linear transformations. **Figure 4** shows the comparison of a sample before and after spatial standardization.

Spatial Smoothing

After completing the above series of processing, it is also necessary to perform a smoothing process on the image to suppress the noise of the functional image. Additionally, the signal-to-noise ratio needs to be improved to reduce anatomical

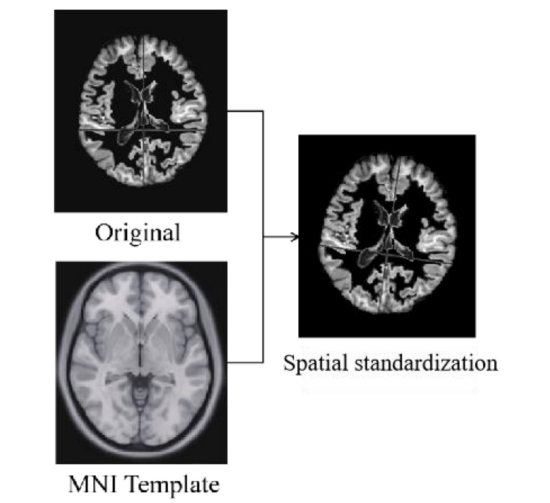


FIGURE 4 | Results of spatial standardization.

or functional differences between images. Usually, the function used for the smoothing process is the Gaussian kernel function. In addition, based on experience and practical attempts, we use a 64 × 64 × 64 pixel cube to down-sample gray matter density images, and this processing saves computing time and memory consumption with no loss of classification accuracy. **Figure 5** shows the comparison of a sample before and after spatial smoothing.

3D-ResNet Framework

Although ResNet has achieved excellent results on many 2D natural image datasets, it has little success in medical images. The reason is that the convolution kernels and pooling kernels in 2D networks are two-dimensional matrices. It can only move in the two directions of height H and width W of 2D flat images, so only 2D features can be extracted. In contrast, most medical image

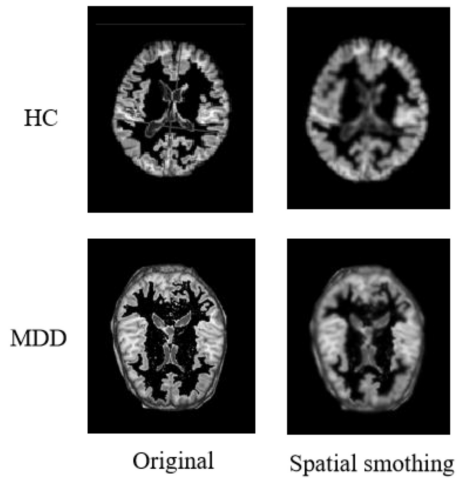


FIGURE 5 | Results of spatial smoothing.

data such as sMRI are 3D stereo data. When using 2D network processing, only 3D images can be input in layers, or one of the dimensions can be used as the channel dimension. But neither of the two methods can make good use of the inter-layer structure information of the data.

Based on this, this paper adds a depth dimension D to the filters such as convolution kernels and pooling kernels in the 2D network, and extends them into 3D matrix. In this way, the filters can be moved in all 3 directions (H, W, D) of the sMRI data, so that the spatial information of the data can be fully exploited. And the output of each filter is also a 3D data. The structure diagram of the 3D-ResNet is shown in **Figure 6**. Let the size of one of the 3D convolution kernels is $k \times k \times k \times channel$, the number is n, the input data size is $h \times w \times d$. And since the sMRI data used in this paper is similar to a grayscale map, the channel dimension is 1. Therefore, the output size of this convolution kernel is as follow:

$$(h-k+1) \times (w-k+1) \times (d-k+1) \times n \tag{1}$$

By a similar method, the pooling layer and batch normalization layer in ResNet can be extended to construct a 3D residual connected network (3D-ResNet). The network can better extract representative features from 3D sMRI data and improve the accuracy of identification in MDD patients.

The 3D-ResNet network structure is shown in **Figure 6**. Due to the small size of the input region feature map, the convolution pooling operation is removed from the bottom layer of the network. And the input map is directly made to enter the residual network consisting of four stacked residual convolution modules.

Figure 7 shows an example of feature extraction from the 3D-ResNet middle layer. At the end of the extraction process, the network learns details such as contour boundaries, position, and orientation, enabling more learning of deeper features in the sMRI and preparing it for the next step.

Feature Map Reconstruction Networks Framework

The feature extractor can produce a feature map. However, the distance metric function requires a vector representation of the whole graph. Therefore, a method needs to be found to convert the feature map into a vector representation. Ideally, this conversion would preserve the granularity of information and details of the feature map without overfitting the model. But existing methods, such as global average pooling, are very crude in discarding some spatial information or flattening a feature map into a long vector, which also loses location information. In order to convert the feature map into a vector representation while preserving the spatial details, Feature Map Reconstruction Networks (FRN) are proposed in this paper.

When there is a single input image x_q, we wish to predict its label y_q. Firstly, let x_q passes through feature extractor to generate a feature map $Q \in \mathrm{R}^{r \times d}$, where r represents the size of the space and d is the number of channels. For each class $c \in C$, we pool all features from the k input images into a feature matrix $S_c \in R^{kr \times d}$.

Then, we try to reconstruct Q as a weighted sum of rows in S_c by finding the matrix $W \in Rr \times kr$ so that $W \times S_c \approx Q$ can be obtained. Finding the optimal $\overline{W}$ is equivalent to solving the linear least squares problem:

$$\overline{W} = \arg\min_{W} ||\mathrm{Q\text{-}WS}_c||^2 + \lambda ||W||^2 \tag{2}$$

where $||\bullet||$ is the Frobenius norm, which λ is a weighted ridge regression penalty term used to ensure the treatability of the linear system when it is over- or under-constrained ($kr \neq d$).

The ridge regression equation leads to the optimal solution W and Q_c.

$$\overline{W} = QS_c^T(S_cS_c^T + \lambda I)^{-1} \tag{3}$$

$$\overline{Q}_c = \overline{W}S_c \tag{4}$$

For a given class c, the distance between Q and Q_c is defined as the Euclidean distance and then deflated by using $\frac{1}{r}$. A learnable temperature factor λ is also introduced. The final predicted probability is thus given by:

$$\langle Q, \overline{Q}_c \rangle = \frac{1}{r}||Q - \overline{Q}_c||^2 \tag{5}$$

$$P(y_q = c|x_q) = \frac{e^{(-\gamma\langle Q,\overline{Q}_c\rangle)}}{\sum_{c' \in C} e^{\left(-\gamma\left\langle Q,\overline{Q}_{c'}\right\rangle\right)}} \tag{6}$$

In order to ensure the stability of the training, we decide to use $\frac{1}{kr}$ to improve λ. This has the additional benefit of making our model somewhat robust, in addition to the parameters that λ should be learned. The change λ has diverse effects: the large one λ avoids over-reliance on the weights of W, but it also reduces the effectiveness of the reconstruction. And it increases the reconstruction errors as well as limit the distinguishability. Therefore, we disentangle the degree of regularization ρ from

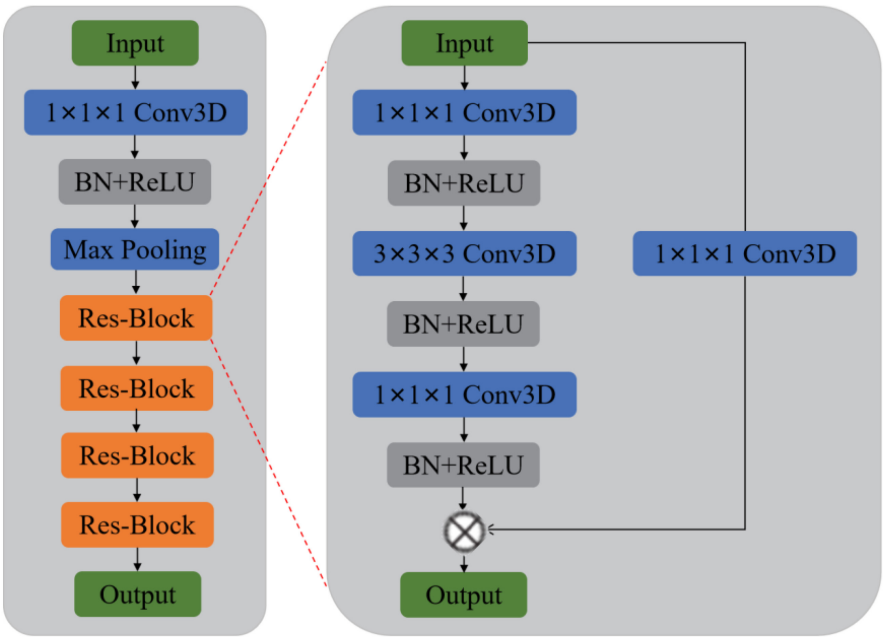


FIGURE 6 | Proposed 3D-ResNet structure.

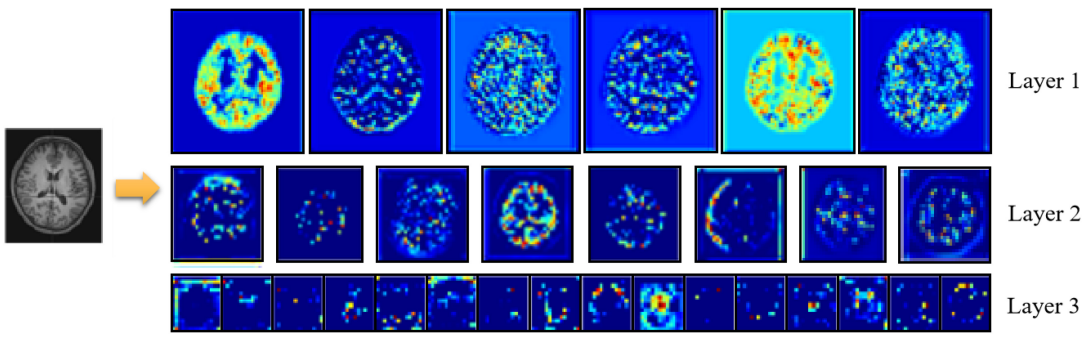


FIGURE 7 | Visualization of extracted features.

the magnitude of Q_c by introducing a learned recalibration term. This leads to the following formula:

$$\overline{Q}_c = \rho \overline{W} S_c \tag{7}$$

λ andρare parameterized as e^{α} and e^{β} to ensure non-negativity and are initialized to zero. In summary, our final prediction is given by the following equation.

$$\lambda = \frac{kr}{d} e^{\alpha} \qquad \rho = e^{\beta} \tag{8}$$

$$\overline{Q}_c = \rho \overline{W} S_c = \rho Q S_c^T (S_c S_c^T + \lambda I)^{-1} S_c \tag{9}$$

$$P(y_q = c | x_q) = \frac{e^{(-\gamma \langle Q, \overline{Q}_c \rangle)}}{\sum_{c' \in C} e^{(-\gamma \langle Q, \overline{Q}_{c'} \rangle)}} \tag{10}$$

The method introduces only three learning parameters:α,β, andγ. The temperature γ appears in previous works (Chen et al., 2020).

Figure 8 is a diagram of the FRN network structure. Support image is converted to a feature map (left) and aggregated to a pool of class conditions (middle). A best-fit reconstruction of the query feature map is computed for each class, and the closest candidate generates the predicted class (right). Among them, $h \times w$ is the feature map resolution, d is the number of channels, and the green triangle represents the convolutional feature extractor.

Loss Function

Medical image classification often faces the problem of minor differences in the appearance of pathological targets and non-targets. We also face this challenge for our MDD brain tumor classification task. For this purpose, our loss function consists

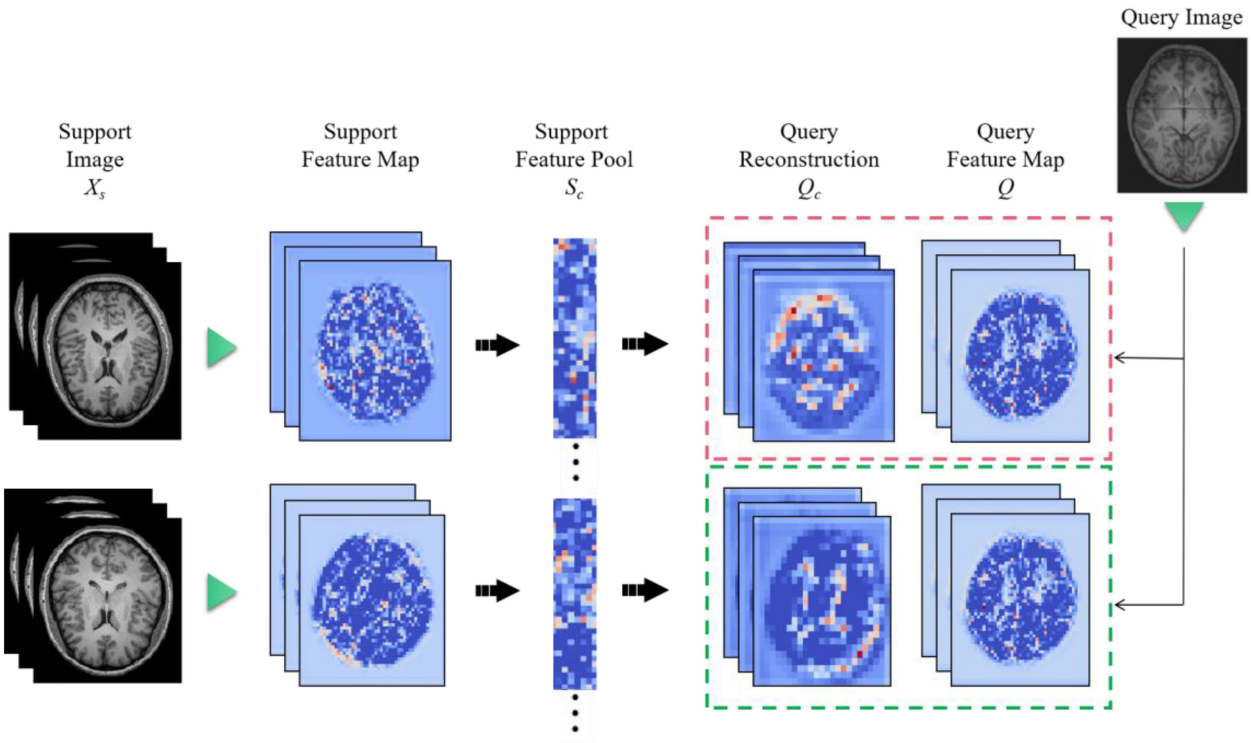


FIGURE 8 | Feature map reconstruction networks network structure diagram.

of two components. The first is the cross-entropy loss function, which can be understood as a composition of two parts. The first part is the calculation of the mutual entropy with label 1, and the second part is the calculation of the mutual entropy with label 0. We sum the two to obtain the overall mutual entropy. The formula is as follows.

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(p_i) + (1-y_i)\log(1-p_i)\right] \tag{11}$$

where N is the total number of samples, y_i is the category to which the *ith* sample belongs, and p_i is the predicted value of the *ith* sample.

In addition to the classification loss, we use an auxiliary loss that encourages support features from different classes to span the potential space.

$$L_{aux} = \sum_{i\in C}\sum_{j\in C, j\neq i} ||\widehat{S}_i\widehat{S}_j^T||^2 \tag{12}$$

Among then, $\widehat{S}$ is line normalized and projects the features onto the unit sphere. This loss encourages orthogonality between features from different classes. Similar to Christian et al. (2020), we reduce this loss by a factor of 0.03. We use L_{aux} as the auxiliary loss in our subspace network implementation, which replaces the SimCLR fragment in the cross-transformer implementation (Carl et al., 2020).

EXPERIMENTS

Dataset

The benchmarking clinical MDD sMRI images dataset is collected at the Seventh Hospital of Hangzhou (SHH) with Institutional Review Board (IRB) approval, and is used to train and test our model. Furthermore, the SHH dataset contains 68 subjects, including 34 MDD patients and 34 healthy controls (HC). All patients with MDD met the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) for MDD. And all healthy controls passed the non-patient version of the structured clinical interview of the DSM-IV. All sMRI images have an imaging field of view (FOV) = 240 mm × 256 mm, a voxel size of 1 mm × 1 mm × 1 mm, a layer thickness of 1 mm, and a scan layer count of 192. sMRI slice images from the MDD and HC in SHH dataset are shown in **Figure 9**.

Evaluation Metrics

A total of 54 samples in SHH dataset are used in the training process, including 27 MDD patients and 27 healthy individuals. In addition, 14 samples are used for validation, including 7 MDD patients and 7 healthy individuals. We use four metrics to evaluate the model performance: Accuracy, Recall, Precision, and $F1$ score. Accuracy is calculated as:

$$Accuracy = \frac{TN + TP}{FP + TN + TP + FN} \tag{13}$$

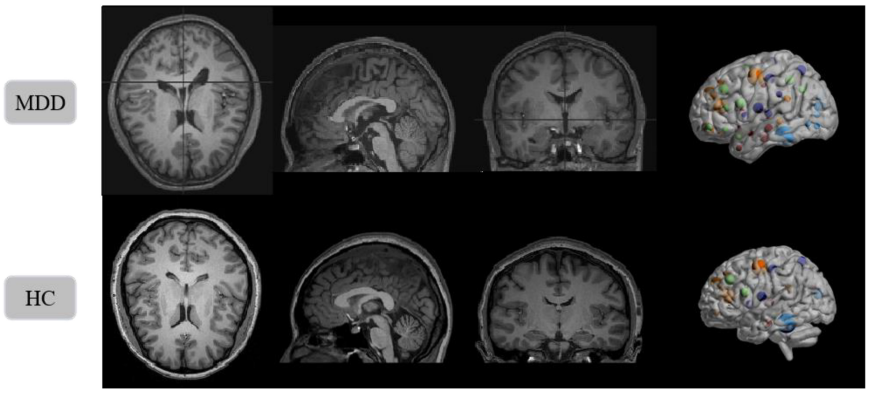


FIGURE 9 | Structural magnetic resonance imaging slice images from the MDD and HC in SHH dataset. Left to right: axial view, sagittal view, coronal view, and 3D presentation.

where TN, TP, FP, and FN are the number of true negative, true positive, false positive, and false negative, respectively. Recall refers to the ability of a classifier to correctly detect positive samples, reflecting the proportion of patients with MDD that are correctly determined as a percentage of the total number of patients, defined as:

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

Precision refers to the proportion of samples with a positive prediction that are correctly predicted, defined as:

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

Precision and Recall are contradictory metrics. In general, Recall tends to be low when Precision is high, while Recall tends to be high when Precision is low. When the classification confidence is high, Precision is high; when the classification confidence is low, Recall is high. To be able to consider these two metrics together, the weighted average F-measure of Precision and Recall is proposed, which reflects the overall metric, defined as:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{16}$$

In disease diagnosis studies, the higher the recall rate, the smaller the missed diagnosis rate. Therefore, the accuracy and recall of models are of most interest.

Experimental Details

In deep learning training, the setting of hyperparameters is critical and determines the performance of our model. In the training of the 3D FRN-ResNet model, the initial learning rate is set to 0.01, the weight decay value is set to 0.001, the number of epochs is 100, and then the learning rate is changed to 0.1 times when the validation set loss value does not drop for 10 consecutive epochs. Considering the sample size limitation and using a fivefold cross-validation method to enhance the model's generalization ability.

All experiments are performed on a CentOS server with NVIDIA TITAN Xp GPU, dual-core Intel(R) Xeon(R) Silver 4210 CPU @ 2.20 GHz processor, Python 3.6 programming language, and PyTorch 1.0 deep learning framework.

RESULTS

We use four metrics, Accuracy, Recall, Precision, and *F*1 value, to measure model performance. The average results of the metrics obtained on the training and validation sets are shown in **Table 1**. The experimental results show that the model has good robustness. We can observe that the Recall is at a high value, which indicates that the model is quite comprehensive in MDD patient identification. Furthermore, we can see that the Precision is also at a high value, demonstrating that the model has a good ability in MDD patient identification. In addition, the Recall of the training set is 0.84, and the *F*1 value of the training set is 0.85, which is very close. The same is true in the validation set, suggesting the ability to discriminate between healthy and MDD patients in our model is about the same.

Figures 10, **11** respectively show the composite plot of the scatter plot and box plot of the evaluation index results of the training set and the validation set. It can be seen from the box plots that the fluctuations of the results are tiny, and only a few outliers appear. In the box plot, the horizontal line in the middle of the box indicates the median of a dataset. It can also be observed in the scatter plot that the recall rate reaches a high range, and the recall rate represents the ability of the model to diagnose patients who suffer from MDD. The smaller the difference between the Recall and *F*1 values, the better the model's performance in resolving class imbalance. It can be seen

TABLE 1 | Test results on the training and validation sets.

	Accuracy	Recall	Precision	*F*1
Training	0.86	0.84	0.85	0.85
Validation	0.78	0.76	0.77	0.76

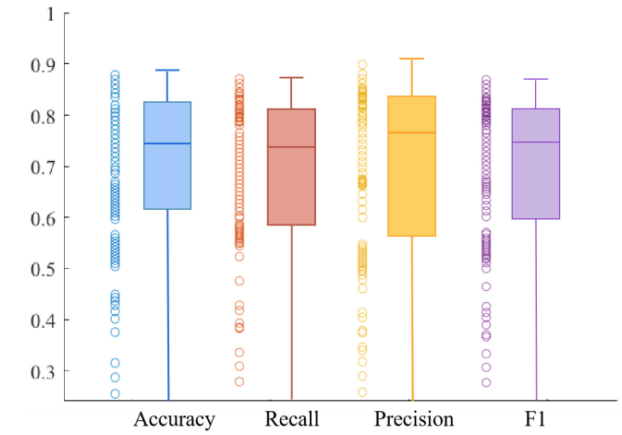


FIGURE 10 | Combination of scatter plot and box plot of training set.

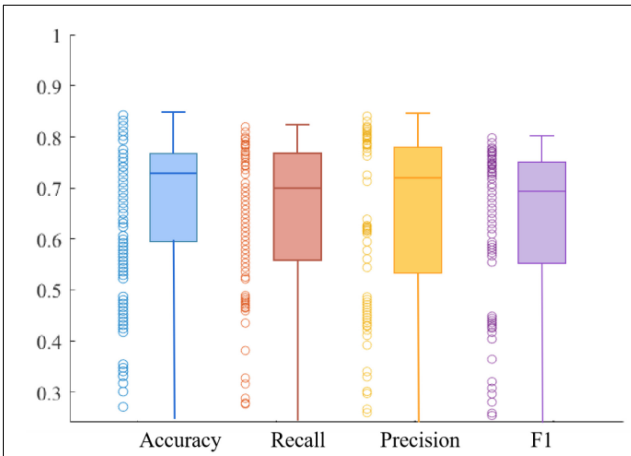


FIGURE 11 | Combination of scatter plot and box plot of validation set.

from the figure below that the recall fluctuation range is not large, indicating that the model has the similar ability to predict the MDD patients and healthy individuals. After validation, the overall performance of the model reached a high level.

In order to explore the influence of different network structures on the performance of the MDD identification algorithm, firstly we use five feature extraction networks with different structures for training based on the FRN structure in the classification layer. After that, the 3D-Resnet structure with the best effect is used as the feature extraction network, and the FRN structure is replaced with a general fully connected layer for classification. The experimental results show that compared with ordinary convolutional networks, ResNet and DenseNet structures can extract and retain richer detail information, and learn feature representations with strong discriminative power, thereby effectively improving the identification accuracy of the network.

From **Table 2**, we can see that the structural model combining 3D-ResNet and FRN has the highest classification accuracy, with the correct rate and recall rate achieving 85 and 84%, respectively. We can also observe that accuracy and recall have been significantly improved after the 3D operation of the network. For example, the identification accuracy of 3D-ResNet is 6% higher than that of 2D DenseNet, which shows that the 3D-ResNet proposed in this paper can mine effective information, providing more effective features than the general ResNet and the traditional 2D networks. Meanwhile, it can be seen from **Table 2** that the FRN network can effectively improve the high heterogeneity problem in the sMRI images of MDD patients and thus is applicable in MDD sMRI images identification.

Figure 12 shows the ROC curves of different algorithms using FRN-net on the SHH dataset. It can be seen that from the figure out algorithms outperforms others, which further confirms the effectiveness of our algorithm. The main reason is that we exploit both multi-scale layers and contextual spatial information to reduce the semantic gap to a large extent.

The results of the ROC curve in **Figure 13** are consistent with those in **Figure 12**, indicating that our algorithm does improve image identification accuracy. On the one hand, our algorithm proposes a 3D residual connection network, which extends the idea of residual connections to three dimensions. It makes full use of the spatial and contextual information of the image, and preserves the spatial details when converting the extracted features into vectors and location information. Thus, higher average accuracy than other methods is achieved, which also demonstrate the effectiveness of the 3D residual connection network and classification based on feature map reconstruction. On the other hand, since we decompose the image into multiple-scale layers, sufficient scale information is used when generating multi-scale visual histograms. Therefore, our method has the best classification specificity and sensitivity.

To further illustrate that the feature map reconstruction method proposed in this paper is informative for correct classification, we obtain experimental results for each query image. In **Table 3**, all networks are trained with 3D-ResNet as the backbone. The results in their tables validate the effectiveness of the classification method based on the feature map reconstruction proposed in this paper.

Figure 14 illustrates the algorithm's performance based on the above test parameters. The proposed FRN can be predicted to be the best due to its property of classifying affected regions spread over a given image from a performance overview. 3D ResNet guarantees its performance in computation time and average

TABLE 2 | Results comparison with different network structures.

Model	Backbone	Accuracy	Recall	Precision	*F*1
FRN(ours)	3D-ResNet	**0.85**	**0.84**	0.86	**0.84**
	ResNet	0.79	0.78	0.80	0.79
	3D-DenseNet	0.84	0.82	**0.87**	0.84
	DenseNet	0.78	0.78	0.79	0.78
	SimpleCNN	0.60	0.58	0.61	0.60
Full connected	3D-ResNet	0.82	0.80	0.82	0.81

Bold values mean the best performance.

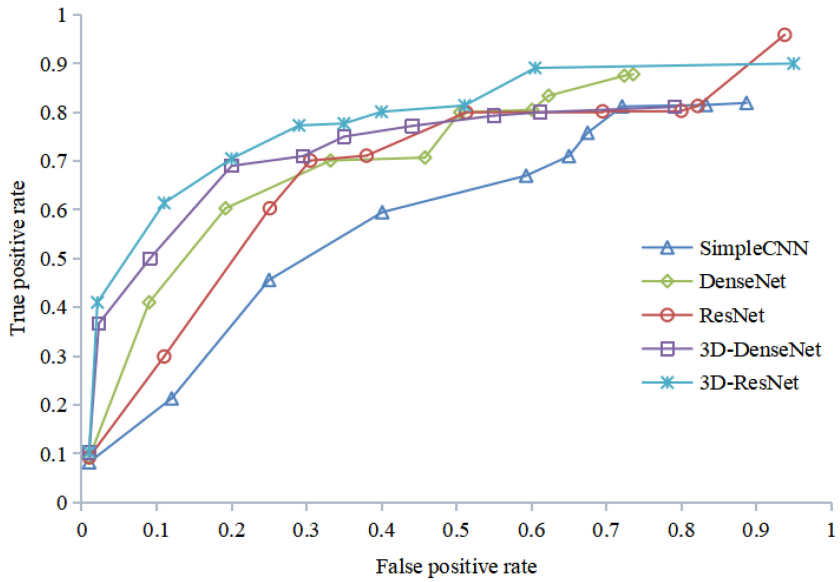


FIGURE 12 | ROC curves with FRN-Net for different backbones of the training set.

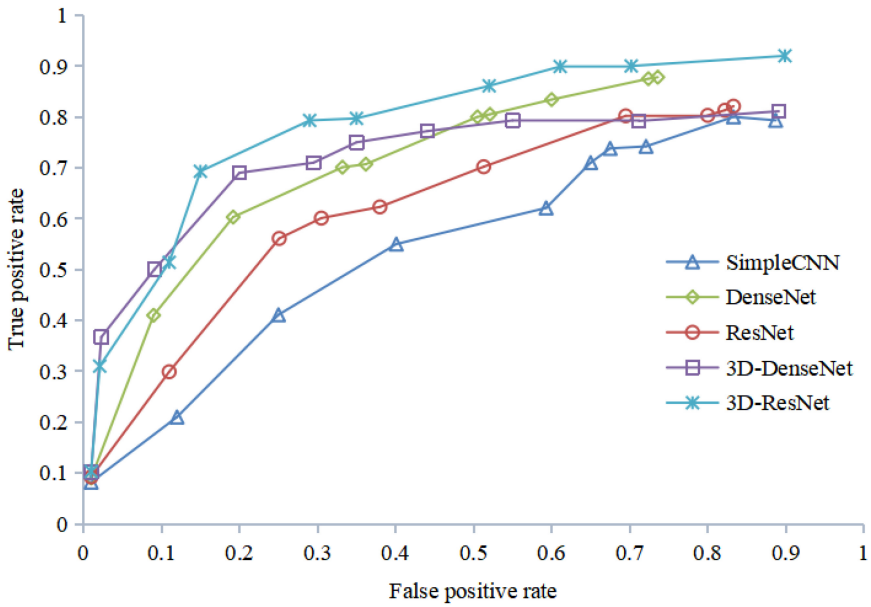


FIGURE 13 | ROC curves with FRN-Net for different backbones of the validation set.

accuracy for medical image datasets, with the highest recall and satisfying precision. Statistical, visual, and experimental evidence is provided through comparisons with other algorithms.

To sum up, through the above experiments, we can see that the performance of the ProtoNet method is not as good as other methods. Because traditional ProtoNet algorithms extract feature histograms through direct statistical methods, which are linear features that need to be combined with non-linear classifiers to perform well. The DSN method outperforms the ProtoNet method, probably because the DSN algorithm predicts class membership by computing the distances between query points and their projections into the latent subspace formed by the supporting images of each class, which improves methods for image predictive classification. Whereas the CTX method explicitly produces class-level linear reconstructions and outperforms the DSN method. Our algorithm decomposes the image into multi-scale layers and performs 3D residual network feature extraction and feature map reconstruction to predict classification, greatly enhancing the discrimination of image feature representation. Therefore, our algorithm has the

TABLE 3 | Results comparison with different classifiers.

	Model	Accuracy	Recall	Precision	*F*1
Train	ProtoNet	0.82	0.81	0.83	0.82
	DSN	0.81	0.79	0.82	0.81
	CTX	0.80	0.79	0.81	0.80
	FRN(ours)	**0.86**	**0.83**	**0.84**	**0.83**
Validation	ProtoNet	0.76	0.75	0.77	0.76
	DSN	0.74	0.72	0.75	0.74
	CTX	0.75	0.74	0.76	0.75
	FRN(ours)	**0.80**	**0.78**	**0.76**	**0.77**

TABLE 4 | Results comparison with typical methods.

Method	Accuracy	Recall	Precision	*F*1
Jiao et al., 2017	0.82	0.79	**0.84**	0.81
An et al., 2021	0.81	0.79	0.82	0.80
Ben et al., 2020	0.79	0.77	0.81	0.79
Cheng et al., 2022	0.83	0.80	0.82	0.81
Abdar et al., 2021	0.81	0.80	0.81	0.79
Proposed	**0.85**	**0.82**	0.82	**0.82**

Bold values mean the best performance.

best average classification accuracy, specificity, and sensitivity, which indicates that 3D FRN-ResNet indeed improves image classification accuracy. On the one hand, our algorithm proposes a 3D residual connection network, which extends the idea of residual connections to three dimensions. It makes full use of the spatial and contextual information of the image and preserves the spatial details when converting the extracted features into vectors and location information. Thus, higher average accuracy than other methods is achieved, demonstrating the effectiveness of the 3D residual connection network and classification based on feature map reconstruction. On the other hand, since we decompose the image into multiple-scale layers, sufficient scale information is used when generating multi-scale visual histograms. Therefore, our algorithm has the best classification specificity and sensitivity.

DISCUSSION

The 3D FRN-ResNet proposed in this paper can effectively improve the identification accuracy and recall rate of sMRI data from MDD patients and healthy controls, and verifies its effectiveness and feasibility. The proposed model can assist physicians to complete the diagnosis, and has significant significance in research value.

The method is compared with some typical medical image classification algorithms, and the results are shown in **Table 4**. All of these methods use the private SHH dataset. These results can be compared with those obtained using the proposed method. Our proposed method is one of the best and achieves better performance than other methods evaluated under the same conditions. Jiao et al. (2017) introduced a joint model with a CNN layer and a parasitic metric layer. Where the CNN layer provides the essential discriminative representation, and the metric learning layer enhances the classification performance for that particular task (Jiao et al., 2017). An et al. (2021) proposed a multi-scale convolutional neural network, a medical classification algorithm based on a visual attention mechanism, which automatically extracts high-level discriminative appearance features from the original image. In the method proposed by Ben et al. (2020), a new classification framework was developed to classify medical images using sparse coding and wavelet analysis, which showed a significant improvement in identification accuracy. Cheng et al. (2022)

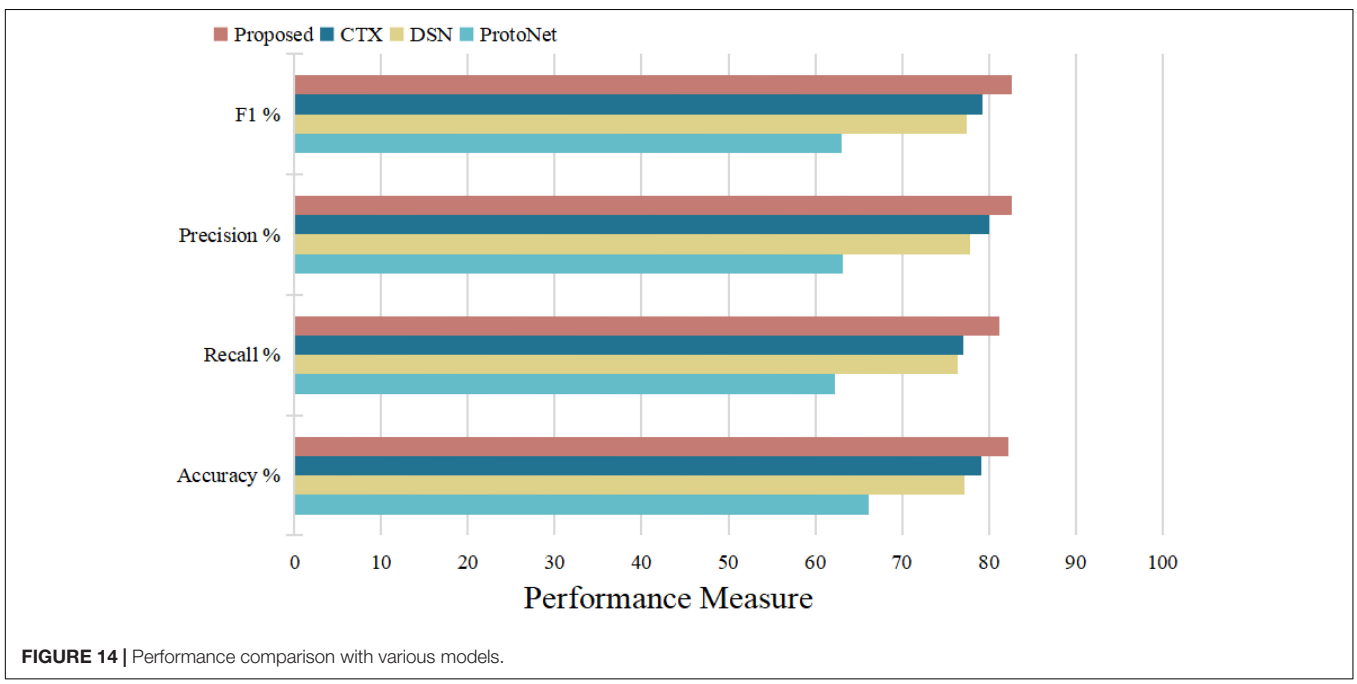


FIGURE 14 | Performance comparison with various models.

proposed a modular group attention block that captures feature dependencies in medical images in both channel and spatial dimensions for resulting in improved classification accuracy. Abdar et al. (2021) proposed a novel, simple and effective fusion model with uncertainty-aware module for medical image classification called Binary Residual Feature fusion (BARF).

Table 4 shows that the model has some advantages in classification. The bold text in the table represents the best performance. But there are still differences in accuracy, and the model has limitations. In future work, solutions can be proposed for this situation, such as designing a network structure more suitable for small samples to maximize the neural network learning ability. In addition, many of the algorithms proposed in the top methods have excellent performance. How to combine the advantages of these algorithms and integrating them into models is the focus of future work. In clinical care, it helps experts understand patients' current situation faster and more accurately, saving experts' time and achieving a leap in the quality of automatic medical classification.

CONCLUSION

This paper proposes an automated MDD sMRI data identification framework and performs a performance validation on the private SHH dataset with satisfactory results. The framework comprises a feature extractor and a feature map reconstruction network. 3D-ResNet acts as a feature extractor to ensure that MDD sMRI data with depth features can be learned. Then, the feature map reconstruction network solving the reconstruction problem in a closed-form produces a class of simple and powerful characters, which contains fine spatial details without overfitting the position or pose. Furthermore, we use an auxiliary loss that encourages support features from different classes to span the potential space to more clearly distinguish between classes. Additionally, a benchmarking clinical MDD sMRI images dataset with 68 subjects (SHH) is collected to train and test the model, and we evaluate the proposed 3D FRN-ResNet on the SHH dataset. Experimental results show that the proposed model exhibits promising performance and outperforms the typical other methods, achieving the accuracy, recall, precision, and $F1$ values of 0.86776, 0.84237, 0.85333, and 0.84781, respectively. Compared with some benchmark methods, the method proposed in this paper can effectively improve the identification accuracy and recall of MDD and healthy controls, and then assist doctors to complete the diagnosis in medicine, which has great value in practical clinical computer-aided diagnosis applications.

Even though the 3D FRN-ResNet framework has demonstrated its potential within the automated identification for MDD sMRI data, some limitations still need to be improved. For example, the model performance cannot be well exploited due to sample size limitations. Thus, we can use better data enhancement methods, which provide a good starting point for further research.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study. Requests to access the datasets should be directed to XL, dmia_lab@zcmu.edu.cn.

AUTHOR CONTRIBUTIONS

JH, YH, and XL conceived and designed the study. JH, JY, JW, XX, and XL contributed to the literature search. JH, YH, YW, and YL contributed to data analysis and data curation. JH, JY, JZ, RL, and XL contributed to data visualization. JH and JK contributed to software implementation. JH, YH, JY, JW, XX, and XL contributed to the tables and figures. JH, YW, YL, RL, JK, and XL contributed to the writing of the report. JH, JY, YW, RL, and XL contributed to review and editing. All authors have read and approved the publication of this work.

FUNDING

This work was funded in part by the National Natural Science Foundation of China (Grant No. 62072413).

ACKNOWLEDGMENTS

We would like to thank YH from the Seventh Hospital of Hangzhou to collect the private MDD sMRI data dataset for our study.

REFERENCES

Abdar, M., Fahami, M. A., Chakrabarti, S., Khosravi, A., Plawiak, P., Acharya, U., et al. (2021). BARF: a new direct and cross-based binary residual feature fusion with uncertainty-aware module for medical image classification. *Inform. Sci.* 577, 353–378. doi: 10.1016/j.ins.2021.07.024

An, F., Li, X., and Ma, X. (2021). Medical image classification algorithm based on visual attention Mechanism-MCNN. *Oxid. Med. Cell Longev.* 2021:6280690. doi: 10.1155/2021/6280690

Arnone, D., McKie, S., Elliott, R., Juhasz, G., Thomas, E. J., Downey, D., et al. (2013). State-dependent changes in hippocampal grey matter in depression. *Mol. Psychiatry* 18, 1265–1272. doi: 10.1038/mp.2012.150

Ashburner, J., Barnes, G., Chen, C., Daunizeau, J., Flandin, G., Friston, K. J., et al. (2021). SPM12 Manual. London: UCL Queen Square Institute of Neurology.

Bachmann, M., Lass, J., and Hinrikus, H. (2017). Single channel EEG analysis for detection of depression. *Biomed. Signal Proc. Control.* 31, 391–397.

Baek, J., and Chung, K. (2020). Context deep neural network model for predicting depression risk using multiple regression. *IEEE Access.* 8, 18171–18181. doi: 10.1109/access.2020.2968393

Ben, R., Ejbali, R., and Zaied, M. (2020). Classification of medical images based on deep stacked patched auto-encoders. *Multimed Tools Appl.* 79, 25237–25257. doi: 10.1007/s11042-020-09056-5

Carl, D., Ankush, G., and Andrew, Z. (2020). CrossTransformers: spatially-aware few-shot transfer. *arXiv [preprint]* doi: 10.48550/arXiv.2007.11498

Chen, X., Xie, L., Wu, J., and Tian, Q. (2021). Cyclic CNN: image classification with multiscale and multilocation contexts. *IEEE Internet Things J.* 8, 7466–7475. doi: 10.1109/jiot.2020.3038644

Chen, Y., Wang, X., Liu, Z., Xu, H., and Darrell, T. (2020). A new meta-baseline for few-shot learning. *arXiv [preprint]* doi: 10.48550/arXiv.2003.04390

Cheng, J., Tian, S., Yu, L., Gao, C., Kang, X., Ma, X., et al. (2022). ResGANet: residual group attention network for medical image classification and segmentation. *Med. Image Anal.* 76:102313. doi: 10.1016/j.media.2021.102313

Christian, S., Piotr, K., Richard, N., and Mehrtash, H. (2020). "Adaptive subspaces for few-shot learning," in *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. (Piscataway, NJ: IEEE). doi: 10.1016/j.isatra.2021.10.025

Dai, Y., Gao, Y., and Liu, F. (2021). TransMed: transformers advance multi-modal medical image classification. *Diagnostics (Basel)* 11:1384. doi: 10.3390/diagnostics11081384

Ding, J., and Fu, L. (2018). A hybrid feature selection algorithm based on information gain and sequential forward floating search. *J. Intell. Comp. Vol.* 9:93. doi: 10.6025/jic/2018/9/3/93-101

Eslami, T., Mirjalili, V., Fong, A., Laird, A. R., and Saeed, F. (2019). ASD-DiagNet: a hybrid learning approach for detection of autism spectrum disorder using fMRI data. *Front. Neuroinform.* 13:70. doi: 10.3389/fninf.2019.00070

Farooq, A., Anwar, S. M., Awais, M., and Rehman, S. (2017). "A deep CNN based multi-class classification of Alzheimer's disease using MRI," in *Proceedings of the 2017 IEEE International Conference on Imaging Systems and Techniques (IST)*. (Piscataway, NJ: IEEE).

GBD 2017 Disease and Injury Incidence and Prevalence Collaborators (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 392, 1789–1858. doi: 10.1016/S0140-6736(18)32279-7

Gold, P. W., Machado-Vieira, R., and Pavlatou, M. G. (2015). Clinical and biochemical manifestations of depression: relation to the neurobiology of stress. *Neural Plast.* 2015:581976.

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46, 389–422.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *arXiv [preprint]* doi: 10.48550/arXiv.1512.03385

Hidalgo-Muñoz, A. R., Ramírez, J., Górriz, J. M., and Padilla, P. (2014). Regions of interest computed by SVM wrapped method for Alzheimer's disease examination from segmented MRI. *Front. Aging Neurosci.* 6:20. doi: 10.3389/fnagi.2014.00020

Jiao, Z., Gao, X., Wang, Y., and Li, J. (2017). A parasitic metric learning net for breast mass classification based on mammography. *Pattern Recogn.* 75, 292–301. doi: 10.1016/j.patcog.2017.07.008

Karthikeyan, D., Varde, A. S., and Wang, W. (2020). Transfer learning for decision support in Covid-19 detection from a few images in big data. *IEEE Int. Conf. Big Data* 2020, 4873–4881.

Lazli, L., Boukadoum, M., and Ait Mohamed, O. (2019). Computer-Aided diagnosis system of alzheimer's disease based on multimodal fusion: tissue quantification based on the hybrid fuzzy-genetic-possibilistic model and discriminative classification based on the SVDD model. *Brain Sci.* 9:289. doi: 10.3390/brainsci9100289

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Comp.* 1, 541–551. doi: 10.1162/neco.1989.1.4.541

Li, F., and Liu, M. (2018). Alzheimer's disease neuroimaging initiative. Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks. *Comp. Med. Imag. Graph.* 70, 101–110. doi: 10.1016/j.compmedimag.2018.09.009

Liang, G., and Wang, H. (2022). I-CNet: leveraging involution and convolution for image classification. *IEEE Access.* 10, 2077–2082.

Mwangi, B., Ebmeier, K. P., Matthews, K., and Steele, J. D. (2012). Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder. *Brain* 135, 1508–1521. doi: 10.1093/brain/aws084

Pathak, K., Pavthawala, M., Patel, N., Malek, D., Shah, V., and Vaidya, B. (2019). "Classification of brain tumor using convolutional neural network," in *Proceedings of the 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*. (Coimbatore).

Payan, A., and Montana, G. (2015). Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. *arXiv [preprint]* doi: 10.48550/arXiv.1502.02506

Sakai, K., and Yamada, K. (2019). Machine learning studies on major brain diseases: 5-year trends of 2014-2018. *Japanese J. Radiol.* 37, 34–72.

Seal, A., Bajpai, R., Agnihotri, J., Yazidi, A., Herrera-Viedma, E., and Krejcar, O. (2021). DeprNet: a deep convolution neural network framework for detecting depression using EEG. *IEEE Trans. Instrumentation Measurement* 70, 1–13.

Segall, J. M., Turner, J. A., van Erp, T. G., White, T., Bockholt, H. J., Gollub, R. L., et al. (2009). Voxel-based morphometric multisite collaborative study on schizophrenia. *Schizophrenia Bull.* 35, 82–95. doi: 10.1093/schbul/sbn150

Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *Comp. Sci.* 18, 178–182. doi: 10.3390/s21082852

Talaat, A., Yousri, D., Ewees, A., Al-Qaness, M. A. A., Damasevicius, R., and Elaziz, M. A. (2020). COVID-19 image classification using deep features and fractional-order marine predators' algorithm. *Sci. Rep.* 10:15364. doi: 10.1038/s41598-020-71294-2

Ulloa, A., Plis, S., and Calhoun, V. (2015). "Synthetic structural magnetic resonance image generator improves deep learning prediction of schizophrenia," in *Proceedings of the IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*. (Piscataway, NJ: IEEE).

Wang, L., Sun, Y., and Wang, Z. (2021). CCS-GAN: a semi-supervised generative adversarial network for image classification. *Vis. Comp.* 4.

Wertheimer, D., Tang, L., and Hariharan, B. (2021). "Few-Shot classification with feature map reconstruction networks," in *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. (Piscataway, NJ: IEEE).

Yang, P., Zhao, C., Yang, Q., Wei, Z., Xiao, X., Shen, L., et al. (2022). Diagnosis of obsessive-compulsive disorder via spatial similarity-aware learning and fused deep polynomial network. *Med. Image Anal.* 75:102244. doi: 10.1016/j.media.2021.102244

Zadeh Shirazi, A., Fornaciari, E., McDonnell, M. D., Yaghoobi, M., Cevallos, Y., Tello-Oquendo, L., et al. (2020). The application of deep convolutional neural networks to brain cancer images: a survey. *J. Personal. Med.* 10:224. doi: 10.3390/jpm10040224

Zhang, Y., Sun, Y., Wang, N., Gao, Z., Chen, F., Wang, C., et al. (2021). MSEC: multi-scale erasure and confusion for fine-grained image classification. *Neurocomputing* 449, 1–14. doi: 10.1016/j.neucom.2021.03.114

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 16 May 2022
doi: 10.3389/fnagi.2022.911220

Multi-Modal Neuroimaging Neural Network-Based Feature Detection for Diagnosis of Alzheimer's Disease

Xianglian Meng[1], Junlong Liu[1], Xiang Fan[1], Chenyuan Bian[2], Qingpeng Wei[1], Ziwei Wang[1], Wenjie Liu[1] and Zhuqing Jiao[3]**

[1] School of Computer Information and Engineering, Changzhou Institute of Technology, Changzhou, China, [2] Shandong Provincial Key Laboratory of Digital Medicine and Computer-Assisted Surgery, Affiliated Hospital of Qingdao University, Qingdao, China, [3] School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, China

Alzheimer's disease (AD) is a neurodegenerative brain disease, and it is challenging to mine features that distinguish AD and healthy control (HC) from multiple datasets. Brain network modeling technology in AD using single-modal images often lacks supplementary information regarding multi-source resolution and has poor spatiotemporal sensitivity. In this study, we proposed a novel multi-modal LassoNet framework with a neural network for AD-related feature detection and classification. Specifically, data including two modalities of resting-state functional magnetic resonance imaging (rs-fMRI) and diffusion tensor imaging (DTI) were adopted for predicting pathological brain areas related to AD. The results of 10 repeated experiments and validation experiments in three groups prove that our proposed framework outperforms well in classification performance, generalization, and reproducibility. Also, we found discriminative brain regions, such as Hippocampus, Frontal_Inf_Orb_L, Parietal_Sup_L, Putamen_L, Fusiform_R, etc. These discoveries provide a novel method for AD research, and the experimental study demonstrates that the framework will further improve our understanding of the mechanisms underlying the development of AD.

Keywords: multi-modal, LassoNet, resting state functional magnetic resonance imaging, diffusion tensor imaging, feature detection

OPEN ACCESS

Edited by:
Yu-Dong Zhang, University of Leicester, United Kingdom

Reviewed by:
Yang Zhou, Harbin Medical University Cancer Hospital, China
Ying Wang, Guangdong Provincial People's Hospital, China

***Correspondence:**
Wenjie Liu
liuwj@czust.edu.cn
Zhuqing Jiao
jzq@cczu.edu.cn

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

Received: *02 April 2022*
Accepted: *19 April 2022*
Published: *16 May 2022*

Citation:
Meng X, Liu J, Fan X, Bian C, Wei Q, Wang Z, Liu W and Jiao Z (2022) Multi-Modal Neuroimaging Neural Network-Based Feature Detection for Diagnosis of Alzheimer's Disease. Front. Aging Neurosci. 14:911220. doi: 10.3389/fnagi.2022.911220

INTRODUCTION

Alzheimer's disease (AD) is a neurodegenerative brain disease that leads to the damage and death of brain nerve cells in disease progression. It destroys people's memory, learning, language, cognition, life, and other abilities, and seriously affects the quality of life of patients and families (Zhang and Wang, 2015; Lam et al., 2021; Lim et al., 2021). AD risk is also greater later in life for people with cardiovascular disease, high blood pressure, and diabetes. The Alzheimer's Association published a "2021 Alzheimer's Disease Facts and Figures," reporting a significant increase in AD deaths worldwide due to the COVID-19 pandemic. According to the clinical symptoms of patients, Alzheimer's disease is divided into a normal state (normal control, NC), mild cognitive impairment (mild cognitive impairment, MCI) state, and diseased AD state. MCI manifests as a decline in memory and thinking abilities at a rate greater than the decline in perception caused by normal aging, but this decline does not interfere with normal social interaction and work. However, patients with MCI have a high probability of further deterioration to AD (Zhang et al., 2016; Wang et al., 2017). It is currently difficult to distinguish MCI from

memory decline due to normal aging, and MCI involves very subtle brain changes. Therefore, the early diagnosis of MCI/AD is extremely challenging (Davis et al., 2018; Wang et al., 2018; Zhang et al., 2018; Potashman et al., 2021).

Magnetic resonance imaging (MRI) has become a hot spot in the field of AD and MCI disease research due to its non-invasiveness, multi-sequence imaging, high resolution, and strong repeatability (Zhang Y.-D. et al., 2014; Zhang et al., 2015a). Resting-state functional MRI (rs-fMRI) and MRI diffusion tensor imaging (DTI) are imaging techniques that can study brain mechanisms from the perspective of human brain functional connectivity and structural connectivity. They provide imaging evidence for the pathological studies on AD and MCI. Many studies have found the network structure related to the resting state in the cerebral cortex, which covers the brain regions that show a decline in metabolic function in the early stages of AD, including the posterior cingulate cortex and the internal parietal region (Choo et al., 2010; Hu et al., 2014; Zhang et al., 2015b; Shim et al., 2017; Wang et al., 2021). Neuroimaging data from a single modality usually can only reflect part of the brain characteristics, but many current research studies show that the fusion of information from multiple imaging modalities can reflect the brain activity mechanism more comprehensively (Zhang Q et al., 2014; Zhang and Shi, 2020; Lei et al., 2021; Jiao et al., 2022).

Functional MRI quantifies the temporal correlation between brain regions by detecting the blood oxygen level dependence (BOLD) in the human brain (Zhang and Shi, 2020; Wang et al., 2017), while DTI can track the spatial correlation of white matter fiber tracts by exploiting the kinetic mechanism of water molecule diffusion. Combining the spatiotemporal high-resolution information reflected by fMRI and DTI can comprehensively describe biological brain characteristics from a spatiotemporal perspective and improve the accuracy of brain network modeling, which is of great scientific significance for studying the neurophysiological mechanisms of AD/MCI diseases (Dyrba et al., 2015; Aderghal et al., 2020; Xu et al., 2021). Wee et al. considered the information regarding the complementary features of multiple imaging techniques, integrated multi-modal information from DTI and rs-fMRI, and used multi-kernel support vector machines to build a classifier for the study of disease classification and early prediction of MCI (Dai et al., 2019). Schonberg et al. used fMRI to define the regions of interest for DTI, providing a more comprehensive and functionally relevant white matter mapping map for preoperative preparation of brain tumors (Schonberg et al., 2006). Qi et al. propose a framework that combines DTI and fMRI multimodal imaging data to accurately identify potential neurological markers responsible for working memory deficits (Qi et al., 2018). Li et al. integrated the image information of rs-fMRI and DTI into a Lasso modeling framework for the accurate diagnosis of brain network lesions in early AD, further demonstrating that fusion of multi-modal information can effectively identify brain network features (Li et al., 2020). The above-mentioned finding proves that compared with single-modal data, more valuable features can be obtained by using multi-modal data. The multi-modal fusion method may further improve the recognition accuracy of AD/MCI (Zhang et al., 2015b; Wang et al., 2016; Mak et al., 2017).

In multi-modal neuroimaging analysis, since the features extracted from the original images tend to have higher dimensionality, only a few clinical samples contain complete multi-modal data, which will produce the curse of dimensionality. Therefore, we propose a neural network framework with Lasso regression for multi-modal image feature extraction and classification. **Figure 1** illustrates the neural network framework of multi-modal neuroimaging for Alzheimer's disease.

MATERIALS AND METHODS

Data Processing

The images of 85 subjects (33 healthy control, 29 early mild cognitive impairment, and 23 AD) were downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI[1]), including rs-fMRI and DTI. All neuroimaging data were obtained using a SIEMENS 3T MRI scanner. For the rs-fMRI images, the echo time (TE), the repetition time (TR), the flip angle, the slice thickness, and the time points were set as 30.0 ms, 3.0 s, 90, 3.4 mm, and 197, respectively. For the DTI data, the gradient directions, the echo time, the repetition time, the flip angle, and the voxel size were set as 30, 95 ms, 12.4 s, 90, and $2 \times 2 \times 2$ ms^3. For the T1 images, the TE, TR, flip angle, the slice thickness, and the T1 time were set as 3.0 ms, 2.3 s, 9.0, 1.0 mm, and 900 ms, and the collection plane was SAGITTAL. The **Table 1** showed the significant differences among the three groups in terms of gender ($p < 0.001$), age ($p < 0.001$), MMSE ($p < 0.001$), and EDU ($p < 0.001$) by t-test.

Data Acquisition

The rs-fMRI images were processed using SPM12[2] (Han and Glenn, 2018) and DPARBI 6.1[3] (Yan et al., 2016) as follows: (1) The raw DICOM files were converted to NIFITI format. (2) The first 10 time series nodes of each individual subject were removed manually to avoid the magnetic field inhomogeneity problem caused by the startup of the scanner and the influence of the discomfort of the subject's initial state on the results. (3) The interslice scan times were corrected to the same time point. (4) Images with head movement beyond 2.5 mm translation or 2.5-degree rotation were removed to correct head movement during scanning. (5) The head motion, white matter signal, and cerebrospinal fluid signal were set as the main noise covariates to reduce the influence of noisy covariate signals on scan results and reduce biological artifacts. (6) Different morphological brains were standardized to the same standard template and were registered to T1 images. (7) The $4 \times 4 \times 4$ mm^3 Gaussian kernel was applied for spatial smoothing to reduce spatial noise. (8) The linear trend was removed, and 0.01–0.1 Hz

[1]www.adni-info.org

[2]https://www.fil.ion.ucl.ac.uk/spm/software/spm12/

[3]http://rfmri.org/DPABI

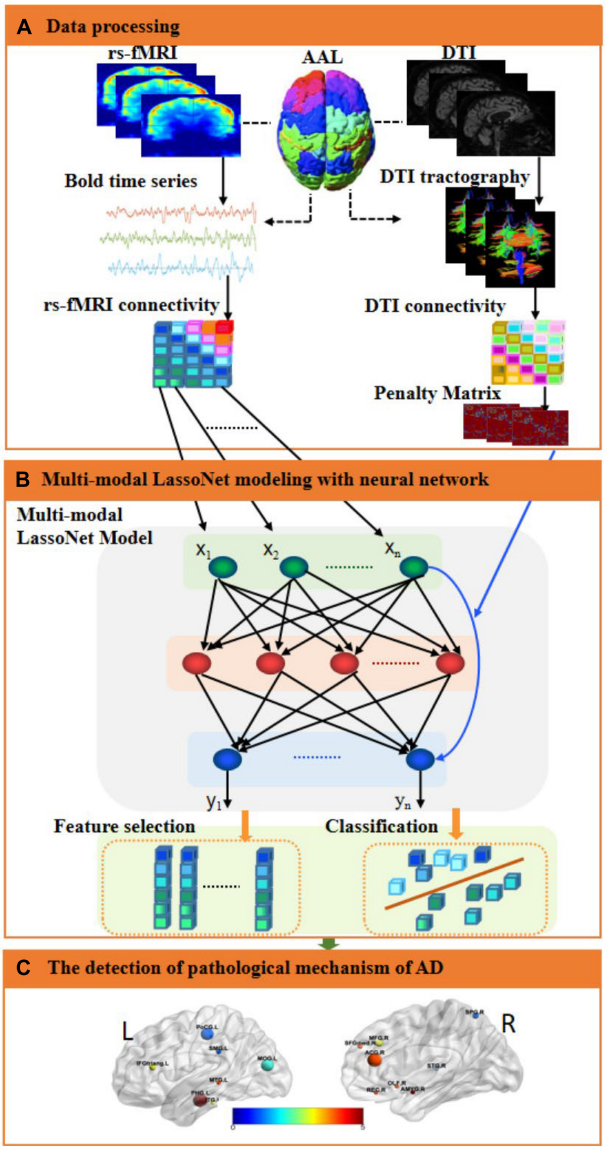


FIGURE 1 | An illustration of the proposed multi-modal framework for AD. **(A)** Data processing. The fMRI and DTI images were preprocessed, and then the regions of interest were extracted as fMRI and DTI features through the AAL template, and the corresponding brain networks of fMRI and DTI were obtained, respectively. Then, computed the inverse proportional function of the structural brain network as a penalty matrix. **(B)** Multi-modal LassoNet Modeling with a neural network. We constructed a multi-modal network framework for feature selection and classification based on the LassoNet model. It consisted of residual connection and an arbitrary feed-forward neural network. The input to the network was the fMRI feature information. The penalty matrix was introduced to the residual connection to sparse features. **(C)** The detection of the pathological mechanism of AD. We visualized brain regions for selected features to analyze the affected discriminative brain regions.

filtering was applied to reduce the interference due to low-frequency and high-frequency noise. The automated anatomical labeling (AAL; Tzourio-Mazoyer et al., 2002) atlas was applied to segment the brain into 90 regions, and the time series of BOLD signals were extracted.

TABLE 1 | Participant characteristics.

Subjects	HC	EMCI	AD	*P*
Number	33	29	23	
Gender (M/F)	12/21	14/15	14/9	<0.001
Age (Mean ± sd)	73.88 ± 7.15	74.52 ± 7.30	74.34 ± 8.14	<0.001
MMSE (Mean ± sd)	29.15 ± 1.13	28.52 ± 1.45	21.78 ± 1.89	<0.001
EDU (Mean ± sd)	16.55 ± 2.34	16.31 ± 2.56	14.96 ± 1.90	<0.001

HC, healthy control; EMCI, early mild cognitive impairment; AD, Alzheimer's disease; MMSE, Mini-mental status examination; M/F, male/female; Edu, education; sd, standard deviation.

The DTI data were processed using FSL[4] (Woolrich et al., 2009), PANDA[5] (Abbasi et al., 2021), and MRIcron (NITRC: MRIcron: Tool/Resource Info) software in Ubuntu18.04 as follows: (1) The raw DICOM files were converted to NIFITI format (*.nii.gz). (2) The brain templates were estimated based on non-diffusion-weighted b0 images using the bet command. (3) The non-brain space was removed using the fslroi command and eddy current correction. (4) The diffusion tensor metric was calculated using the dtifit command. (5) Deterministic white matter tract in the brain was tracked using the dti_recon and dti_tracker commands. (6) A part of the skull tissue in the T1 images was removed using the bet command. (7) The fractional anisotropy (FA) value of each subject was registered to its corresponding T1 image using the flirt command of FSL. When DTI images were registered with other images, DTI data causing significant deformities were removed. It should be noted that the DTI images and rs-fMRI images were registered with the same T1 imaging.

Multi-Modal LassoNet Framework Construction

The rs-fMRI functional brain networks can measure temporal correlations between anatomically segmented brain regions; DTI-based structural brain networks can characterize and track spatial white matter tracts in the brain. Herein, it is considered to unify the multi-modal image information of rs-fMRI and DTI in a brain network modeling framework, combining the respective advantages of the two modalities, which can describe the dynamic mechanism of the brain network from the perspective of time and space, and realize the construction of the brain network model.

After preprocessing of fMRI images, we obtained 187 time series (BOLD signal) of 85 participants, and there were 90 ROIs in each image. Let us assume that we have n participants and i ROIs. We explored a multi-modal network framework for feature selection and classification based on the LassoNet (Yan and Bien, 2017; Chen et al., 2019). For n participant, we assumed that the fMRI time series of the i-th ROI was $x_i = \{x_{1i}, x_{2i}, ..x_{di}\} \in R^{n \times d}$, $(i = 185)$, where d was the number of time points. Our goal was to find the best function $f^*(x_i)$ for predicting y_i (the type

[4]https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/

[5]https://www.nitrc.org/frs/?group_id=582

of Alzheimer's diagnosis). As the problem of learning $f^*(x_i)$ is non-parametric, we assumed that there was no linear or quadratic restriction. The multi-modal network consisted of two parts: residual connection and arbitrary feed-forward neural network. The penalty was introduced to the residual connection to sparse features. We define G to be the class of residual feed-forward neural networks:

$$G = \left\{ f \equiv f_\theta, w : x \longmapsto \theta^T X + gw(X) \right\} \quad (1)$$

where $gW(X)$ denotes a feed-forward network with weights W, $W^{(1)} \in \mathbb{R}^{d \times K}$ represents the weights in the first hidden layer, and $\theta \epsilon R^d$ represents the weights in the residual layer.

Let L be the empirical loss[2] on the training set with fMRI time series, then L is defined as Equation 2.

$$L(\theta, W) = \frac{1}{n} \sum_{i=1}^{n} \ell(f_\theta, w(x_i), \mathcal{Y}_i) \quad (2)$$

where W is the weight of the first hidden part; θ is the weight of the residual part; n is the number of participants as training observations size, and ℓ is the loss function. The LassoNet model objective function is defined as Equation 3.

$$\underset{\theta,\ W}{minimize}\ L(\theta, W) + \lambda ||\theta||_1 \ subject\ to \left|\left| W_i^{(1)} \right|\right| \leq M |\theta_i|,$$

$$i = 1, \cdots d \quad (3)$$

where $W_i^{(1)}$ is the weight of feature i and d is the data dimension.

The coupling strength of human brain functional connectivity and structural connectivity is closely related to the brain excitation process, and stronger structural brain connectivity is likely to lead to the enhancement of corresponding functional connectivity. Here, we introduced a parameter named the punishment factor to improve the LassoNet model. The punishment matrix of each DTI image is defined as the inverse proportional function of structural brain networks (Equation 4).

$$D_{ji} = e^{-\frac{\rho_{ji}^2}{\sigma}} \quad (4)$$

where ρ_{ji} is the FA information between j-th brain region and i-th brain region in the DTI network, and σ is the mean of the standard deviation of all elements in the structural brain network of all participants. Equation 4 is used to penalize the estimated connection strength value between the j-th ROI and the i-th ROI.

Since each participant had a corresponding DTI structure network information D, we calculated the max feature λ_{max} of each D using Equation 5.

$$(\lambda_a E - D)\, x = 0 \quad (5)$$

where $\lambda_{max} = \max(\lambda_a)$, E is the unity matrix, and x is the eigenvector. The DTI feature matrix is defined as Equation 6.

$$DTI_{vector} = [\lambda_1, \lambda_2 \cdot \cdots \cdot \lambda_n],\ n \in [1,\ 85] \quad (6)$$

Then, we modify the LassoNet objective function to Equation 7.

$$\underset{\theta,\ W}{minimize}\ L(\theta, W) + \lambda \cdot DTI_{vector} ||\theta||_1$$

$$subject\ to \left|\left| W_j^{(1)} \right|\right| \leq M |\theta_j|,\ j = 1, \cdots d \quad (7)$$

So, the multi-modal LassoNet framework was constructed. We summarize the training algorithm of multi-modal LassoNet, as shown **Table 2**.

Feature Detection and Model Comparison

Using the resulting images, we obtained the initial dataset of 85 participants and 187 × 90 features in each participant. We extracted three groups from the dataset, namely, AD-HC, AD-EMCI, and EMCI-HC. For each group, the train set, validation set, and test set were selected randomly using the ratio $S_{train} : S_{valid} : S_{test} = 6 : 2 : 2$. Integrating with DTI structure network information, the S_{train} and S_{valid} were applied to filter the optimal λ and integrating with DTI structure network information. With the resulting λ, the S_{train} and S_{test} were used to detect features and get the sparse feature matrix that classified well in AD-HC, AD-EMCI, or EMCI-HC.

Since the multi-modal framework was optimized based on the LassoNet model, to determine the superiority of our proposed framework, we used the classic Lasso, Group Lasso, Sparse Group Lasso, and ElasticNet to compare the classification accuracy.

Given n data samples $\{(x_1, y_1), (x_2, y_2), \cdots (x_n, y_n)\}$, $x_i \in R^d$, x_i was a d dimensional vector, that is, each observed data were composed of the values of d variables, and each $y_i \in R$ was a real value. Let the mapping $f : R^d \to R$ that minimize the sum of squared errors, and the optimization objective is defined as Equation 8.

$$W^* = argmin_\beta \frac{1}{n} ||y - XW||_2^2 \quad (8)$$

The optimization objective of Lasso (Equation 8) was obtained by introducing the L1 regularization term in Equation 9.

$$W^* = argmin_\beta \frac{1}{n} ||y - XW||_2^2 + \lambda ||W||_1 \quad (9)$$

The Lasso was applied to the group and the Group Lasso was obtained as Equation 10.

$$\underset{W \in R^p}{min} \left(\left|\left| y - \sum_{l=1}^{L} X_l W_l \right|\right|_2^2 + \lambda \sum_{l=1}^{L} \sqrt{p_l} ||W_l||_2 \right) \quad (10)$$

The Sparse Group Lasso was obtained by integrating the original Lasso into the Group Lasso, as Equation 11.

$$\underset{W \in R^p}{min} \left(\left|\left| y - \sum_{l=1}^{L} X_l W_l \right|\right|_2^2 + \lambda_1 \sum_{l=1}^{L} ||W_l||_2 + \lambda_2 ||W||_1 \right) \quad (11)$$

TABLE 2 | Training algorithm of multi-modal LassoNet.

Algorithm: Multi-Modal LassoNet with neural network

1: **Input:** $X \in R^{n \times d}$ represents fMRI time series (BOLD signal), B represents Number of epochs, M represents hierarchy multiplier, ϵ represents path multiplier, α represents learning rate, D represents penalty matrix from DTI network.
2: **Initialize:** $L(\theta, W)$ represents the feed-forword network on the loss, λ represents the penalty, k represents the number of activate features, $DTI_V = [\lambda_1, \lambda_2, \lambda_3 \cdots\cdots \lambda_n$ represents the multimodal matrix calculated from the penalty matrix D, d represents the number of features, $\theta \in R^d$ represents the weights in the residual layer, K is the number of units in the first hidden layer, θ^* and W^* are the optimal parameters after iteration.
3: while k > 0 do
4: Update $\lambda \leftarrow (1+\epsilon)\lambda DTI_V$
5: for b∈ (1...B) do
6: Compute gradient of the w.r.t to (θ, W) with back-propagation
Update $\theta \leftarrow \theta - \alpha\nabla_\theta L$ *and* $W \leftarrow W - a\nabla_W L$
7: for $j \in \{1 \ldots d\}$ do
8: Sort the entries of $W_j^{(1)}$ into $\left|W_j^{(1)}\right| \geq \ldots \geq \left|W_{(j,K)}^{(1)}\right|$
9: Compute $w_n := \frac{M}{1+nM^2} \cdot S_{\lambda D}\left(|\theta_j| + M \cdot \sum_{i=1}^{n} \left|W_{(j,i)}^{(1)}\right|\right)$
10: Find n^*, the first $n \in \{0, \ldots, K\}$ such that $\left|W_{(j,n+1)}^{(1)}\right| \leq w_n \leq \left|W_{(j,n)}^{(1)}\right|$
11: Update $\theta_j^* \leftarrow \frac{1}{M} \cdot sign(\theta_j) \cdot w_{n^*}, W_j^{(1)^*} \leftarrow sign(W_j^{(1)}) \cdot \min\left(w_{n^*}, \left|W_j^{(1)}\right|\right)$
12: end for
end for
13: **return** $(\theta^*, W^{(1)^*})$
14: end while

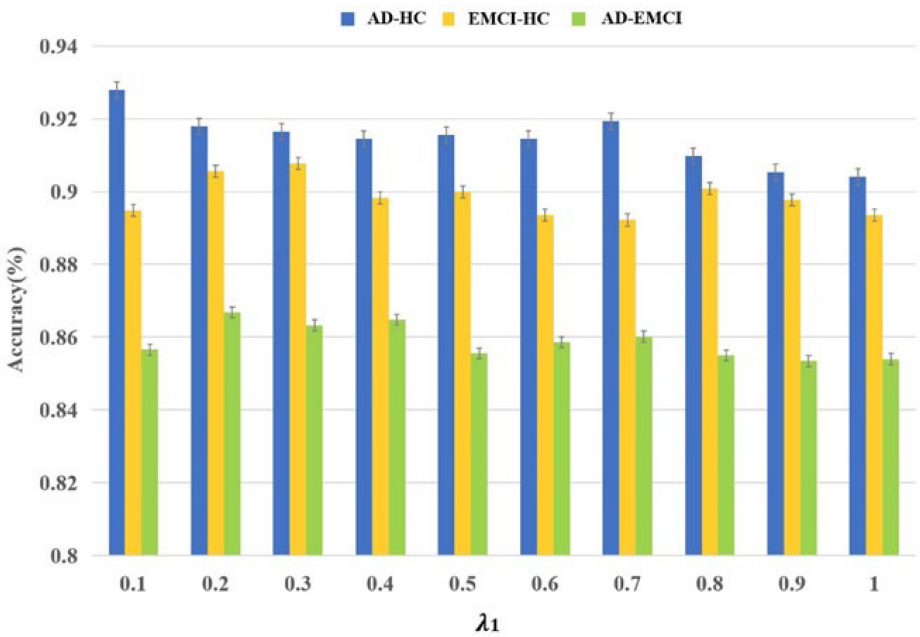


FIGURE 2 | The relationship between the accuracies and λ.

The definition of ElasticNet was obtained by combining L1 and L2 regularization and Lasso (Equation 12).

$$\min_{W \in R^p} \left(\left\| y - \sum_{l=1}^{L} X_l W_l \right\|_2^2 + \lambda_2 \sum_{l=1}^{L} ||W_l||_2 + \lambda_1 \sum_{l=1}^{L} ||W_l||_1 \right) \tag{12}$$

The same S_{train} and S_{valid} were applied to filter the optimal parameters. Using the same S_{train} and S_{tes}, the experiments were repeated 10 times in all five frameworks with the optimal parameters.

Evaluation Metrics

In this study, the samples were positive and negative, and the results classified had the following cases:

True Positive (TP): the positive sample was predicted as a positive sample.

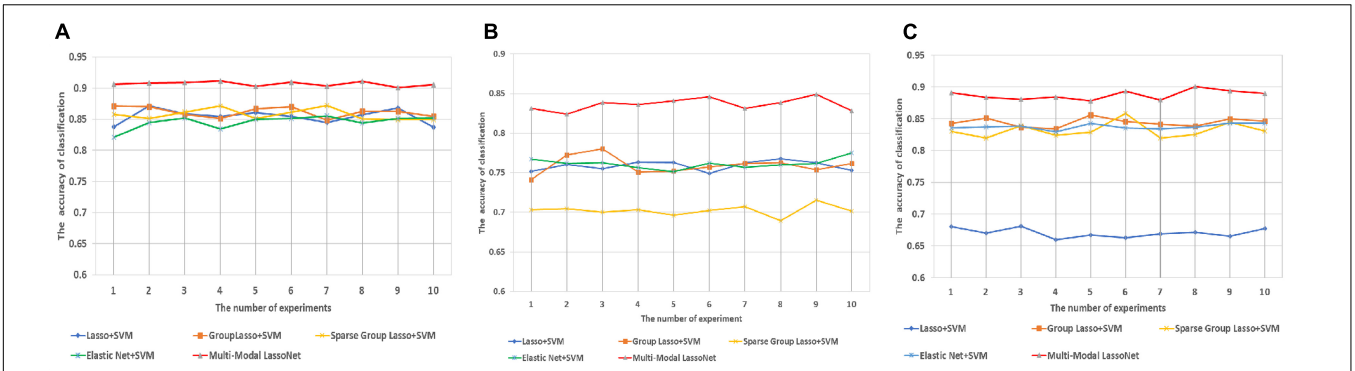


FIGURE 3 | The prediction accuracy was obtained through 10 experiments for five methods in three groups. **(A)** Prediction accuracy of AD-HC group. **(B)** Prediction accuracy of AD-EMCI group. **(C)** Prediction accuracy of EMCI-HC group.

TABLE 3 | The classification performance comparison of the five methods.

Group	Methods	ACC (%) ± SD	SEN (%) ± SD	SPE (%) ± SD	GMean (%) ± SD	F1 (%) ± SD
AD-HC	Lasso + SVM	85.45 ± 1.10	75.34 ± 2.87	91.89 ± 0.81	83.19 ± 1.49	80.05 ± 1.66
	GroupLasso + SVM	86.16 ± 0.78	77.13 ± 1.73	91.97 ± 1.00	84.24 ± 0.93	80.75 ± 2.75
	ElasticNet + SVM	84.56 ± 1.00	74.90 ± 2.14	90.72 ± 1.24	82.41 ± 1.13	79.04 ± 1.24
	Sparse Group Lasso + SVM	85.75 ± 0.83	75.99 ± 2.29	**92.20 ± 0.88**	83.69 ± 1.11	80.89 ± 1.18
	Multi-modal LassoNet	**90.68 ± 0.34**	**88.81 ± 0.68**	91.91 ± 0.55	**90.34 ± 0.36**	**88.25 ± 0.52**
AD-EMCI	Lasso + SVM	75.88 ± 0.58	93.06 ± 0.87	54.22 ± 0.95	71.03 ± 0.61	81.15 ± 0.56
	GroupLasso + SVM	75.92 ± 1.04	**93.12 ± 0.63**	54.33 ± 1.77	71.12 ± 1.14	81.16 ± 0.87
	ElasticNet + SVM	76.13 ± 0.61	92.91 ± 0.94	54.98 ± 1.96	71.45 ± 1.01	81.27 ± 0.60
	Sparse Group Lasso + SVM	70.23 ± 0.63	90.44 ± 1.05	44.69 ± 1.82	63.56 ± 1.04	77.05 ± 0.67
	Multi-modal LassoNet	**83.63 ± 0.74**	87.32 ± 1.22	**79.00 ± 1.33**	**83.05 ± 0.76**	**85.70 ± 0.84**
EMCI-HC	Lasso + SVM	67.04 ± 0.69	78.67 ± 1.98	57.61 ± 1.49	67.30 ± 0.68	68.12 ± 0.90
	GroupLasso + SVM	84.42 ± 0.65	**96.62 ± 0.65**	74.43 ± 0.98	84.80 ± 0.52	84.08 ± 0.57
	ElasticNet + SVM	83.76 ± 0.42	94.69 ± 0.69	74.72 ± 0.92	84.11 ± 0.44	84.07 ± 0.41
	Sparse Group Lasso + SVM	83.20 ± 1.15	95.83 ± 0.88	72.74 ± 1.43	83.49 ± 1.14	70.86 ± 1.16
	Multi-modal LassoNet	**88.77 ± 0.70**	90.87 ± 1.05	**87.06 ± 0.95**	**88.94 ± 0.69**	**87.92 ± 0.83**

True Negative (TN): the negative sample was predicted as a negative sample.
False Positive (FP): the negative sample was predicted as a positive sample.
False Negative (FN): the positive sample was predicted as a negative sample.

ACC (accuracy) is the number of correctly classified samples divided by the total number of samples (Equation 13).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

SEN (sensitivity) is the proportion of pairs of all positive samples (Equation 14).

$$SEN = \frac{TP}{TP + FN} \tag{14}$$

SPE (specificity) is the proportion of pairs of all negative samples (Equation 15).

$$SPE = \frac{TN}{TN + FP} \tag{15}$$

GMean is the geometric mean (Equation 16).

$$GMean = \sqrt{SEN + SPE} \tag{16}$$

F1 is a comprehensive evaluation indicator. Sometimes, accuracy and sensitivity needed to be considered together as Equation 17.

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{17}$$

The receiver operating characteristic (ROC) curve and the area under curve (AUC) value are also used to evaluate the performance of the classifier.

RESULTS

The Results of Parameter Optimization

Initially, $187 \times 90 = 16,830$ features were obtained and S_{train} and S_{valid} were applied to filter the optimal parameters. The λ was the interval of (0.1, 1), and the corresponding accuracy was calculated in each group. As shown in **Figure 2**, the best accuracy of the AD-HC group is 92.79% and λ is 0.1. The peak value

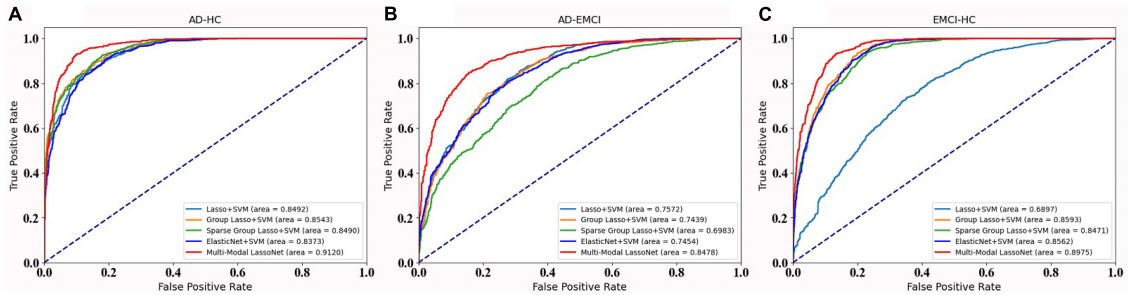


FIGURE 4 | The ROC curve of the five methods in three groups. **(A)** Prediction accuracy of AD-HC group. **(B)** Prediction accuracy of AD-EMCI group. **(C)** Prediction accuracy of EMCI-HC group.

TABLE 4 | Discriminative brain regions.

Group	ID	Regions	Abbreviation	ID	Regions	Abbreviation
AD-HC	61	Parietal_Inf_L	IPL.L	19	Supp_Motor_Area_L	SMA.L
	24	Frontal_Sup_Medial_R	SFGmed.R	59	Parietal_Sup_L	SPG.L
	37	Hippocampus_L	HIP.L	83	Temporal_Pole_Sup_L	TPOsup.L
	79	Heschl_L	HES.L	64	SupraMarginal_R	SMG.R
	7	Frontal_Mid_L	MFG.L	81	Temporal_Sup_L	STG.L
	73	Putamen_L	PUT.L	52	Occipital_Mid_R	MOG.R
	15	Frontal_Inf_Orb_L	ORBinf.L	32	Cingulum_Ant_R	ACG.R
	56	Fusiform_R	PoCG.L			
EMCI-HC	37	Hippocampus_L	HIP.L	14	Frontal_Inf_Tri_R	IFGtriang.R
	27	Rectus_L	REC.L	59	Parietal_Sup_L	SPG.L
	17	Rolandic_Oper_L	ROL.L	88	Temporal_Pole_Mid_R	TPOmid.R
	30	Insula_R	INS.R	44	Calcarine_R	CAL.R
	6	Frontal_Sup_Orb_R	ORBsup.R	49	Occipital_Sup_L	SOG.L
	8	Frontal_Mid_R	MFG.R	31	Cingulum_Ant_L	ACG.L
	38	Hippocampus_R	HIP.R	7	Frontal_Mid_L	MFG.L
	15	Frontal_Inf_Orb_L	ORBinf.L			
AD-EMCI	22	Olfactory_R	OLF.R	63	SupraMarginal_L	SMG.L
	32	Cingulum_Ant_R	ACG.R	57	Postcentral_L	PoCG.L
	89	Temporal_Inf_L	ITG.L	51	Occipital_Mid_L	MOG.L
	82	Temporal_Sup_R	STG.R	24	Frontal_Sup_Medial_R	SFGmed.R
	85	Temporal_Mid_L	MTG.L	39	ParaHippocampal_L	PHG.L
	42	Amygdala_R	AMYG.R	13	Frontal_Inf_Tri_L	IFGtriang.L
	28	Rectus_R	REC.R	60	Parietal_Sup_R	SPG.R
	8	Frontal_Mid_R	MFG.R			

of the EMCI-HC group is at the node of 0.3. The prediction accuracy reaches a peak with a λ value of 0.2. We can also observe that the accuracy of the AD-HC group is much higher than the other two groups. This may be caused by the large difference between AD and HC. An interesting finding is that the accuracy of the AD-EMCI group is the lowest and the gap in this group is also the lowest. This proves that the similarity between AD and EMCI is higher, and the similar features make the classification more stable.

Comparison With Other Methods

We applied the same S_{train} and S_{test} to assess the performance of the five models, and 10 independent experiments were conducted to evaluate the universality of these models. As shown in **Figure 3**, the Multi-modal LassoNet has good prediction accuracy, and in three groups, the accuracy of the Multi-modal LassoNet is the highest, far exceeding the other four models. The peaks of the Multi-modal LassoNet are above 90% in the AD-HC and EMCI-HC groups, and in the other four models, they are all below 90%. In the AD-EMCI group, the best accuracy is above 85%, and in the other four models, it is below 80%. Additionally, the gap of the Multi-modal LassoNet in 10 experiments is less than 2%. It can be seen from **Figure 3** that the Multi-modal LassoNet framework has satisfactory classification accuracy in different groups only by adjusting the λ. The curves of the Multi-modal LassoNet also prove that the proposed framework has good stability, and the introduction of DTI information improves the classification performance of the LassoNet model.

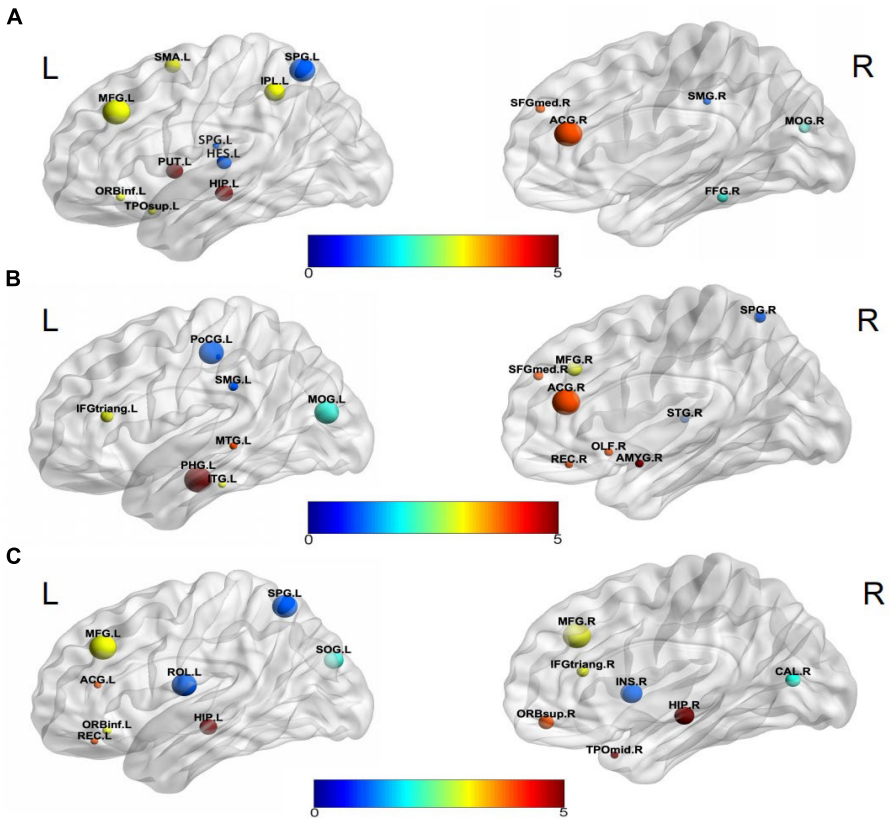


FIGURE 5 | Visualization of discriminative brain regions. **(A)** AD-HC, **(B)** AD-EMCI, and **(C)** EMCI-HC.

The classification information of the five methods is presented in **Table 3**. Multi-modal LassoNet classifiers reported very good performance. In the AD-HC group classification, ACC, SEN, SPE, GMean, and F1 were 90.68 ± 0.34, 88.81 ± 0.68, 91.91 ± 0.55, 90.34 ± 0.36, and 88.25 ± 0.52, respectively. In the AD-EMCI group classification, ACC, SEN, SPE, GMean, and F1 were 83.63 ± 0.74, 87.32 ± 1.22, 79.00 ± 1.33, 83.05 ± 0.76, and 85.70 ± 0.84, respectively. In the EMCI-HC group classification, ACC, SEN, SPE, GMean, and F1 were 88.77 ± 0.70, 90.87 ± 1.05, 87.06 ± 0.95, 88.94 ± 0.69, and 87.92 ± 0.83, respectively.

For further validation of our framework and results, we plot the ROC curves of five methods for the AD-HC, AD-EMCI, and EMCI-HC groups, as shown **Figure 4**. The AUC values of our proposed Multi-modal LassNet for AD-HC, AD-EMCI, and EMCI-HC groups were 0.9120, 0.8478, and 0.8975, respectively.

DISCUSSION

Modeling techniques based on a single neuroimaging modality lacked the spatial and temporal high-resolution information brought by different modalities in characterizing the brain network structure, and could not fully reflect the dynamic mechanism of brain network connections (Tulay et al., 2019; Zhuang et al., 2019; Lei et al., 2020). Therefore, we proposed a multi-modal LassoNet model that was a Lasso neural network modeling framework using multi-modal information fusion. This method fused two modalities of fMRI and DTI in a sparse Lasso neural network framework and introduced connection strength and subject structure to complete the construction of a multi-modal brain network. Our proposed method mainly addresses two issues, which include the selection of AD-related brain ROIs and the classification and diagnosis of AD. The experimental results showed that the multi-LassoNet modeling of multi-modal information could facilitate higher sensitivity of disease diagnosis and effectively improved the accuracy of model classification. The good classification performance also revealed that the detected features of the multi-modal model based on fMRI and DTI reflected that the brain atrophy caused by the disease process would lead to the decrease of white matter fiber connectivity (Gupta et al., 2020). It also proved that structural connectivity and functional brain network features between connections had coupling effects.

Compared with the current popular Lasso method, Group Lasso, Sparse Lasso, and elastic network method, it was proved that the proposed multi-modal Lasso-based neural network method was higher than other methods in classification

performance and had strong regularization parameter stability. It proved that fusion of multi-modal information more effectively identified brain network features. Moreover, the results indicated that the constraint effect of the DTI structural network and the introduction of the strength of the brain area connection had a certain degree of influence on the validity of the multi-modal brain network model. **Table 4** shows the top 15 important brain regions with different classification results.

Visualization selected discriminative brain regions using the BrainNet Viewer toolbox (Xia et al., 2013), as shown in **Figure 5**. By analyzing the brain regions classifying AD-HC,AD-EMCI, and EMCI-HC, we found that the brain regions belonging to Hippocampus, Frontal_Inf_Orb_L, and Parietal_Sup_L were among the top 15 brain regions. Previous studies had found that the hippocampus of the brain was responsible for human memory and spatial activities and was closely related to AD pathology (Douaud et al., 2011; Fares et al., 2019). In addition, some studies had also shown that functional atrophy in the parahippocampal gyrus is an early marker of AD/MCI disease (Wang et al., 2016), and the parahippocampal gyrus shows a more distinct ability than the hippocampus in the early stage of the disease (Zhu et al., 2019). Frontal_Inf_Orb_L corresponded to the region of interest recommended by physicians for the clinical diagnosis of AD (Jiang et al., 2015). Parietal_Sup_L may be associated with the underlying mechanism of its clinical effect, and it may play a role in the potential compensatory mechanism of mobilizing more regions to complete the function after a functional decline (He et al., 2021). The Hippocampus_L and Hippocampus_R found in the AD-HC and EMCI-HC groups were reported as the pathogenic regions of AD. Chik et al. (Yuan et al., 2022) found that the neurosteroids in the hippocampus were changed during the progression of Lv et al. (2022) found that compared to the healthy mouse, the mice having TYRO protein kinase-binding protein had insufficient learning and memory abilities, and the amyloid β in the hippocampus was increased, which worsened with aging. Liu et al. (2022) proved that memory could be improved by enhancing the functional activity in the hippocampus and the medial prefrontal cortex. Moreover, the hippocampus region was not found in AD-EMCI. This gives a message that the difference in the hippocampus between AD and EMCI is not obvious, and their main difference is found in other brain regions, such as the Amygdala_R, which is not found in the other two groups. Hong et al. (2022) reported tau deposition in the parahippocampus and amygdala by studying positron emission tomography (PET) images in patients with AD. The amygdala atrophy was found in mild AD subjects and could be used to predict the Mini-Mental State Examination scores and hippocampal atrophy (Poulin et al., 2011).

In addition, the Putamen_L was reported to be the earliest brain region to show increased Aβ deposition and is a marker of cognitive decline (Zammit et al., 2020; Cogswell et al., 2021). The Fusiform_R was confirmed to be a characteristic region of AD (Guo et al., 2017; Sprung et al., 2021). Brain network analysis results generally had a high sensitivity to segmentation template selection. Different segmentation templates produced different brain network topology structures, which might potentially affect the reproducibility of model classification performance. The segmentation template used in this paper was the AAL structure of 90 brain regions. However, in the future, the robustness value of the proposed method would be further verified from the perspectives of multiple segmentation scales.

In this study, a deterministic fiber tracking technique derived from DTI images was used to construct a structural brain network in a multi-modal modeling framework. But this tracking method only considered the trajectories where white matter fibers cross or diverge (Lei et al., 2021). Therefore, there may be biases in determining the most reasonable fiber configuration, affecting the accuracy of structural network construction. Future research work will consider adopting a more efficient probabilistic fiber tract-tracing strategy to obtain the probability value of brain area connection to complete accurate multi-modal brain network construction.

In this study, we proposed a novel multi-modal LassoNet framework for the discriminant analysis of features. This research is an attempt to apply fMRI and DTI multi-modal information and sparse representation technology to the research of neural network framework, and provides a new idea for designing a brain network modeling framework that integrates more modal information in the future. The features of multi-modal data can be fused to obtain more comprehensive pathological information. Compared to the conventional methods, the proposed method seeks to identify AD-related brain ROIs and in the classification and diagnosis of AD. The high-performance classification implied that the proposed multi-modal LassoNet framework was beneficial for the early diagnosis and prediction of AD disease.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The data is available at http://adni.loni.usc.edu/.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants. Data collection and sharing for this project were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI).

AUTHOR CONTRIBUTIONS

XM, CB, WL, and ZJ led and supervised the research. XM, WL, and ZJ designed the research and wrote the article. XF, QW, and ZW performed data preprocessing. XM, JL, WL, and ZJ performed feature extraction and selection by the multi-modal LassoNnet framework. XF, QW, and CB did a discriminative analysis of brain regions. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the National Natural Science Foundation of China (61901063 and 51877013), MOE (Ministry of Education in China) Project of Humanities and Social Sciences (19YJCZH120), the Science and Technology Plan Project of Changzhou (CE20205042 and CJ20210155), and by Jiangsu Provincial Key Research and Development Program (BE2021636). This work was also sponsored by Qing Lan Project of Jiangsu Province.

ACKNOWLEDGMENTS

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

REFERENCES

Abbasi, W. A., Abbas, S. A., and Andleeb, S. (2021). PANDA: Predicting the change in proteins binding affinity upon mutations by finding a signal in primary structures. *J. Bioinform. Comput. Biol.* 19:2150015. doi: 10.1142/S0219720021500153

Aderghal, K., Afdel, K., Benois-Pineau, J., and Catheline, G. (2020). Improving Alzheimer's stage categorization with Convolutional Neural Network using transfer learning and different magnetic resonance imaging modalities. *Heliyon* 6:e05652. doi: 10.1016/j.heliyon.2020.e05652

Chen, Z., Zeng, W., Yang, Z., Yu, L., Fu, C. W., and Qu, H. (2019). LassoNet: deep Lasso-Selection of 3D Point Clouds. *IEEE Trans. Vis. Comput. Graph.* 2019:2934332. doi: 10.1109/tvcg.2019.2934332

Choo, I. H., Lee, D. Y., Oh, J. S., Lee, J. S., Lee, D. S., Song, I. C., et al. (2010). Posterior cingulate cortex atrophy and regional cingulum disruption in mild cognitive impairment and Alzheimer's disease. *Neurobiol. Aging* 31, 772–779. doi: 10.1016/j.neurobiolaging.2008.06.015

Cogswell, P. M., Wiste, H. J., Senjem, M. L., Gunter, J. L., Weigand, S. D., Schwarz, C. G., et al. (2021). Associations of quantitative susceptibility mapping with Alzheimer's disease clinical and imaging markers. *Neuroimage* 224:117433. doi: 10.1016/j.neuroimage.2020.117433

Dai, Z., Lin, Q., Li, T., Wang, X., Yuan, H., Yu, X., et al. (2019). Disrupted structural and functional brain networks in Alzheimer's disease. *Neurobiol. Aging* 75, 71–82. doi: 10.1016/j.neurobiolaging.2018.11.005

Davis, M. T. O. C., Johnson, S., Cline, S., Merikle, E., Martenyi, F., et al. (2018). Estimating Alzheimer's Disease Progression Rates from Normal Cognition Through Mild Cognitive Impairment and Stages of Dementia. *Curr. Alzheimer. Res.* 15, 777–788. doi: 10.2174/1567205015666180119092427

Douaud, G., Jbabdi, S., Behrens, T. E., Menke, R. A., Gass, A., Monsch, A. U., et al. (2011). DTI measures in crossing-fibre areas: increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *Neuroimage* 55, 880–890. doi: 10.1016/j.neuroimage.2010.12.008

Dyrba, M., Grothe, M., Kirste, T., and Teipel, S. J. (2015). Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM. *Hum. Brain Mapp.* 36, 2118–2131. doi: 10.1002/hbm.22759

Fares, J., Bou Diab, Z., Nabha, S., and Fares, Y. (2019). Neurogenesis in the adult hippocampus: history, regulation, and prospective roles. *Int. J. Neurosci.* 129, 598–611. doi: 10.1080/00207454.2018.1545771

Guo, Z., Liu, X., Li, J., Wei, F., Hou, H., Chen, X., et al. (2017). Fractional amplitude of low-frequency fluctuations is disrupted in Alzheimer's disease with depression. *Clin. Neurophysiol.* 128, 1344–1349. doi: 10.1016/j.clinph.2017.05.003

Gupta, Y., Kim, J. I., Kim, B. C., and Kwon, G. R. (2020). Classification and Graphical Analysis of Alzheimer's Disease and Its Prodromal Stage Using Multimodal Features From Structural, Diffusion, and Functional Neuroimaging Data and the APOE Genotype. *Front. Aging Neurosci.* 12:238. doi: 10.3389/fnagi.2020.00238

Han, H., and Glenn, A. L. (2018). Evaluating methods of correcting for multiple comparisons implemented in SPM12 in social neuroscience fMRI studies: an example from moral psychology. *Soc. Neurosci.* 13, 257–267. doi: 10.1080/17470919.2017.1324521

He, F., Li, Y., Li, C., Fan, L., Liu, T., and Wang, J. (2021). Repeated anodal high-definition transcranial direct current stimulation over the left dorsolateral prefrontal cortex in mild cognitive impairment patients increased regional homogeneity in multiple brain regions. *PLoS One* 16:e0256100. doi: 10.1371/journal.pone.0256100

Hong, J., Kang, S. K., Alberts, I., Lu, J., Sznitman, R., Lee, J. S., et al. (2022). Image-level trajectory inference of tau pathology using variational autoencoder for Flortaucipir PET. *Eur. J. Nucl. Med. Mole. Imag.* 2022:5662. doi: 10.1007/s00259-021-05662-z

Hu, Z., Wu, L., Jia, J., and Han, Y. (2014). Advances in longitudinal studies of amnestic mild cognitive impairment and Alzheimer's disease based on multi-modal MRI techniques. *Neurosci. Bull* 30, 198–206. doi: 10.1007/s12264-013-1407-y

Jiang, J., Duan, H., Huang, Z., and Yu, Z. (2015). Study of amyloid-β peptide functional brain networks in AD, MCI and HC. *Biomed. Mater. Eng.* 26(Suppl. 1), S2197–S2205. doi: 10.3233/bme-151525

Jiao, Z., Chen, S., Shi, H., and Xu, J. (2022). Multi-Modal Feature Selection with Feature Correlation and Feature Structure Fusion for MCI and AD Classification. *Brain Sci.* 12:80. doi: 10.3390/brainsci12010080

Lam, J., Lee, J., Liu, C. Y., Lozano, A. M., and Lee, D. J. (2021). Deep Brain Stimulation for Alzheimer's Disease: tackling Circuit Dysfunction. *Neuromodulation* 24, 171–186. doi: 10.1111/ner.13305

Lei, B., Cheng, N., Frangi, A. F., Wei, Y., Yu, B., Liang, L., et al. (2021). Auto-weighted centralised multi-task learning via integrating functional and structural connectivity for subjective cognitive decline diagnosis. *Med. Image Anal.* 74:102248. doi: 10.1016/j.media.2021.102248

Lei, D., Pinaya, W. H. L., Young, J., van Amelsvoort, T., Marcelis, M., Donohoe, G., et al. (2020). Integrating machining learning and multimodal neuroimaging to detect schizophrenia at the level of the individual. *Hum. Brain Mapp.* 41, 1119–1135. doi: 10.1002/hbm.24863

Li, J., Bian, C., Chen, D., Meng, X., Luo, H., Liang, H., et al. (2020). Persistent Feature Analysis of Multimodal Brain Networks Using Generalized Fused Lasso for EMCI Identification. *Med. Image Comp. Comp. Assist. Interv.* 12267, 44–52. doi: 10.1007/978-3-030-59728-3_5

Lim, B., Prassas, I., and Diamandis, E. P. (2021). Alzheimer Disease Pathogenesis: The Role of Autoimmunity. *J. Appl. Lab. Med.* 6, 756–764. doi: 10.1093/jalm/jfaa171

Liu, W., Li, J., Li, L., Zhang, Y., Yang, M., Liang, S., et al. (2022). Enhanced Medial Prefrontal Cortex and Hippocampal Activity Improves Memory Generalization in APP/PS1 Mice: a Multimodal Animal MRI Study. *Front. Cell. Neurosci.* 16:848967–848967. doi: 10.3389/fncel.2022.848967

Lv, Z., Xu, T., Li, R., Zheng, D., Li, Y., Li, W., et al. (2022). Downregulation of m6A Methyltransferase in the Hippocampus of Tyrobp (-/-) Mice and Implications for Learning and Memory Deficits. *Front. Neurosci.* 16:739201–739201. doi: 10.3389/fnins.2022.739201

Mak, E., Gabel, S., Su, L., Williams, G. B., Arnold, R., Passamonti, L., et al. (2017). Multi-modal MRI investigation of volumetric and microstructural changes in the hippocampus and its subfields in mild cognitive impairment, Alzheimer's disease, and dementia with Lewy bodies. *Int. Psychogeriatr.* 29, 545–555. doi: 10.1017/s1041610216002143

Potashman, M., Buessing, M., Levitchi Benea, M., Cummings, J., Borson, S., Pemberton-Ross, P., et al. (2021). Estimating Progression Rates Across the Spectrum of Alzheimer's Disease for Amyloid-Positive Individuals Using National Alzheimer's Coordinating Center Data. *Neurol. Ther.* 10, 941–953. doi: 10.1007/s40120-021-00272-1

Poulin, S. P., Dautoff, R., Morris, J. C., Barrett, L. F., and Dickerson, B. C. (2011). Amygdala atrophy is prominent in early Alzheimer's disease and relates to

symptom severity. *Psychiatry Res.* 194, 7–13. doi: 10.1016/j.pscychresns.2011.06.014

Qi, S., Calhoun, V. D., van Erp, T. G. M., Bustillo, J., Damaraju, E., Turner, J. A., et al. (2018). Multimodal Fusion With Reference: searching for Joint Neuromarkers of Working Memory Deficits in Schizophrenia. *IEEE Trans. Med. Imag.* 37, 93–105. doi: 10.1109/tmi.2017.2725306

Schonberg, T., Pianka, P., Hendler, T., Pasternak, O., and Assaf, Y. (2006). Characterization of displaced white matter by brain tumors using combined DTI and fMRI. *Neuroimage* 30, 1100–1111. doi: 10.1016/j.neuroimage.2005.11.015

Shim, G., Choi, K. Y., Kim, D., Suh, S. I., Lee, S., Jeong, H. G., et al. (2017). Predicting neurocognitive function with hippocampal volumes and DTI metrics in patients with Alzheimer's dementia and mild cognitive impairment. *Brain Behav.* 7:e00766. doi: 10.1002/brb3.766

Sprung, J., Warner, D. O., Knopman, D. S., Petersen, R. C., Mielke, M. M., Jack, C. R. Jr., et al. (2021). Brain MRI after critical care admission: a longitudinal imaging study. *J. Crit. Care* 62, 117–123. doi: 10.1016/j.jcrc.2020.11.024

Tulay, E. E., Metin, B., Tarhan, N., and Arıkan, M. K. (2019). Multimodal Neuroimaging: Basic Concepts and Classification of Neuropsychiatric Diseases. *Clin. EEG Neurosci.* 50, 20–33. doi: 10.1177/1550059418782093

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Wang, S., Zhang, Y., Liu, G., Phillips, P., and Yuan, T. F. (2016). Detection of Alzheimer's Disease by Three-Dimensional Displacement Field Estimation in Structural Magnetic Resonance Imaging. *J. Alzheimers Dis.* 50, 233–248. doi: 10.3233/jad-150848

Wang, S. H., Du, S., Zhang, Y., Phillips, P., Wu, L. N., Chen, X. Q., et al. (2017). Alzheimer's Disease Detection by Pseudo Zernike Moment and Linear Regression Classification. *CNS Neurol. Disord. Drug Targets* 16, 11–15. doi: 10.2174/1871527315666161111123024

Wang, S.-H., Zhang, Y., Li, Y.-J., Jia, W.-J., Liu, F.-Y., Yang, M.-M., et al. (2018). Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. *Multimedia Tools Appl.* 77, 10393–10417. doi: 10.1007/s11042-016-4222-4

Wang, S. H., Zhou, Q., Yang, M., and Zhang, Y. D. (2021). ADVIAN: Alzheimer's Disease VGG-Inspired Attention Network Based on Convolutional Block Attention Module and Multiple Way Data Augmentation. *Front. Aging Neurosci.* 13:687456. doi: 10.3389/fnagi.2021.687456

Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., et al. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 1(Suppl.), S173–S186. doi: 10.1016/j.neuroimage.2008.10.055

Xia, M., Wang, J., and He, Y. (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One* 8:e68910. doi: 10.1371/journal.pone.0068910

Xu, X., Wang, T., Li, W., Li, H., Xu, B., Zhang, M., et al. (2021). Morphological, Structural, and Functional Networks Highlight the Role of the Cortical-Subcortical Circuit in Individuals With Subjective Cognitive Decline. *Front. Aging Neurosci.* 13:688113. doi: 10.3389/fnagi.2021.688113

Yan, C. G., Wang, X. D., Zuo, X. N., and Zang, Y. F. (2016). DPABI: data Processing & Analysis for (Resting-State) Brain Imaging. *Neuroinformatics* 14, 339–351. doi: 10.1007/s12021-016-9299-4

Yan, X., and Bien, J. J. S. S. (2017). Hierarchical Sparse Modeling: a Choice of Two Group Lasso Formulations. *Statist. Sci.* 32, 531–560.

Yuan, K., Zeng, T., and Chen, L. (2022). Interpreting Functional Impact of Genetic Variations by Network QTL for Genotype–Phenotype Association Study. *Front. Cell Dev. Biol.* 9:2022. doi: 10.3389/fcell.2021.720321

Zammit, M. D., Laymon, C. M., Tudorascu, D. L., Hartley, S. L., Piro-Gambetti, B., Johnson, S. C., et al. (2020). Patterns of glucose hypometabolism in Down syndrome resemble sporadic Alzheimer's disease except for the putamen. *Alzheimers Dement* 12:e12138. doi: 10.1002/dad2.12138

Zhang, Q., Alexander, M., and Ryner, L. (2014). Multimodality Neurological Data Visualization with Multi-VOI Based DTI Fiber Dynamic Integration. *IEEE J. Biomed. Health Inform.* 20:2367026. doi: 10.1109/JBHI.2014.2367026

Zhang, Y.-D., Wang, S., and Dong, Z. (2014). Classification of Alzheimer Disease Based on Structural Magnetic Resonance Imaging by Kernel Support Vector Machine Decision Tree. *Prog. Electrom. Res.* 144, 185–191. doi: 10.2528/PIER13121310

Zhang, T., and Shi, M. (2020). Multi-modal neuroimaging feature fusion for diagnosis of Alzheimer's disease. *J. Neurosci. Methods* 341:108795. doi: 10.1016/j.jneumeth.2020.108795

Zhang, Y., Dong, Z., Phillips, P., Wang, S., Ji, G., Yang, J., et al. (2015a). Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. *Front. Comput. Neurosci.* 9:66. doi: 10.3389/fncom.2015.00066

Zhang, Y., Wang, S., Phillips, P., Dong, Z., Ji, G., and Yang, J. (2015b). Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC. *Biomed. Signal Proc. Control* 21, 58–73. doi: 10.1016/j.bspc.2015.05.014

Zhang, Y., and Wang, S. (2015). Detection of Alzheimer's disease by displacement field and machine learning. *PeerJ.* 3:e1251. doi: 10.7717/peerj.1251

Zhang, Y., Wang, S., Phillips, P., Yang, J., and Yuan, T.-F. (2016). Three-Dimensional Eigenbrain for the Detection of Subjects and Brain Regions Related with Alzheimer's Disease. *J. Alzheimer's Dis.* 50, 1163–1179. doi: 10.3233/JAD-150988

Zhang, Y., Wang, S., Sui, Y., Yang, M., Liu, B., Cheng, H., et al. (2018). Multivariate Approach for Alzheimer's Disease Detection Using Stationary Wavelet Entropy and Predator-Prey Particle Swarm Optimization. *J. Alzheimers Dis.* 65, 855–869. doi: 10.3233/jad-170069

Zhu, J., Zhao, Y. P., and Zhang, Y. Q. (2019). The rs-fMRI study of effects of fornix and hippocampus-related brain function after the transcallosal interforniceal approach. *Brain Res. Bull.* 150, 207–215. doi: 10.1016/j.brainresbull.2019.05.014

Zhuang, H., Liu, R., Wu, C., Meng, Z., Wang, D., Liu, D., et al. (2019). Multimodal classification of drug-naïve first-episode schizophrenia combining anatomical, diffusion and resting state functional resonance imaging. *Neurosci. Lett.* 705, 87–93. doi: 10.1016/j.neulet.2019.04.039

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 25 May 2022
doi: 10.3389/fnagi.2022.916020

Large Margin and Local Structure Preservation Sparse Representation Classifier for Alzheimer's Magnetic Resonance Imaging Classification

Runmin Liu[1], Guangjun Li[1], Ming Gao[2], Weiwei Cai[3,4] and Xin Ning[5]*

[1] College of Sports Engineering and Information Technology, Wuhan Sports University, Wuhan, China, [2] College of Sports Science and Technology, Wuhan Sports University, Wuhan, China, [3] School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, China, [4] AiTech Artificial Intelligence Research Institute, Changsha, China, [5] Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

Alzheimer's disease (AD) is a progressive dementia in which the brain shrinks as the disease progresses. The use of machine learning and brain magnetic resonance imaging (MRI) for the early diagnosis of AD has a high probability of clinical value and social significance. Sparse representation classifier (SRC) is widely used in MRI image classification. However, the traditional SRC only considers the reconstruction error and classification error of the dictionary, and does not consider the global and local structural information between images, which results in unsatisfactory classification performance. Therefore, a large margin and local structure preservation sparse representation classifier (LMLS-SRC) is developed in this manuscript. The LMLS-SRC algorithm uses the classification large margin term based on the representation coefficient, which results in compactness between representation coefficients of the same class and a large margin between representation coefficients of different classes. The LMLS-SRC algorithm uses local structure preservation term to inherit the manifold structure of the original data. In addition, the LMLS-SRC algorithm imposes the $\ell_{2,1}$-norm on the representation coefficients to enhance the sparsity and robustness of the model. Experiments on the KAGGLE Alzheimer's dataset show that the LMLS-SRC algorithm can effectively diagnose non AD, moderate AD, mild AD, and very mild AD.

Keywords: Alzheimer's disease, sparse representation classifier, image classification, magnetic resonance imaging, KAGGLE Alzheimer's dataset

OPEN ACCESS

Edited by:
Yuanpeng Zhang, Nantong University, China

Reviewed by:
Lijun Xu, Nanjing Institute of Technology (NJIT), China
Min Shi, Fuzhou University of International Studies and Trade, China

****Correspondence:***
Guangjun Li
liguangjun@whsu.edu.cn

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

Received: *08 April 2022*
Accepted: *09 May 2022*
Published: *25 May 2022*

Citation:
Liu R, Li G, Gao M, Cai W and Ning X (2022) Large Margin and Local Structure Preservation Sparse Representation Classifier for Alzheimer's Magnetic Resonance Imaging Classification. Front. Aging Neurosci. 14:916020. doi: 10.3389/fnagi.2022.916020

INTRODUCTION

Alzheimer's disease (AD) is a chronic progressive neurodegenerative disease that usually progresses slowly in the early stages and gets worse over time (Katabathula et al., 2021). AD often occurs in the elderly. The initial symptoms are easy to forget recent events. With the development of the disease, the symptoms may include language problems, disorientation, mood swings, loss of self-care ability, etc., which will eventually seriously affect the daily life of the elderly. Currently, about 90 million people worldwide suffer from AD of varying degrees. It is estimated that by 2050, the number of AD patients will reach 300 million (Wong, 2020). The specific symptoms of very mild AD are progressive decline in memory or other cognitive functions, but do not affect the ability of daily living. According to statistics, about 10–15% of very mild AD will eventually transform into AD (Porsteinsson et al., 2021). Current scientific and clinical research has not yet clearly identified the

pathogenesis and etiology of AD, and there is no fully effective treatment drug. AD is uncontrollable and irreversible after being diagnosed. However, if patients can be intervened and treated in the early stage of mild cognitive impairment (MCI), it is hoped that the onset of AD will be delayed by 5 years, and even stop the progression of AD in the stage of MCI, and no longer worsen into AD, reducing the number of patients with AD by 40% (Venugopalan et al., 2021).

In the past decade, neuroimaging techniques have been widely used in the classification and prediction of AD. Among them, magnetic resonance imaging (MRI) is a non-contact imaging technology that can provide detailed three-dimensional anatomical images of the brain and provide effective information for the classification and prediction of AD (Al-Khuzaie and Duru, 2021). The AD classification algorithms based on machine learning usually extract the required features from the collected medical images by manual or semi-manual methods. Various parts of the brain regions of AD patients will atrophy to varying degrees due to the progression of the disease process. The volume, shape and texture information of the hippocampus, gray matter, white matter, and cerebral cortex of the brain are important features to distinguish AD and healthy people (Lee et al., 2020; Gao, 2021). To classify AD MRI images, some studies extract the volume information of the whole brain or part of the brain. Some scholars segment different regions of the brain and take the volume of each segment as features. According to the anatomical automatic labeling brain region template, some researchers divide the entire brain or part of the brain region into multiple regions and then obtain the features for each region. AD Patients often experience cerebral cortex atrophy and ventricular enlargement, and early AD patients usually have hippocampal atrophy (van Oostveen and de Lange, 2021). Therefore, some scholars use the volume information of different regions of interest such as the hippocampus as features based on medical prior knowledge. Another common feature extraction method is the morphometric measurement method, which is often implemented based on MRI images and PET images. For example, Al-Khuzaie and Duru (2021) took the overall shape of the brain in MRI images as features. Katabathula et al. (2021) used the shape information of the hippocampus as features. Brain gully dilation is often seen in AD patients. Furthermore, texture features are also widely used in MRI images. Gao (2021) extracted the grayscale co-occurrence matrix of images as features. Hett et al. (2018) used 3D Gabor filter to extract and classify multi-directional texture features of MRI images.

Classifiers such as sparse representation classifier (SRC), logistic regression (LR), support vector machine (SVM), and decision tree (DT) are widely used in AD MRI image classification. For example, Kruthika et al. (2019) used a multi-level classifier to classify AD MRI images. They first used a naive Bayes classifier, and then used SVM as secondary classification to classify the data with confidence lower than the threshold. Liu et al. (2015) proposed a multi-view learning algorithm based on inherent structure of mild cognitive impairment (MCI) MRI images, which used the multi-view features of MCI images to train multiple SVMs, and then fused and discriminated each classifier result. Altaf et al. (2018) used SVM, random forest, and K-nearest neighbor (KNN) to train AD classifiers, respectively, and the final classification result was the weighted sum of the results of each classifier. Yao et al. (2018) used the idea of hierarchical classification to classify AD MRI images. They initially classified samples into four classes (AD, healthy, MCI, converted MCI), then they trained several binary classifiers (AD and converted MCI, healthy and MCI), and finally got a classifier that can classify all samples into four classes. Pan et al. (2019) proposed an algorithm to integrate multi-level features based on FDG-PET images, and simultaneously considered the region features and connectivity between regions to classify AD or MCI from healthy people. Finally, multiple SVMs were used for voting classification, and good results had been achieved in multiple binary classification tasks.

Magnetic resonance imaging image features usually suffer from high dimensionality and small sample size, which may lead to overfitting in data-driven machine learning methods (Jiang et al., 2019). To solve this problem, most existing methods adopt feature selection or feature representation to exploit the potential knowledge of data. Sparse representation is one of the widely used feature representation methods. Sparse representation can explore potential relationships within the data (Gu et al., 2021). Chang et al. (2015) proposed a dictionary learning algorithm based on sparse decomposition of stacked prediction. They used the spatial pyramid matching method to encode representation coefficients, and used SVM to classify the pathological state of tumors. Shi et al. (2013) developed a multi-modal SRC algorithm for lung histopathological image classification, which used genetic algorithm to guide the learning of three sub-dictionaries of color, shape and texture, and then combined sparse reconstruction error and majority voting algorithm for classification of lung histopathology images. He (2019) proposed a spatial pyramid matching algorithm based on joint representation coefficient, which utilized the three color channel information of RGB, and converted the grayscale description operator into a color description operator, which improved the image classification performance. Jiang et al. (2019) extracted features from breast cancer histopathological images based on stacked sparse autoencoder, and used Softmax function to detect cell nuclei in histopathological images. Zhang et al. (2016) realized the fusion of global and local features of the nuclear image, and then combined the ranking and majority voting algorithm to classify the histopathological images of breast cancer. The above algorithms can effectively extract image features by introducing the sparsity of the image, and the extracted features have good reconstruction properties, but they do not have good discriminative ability.

To improve the diagnosis of MCI and AD based on MRI images, we propose large margin and local structure preservation sparse representation classifier (LMLS-SRC) in this manuscript. The traditional SRC only uses the classification error term to control the classification accuracy, and does not fully consider the class label information of the representation coefficients. Different from the traditional SRC, the LMLS-SRC algorithm introduces the classification margin term of representation coefficients into the sparse representation classifier, so that the similar representation coefficients are

compact in the representation space, and the dissimilar representation coefficients are separated as much as possible in the representation space. Experiments on the KAGGLE Alzheimer's dataset verify the advantages of our algorithm. Major contributions of this manuscript are highlighted below: (1) Considering the global information of the data by using the large margin term, the obtained dictionary is discriminative, and the representation coefficient has the small intra-class distance and large inter-class distance. (2) The local structure preservation term is introduced, which can inherit the manifold structure of the original data. (3) The $\ell_{2,1}$-norm term on the representation coefficients is used, which can enhance the sparsity and robustness of the representation coefficients.

BACKGROUNDS

Dictionary-Based Sparse Representation Classifier

Using SRC algorithm in image classification, how to design effective dictionary and representation coefficient for feature representation is the key factor to determine the algorithm performance (Wright et al., 2009). There are three aspects considered in the design of SRC algorithm: (1) The reconstruction error of the representation coefficients is small, so that the samples are as close to the original samples as possible in the sparse representation; (2) The representation coefficients are constrained to make the representation coefficients as sparse as possible; (3) The discrimination term should be considered to better extract more discriminative information of data (Jiang et al., 2013).

Let $\mathbf{X} = [\mathbf{X}_1, ..., \mathbf{X}_K] \in \mathbf{R}^{d\times N}$ be the K-classes training sample set, $\mathbf{X}_k = [\mathbf{x}_1, ..., \mathbf{x}_{N_k}]$ be the k-th class training sample subset, $k = 1, 2,..., K, N = N_1 + N_2 + \cdots + N_K$. d is the dimensional of samples. The SRC algorithm for image classification can be represented as,

$$\min_{D,A} \ ||\mathbf{X}\text{-}\mathbf{DA}||_F^2 + \lambda g(\mathbf{A}) + \eta f(\mathbf{D},\mathbf{A},\mathbf{Y}), \tag{1}$$

where $\mathbf{Y}$ is the class label matrix of $\mathbf{X}$. $\mathbf{D} \in \mathbf{R}^{d\times m}$ is the learned dictionary, and $\mathbf{A} \in \mathbf{R}^{m\times N}$ is the representation coefficient matrix of $\mathbf{X}$. m is the size of dictionary. In model training, the data reconstruction item $||\mathbf{X}\text{-}\mathbf{DA}||_F^2$ is to ensure the representation ability of the dictionary $\mathbf{D}$, so that the reconstruction error of the training data is minimized, and the reconstructed image is as close to the original sample as possible. The regularization term is used to constrain the sparsity of the representation coefficients, which is usually represented as,

$$g(\mathbf{A}) = ||\mathbf{A}||_{\mathrm{p}}. \tag{2}$$

where $||\cdot||_p$ is the regularization term of the representation coefficient $\mathbf{A}$ ($p < 2$), which makes the representation coefficient as sparse as possible. $f(\mathbf{D},\mathbf{A},\mathbf{Y})$ is the discriminative function term of representation coefficient for classification to ensure the discriminative ability of $\mathbf{D}$ and $\mathbf{A}$.

To obtain a discriminative dictionary, Yang et al. (2017) developed a supervised Fisher discrimination dictionary learning (FDDL), which associated the elements in the dictionary with the class labels of the samples based on the Fisher discrimination criterion. Jiang et al. (2013) proposed the discriminative Label consistent K-SVD (LC-KSVD) algorithm. Zhang et al. (2019) proposed a robust flexible discriminative dictionary learning (RFDDL) algorithm based on subspace recovery and enhanced locality. This algorithm improved image representation and classification by enhancing representation coefficient robustness. The computational complexity of the SRC representation coefficient is usually high. To quickly obtain the representation coefficients, Ma et al. (2017) proposed the local sparse representation algorithm, which used the KNN criterion to select k samples adjacent to the current sample to build a dictionary matrix. In this way, the size of the dictionary is reduced and the process of representation coefficient is greatly accelerated. Similarly, inspired by the KNN criterion, Zheng and Ding (2020) developed a sparse KNN classifier based on group lasso strategy and KSVD algorithm. Wang et al. (2018) proposed a SRC algorithm based on the ℓ_2-norm, which replaced the ℓ_1-norm with the ℓ_2-norm to constrain the coefficients. Ortiz and Becker (2014) proposed an approximate linear SRC algorithm. Authors used least square algorithm to select the training samples corresponding to the absolute values of the k largest coefficients to build a sub-dictionary.

KAGGLE Alzheimer's Image Dataset

The experiments in this manuscript are carried out on the KAGGLE Alzheimer's image dataset (Loddo et al., 2022). The KAGGLE Alzheimer's dataset contains a total of four types of MRI images: non AD (3,200 images), very mild AD (2,240 images), mild AD (896 images) and moderate AD (64 images), with the resolution of 176 × 208. The KAGGLE Alzheimer's dataset does not provide detailed information on patient status. **Figure 1** shows some example images of the KAGGLE Alzheimer's dataset.

THE PROPOSED ALGORIHTM

Objective Function

The purpose of sparse representation is to represent the sample with as few elements as possible on a given dictionary, so that a more concise representation of the sample can be obtained, and the useful information contained in the sample can be easily obtained. Thus the core problem of sparse representation is how to compute sparse coding coefficients on a given learned dictionary. Compared with the commonly used ℓ_1-norm and ℓ_2-norm, $\ell_{2,1}$-norm can improve the robustness of the model and reduce the computational complexity. Thus, we introduce $\ell_{2,1}$-norm constraint on representation coefficients in LMLS-SRC, i.e.,

$$\Pi_1 = \arg\min_{D,A}\{||\mathbf{X}\text{-}\mathbf{DA}||_F^2 + \lambda_1||\mathbf{A}||_{2,1}^2\}, \tag{3}$$

where λ_1 is a constant.

We define a large margin term on representation coefficient that relies on a specific neighborhood size for intra-class and inter-class representation coefficients. The large margin

term minimizes the intra-class distance of the representation coefficient and maximizes the inter-class distance of the representation coefficient, so as to improve the difference between the representation coefficients of different classes. The large margin term on representation coefficient can be written as,

$$f(\mathbf{a}_i) = \arg\min\{\sum_{t\in C_k}\frac{||\mathbf{a}_i-\mathbf{a}_t||^2}{N_k} - \sum_{j\notin C_k}\frac{||\mathbf{a}_i-\mathbf{a}_j||^2}{N-N_k}\}, \quad (4)$$

where $\sum_{t\in C_k}\frac{||\mathbf{a}_i-\mathbf{a}_t||^2}{N_k}$ represents the distance between a_i and the sparse representation of the same class. $\sum_{j\notin C_k}\frac{||a_i-a_t||^2}{N-N_k}$ represents the distance between a_i and the sparse representation of the different class. C_k is the index set of the k-th class sample.

We build the intra-class similarity matrix Q^w and inter-class similarity matrix Q^b based on representation coefficient. The elements of the matrix Q^w and matrix Q^b are expressed as,

$$q_{ij}^w = \begin{cases} 1/N_k, \textit{ if } i,j \in C_k \\ 0, \quad \textit{otherwise} \end{cases} \quad (5)$$

$$q_{ij}^b = \begin{cases} 1/(N-N_k), \textit{ if } i \in C_k, j \notin C_k \\ 0, \quad \textit{otherwise} \end{cases} \quad (6)$$

Then the large margin term on representation coefficient can be expressed as,

$$\begin{aligned} \Pi_2 &= \arg\min_A \tfrac{1}{N}\textstyle\sum_{i=1}^N f(\mathbf{a}_i) \\ &= \tfrac{1}{N}\textstyle\sum_{i=1}^N\sum_{j=1}^N (q_{ij}^w||\mathbf{a}_i-\mathbf{a}_j||^2 - q_{ij}^b||\mathbf{a}_i-\mathbf{a}_j||^2) \\ &= \tfrac{1}{N}\left(2\textstyle\sum_{i=1}^N \mathbf{a}_i^2 - 2\sum_{i=1}^N\sum_{j=1}^N \mathbf{a}_i q_{ij}^w \mathbf{a}_j\right) - \\ &\tfrac{1}{N}\left(\textstyle\sum_{i=1}^N \mathbf{a}_i^2 + \sum_{j=1}^N \mathbf{a}_j^2 q_{jj}^b - 2\sum_{i=1}^N\sum_{j=1}^N \mathbf{a}_i q_{ij}^b \mathbf{a}_j\right) \\ &= tr(\tfrac{1}{N}\mathbf{A}^T(2\mathbf{I}-2\mathbf{Q}^w)\mathbf{A}) - tr(\tfrac{1}{N}\mathbf{A}^T(\mathbf{I}+\tilde{\mathbf{Q}}^b-2\mathbf{Q}^b)\mathbf{A}) \\ &= tr(\tfrac{1}{N}\mathbf{A}^T(\mathbf{I}-2\mathbf{Q}^w-\tilde{\mathbf{Q}}^b+2\mathbf{Q}^b)\mathbf{A}) \\ &= tr(\mathbf{A}^T\mathbf{S}\mathbf{A}) \end{aligned} \quad (7)$$

where $\mathbf{S} = \frac{1}{N}(\mathbf{I}-2\mathbf{Q}^w-\tilde{\mathbf{Q}}^b+2\mathbf{Q}^b)$. The matrix $\tilde{\mathbf{Q}}^b$ is the diagonal matrix with the element being the column-sum of $\mathbf{Q}^b$.

Following the principle of local structure preservation, if two images are close in the original space, they should also have similar representation coefficients. To this end, we construct a similarity matrix $\mathbf{P}$ that reflects the intrinsic local structure between images. The element of matrix $\mathbf{P}$ is defined as,

$$p_{ij} = \begin{cases} \exp\left(-\frac{||\mathbf{x}_i-\mathbf{x}_j||_2^2}{2\sigma^2}\right), \textit{ if } \mathbf{x}_i \in N(\mathbf{x}_j) \textit{ or } \mathbf{x}_j \in N(\mathbf{x}_i), \\ 0, \quad \textit{otherwise}, \end{cases} \quad (8)$$

where $N(\mathbf{x}_j)$ represents the k nearest neighbors of $\mathbf{x}_j$.

The local structure preservation term on representation coefficient is expressed as,

$$\begin{aligned} \Pi_3 &= \arg\min_A \textstyle\sum_{i,j} p_{ij}||\mathbf{a}_i-\mathbf{a}_j||_2^2 \\ &= tr(\mathbf{A}^T(\mathbf{P}-\tilde{\mathbf{P}})\mathbf{A}) \\ &= tr(\mathbf{A}^T\mathbf{L}\mathbf{A}), \end{aligned} \quad (9)$$

where the graph Laplacian matrix $\mathbf{L}$ is $\mathbf{L} = \mathbf{P}-\tilde{\mathbf{P}}$,$\tilde{\mathbf{P}}$ is the diagonal matrix with the element being the row-sum of $\mathbf{P}$.

The LMLS-SRC algorithm is a supervised learning model. Using the class labels of all training samples, we use a linear classifier $\mathbf{W}$ for representation coefficient $\mathbf{A}$ and dictionary $\mathbf{D}$, i.e.,

$$\Pi_4 = \arg\min_{W,A} ||\mathbf{W}\mathbf{A}-\mathbf{Y}||_F^2 + \lambda_5||\mathbf{W}||_F^2. \quad (10)$$

In summary, the objective function of the LMLS-SRC algorithm can be written as,

$$F(\mathbf{D},\mathbf{A},\mathbf{W}) = \min_{D,A,W} \Pi_1 + \Pi_2 + \Pi_3 + \Pi_4, \quad (11)$$

i.e.,

$$\begin{aligned} &\min_{D,A,W} ||\mathbf{X}-\mathbf{D}\mathbf{A}||_F^2 + \lambda_1||\mathbf{A}||_{2,1}^2 + \lambda_2 tr(\mathbf{A}^T\mathbf{S}\mathbf{A}) \\ &+\lambda_3 tr(\mathbf{A}^T\mathbf{L}\mathbf{A}) + \lambda_4||\mathbf{W}\mathbf{A}-\mathbf{Y}||_F^2 + \lambda_5||\mathbf{W}||_F^2, \\ &s.t.\ ||d_i||_2^2 \le 1, \forall i \end{aligned} \quad (12)$$

where λ_1,λ_2,λ_3,λ_4, and λ_5 are trade-off parameters.

By alternately optimizing the representation coefficient $\mathbf{A}$, dictionary $\mathbf{D}$ and classifier parameter $\mathbf{W}$, the following performance can be obtained as: (1) the dictionary $\mathbf{D}$ has more sparse representation performance, which enhances the reconstruction of the sample by the dictionary. (2) LMLS-SRC maximizes the distance between different classes of representation coefficients and greatly reduces the similarity between different classes of representation coefficients. (3) The representation coefficient is more discriminative, which is beneficial to the performance of image classification.

Optimization

(1) Fix $\mathbf{D}$, $\mathbf{W}$, and update $\mathbf{A}$. Eq. (12) can be written by,

$$\begin{aligned} \min F(A) = &||\mathbf{X}-\mathbf{D}\mathbf{A}||_F^2 + \lambda_1||\mathbf{A}||_{2,1}^2 + \lambda_2 tr(\mathbf{A}^{\mathrm{T}}\mathbf{S}\mathbf{A}) \\ &+\lambda_3 tr(\mathbf{A}^{\mathrm{T}}\mathbf{L}\mathbf{A}) + \lambda_4||\mathbf{W}\mathbf{A}-\mathbf{Y}||_F^2. \end{aligned} \quad (13)$$

According to the definition of $\ell_{2,1}$-norm, $||\mathbf{A}||_{2,1}^2 = tr(\mathbf{A}^{\mathrm{T}}\mathbf{\Omega}\mathbf{A})$. $\mathbf{\Omega}$ is a diagonal matrix whose elements are setting by $\mathbf{\Omega}_{ii} = 1/(2||\mathbf{A}_i||_2)$ where $\mathbf{A}_i$ represents the i-th row of $\mathbf{A}$.

Equation (12) can be re-written by,

$$\begin{aligned} \min F(\mathbf{A}) = &||\mathbf{X}-\mathbf{D}\mathbf{A}||_F^2 + \lambda_1 tr(\mathbf{A}^{\mathrm{T}}\mathbf{\Omega}\mathbf{A}) + \lambda_2 tr(\mathbf{A}^{\mathrm{T}}\mathbf{S}\mathbf{A}) \\ &+\lambda_3 tr(\mathbf{A}^{\mathrm{T}}\mathbf{L}\mathbf{A}) + \lambda_4||\mathbf{W}\mathbf{A}-\mathbf{Y}||_F^2. \end{aligned} \quad (14)$$

Setting $\partial F(\mathbf{A})/\partial\mathbf{A} = 0$, we can obtain,

$$\begin{aligned} \frac{\partial L}{\partial A} = &2\mathbf{D}^{\mathrm{T}}\mathbf{D}A - 2\mathbf{D}^{\mathrm{T}}\mathbf{X} + (2\lambda_1\mathbf{\Lambda} + 2\lambda_2\mathbf{S} + 2\lambda_3\mathbf{L})\mathbf{A} \\ &+2\lambda_4(\mathbf{W}^{\mathrm{T}}\mathbf{W}\mathbf{A}-\mathbf{W}^{\mathrm{T}}\mathbf{Y}). \end{aligned} \quad (15)$$

$\mathbf{A}$ can obtained by the updated by,

$$\begin{aligned} \mathbf{A}^* = &(\mathbf{D}^{\mathrm{T}}\mathbf{D} + \lambda_1\mathbf{\Lambda} + \lambda_4\mathbf{W}^{\mathrm{T}}\mathbf{W} + \lambda_2\mathbf{S} + \lambda_3\mathbf{L})^{-1} \\ &(\lambda_4\mathbf{W}^{\mathrm{T}}\mathbf{Y} + \mathbf{D}^{\mathrm{T}}\mathbf{X}). \end{aligned} \quad (16)$$

(2) Fix $\mathbf{A}$, $\mathbf{W}$, and update $\mathbf{D}$. Equation (12) can be written by,

$$\begin{aligned} &\min F(\mathbf{D}) = ||\mathbf{X}-\mathbf{D}\mathbf{A}||_F^2, \\ &s.t.\ ||d_i||_2^2 \le 1, \forall i \end{aligned} \quad (17)$$

We can solve Eq. (17) by the following Lagrangian dual function,

$$\min F(\mathbf{D}, \boldsymbol{\sigma}) = ||\mathbf{X}-\mathbf{DA}||_F^2 + \sum_{i=1}^{m} \gamma_i \left(||d_i||_2^2 - 1 \right), \quad (18)$$

where γ_i is the Lagrange multiplier of i-th atoms.

We build a diagonal matrix $\boldsymbol{\Theta}$ with the element $\Theta_{ii} = \gamma_i$. Equation (18) can be written by,

$$\min F(\mathbf{D}, \boldsymbol{\Theta}) = ||\mathbf{X}-\mathbf{DA}||_F^2 + tr(\mathbf{D}^{\mathrm{T}}\mathbf{D}\boldsymbol{\Theta}) - tr(\boldsymbol{\Theta}). \quad (19)$$

Setting $\partial F(\mathbf{D}, \boldsymbol{\Theta}) \big/ \partial D = 0$, we can obtain,

$$\mathbf{D}^* = \mathbf{X}\mathbf{A}^{\mathrm{T}}(\mathbf{A}\mathbf{A}^{\mathrm{T}} + \boldsymbol{\Theta})^{-1}. \quad (20)$$

(3) Fix $\mathbf{A}$ and $\mathbf{D}$, and update $\mathbf{W}$. Equation (12) can be written by,

$$\min F(\mathbf{W}) = \lambda_4 ||\mathbf{W}\mathbf{A}-\mathbf{Y}||_F^2 + \lambda_5 ||\mathbf{W}||_F^2. \quad (21)$$

Setting $\partial F(\mathbf{W}) \big/ \partial W = 0$, we can obtain,

$$\mathbf{W}^* = \lambda_4 \mathbf{Y}\mathbf{A}^{\mathrm{T}}(\lambda_4 \mathbf{A}\mathbf{A}^{\mathrm{T}} + \lambda_5 \mathbf{I})^{-1}. \quad (22)$$

The optimization steps of LMLS-SRC algorithm are shown in Algorithm 1.

Input: training set $\mathbf{X}$ and its label matrix $\mathbf{Y}$, tolerance error δ, maximum number of iterations *maxiter*, parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, and λ_5,

Output: parameters $\mathbf{D}$, $\mathbf{A}$, and $\mathbf{W}$.

Initialize: initialize $\mathbf{D}$ and $\mathbf{A}$ using the LC-KSVD algorithm, $\mathbf{W} = \mathbf{I}$, $m = 1$,

Calculate matrices Q^w, Q^b, and $\mathbf{P}$;

While not converged and $m \leq$ *maxiter* do

  Calculate $\mathbf{D}(m)$ by Eq. (20);

  Calculate $\mathbf{A}(m)$ by Eq. (16);

  Calculate $\mathbf{W}(m)$ by Eq. (22);

  Check the convergence condition $\frac{|F(\mathbf{D}(m),\mathbf{A}(m),W(m)) - F(\mathbf{D}(m-1),A(m-1),\mathbf{W}(m-1))|}{F(\mathbf{D}(m-1),\mathbf{A}(m-1),\mathbf{W}(m-1))} < \delta$

  $m = m + 1$

end while

EXPERIMENTS

Experimental Settings

In clinical diagnosis, AD classification tasks consist of two categories. The first is the AD binary classification task, which extracts features based on MRI images and uses machine learning models to classify normal individuals and AD patients, which can help doctors diagnose AD patients. The second is the classification of various ADs, especially the diagnosis and identification of mild AD and very mild AD. Early prediction of AD can help to take treatment and intervention measures in the early stage of AD. Therefore, in this manuscript, we design binary, three-class and four-class classification tasks on the KAGGLE Alzheimer's dataset.

Volume analysis is the commonly used feature extraction method in AD classification. Volumetric feature extraction is divided into two categories: density maps and predefined area methods. AD MRI image is mainly related to the volume of the density map structure, cortical structure, subcortical structure and other regions. In this manuscript, we use FSL (FMRIB software library) toolbox to extract MRI features (Jenkinson et al., 2012). FSL is a library of comprehensive analysis tools for brain imaging data such as MRI, developed by the FMRIB Centre in Oxford. We use the FSL toolbox to calculate the volume, area and thickness characteristics of various brain tissues in brain MRI images. In the comparison experiment, the LMLS-SRC algorithm is compared with SRC (Wright et al., 2009), logistic regression (LR) (Tsangaratos and Ilia, 2016), linear discriminant (LD) (Kim et al., 2011), LC-KSVD, FDDL, and sparse representation-based discriminative metric learning (SRDML) (Zhou et al., 2022). The radial basis function (RBF) kernel is used in LR. The default settings are used to produce test results from these classifiers using the MATLAB classification learner toolbox. The RBF kernel and the regularization parameters for all comparison algorithms range from 10^{-3} to 10^3. The number of dictionary atoms in SRC and dictionary learning is set as the number of training samples. Indicators of classification performance include classification accuracy, sensitivity, specificity, precision, F1-score, and G-mean. We carry out 5-fold cross-validation strategy and record the experimental results.

Experimental Results

(1) Binary classification task. The main goal of this work is to classify brain MRI into AD and non AD classes. We utilized 3,200 and 62 MRI images for non AD and AD classes, respectively. We randomly selected 1,000 MRI images from the non AD class images to increase the moderate AD class dataset to 620 MRI images using data augmentation techniques. The comparative training and test results in binary classification task are shown in **Tables 1**, **2**, respectively.

(2) Three-class classification tasks. The main goal of this work was to classify brain MRI into three classes: non AD, mild AD, and moderate AD. Using data augmentation techniques, these three classes of datasets contain 3,200, 700, and 620 images, respectively. We randomly selected 1,000 MRI images from the non AD class. The comparative training and test results in three-class classification task are shown in **Tables 3**, **4**, respectively.

(3) Four-class classification tasks. The main goal of this work is to classify brain MRI images into four classes: very mild AD, non AD, mild AD, and moderate AD. Similar to the three-class classification task described, we randomly selected 1,000 MRI images each from non AD class images and very mild AD, respectively, and used data augmentation to increase the moderate dementia dataset to 520 MRI images. The number of images in the four categories of very mild AD, non AD, mild AD, and moderate AD are 1,000, 1,000, 700, and 520, respectively. The comparison training and test results in four-class classification task are shown in **Tables 5**, **6**, respectively.

We can see that all the comparison algorithms have the highest classification accuracy in the binary classification task (AD and non AD). It shows that these machine learning algorithms have excellent performance in the classification and diagnosis of AD. It is more practical to classify patients, very mild AD, non

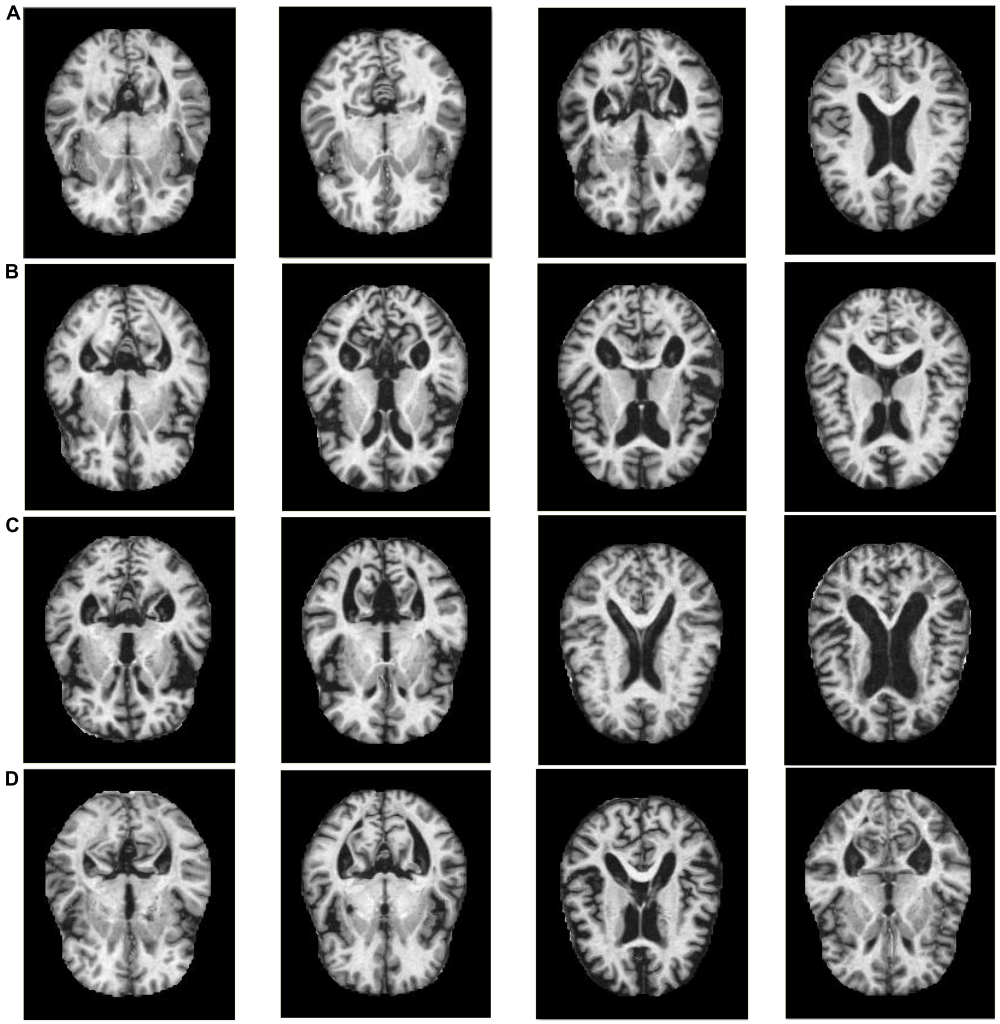


FIGURE 1 | Example samples of the KAGGLE Alzheimer's dataset, **(A)** Non AD, **(B)** Moderate AD, **(C)** Mild AD, **(D)** Very Mild AD.

AD, mild AD, and moderate AD into four classes, and this classification task is more difficult. The classification accuracy of all the comparison algorithms on the four-class task is slightly lower than that on the two-class task. However, the LMLS-SRC algorithm achieves the best results in these tables, indicating that our algorithm has a great improvement in the diagnosis of AD.

In **Tables 2**, **4**, **6**, the LMLS-SRC algorithm improves the classification accuracy of the second best algorithm by 2.84, 3.00, and 2.41%, respectively. This shows that the dictionary learned in this study has better reconstruction performance for the samples of same class and better discriminative performance for samples of different classes. KSVD, LC-KSVD, and LMLS-SRC are SRC algorithms. The KSVD and LC-KSVD algorithms only constrain the discriminative ability of the representation coefficients, and do not take into account the large margin between the representation coefficients of different classes. Therefore, the discriminative ability of the learned dictionary obtained by KSVD and LC-KSVD is still weak. The dictionary learned by the LMLS-SRC algorithm in this manuscript is combined with the classification large margin criterion, which directly constrains the intra-class distance and inter-class distance of the representation coefficients. Compared with the other three algorithms, the inter-class differences of the dictionary learned by our algorithm are more discriminative.

Parameter Analysis

(1) Convergence analysis. The update of {(**D**), (**A**), (**W**)} in the objective function are three convex optimization problems. That is, when other parameters are fixed, the iterative solution of dictionary **D**, representation coefficient **A** and classifier parameter **W** is the convex problem. The solution of dictionary **D** is obtained by Eq. (20). The solution of dictionary **A** is obtained by Eq. (16). The solution of dictionary **W** is obtained by Eq. (22). **Figure 2** shows the convergence of the LMLS-SRC algorithm. As shown in **Figure 2**, it can be seen that the classification accuracy of the LMLS-SRC algorithm tends to be parallel to the X-axis from the 10th iteration. Here, it can be considered that our algorithm converges after 12 iterations.

TABLE 1 | The comparative training results (with standard deviation) in binary classification task.

Algorithms	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-mean
LD	81.30	81.99	80.06	80.58	82.26	81.02
	(2.84)	(3.15)	(2.80)	(3.32)	(3.27)	(2.97)
LR	82.15	82.62	81.79	82.68	82.51	82.20
	(2.55)	(2.66)	(2.70)	(2.35)	(2.56)	(2.60)
SRC	82.10	78.97	77.33	77.63	77.55	78.15
	(2.35)	(2.01)	(2.64)	(1.62)	(1.43)	(2.28)
LC-KSVD	80.27	81.34	78.94	80.85	79.93	80.13
	(2.54)	(2.12)	(2.63)	(1.82)	(2.07)	(1.59)
FDDL	83.16	84.47	81.38	85.20	82.86	82.91
	(2.64)	(**??**)	(1.83)	(1.45)	(1.69)	(1.54)
SRDML	85.71	85.91	85.09	84.10	85.08	85.50
	(2.15)	(2.23)	(1.75)	(1.88)	(1.74)	(1.96)
LMLS-SRC	**89.80**	**90.39**	**87.87**	**88.89**	**90.43**	**89.12**
	(2.02)	(1.35)	(2.06)	(1.35)	(1.28)	(1.19)

*The bold values in **Tables 1–6** are the best experiment results.*

TABLE 2 | The comparative test results (with standard deviation) in binary classification task.

Algorithms	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-mean
LD	80.92	81.64	80.44	81.45	80.70	81.04
	(2.26)	(1.69)	(2.10)	(2.06)	(1.62)	(1.37)
LR	81.61	82.28	80.96	82.81	80.79	81.62
	(1.71)	(2.58)	(2.70)	(1.04)	(1.88)	(2.64)
SRC	82.91	83.18	82.86	83.07	82.94	83.02
	(1.75)	(2.46)	(2.28)	(1.16)	(1.87)	(2.37)
LC-KSVD	82.15	82.59	80.51	82.78	82.56	81.54
	(2.74)	(1.38)	(2.80)	(2.55)	(1.96)	(1.93)
FDDL	82.89	84.26	81.71	84.35	83.23	82.98
	(2.23)	(1.50)	(1.43)	(1.14)	(2.02)	(1.46)
SRDML	85.44	87.13	84.35	86.42	85.42	85.73
	(2.14)	(2.20)	(2.10)	(2.74)	(2.05)	(2.15)
LMLS-SRC	**88.28**	**90.15**	**86.75**	**90.08**	**88.31**	**88.43**
	(2.07)	(2.06)	(1.67)	(1.92)	(1.18)	(1.68)

TABLE 3 | The comparative training results (with standard deviation) in three-class classification task.

Algorithms	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-mean
LD	80.13	80.70	80.52	79.57	80.94	80.61
	(2.72)	(1.92)	(2.24)	(2.28)	(2.34)	(2.36)
LR	81.31	82.55	80.25	81.17	81.82	81.39
	(2.55)	(2.30)	(2.03)	(2.62)	(2.19)	(2.16)
SRC	81.94	82.20	80.46	81.09	81.17	81.33
	(2.20)	(2.49)	(2.59)	(2.10)	(2.21)	(2.54)
LC-KSVD	83.80	85.54	81.64	83.94	83.32	83.57
	(1.76)	(1.68)	(2.98)	(2.23)	(1.80)	(2.24)
FDDL	84.04	86.12	81.13	83.93	84.32	83.59
	(2.30)	(2.61)	(2.33)	(2.24)	(2.36)	(2.47)
SRDML	85.39	86.82	84.88	86.32	86.86	85.85
	(2.33)	(2.00)	(2.37)	(2.05)	(2.33)	(2.02)
LMLS-SRC	**89.32**	**91.38**	**86.81**	**88.86**	**89.00**	**89.07**
	(1.84)	(1.20)	(2.81)	(2.12)	(1.53)	(1.83)

TABLE 4 | The comparative test results (with standard deviation) in three-class classification task.

Algorithms	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-mean
LD	78.47	79.40	77.83	78.94	78.76	78.61
	(2.16)	(1.99)	(2.50)	(2.29)	(1.60)	(2.23)
LR	79.43	80.38	78.75	79.50	78.99	79.56
	(2.02)	(2.56)	(2.19)	(1.95)	(2.18)	(1.75)
SRC	80.23	80.22	79.26	79.31	79.47	79.74
	(1.79)	(2.53)	(2.30)	(2.54)	(1.39)	(2.31)
LC-KSVD	81.72	82.22	80.59	81.19	81.05	81.40
	(1.31)	(2.34)	(2.41)	(2.22)	(1.35)	(2.40)
FDDL	82.26	83.12	80.87	82.53	82.39	81.98
	(2.20)	(2.37)	(1.42)	(2.56)	(2.44)	(1.84)
SRDML	84.90	85.66	83.86	85.27	85.11	84.76
	(2.27)	(2.49)	(1.80)	(2.83)	(2.13)	(2.12)
LMLS-SRC	**87.90**	**89.25**	**86.53**	**88.71**	**88.44**	**87.88**
	(1.81)	(2.02)	(2.04)	(1.74)	(1.81)	(2.27)

TABLE 5 | The comparative training results (with standard deviation) in four-class classification task.

Algorithms	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-mean
LR	79.70	80.06	78.70	81.49	79.23	79.37
	(1.47)	(2.15)	(2.20)	(2.73)	(1.48)	(2.10)
LR	80.81	81.71	79.41	80.87	80.40	80.55
	(1.88)	(1.47)	(2.09)	(2.11)	(1.22)	(1.27)
SRC	80.86	82.38	79.92	78.97	80.41	81.14
	(2.02)	(2.29)	(1.84)	(1.37)	(1.62)	(2.05)
LC-KSVD	82.61	84.10	80.92	82.36	83.52	82.50
	(2.16)	(1.58)	(1.55)	(2.02)	(2.32)	(1.59)
FDDL	83.85	84.56	82.70	83.46	84.09	83.63
	(1.56)	(2.80)	(2.29)	(3.09)	(2.07)	(2.53)
SRDML	85.91	86.46	83.28	83.39	84.97	84.85
	(2.05)	(2.63)	(2.34)	(2.16)	(1.55)	(2.48)
LMLS-SRC	**86.58**	**87.64**	**85.93**	**86.93**	**86.21**	**86.78**
	(1.59)	(1.13)	(2.45)	(2.00)	(1.49)	(1.66)

(2) Training set size. The size of the training set usually directly determines the performance of machine learning algorithms. **Figure 3** shows the classification accuracy of the LMLS-SRC algorithm on binary-class, three-class and four-class classification tasks under different training sets of each subclass. The X-axis represents the training sample size N of each subclass, N = [50, 100,. . ., 400]. From **Figure 3**, we can see that the accuracy of LMLS-SRC increases with the increase of training samples. When the training sample size of each subset reaches 200, the performance of the LMLS-SRC algorithm is basically stable, indicating that the LMLS-SRC algorithm can achieve better performance without too many training samples.

(3) Regularization parameters. The LMLS-SRC algorithm has five regularization parameters λ_1,λ_2,λ_3,λ_4, and λ_5, and the regularization parameters are all obtained in [1.0E-3. . ., 1.0E+3]. λ_2 controls the role of the large margin term. λ_3 controls the role of the local structure preservation term. λ_4 controls the role of the linear classifier. **Figure 4** shows the classification accuracy of the LMLS-SRC algorithm in the binary, three-class and four-class

TABLE 6 | The comparative test results (with standard deviation) in four-class classification task.

Algorithms	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-mean
LD	77.67	79.69	77.51	78.40	77.50	78.59
	(2.22)	(1.51)	(2.15)	(2.52)	(2.08)	(1.80)
LR	78.56	79.60	78.47	78.59	78.23	79.03
	(1.89)	(2.51)	(1.60)	(2.74)	(1.43)	(2.01)
SRC	79.40	79.77	79.06	80.25	79.15	79.41
	(2.13)	(2.33)	(2.68)	(1.44)	(2.32)	(2.50)
LC-KSVD	81.25	81.77	81.34	80.87	81.55	81.55
	(2.40)	(2.19)	(1.59)	(2.36)	(2.08)	(1.86)
FDDL	81.45	81.06	80.02	80.67	80.69	80.54
	(1.33)	(2.00)	(2.09)	(2.73)	(1.25)	(2.05)
SRDML	83.13	82.10	82.94	83.56	83.17	82.52
	(2.06)	(2.26)	(2.04)	(1.49)	(1.99)	(2.15)
LMLS-SRC	**85.54**	**86.19**	**84.51**	**86.15**	**85.97**	**85.34**
	(1.59)	(2.03)	(2.12)	(1.63)	(1.06)	(2.07)

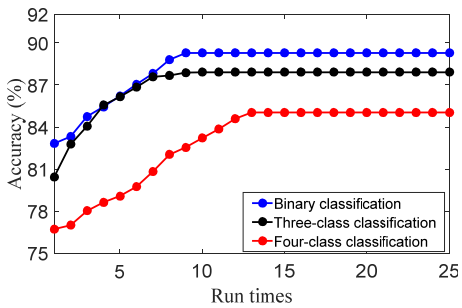


FIGURE 2 | Convergence of the LMLS-SRC algorithm.

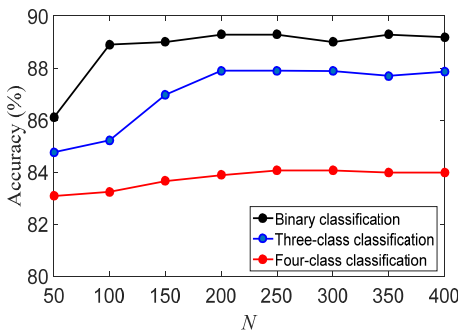


FIGURE 3 | Classification accuracy of local structure preservation sparse representation classifier (LMLS-SRC) under different training sets of each subclass.

tasks with different λ_2, λ_3, and λ_4, respectively. **Figure 4** shows that the performance of the LMLS-SRC algorithm varies greatly with different λ_2, λ_3, and λ_4, while fixing the other parameters. Therefore, it is reasonable to use a grid search strategy to optimize the regularization parameters.

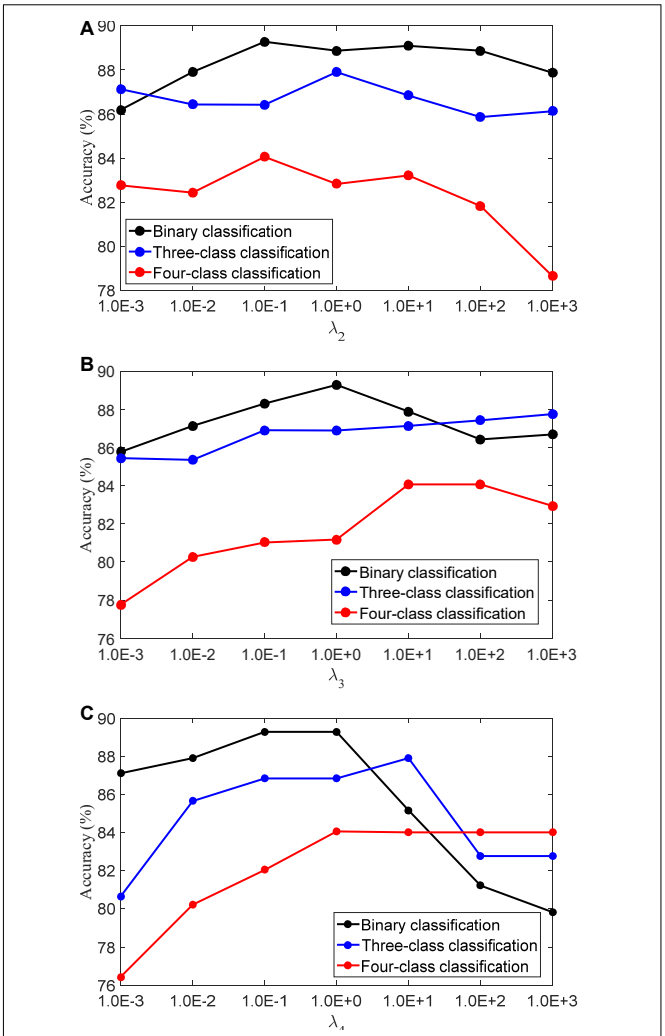


FIGURE 4 | Classification accuracy of LMLS-SRC with different regularization parameters, **(A)**λ_2, **(B)**λ_3, and **(C)** λ_4.

CONCLUSION

With the acceleration of the global aging trend, one of the problems brought about is the rapid increase in the number of AD patients. The pathogenesis and effective treatment of AD are still unclear at present. Early detection, classification, and prediction of AD, and targeted care and treatment of patients on this basis can delay the progression of AD. Machine learning algorithms that can automatically extract information and complete inference have good application prospects in AD classification and prediction. Therefore, this manuscript conducts research based on the application of SRC algorithm in AD classification. The research content mainly includes two aspects: model construction and model performance evaluation. The proposed LMLS-SRC algorithm introduces the large margin term and local constraint term in the traditional SRC model, and obtains the dictionary and representation coefficients with discriminative ability while maintaining the data manifold

structure. The effectiveness of the LMLS-SRC algorithm is validated on the KAGGLE Alzheimer's dataset.

Although the LMLS-SRC algorithm shows the advantages compared with some excellent algorithms, there are still some problems to be solved. In the future, we will mainly focus on the following aspects: (1) The LMLS-SRC algorithm belongs to the shallow model. How to design the deep model of the sparse representation algorithm needs to be further studied. (2) In this manuscript, brain MRI images are used as the basic data to study the application of AD classification. Multimodal data can provide richer information, and how to extract AD-related features from multimodal data can be studied in the future. (3) This manuscript uses the volume features extracted by using FSL tool. Extracting various features for AD classification can be done in the next future. (4) In practical applications, image classification often encounters small samples or even a single training sample, and traditional SRC algorithms cannot effectively handle such situations. How to deal with the single training sample is the work to be further studied in the future.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RL conceived and developed the theoretical framework of the manuscript. All authors carried out experiment and data process, drafted the manuscript, and approved the submitted version.

REFERENCES

Al-Khuzaie, F., and Duru, A. D. (2021). Diagnosis of Alzheimer disease using 2D MRI slices by convolutional neural network. *Appl. Bion. Biomech.* 2:6690539. doi: 10.1155/2021/6690539

Altaf, T., Anwar, S. M., Gul, N., Majeed, M. N., and Majid, M. (2018). Multi-class Alzheimer's disease classification using image and clinical features. *Biomed. Signal Process. Control* 43, 64–74. doi: 10.1016/j.bspc.2018.02.019

Chang, H., Zhou, Y., Borowsky, A., Barner, K., Spellman, P., and Parvin, B. (2015). Stacked predictive sparse decomposition for classification of histology sections. *Int. J. Comput. Vis.* 113, 3–18. doi: 10.1007/s11263-014-0790-9

Gao, S. S. (2021). Gray level co-occurrence matrix and extreme learning machine for Alzheimer's disease diagnosis. *Int. J. Cogn. Comput. Eng.* 2, 116–129. doi: 10.1016/j.ijcce.2021.08.002

Gu, X., Zhang, C., and Ni, T. G. (2021). A hierarchical discriminative sparse representation classifier for EEG signal detection. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 18, 1679–1687. doi: 10.1109/TCBB.2020.3006699

He, T. (2019). Image classification based on sparse coding multi-scale spatial latent semantic analysis. *EURASIP J. Image Video Process.* 2019:38. doi: 10.1186/s13640-019-0425-8

Hett, K., Ta, V. T., Manjón, J. V., and Coupé, P. (2018). Adaptive fusion of texture-based grading for Alzheimer's disease classification. *Comput. Med. Imag. Graph.* 70, 8–16. doi: 10.1016/j.compmedimag.2018.08.002

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). FSL. *Neuroimage* 62, 782–790. doi: 10.1016/j.neuroimage.2011.09.015

Jiang, Y., Chen, L., Zhang, H., and Xiao, X. (2019). Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. *PLoS One* 14:e0214587. doi: 10.1371/journal.pone.0214587

Jiang, Z., Lin, Z., and Davis, L. S. (2013). Label consistent K-SVD: learning a discriminative dictionary for recognition. *IEEE Trans. Pattern Analys. Mach. Intell.* 35, 2651–2664. doi: 10.1109/TPAMI.2013.88

Katabathula, S., Wang, Q., and Xu, R. (2021). Predict Alzheimer's disease using hippocampus MRI data: a lightweight 3D deep convolutional network model with visual and global shape representations. *Alzheimers Res. Ther.* 13:104. doi: 10.1186/s13195-021-00837-0

Kim, K. S., Choi, H. H., Moon, C. S., and Mun, C. W. (2011). Comparison of *k*-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of Electromyogram signals based on the wrist-motion directions. *Curr. Appl. Phys.* 11, 740–745. doi: 10.1016/j.cap.2010.11.051

Kruthika, K. R., Maheshappa, H. D., Rajeswari, and Alzheimer's Disease Neuroimaging Initiative (2019). Multistage classifier-based approach for Alzheimer's disease prediction and retrieval. *Inform. Med. Unlocked* 14, 34–42. doi: 10.1016/j.imu.2018.12.003

Lee, S., Lee, H., and Kim, K. W. (2020). Magnetic resonance imaging texture predicts progression to dementia due to Alzheimer disease earlier than hippocampal volume. *J. Psychiatry Neurosci.* 45, 7–14. doi: 10.1503/jpn.180171

Liu, M., Zhang, D., Adeli, E., and Shen, D. G. (2015). Inherent structure-based multiview learning with multitemplate feature representation for Alzheimer's disease diagnosis. *IEEE Trans. Biomed. Eng.* 63, 1473–1482. doi: 10.1109/TBME.2015.2496233

Loddo, A., Buttau, S., and Ruberto, C. D. (2022). Deep learning based pipelines for Alzheimer's disease diagnosis: a comparative study and a novel deep-ensemble method. *Comput. Biol. Med.* 141:105032. doi: 10.1016/j.compbiomed.2021.105032

Ma, H., Gou, J., Wang, X., Ke, J., and Zeng, S. (2017). Sparse coefficient-based *k*-nearest neighbor classification. *IEEE Access* 5, 16618–16634. doi: 10.1109/ACCESS.2017.2739807

Ortiz, E. G., and Becker, B. C. (2014). Face recognition for web-scale datasets. *Comput. Vis. Image Understand.* 118, 153–170. doi: 10.1016/j.cviu.2013.09.004

Pan, X., Adel, M., Fossati, C., Gaidon, T., and Guedj, E. (2019). Multilevel Feature Representation of FDG-PET Brain Images for Diagnosing Alzheimer's Disease. *IEEE J. Biomed. Health Inform.* 23, 1499–1506. doi: 10.1109/JBHI.2018.2857217

Porsteinsson, A. P., Isaacson, R. S., Knox, S., Sabbagh, M. N., and Rubino, I. (2021). Diagnosis of Early Alzheimer's Disease: clinical Practice in 2021. *J. Prevent. Alzheimers Dis.* 8, 371–386. doi: 10.14283/jpad.2021.23

Shi, Y. H., Gao, Y., Yang, Y. B., Zhang, Y., and Wang, D. (2013). Multimodal sparse representation-based classification for lung needle biopsy images. *IEEE Trans. Biomed. Eng.* 60, 2675–2685. doi: 10.1109/TBME.2013.2262099

Tsangaratos, P., and Ilia, I. (2016). Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: the influence of models complexity and training dataset size. *Catena* 145, 164–179. doi: 10.1016/j.catena.2016.06.004

van Oostveen, W. M., and de Lange, E. (2021). Imaging Techniques in Alzheimer's Disease: A Review of Applications in Early Diagnosis and Longitudinal Monitoring. *Int. J. Mol. Sci.* 22:2110. doi: 10.3390/ijms22042110

Venugopalan, J., Tong, L., Hassanzadeh, H. R., and Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci. Rep.* 11:3254. doi: 10.1038/s41598-020-74399-w

Wang, K., Hu, H., and Liu, T. (2018). Discriminative kernel sparse representation *via* *l*2 regularisation for face recognition, Image and vision processing and

display technology. *IET Inst. Eng. Technol.* 54, 1324–1326. doi: 10.1049/el.2018.6727

Wong, W. (2020). Economic burden of alzheimer disease and managed care considerations. *Am. J. Manag. Care* 26, S177–S183. doi: 10.37765/ajmc.2020.88482

Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., and Ma, Y. (2009). Robust face recognition *via* sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 210–227. doi: 10.1109/TPAMI.2008.79

Yang, M., Chang, H., Luo, W., and Yang, J. (2017). Fisher discrimination dictionary pair learning for image classification. *Neurocomputing* 269, 13–20. doi: 10.1016/j.neucom.2016.08.146

Yao, D., Calhoun, V. D., Fu, Z., Du, Y., and Shi, J. (2018). An ensemble learning system for a 4-way classification of Alzheimer's disease and mild cognitive impairment. *J. Neurosci. Methods* 302, 75–81. doi: 10.1016/j.jneumeth.2018.03.008

Zhang, X. F., Dou, H., Ju, T., Xu, J., and Zhang, S. T. (2016). Fusing heterogeneous features from stacked sparse autoencoder for histopathological image analysis. *IEEE J. Biomed. Health Inform.* 20, 1377–1383. doi: 10.1109/JBHI.2015.2461671

Zhang, Z., Ren, J., Jiang, W., Zhang, Z., Hong, R., Yan, S., et al. (2019). Joint subspace recovery and enhanced locality driven robust flexible discriminative dictionary learning. *IEEE Trans. Circuits Syst. Video Technol.* 30, 2430–2446. doi: 10.1109/TCSVT.2019.2923007

Zheng, S., and Ding, C. (2020). A group lasso based sparse KNN classifier. *Pattern Recogn. Lett.* 131, 227–233. doi: 10.1016/j.patrec.2019.12.020

Zhou, G., Lu, B., Hu, X., and Ni, T. (2022). Sparse representation-based discriminative metric learning for brain MRI image retrieval. *Front. Neurosci.* 15:829040. doi: 10.3389/fnins.2021.829040

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 30 May 2022
doi: 10.3389/fnagi.2022.893250

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

Chunting Cai[1], Jiangsheng Cao[1], Chenhui Yang[1] and E. Chen[2]**

[1] School of Informatics, Xiamen University, Xiamen, China, [2] 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 impressible 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[1].

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)[2], which is based on math software MATLAB2012a[3] and Statistical Parametric Mapping software (SPM12)[4]. 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 (FWHW) 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,

[1]http://adni.loni.usc.edu/

[2]http://restfmri.net/forum/RESTplusV1.2

[3]http://www.mathworks.com/products/matlab/

[4]https://www.fil.ion.ucl.ac.uk/spm/software/spm12/

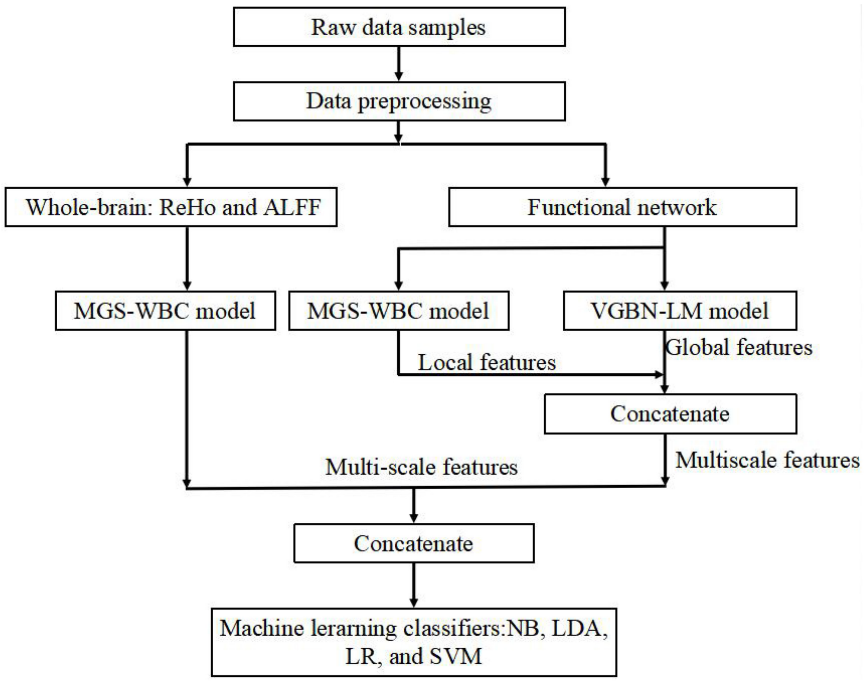


FIGURE 1 | Overall framework of amnesic mild cognitive impairment (aMCI) diagnosis system.

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

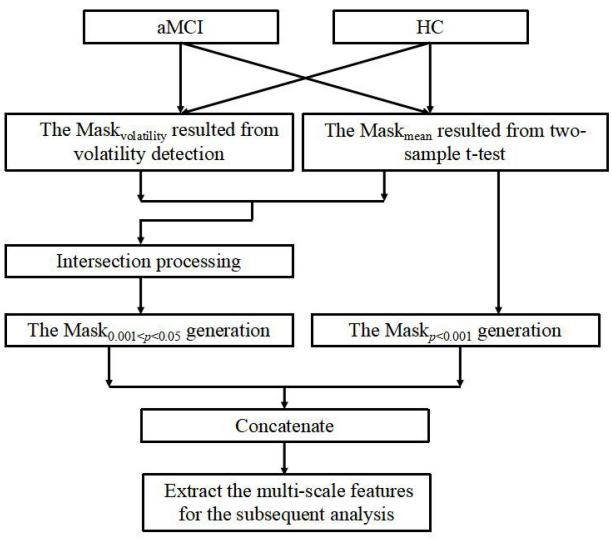


FIGURE 2 | The description of mask generation strategy based on within-class and between-class criterion (MGS-WBC) model.

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 × 90 functional network, which is defined as $\text{Mask}_{volatility}$, 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 ith 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 $x_1, x_2..., x_n$ 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 $y^T Ly$. 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).

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

$$Y_{opt} = \arg\min_{Y^T DY = I} tr(Y^T LY) \quad (5)$$

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

$$Ly = \lambda Dy \quad (6)$$

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 × 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 × 90 functional network in terms of reduced dimension vector.

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

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

where σ 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:

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

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

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

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

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 90×90 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:

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

Significantly, the solution of (15) is reached by use of the backward induction. Let $(\lambda_{\xi_i}, y_{\xi_i})$ represents the solutions of $\begin{cases} y^T Dy = \frac{y^T Wy}{1 - \lambda} \Leftrightarrow Ly = \lambda Dy \\ y^T Dy = \sigma \end{cases}$ *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:

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

where d represents the number of solutions in $Ly = \lambda Dy$. Concurrently, we assume that $Dy = Ly$ have the solutions, defined as y_{β_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:

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

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

$$\begin{aligned} &\arg\min_{y^T Dy = \sigma} y^T Ly = \min(\min \lambda_{\xi_i} \frac{y_{\xi_i}^T Wy_{\xi_i}}{1 - \lambda_{\xi_i}}, \min y_{\beta_j}^T Dy_{\beta_j}) \\ &\qquad , i \in (1, 2, ...d) \ and \ j \in (1, 2, ..., h) \end{aligned} \tag{18}$$

Remarkably, the value of σ 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 $a_1, a_2, a_3, ..., a_m$.

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

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

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

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

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 $D = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}, y \in \{0, 1\}$, 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:

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

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

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

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

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

where u denotes mean value.

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

$$S_b = (u_0 - u_1)(u_0 - u_1)^T \tag{24}$$

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

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

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

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

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:

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

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

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

Then, we take the derivative concerning (27).

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

where m represents the number of data samples, and ηrepresents the regularization coefficient. Finally, the parameters θ_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).

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

The Gaussian kernel is described as follows:

$$K(x, x_i) = \exp(-\frac{||\mathrm{x}\text{-}\mathrm{x}_i||^2}{\sigma^2}) \tag{31}$$

where C represents the penalty coefficient, x_i represents the data sample, y_i represents class corresponding to ith data sample, w and b represent the optimal parameters of the model, N represents the number of data, and ψ_i represents the relaxation variable that corresponds to the ith 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).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{32}$$

$$Precision = \frac{TP}{TP + FP} \tag{33}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{34}$$

$$F1 - Score = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \tag{35}$$

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**.

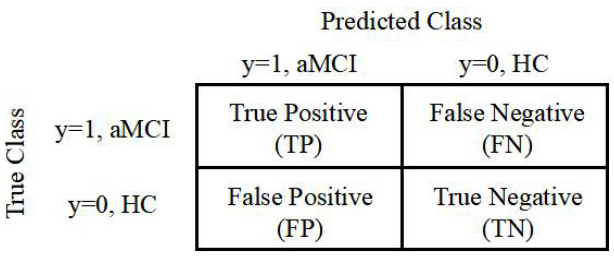


FIGURE 3 | Confusion matrix used to measure the binary classification problems.

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 ≥ 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.

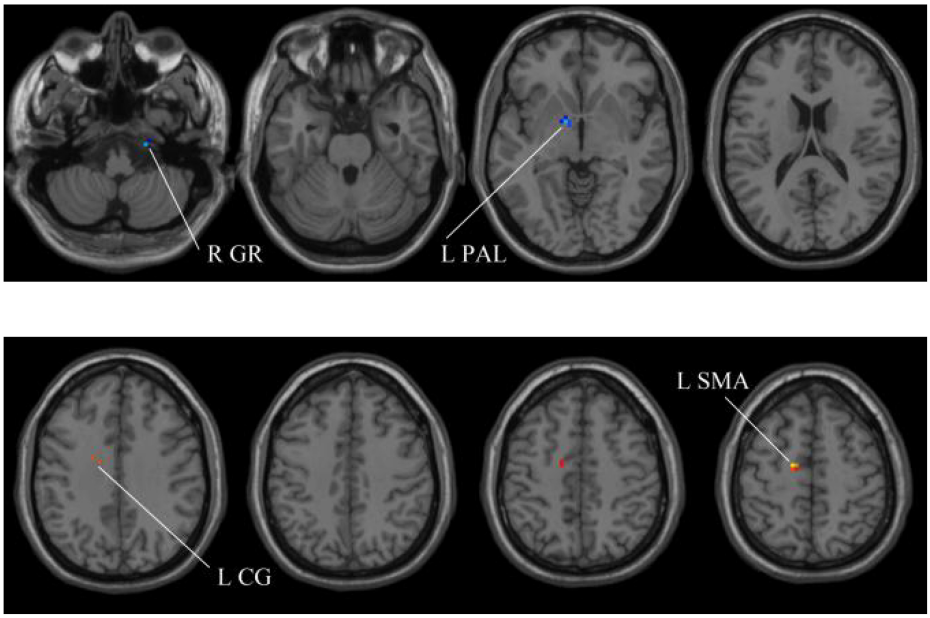


FIGURE 4 | Compared with the HC group, the ReHo in aMCI group exhibits prominent differences based on MGS-WBC[2]. L, left; R, right; GR, gyrus rectus; PAL, pallidum; CG, cingulate gyrus; SMA, supplementary motor area.

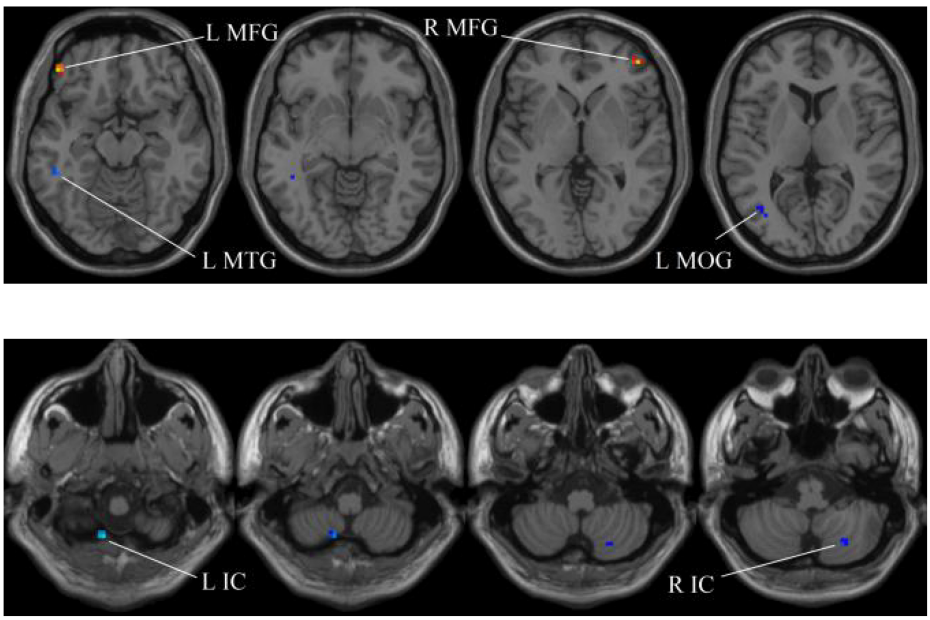


FIGURE 5 | Compared with the HC group, the ReHo in aMCI group exhibits prominent differences based on SSW[1]. L, left; R, right; IC, inferior cerebellum; MTG, middle temporal gyrus; MFG, middle frontal gyrus; MOG, middle occipital gyrus.

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[2] all exceed 66.67% in the four classifiers. Besides, compared with the single significant threshold in the whole-brain structures (SSW[1]) (Zhang et al., 2015; Yang et al., 2018), the results of MGS-WBC[3] 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[1]. 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[1]	0.8358	0.8358	0.9189
	MGS-WBC[2]	0.6866	0.6667	0.7460
	MGS-WBC[3]	0.8806	0.8788	0.9412
LDA	SSW[1]	0.8657	0.8732	0.9412
	MGS-WBC[2]	0.7164	0.6885	0.7799
	MGS-WBC[3]	0.8955	0.8955	0.9608
LR	SSW[1]	0.8060	0.8169	0.9135
	MGS-WBC[2]	0.7164	0.7077	0.7861
	MGS-WBC[3]	0.9104	0.9118	0.9688
SVM	SSW[1]	0.8657	0.8696	0.9421
	MGS-WBC[2]	0.7612	0.7419	0.7709
	MGS-WBC[3]	0.9104	0.9091	0.9572

SSW[1] 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[2] refers to the combination of volatility detection and significant threshold ($0.001 < p < 0.05$) in the MGS-WBC model. MGS-WBC[3] 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 ≥ 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[1].

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[1]. As shown in **Table 4**, the results of MGS-WBC[2] or MGS-WBC[3] 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[2]	0.7313	0.7273	0.7772
LR	/MGS-WBC[3]	0.7164	0.6984	0.7718
SVM		0.6716	0.6071	0.7522

MGS-WBC[2] refers to the combination of volatility detection and significant threshold ($0.001 < p < 0.05$) in MGS-WBC model. MGS-WBC[3] 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[1]. 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[2], 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[1]. 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.

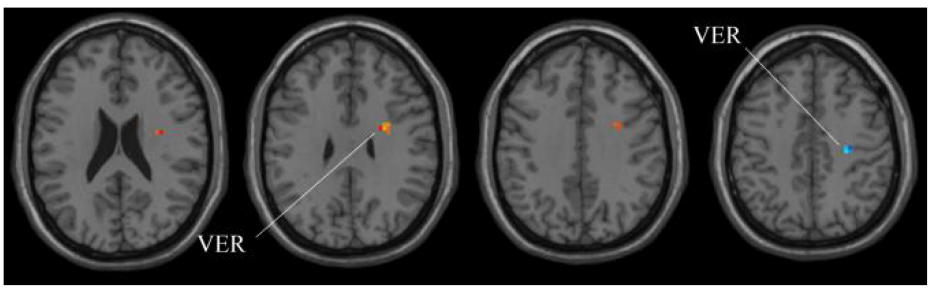


FIGURE 6 | Compared with HC group, the ALFF in aMCI group exhibits prominent differences based on volatility detection in the MGS-WBC model. VER, Vermis.

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

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

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

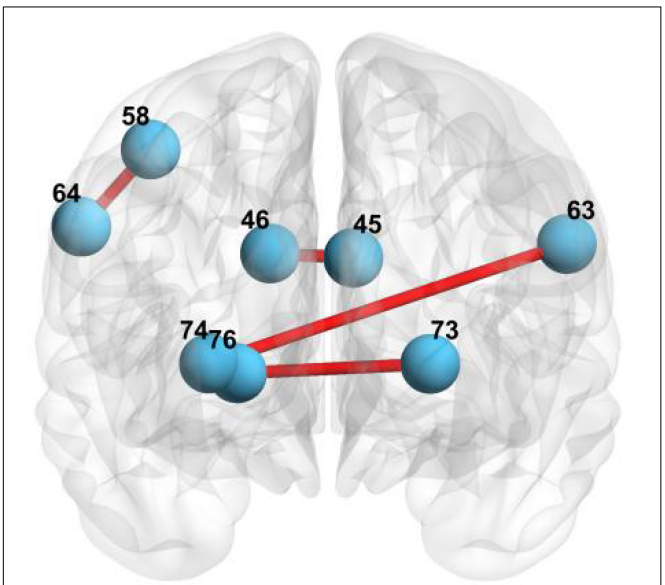


FIGURE 7 | The brain regions involved in local feature selection: 45 represents the left cuneus; 46 represents right cuneus; 73 represents left putamen; 76 represents right pallidum; 74 represents right putamen; 64 represents right supramarginal gyrus; 58 represents postcentral gyrus; and 63 represents 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[1]. When the features generated from MGS-WBC[2] 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[3] compared with SSF[1]. 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[1]	0.4478	0.3934	0.5116
	VGBN-LM	0.6418	0.6757	0.6996
LDA	VGBN-LM[1]	0.5672	0.5246	0.3137
	VGBN-LM	0.6269	0.6377	0.6185
LR	VGBN-LM[1]	0.5224	0.5000	0.5330
	VGBN-LM	0.6567	0.6567	0.6979
SVM	VGBN-LM[1]	0.4627	0.4375	0.3966
	VGBN-LM	0.6119	0.6286	0.6729

VGBN-LM[1] 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[1]. Also, compared to VGBN-LM[1], 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[1]. 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[1] and MGS-WBC[2]. As summarized in **Table 8**, using MGS-VGBN[1], each of NB, LDA,

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

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

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

TABLE 9 | The fused features analysis.

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

$SSWF^1$ 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-bran structures and 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 $MGS\text{-}VGBN^2$ 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^1. 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^1. 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^1.

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^1, 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 ($SSWF^1$) (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 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 ($SSWF^1$) (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. 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.

REFERENCES

Association, A. S. (2018). 2018 Alzheimer's disease facts and figures. *Alzheimer's Dement.* 14, 367–429. doi: 10.21926/obm.geriatr.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: 10.1162/089976604173 2396

Bian, Z., Li, Q., Wang, L., Lu, C., Yin, S., and Li, X. (2014). Relative power and coherence of EEG series are related to amnestic mild cognitive impairment in diabetes. *Front. Aging Neurosci.* 6:11. doi: 10.3389/fnagi.2014.0 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

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., Gossmann, 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. Rehabilit. 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 Amnestic 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 amnestic 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 Amnestic 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 amnestic 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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 30 May 2022
doi: 10.3389/fnagi.2022.883542

Study on Low-Frequency Repetitive Transcranial Magnetic Stimulation Improves Speech Function and Mechanism in Patients With Non-fluent Aphasia After Stroke

Guangtao Bai[1†], Liang Jiang[2†], Sai Huan[3], Pingping Meng[1], Yuyang Wang[1], Xiaona Pan[1], Shuai Yin[1], Yuyang Zhao[1] and Qiang Wang[1]*

[1] Department of Rehabilitation Medicine, The Affiliated Hospital of Qingdao University, Qingdao, China, [2] Department of Otolaryngology, Qingdao Women's and Children's Hospital, Qingdao, China, [3] Department of Rehabilitation Medicine, Qingdao Women's and Children's Hospital, Qingdao, China

OPEN ACCESS

Edited by:
Yuanpeng Zhang,
Nantong University, China

Reviewed by:
Haijie Liu,
Capital Medical University, China
Hongru Zhao,
Soochow University, China

***Correspondence:**
Qiang Wang
wangqiang1964@qdu.edu.cn

[†] These authors have contributed equally to this work

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

Received: 25 February 2022
Accepted: 14 April 2022
Published: 30 May 2022

Citation:
Bai G, Jiang L, Huan S, Meng P, Wang Y, Pan X, Yin S, Zhao Y and Wang Q (2022) Study on Low-Frequency Repetitive Transcranial Magnetic Stimulation Improves Speech Function and Mechanism in Patients With Non-fluent Aphasia After Stroke. Front. Aging Neurosci. 14:883542. doi: 10.3389/fnagi.2022.883542

Objective: To explore the therapeutic effect and mechanism of low-frequency repetitive transcranial magnetic stimulation on the speech function of patients with non-fluent aphasia after stroke.

Methods: According to the inclusion and exclusion criteria, 60 patients with post-stroke non-fluent aphasia were included and randomly divided into treatment group (rTMS group) and sham stimulation group (S-rTMS group). Patients in rTMS group were given low-frequency rTMS + ST training. Patients in the S-rTMS group were given sham low-frequency rTMS + ST training. Once a day, 5 days a week, for a total of 4 weeks. The Western Aphasia Battery and the short-form Token test were used to evaluate the language function of the patients in the two groups before and after treatment. Part of the enrolled patients were subjected to functional magnetic resonance imaging examination, and the morning fasting venous blood of the enrolled patients was drawn before and after treatment to determine the content of BDNF and TNF-α.

Results: In the comparison before and after treatment within the group, all dimensions of the WAB scale of the patients in the rTMS group increased significantly. Only two dimensions of the WAB scale of the patients in the S-rTMS group improved significantly after treatment. The results of the short-form Token test showed that patients in the rTMS group improved significantly before and after treatment. The resting state functional magnetic resonance imaging of the two groups of patients before and after treatment showed: the activation of multiple brain regions in the left hemisphere of the rTMS group increased compared with the control group. The serum BDNF content of the patients in the rTMS group was significantly higher than that of the patients in the S-rTMS group after treatment.

Conclusion: Low-frequency rTMS combined with conventional speech training can significantly improve the speech function of patients with non-fluent aphasia after stroke.

Keywords: repetitive transcranial magnetic stimulation, aphasia, functional magnetic resonance, BDNF, TNF-α

HIGHLIGHTS

- Low-frequency rTMS can improve the expression and other language functions of patients with non-fluent aphasia after stroke.
- Low-frequency rTMS can promote brain plasticity changes in patients with non-fluent aphasia after stroke.
- Low-frequency rTMS can promote the secretion of BDNF by the central nervous system of stroke patients.

INTRODUCTION

Aphasia refers to a type of language disorder syndrome in which organic brain diseases are caused by various reasons, which cause damage to related brain areas that dominate brain language expression and listening comprehension, so that patients cannot perform normal speech expression and understand the other party's words. It is very common in patients with cerebrovascular disease. According to research statistics, the incidence of aphasia in stroke patients is about 20–40% (Menichelli et al., 2019).

Aphasia Recovery Mechanism

Regarding the mechanism of aphasia recovery, when the language hub of the dominant hemisphere is damaged in the acute phase, its inhibition of the surrounding brain areas will be weakened, which promotes the activation of the brain areas around the damaged brain area and the functional reconstruction of plasticity, and promotes the recovery of the patient's language function. In the subacute phase, the mirror brain area of the language hub of the right hemisphere is activated due to the weakening of the inhibition of the dominant hemisphere, which is beneficial to the recovery of the function of patients with aphasia to a certain extent. In the chronic recovery period, as the function of the dominant hemisphere on the left side of the brain gradually recovers, its activation level gradually increases during language training, and the inhibition to the right hemisphere gradually increases. At the same time, the activation level of the right hemisphere gradually decreased the language hub gradually returns to the left dominant hemisphere. Therefore, in the chronic phase, in order to reduce the inhibitory effect of the non-dominant hemisphere on the dominant hemisphere, it is necessary to inhibit the corresponding brain areas of the non-dominant hemisphere, and at the same time, it can excite the language hub in the dominant hemisphere and promote the recovery of the language function of the patients. Also in clinical practice had showed that cortical stimulation could facilitate functional improvement (Zhang J. et al., 2021).

Application of rTMS in Aphasia

Repetitive transcranial magnetic stimulation technology is one of the main representatives of non-invasive brain stimulation technology that has emerged in recent years. It not only has a temporary inhibitory or excitatory effect on the cerebral cortex, but also has a long-term plasticity change effect. A large number of research results affirm its efficacy in the treatment of aphasia (Rossetti et al., 2019), but the specific mechanism of action is still unclear. Some scholars use the method of functional magnetic resonance to explore the specific mechanism of the rTMS by the specific activated/inhibited brain regions, but the conclusions are very different (Szaflarski et al., 2018; Arheix-Parras et al., 2021; Fahmy and Elshebawy, 2021; Neri et al., 2021). In addition, studies have also found that: after rTMS treatment, the levels of brain-derived neurotrophic factor in peripheral blood of patients with depression was higher than before, which may be one of the mechanisms of rTMS (Zhao et al., 2019).

Research Purposes

In this study, low-frequency repetitive transcranial magnetic stimulation was applied to the posterior inferior frontal gyrus of the right cerebral hemisphere in patients with non-fluent aphasia after stroke. Clarify its therapeutic effect on the language function of patients with aphasia, and some patients were enrolled in the rest state functional magnetic resonance scan before and after treatment, using low-frequency amplitude score, degree centrality method to statistically analyze the scanned image data, to identify specific activated or inhibited brain regions, and combined the method of functional connection to explore the plasticity changes of specific brain regions. At the same time, before the start of treatment and after the end of the treatment course, the early morning venous blood of the enrolled patients was collected to determine the content of BDNF, and to explore the treatment mechanism of rTMS in patients with non-fluid aphasia after stroke from the perspective of cytokines, providing clinical and theoretical support for the clinical treatment of aphasia.

PATIENTS AND METHODS

Research Object

According to the inclusion and exclusion criteria, 60 patients with post-stroke aphasia who were hospitalized in the Rehabilitation Medicine Department of Qingdao University Affiliated Hospital from 2017-12 to 2019-10 were randomly divided into treatment group (rTMS group) and control group (S-rTMS group). This study was reviewed by the ethics committee of the Affiliated Hospital of Qingdao University (qyfykyll 2018-23). Written informed consent was obtained from the individual for the publication of any potentially identifiable images or data included in this article.

Inclusion Criteria

(1) Clinical compliance with the criteria of "Diagnosis Essentials for Various Cerebrovascular Diseases" formulated by the Fourth National Cerebrovascular Disease Conference of the Chinese Medical Association in 1995; CT or MRI confirmed as the first stroke in the left hemisphere (dominant hemisphere); (2) Right Handy (standardized measurement), normal language function before onset; (3) The course of illness is about 2 weeks to 6 months after stroke; (4) Western Aphasia Battery (WAB) aphasia quotient (AQ) < 93.8, non-Fluent aphasia, with a score of 0–4 for speech fluency; (5) Chinese is the first language, and the education level above elementary

school can cooperate to complete the assessment; (6) No epilepsy, severe heart disease, severe physical disease; (7) Clear mind, cooperative physical examination, and orientation Complete, without obvious memory impairment and intellectual impairment; (8) Able to independently maintain a sitting position for more than 30 min; (9) Patients and family members sign informed consent.

Exclusion Criteria

(1) Complicated with other neurodegenerative diseases, such as speech disorder caused by Parkinson's disease, dementia, etc.; (2) Auditory or visual defects may affect assessment and treatment; (3) Application of drugs that change the excitability of the cerebral cortex (antiepileptic drugs), sleeping pills, benzodiazepines, etc.; (4) Combined with epilepsy, severe heart, liver, kidney dysfunction or other serious physical diseases; (5) Unconscious and unable to cooperate with examination and treatment; (6) A history of mental abnormalities; (7) According to safety guidelines, there are contraindications to rTMS and MRI, such as metal foreign bodies in the body or other electronic devices implanted in the body.

Among them, 60 patients completed the initial evaluation of the WAB scale and the short-form Token test. The patients in the rTMS group completed the entire experimental process and completed the final evaluation of the WAB scale and the short-form Token test. One patient in the S-rTMS group was converted due to recurrence of cerebral hemorrhage. He was admitted to neurosurgery for treatment, and another patient was discharged from the trial early due to personal reasons. Therefore, only 28 patients completed the final evaluation of the WAB scale and the short-form Token test. In order to explore the specific mechanism of rTMS, we performed resting functional magnetic resonance scans on some patients before and after treatment, and collected the peripheral blood of the patients, and measured the changes of BDNF and TNF-α in their peripheral serum. The general information of the enrolled patients is shown in the following table (**Table 1**).

Research Methods

Apparatus

All experiments were completed in the Department of Rehabilitation Medicine and the Central Laboratory of the Affiliated Hospital of Qingdao University.

TABLE 1 | General information of the enrolled patients.

	rTMS group	S-rTMS group
Number of cases	30	30
Clinical scale measurement (pre/post)	30/30	30/28
Serum factor determination (pre/post)	30/28	30/24
fMRI (pre/post)	16/13	12/10
Gender: (Male: Female)	17:13	14:16
Course of disease (months, $\overline{\chi} \pm s$)	3.27 ± 1.50	3.75 ± 1.67
Age (years, $\overline{\chi} \pm s$)	63.47 ± 7.81	59.91 ± 8.58
Aphasia quotient-pre	28.16 ± 22.86	22.76 ± 18.81
Stroke type (hemorrhagic: infarct: mixed) group	13: 15: 2	16: 13: 1

Reagents

Human brain-derived neurotrophic factor (BDNF) enzyme-linked immunoassay (ELISA) kit: enzyme-linked, CK-E12065 human tumor necrosis factor alpha (TNF-α) enzyme-linked immunoassay (ELISA) kit: mlbio, ml077385.

Consumables

1.5 ml centrifuge tube (American Axygen), each volume tip (American Axygen).

Equipment

Magnetic field stimulator: Wuhan Yiruide Company CCY-IA type.

Electric heating constant temperature blast drying oven: Shanghai Senxin DGG-9140B.

High-speed refrigerated centrifuge: Thermo Scientific.

Pipette: Eppendorf.

Microplate reader: SPECTCA MAX190 (Molecular Company, United States).

rTMS Treatment Method

Measurement of Motor Threshold

Select the contralateral abductor pollicis brevis muscle as the measuring muscle, and place the recording electrode on the muscle abdomen of the muscle and the reference electrode on the first joint of the thumb of the ipsilateral upper limb. The stimulation coil stimulates the patient's right brain, gradually adjust the position of the stimulation coil, determine the most suitable stimulation position and stimulation intensity (at this time the incubation period is shortest, and the amplitude is the largest), and gradually adjust the output intensity to find out 10 consecutive stimulations that trigger the contralateral thumb short The stimulus intensity that the abductor motor evoked potential appears at least 5 times and the amplitude is not less than 50 μV is the motor threshold.

Stimulation Site

select the patient's non-dominant hemisphere (right hemisphere) at the back of the inferior frontal gyrus as the stimulation site, place the stimulation coil close to the surface of the patient's skull and place it tangentially, the center point of the "8" coil is placed at the mark, and the stimulation coil The handle points vertically to the patient's back occiput. The body surface positioning method is selected according to the electrode positioning map calibrated by the International Electroencephalography Society. Before and after treatment, WAB scale assessment and resting functional magnetic resonance scan were performed on the two groups of patients. The stimulus parameters and stimulus parts of the sham stimulation group were the same as the treatment group, but the stimulation coil was perpendicular to the surface of the skull.

Stimulation Parameters

Set 80% of the motor threshold as the stimulus intensity, the stimulus frequency is 1 Hz, 10 pulses are a sequence, the sequence interval is 2 s, 100 sequences per day (total 1,000 pulses in total), treatment for 5 days a week. The total course of treatment is 4 weeks.

Routine Speech Training and Language Function Scale Assessment

Conventional speech training is conducted by speech therapists including Schuell training method, blocking removal method, de-inhibition method, program training method and other methods to conduct one-to-one speech training for patients, and appropriately combine computer picture naming training, etc. Each training time is about 30 min. The Western aphasia battery (WAB) Chinese version scale and the short-form Token test were used to evaluate the two groups of patients before and after treatment, and the evaluation results were summarized according to each dimension.

Functional Magnetic Resonance Data and Processing Methods

Functional Magnetic Resonance Parameter Setting

Use resting functional magnetic resonance (Rs-fMRI) scan for all subjects. The scan parameters are: TR 2,000 ms, TE 30 ms, slice thickness 5.0 mm, no interval, visual The field angle is 240 mm × 240 mm, and the matrix is 960 mm × 960 mm. The imaging range covers the whole brain as much as possible. There are 25 layers from the base of the skull to the parietal lobe, with 279 frames in each layer, and a total of 6,975 images are collected. The acquisition time is 558 s. During the examination, the patient is required to avoid any purposeful thinking activities as much as possible, lie supine on the examination bed with eyes closed, breathe calmly, and keep consciousness. Start scanning after the patient adapts to the magnet and surrounding environment.

Image Preprocessing

based on Matlab R2017b platform for preprocessing, and then use DPABI v4.0[1] and SPM12 software to process the image data. The processing steps are as follows: format conversion, time layer correction, head movement correction, spatial standardization, de-linear shift, regression covariate, etc. Select the 0.010 ∼ 0.027 Hz (slow5) sub-band to process and analyze the image.

Fractional Amplitude of Low-Frequency Fluctuation

The fALFF value is the ratio of the sum of the amplitude of the preset frequency band to the sum of the amplitude of the whole frequency band, and then normalize the whole brain voxels, that is, divide by the mean value of whole brain f ALFF to get mfALFF, then Gaussian smoothing (FWHM is 4 mm × 4 mm × 4 mm), you can get the smfALFF result.

Degree Centrality Analysis

Each voxel is a node, and the connection between the voxel and the voxel is called an edge. Calculate the Pearson's correlation coefficient between any two voxels (nodes) with obvious functional connection in the brain function connection group, according to the threshold level of $r > 0.25$, you can get a (number of voxels) × (number of voxels) undirected adjacency correlation matrix, get the weighted DC value, and then divide it with the whole brain DC mean, that is, complete the data standardization process, and then perform Gaussian smoothing for statistical analysis between groups.

Functional Connectivity Analysis

Select several speech-related brain regions of interest (ROI) based on previous research at home and abroad, calculate the average time series of each ROI, and then perform pairwise analysis of the above ROI Pearson correlation calculation analysis between, obtain the correlation coefficient between any two ROIs, thus get the correlation matrix, and then normalize it, and then enter the next step of processing and analysis.

Serum Processing and Storage Methods

Collect 3 ml of early morning venous blood from the enrolled patients before the treatment and after the end of the treatment course. After centrifugation at room temperature for 5 min, store the centrifuged serum in a refrigerator at −80°C for future reference. If precipitation occurs during storage, it needs to be centrifuged again. Detection steps: balance reagents, prepare reagents, add samples, develop color, terminate the reaction, determine optical density (OD value), use ELISA calc software to calculate serum factor content, etc.

Statistical Analysis

The scores of the various dimensions of the WAB scale and the short-form Token test score data are analyzed using SPSS19.0 statistical software package. The measurement data obtained in this experiment is expressed as ($\overline{\chi} \pm s$), and the measurement data within and between groups are compared using single Factor analysis of variance, count data using chi-square test, pairwise

[1] http://rfmri.org/dpabi

TABLE 2 | Evaluation results and statistical analysis of the clinical scale for the two groups of patients.

	rTMS (30)		S-rTMS (28)	
	Pre	Post	Pre	Post
Spontaneous language	5.10 ± 5.03	9.93 ± 5.24*	3.73 ± 4.22	6.68 ± 4.46*#
Listening comprehension	3.51 ± 2.50	5.86 ± 2.79*	3.13 ± 2.34	4.69 ± 2.46
Repetition	3.21 ± 2.83	5.56 ± 3.04*	2.69 ± 2.52	3.97 ± 2.91
Naming	2.26 ± 2.41	4.96 ± 2.49*	1.86 ± 2.11	3.39 ± 2.17*#
Aphasia quotient	28.16 ± 22.86	52.62 ± 25.02*	22.76 ± 18.81	37.46 ± 20.51*#
Short-form Token Test	12.77 ± 8.41	20.90 ± 9.97*	11.17 ± 8.13	17.57 ± 9.19*

**There is a statistical difference in the comparison before and after treatment within the group, $P < 0.05$.*
#Comparison of the two groups after treatment between the groups, there is a statistical difference, $P < 0.05$.

correlation analysis using linear correlation analysis, $P < 0.05$ indicates that the difference is statistically significant. Use DPABI v4.0 software to perform two-sample *t*-test on the obtained slow5 band fALFF, DC, and FC image data in the rTMS group (after treatment-before treatment) and S-rTMS group (after treatment-before treatment), and use GRF correction to perform multiple Comparative correction, the threshold is individual level $P < 0.05$, clump level $P < 0.05$. Serum BDNF and TNF-α levels were analyzed using SPSS19.0 statistical software package. The measurement data obtained in this experiment are expressed as

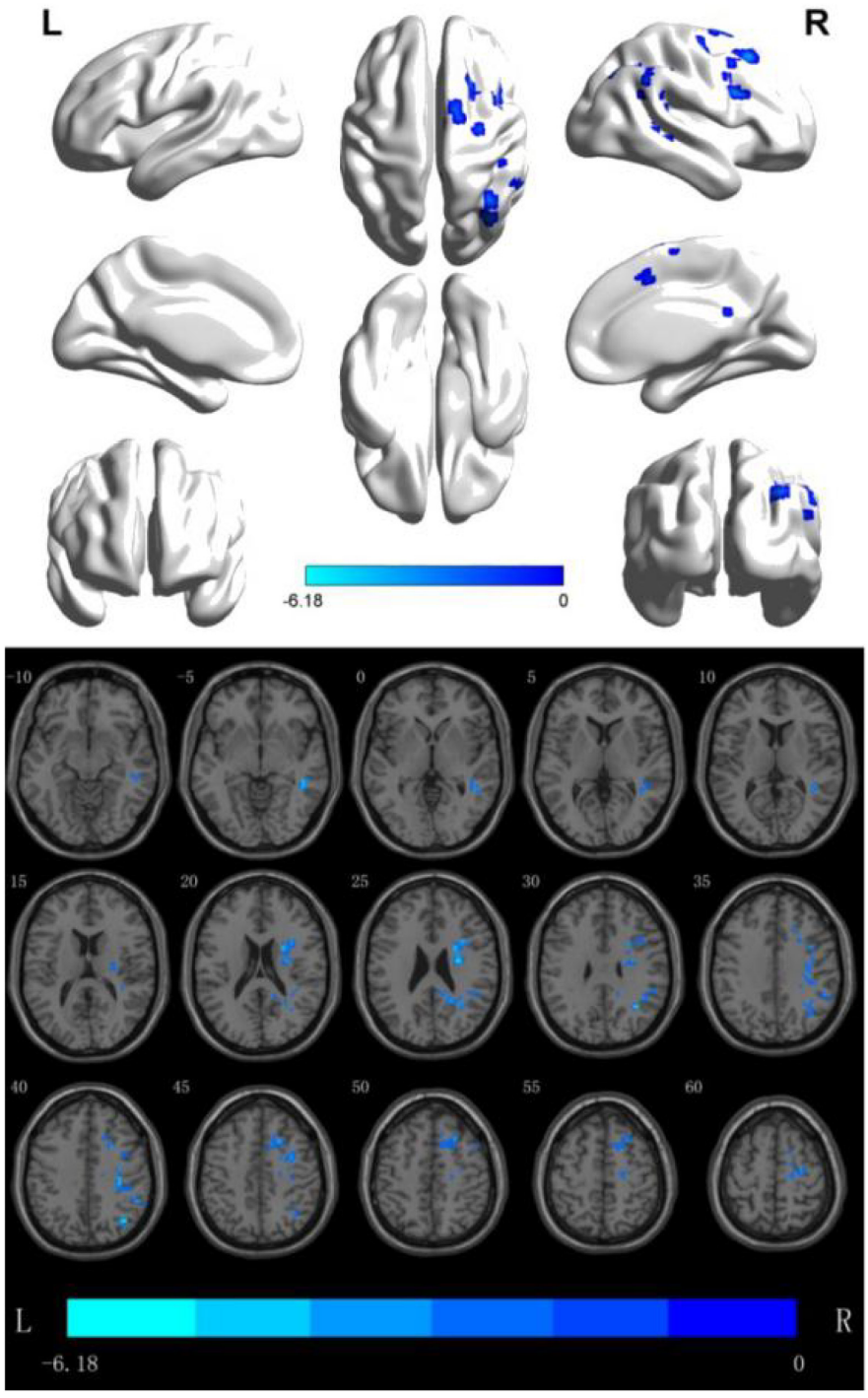


FIGURE 1 | fALFF analysis of the difference between the two groups before and after treatment. The blue area is the brain area where activation is inhibited (threshold: individual level $P < 0.05$, clump level $P < 0.05$).

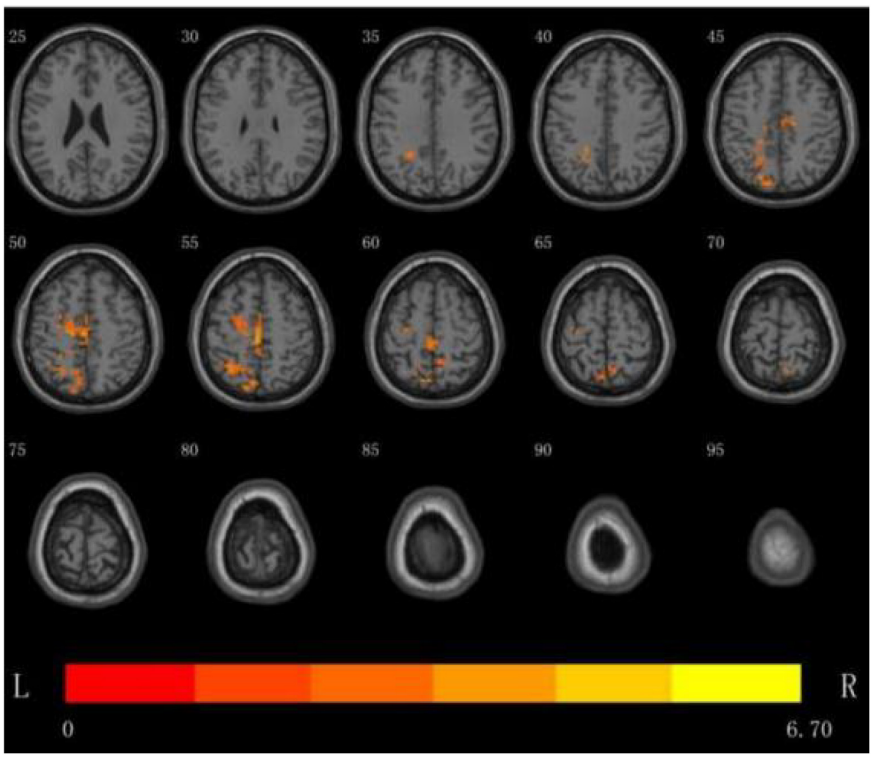


FIGURE 2 | DC analysis of the distribution of brain regions with significant differences between the two groups before and after treatment, the yellow area is the activated brain area (threshold: individual level $P < 0.05$, clump level $P < 0.05$).

$\overline{\chi} \pm s$. The measurement data within and between groups are compared by single-factor analysis of variance, $P < 0.05$ indicates a statistical difference.

RESULTS

(1) The effect of low-frequency rTMS stimulation on the Broca mirror area in the right inferior frontal gyrus on the dimensions of the WAB scale and the short-form Token test scores in patients with non-fluent aphasia after stroke: both the scores of each dimension of the WAB scale in the rTMS group before and after treatment and the short-form Token test scores were significantly improved ($P < 0.05$), while the WAB scale of patients in the S-rTMS group only had three dimensions of spontaneous language, naming, and aphasia quotient, and the short-form Token test scores were significantly improved ($P < 0.05$). After treatment, the scores of the two groups of patients were only statistically different in the three dimensions of WAB scale, spontaneous language, naming, and aphasia quotient ($P < 0.05$), see **Table 2** for details.

(2) fALFF analyzes the brain regions where the difference between the two groups is more meaningful: through data analysis, it can be seen that in the Slow5 subband, there are two Clusters with statistical differences, The fALFF value of multiple brain regions of the patients in the rTMS group was decreased than that of the patients in the S- rTMS group, such as the right dorsolateral superior frontal gyrus, right supplementary motor area, right inferior frontal gyrus pars opercularis (voxel 56, MNI $X = 36$, $Y = -39$, $Z = 15$, $T = -4.76$, $P < 0.05$), right Brodmann area 8, right angular gyrus, right supramarginal gyrus, and right middle temporal gyrus (voxel 19, MNI $X = 27$, $Y = -9$, $Z = 24$, $T = -5.37$, $P < 0.05$) indicating that the activation of the above brain regions in the rTMS group was suppressed than that of the patients in the S-rTMS group. See **Figure 1** for details.

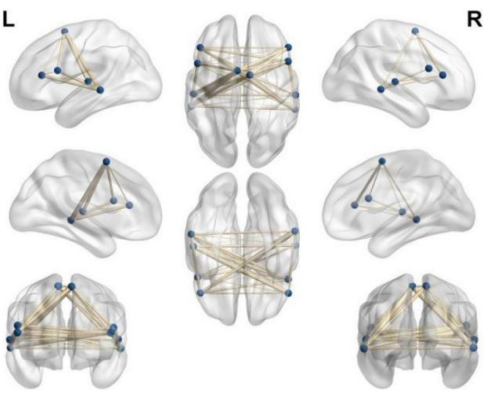


FIGURE 3 | Select ROI function connection pairwise comparison function connection diagram ($p < 0.05$).

(3) DC analysis of the brain regions where the difference between the two groups is more meaningful: the results show that the DC value of multiple brain regions of the patients in the rTMS group was enhanced than that of the patients in the S-rTMS group, such as the left parietal lobe [superior parietal lobule (voxel 78, MNI $X = -12$, $Y = -81$, $Z = 48$, $T = 4.74$, $P < 0.05$), angular gyrus)], left frontal lobe [BA6 area, middle frontal gyrus, superior frontal gyrus, supplementary motor area (voxel 35, MNI $X = -3$, $Y = -24$, $Z = 57$, $T = 6.70$, $P < 0.05$), paracentral lobule], bilateral Limbic lobe (cingulum gyrus) indicating that the activation of the above brain regions in the rTMS group was significantly higher than that of the patients in the S-rTMS group. See **Figure 2** for details.

(4) On the basis of the previous image processing, according to the pre-selected multiple ROIs related to the language function, the pairwise function connection analysis is performed on the basis of the difference between the two groups, and the *t*-test is performed. The result shows: between the left frontal lobe (supplementary motor area) (voxel 35, MNI $X = -3$, $Y = -24$, $Z = 57$, $T = 6.70$, $P < 0.05$) and the right temporal lobe (middle temporal gyrus) (voxel 19, MNI $X = 27$, $Y = -9$, Z = 24, $T = -5.37$, $P < 0.05$) became stronger ($P < 0.05$), indicating that the connection between the two hemispheres of the patients in the rTMS group was strengthened as shown in **Figure 3**.

(5) Changes in the serum BDNF content of the two groups of patients before the treatment and after the end of the treatment course: the serum BDNF content (pg/ml) of the patients in the rTMS group increased from 35.34 to 42.09 ($P < 0.05$), while the serum BDNF of the patients in the S-rTMS group the content (pg/ml) increased from 31.24 to 34.76 ($P > 0.05$). There was no significant difference in serum BDNF content between the two groups before treatment ($P > 0.05$). After the treatment, the serum BDNF content of patients in the rTMS group was significantly higher than that of the patients in the S-rTMS group ($P < 0.05$) and the specific results are shown in **Table 3**.

DISCUSSION

Aphasia is an organic brain disease caused by various reasons, causing damage to the related brain areas that dominate brain language expression and listening comprehension, which leads to a language disorder syndrome with abnormal speech expression and abnormal listening comprehension. Among the various diseases of brain damage, stroke is the most common. According to research statistics, about 30% of stroke patients are accompanied by aphasia (Menichelli et al., 2019).

As a new non-invasive technology that directly acts on the cerebral cortex, rTMS has been proven by many studies to treat patients with aphasia after stroke, but its specific mechanism is still controversial.

In this study, patients with non-fluent aphasia after stroke were treated with low-frequency rTMS for 4 consecutive weeks. The WAB scale and short-form Token test were used to evaluate the aphasia of patients before and after the treatment. The results of the study found that the low-frequency rTMS on the posterior inferior frontal gyrus of the right hemisphere combined with conventional speech training compared with false low-frequency rTMS combined with conventional speech training, the patient's naming, spontaneous language and other expression skills have been significantly improved.

Low-frequency rTMS combined with speech training can significantly improve the expression ability of patients with aphasia. The results of this study are consistent with the conclusions of previous studies. Some scholars use low-frequency rTMS as a single treatment method. After a short-term treatment is given to the patient's non-dominant hemisphere inferior frontal gyrus (the mirror area of the Broca area) for a short period of time, it is found that this single stimulation can improve the accuracy of patient naming, and the patient's reaction time will be significantly shortened (Terao and Ugawa, 2002).

Harvey et al. (2019) used another continuous theta pulse magnetic stimulation (similar to low-frequency rTMS stimulation) to act on the posterior part of the right inferior frontal gyrus of patients with chronic aphasia. After treatment, they found that the patient's picture naming ability was significantly improved, indicating that this treatment plan is beneficial to improve the patient's naming ability (Harvey et al., 2019). Some scholars also use low-frequency rTMS to stimulate the posterior part of the inferior frontal gyrus of the non-dominant hemisphere for 10 times. The results show that this treatment can significantly improve the patient's language fluency (Lopez-Romero et al., 2019). Some scholars have also combined low-frequency rTMS on the posterior part of the right inferior frontal gyrus with speech training. After 2 weeks of treatment, the language fluency of patients with aphasia in the treatment group has improved greatly compared with the control group (Haghighi et al., 2017).

Our study combined low-frequency rTMS with conventional speech training and found that its therapeutic effect was significantly better than that of simple speech training. There are many similar studies. Yoon et al. (2015) combined low-frequency rTMS therapy with speech therapy to explore the combination of the two and the therapeutic effect. The treatment course was 4 weeks. The Korean version of the WAB scale was used to evaluate the two groups of patients before and after treatment. It was found that the naming and retelling ability scores of the patients after low-frequency rTMS stimulation increased

TABLE 3 | Comparison statistical scores of serum BDNF levels before and after treatment in the two groups.

	Number pre/post	Pre-BDNF (pg/ml)	Post-BDNF (pg/ml)	*P*-value in group
rTMS group	30/28	35.34 ± 5.53	42.09 ± 9.16	0.0116
S-rTMS group	30/24	31.14 ± 9.27	34.76 ± 8.79	0.4326
P-value between groups		0.2541	0.0077	

significantly, and rTMS combined with speech training can be used as a treatment for brain Effective treatment for patients with non-fluent aphasia after stroke (Yoon et al., 2015).

In order to explore the specific mechanism of low-frequency rTMS on patients with aphasia, we performed resting functional magnetic resonance examinations on some patients before and after treatment, and used different analysis methods to perform statistical analysis on the obtained image data. First, use the fractional Amplitude of Low-frequency Fluctuation (fALFF) method to analyze. This method is to standardize the ALFF at the individual level. After the standardization at the individual level, the shortcomings of ALFF can be effectively avoided, and the sensitivity and specificity of detection can be improved. Some scholars (Liu, 2016) used functional magnetic resonance to study the brain plasticity of patients with aphasia and found that rehabilitation training can increase the ALFF value of the temporal lobe of the left cerebral hemisphere and the right cerebellum, suggesting that the above brain areas are in the recovery process of aphasia play an important role.

Our study found that the fALFF value of multiple brain regions in the right hemisphere frontal lobe (right inferior frontal gyrus pars opercularis, right supplementary motor area, etc.) decreased. It shows that the activation of the above brain regions was significantly inhibited in the rTMS group of patients. The reason for the analysis is that low-frequency rTMS treatment on the mirror area of the Broca area of the right regions can inhibit the activation of this area, showing the above-mentioned hypoperfusion in the brain regions, and the local blood oxygen is disproportionately reduced due to the decreased oxygen consumption of neurons. Deoxyhemoglobin (paramagnetic) is relatively increased, so it shows a weakened signal. This shows that low-frequency rTMS can indeed significantly inhibit the ROI of the target area, thereby reducing the activation of the right brain area, reducing its inhibition of the Broca area of the dominant hemisphere through the corpus callosum, promoting the activation of the Broca area of the dominant hemisphere, and improve the speech expression functions of patients with aphasia by promote the brain plasticity. Some scholars have found that the degree of language impairment in patients with aphasia is positively correlated with the Pearson correlation test of the right middle frontal gyrus (Zhu et al., 2014), which is consistent with our study that the activation of the right frontal lobe decreases and the activation of the left inferior frontal gyrus increases during the recovery period of speech function.

At the same time, the activation of the right temporal lobe (middle temporal gyrus) and right parietal lobe (corner gyrus, right superior marginal gyrus) and other brain regions in the rTMS group was also inhibited. The analysis reason was considered to be given to the right inferior frontal gyrus with low frequency after rTMS treatment, the activation of this area decreases, and at the same time, there are different degrees of functional connection between this area and the surrounding brain areas. The function decline of a brain region will also affect the function of the surrounding brain regions. The function of the above-mentioned regions in the right hemisphere decreases, and the function of the above-mentioned regions will weaken the inhibitory effect to the corresponding brain regions of the dominant hemisphere, thereby promoting the functional activation and recovery of the above-mentioned brain regions in the dominant hemisphere. As the above brain regions in the dominant hemisphere are the reading center and the naming center, it can also explain why the low frequency rTMS stimulated on the posterior part of inferior frontal gyrus in the right hemisphere can improve naming and dyslexia in patients with aphasia.

Then, we used the Degree Centrality (DC) method to analyze. DC reflects the number of connections in the adjacent areas of the brain. Specifically, it refers to the number of direct connections between a node in the brain and other adjacent nodes, which can be directly quantified (Van den Heuvel and Sporns, 2013). DC can reflect the attributes of important nodes (hub nodes) at the center of the brain network. Because of its high connectivity with the surrounding brain nodes, it has a core dominance, and even has long-distance connections with other nodes, and its functions are the most complex, so its energy consumption (oxygen consumption) is higher than that of general nodes, also it is easy to be damaged in cerebrovascular diseases (Bullmore and Sporns, 2012). Wise (2003) found that speech training can activate the brain areas around the damaged language center in the dominant hemisphere. Other studies believe that inhibiting the activation level of the right cerebral hemisphere and reducing its inhibition to the dominant hemisphere through the corpus callosum can improve the long-term efficacy of patients with aphasia (Breier et al., 2009).

The left triangular of the inferior frontal gyrus in the dominant hemisphere is known as the classic language brain area of "oral expression," which is mainly used for speech planning and execution. In recent years, scholars believe that the scope of the classic Broca area should include other areas of the frontal lobe, such as the frontal middle gyrus of the dominant hemisphere which responsible for participating in language production. At the same time, studies have found that the superior hemisphere superior frontal gyrus is also a key area in the language network, which is related to patients' language fluency and functions such as semantic conversion, retelling, naming, and listening comprehension (Sollmann et al., 2014). The results of our study showed that the DC value of the brain regions was increased such as the left superior parietal lobule, the left angular gyrus, left frontal lobe (BA6 area, middle frontal gyrus, superior frontal gyrus, supplementary motor area, paracentral lobule), bilateral Limbic lobe (cingulum gyrus) indicating that the above brain regions of the patients in the rTMS group were activated compared with that of the patients in the S-rTMS group, which also supports the above view.

Finally, we adopt the method of Functional Connection (FC) to analyze. FC reflects the degree of connection of neuronal activity between different brain regions that are far away. Through rs-fMRI, the functional network and anatomical structure of the entire brain can be studied. Li (2017) studied the brain function of patients with motor aphasia after stroke and found that after 1 month of rehabilitation, the functional connection between the middle temporal gyrus of the left dominant hemisphere and the left frontal lobe, insula and other brain regions increases. At the same time, the functional

connection between the middle temporal gyrus of the left dominant hemisphere, the marginal lobe of the left hemisphere, and the cerebellum decreased. During the recovery period of aphasia, the functional connection between the left middle frontal gyrus and the undamaged brain area around the damaged brain area also increases. In addition, some researchers believe that (Zhu et al., 2017), patients with acute stroke not only have disordered language central function, but also interfere with the default network of the brain, which leads to a decline in the cognitive function of stroke patients.

Some scholars (Sreedharan et al., 2019) found that the recovery of language function in patients with aphasia after stroke is often accompanied by changes in functional connectivity: in the acute phase, the functional connection coefficient of the language neural network is significantly reduced. In the chronic phase, the functional connection coefficient of the language neural network is significantly enhanced. The study also found that even in high-risk patients, there is a decrease in resting functional connectivity. Some scholars (Zhang C. et al., 2021) found in the study of patients with motor aphasia after stroke that the patients were accompanied by a significant decline in language ability, and the average functional connectivity index of the frontal and parietal lobe in the left dominant hemisphere also decreased significantly. As language comprehension improves, and the average connection index of the frontal and parietal lobe of the dominant hemisphere also gradually increases. The above phenomenon shows that the language comprehension ability of patients with aphasia after brain injury improves, it may be achieved by changing the functional connections between brain areas. Moreover, research has confirmed that the improvement of language function in patients with aphasia is also related to the changes in the functional connections of brain regions.

The supplementary motor area is very important for motion control. The front area of the supplementary motor area is mainly responsible for the preparation and selection of sports. The back of the supplementary motor area is responsible for the execution of the movement, and the entire supplementary motor area plays a decisive role in both the low-level execution of the movement and the high-level control of the movement. Studies have confirmed that the treatment of bilateral supplementary motor area in patients with aphasia can improve the naming ability of patients (Naeser et al., 2020). This also indicates that the treatment of low-frequency rTMS on the mirror area of the Broca area in the right hemisphere can improve the language function of patients with aphasia by improving the number and efficiency of functional connections in multiple brain areas. Our research also found that after low-frequency rTMS treatment, the functional connection between the supplementary motor area of patients with aphasia and some brain regions of the bilateral hemispheres was significantly enhanced, and this may be one of the possible mechanisms for low-frequency rTMS to improve patients with aphasia.

At the same time, in order to explore the specific therapeutic mechanism of low-frequency rTMS, we measured the changes in peripheral serum BDNF and TNF-α concentrations of the enrolled patients before and after treatment, and planned to explore the specific mechanism from the perspective of cytokines. Studies have confirmed that there are many nutritional factors in the brain to promote the recovery and improvement of brain function in patients. After rTMS treatment, it may promote the release of some nutritional factors in the brain and promote the repair or improvement of damaged brain function (Arheix-Parras et al., 2021).

Studies have found that after rTMS treatment, the levels of BDNF in peripheral blood of patients with depression are higher than before, which may be one of the mechanisms of rTMS (Zhao et al., 2019). So, we detected the serum BDNF concentration in peripheral blood of the two groups of patients before and after treatment.

It is well known that BDNF is essentially a protein, which plays an important role in the growth and differentiation of nerve cells, and can repair damaged nerve cells, thereby improving advanced cognitive functions (learning, memory, etc.) (Asadi et al., 2018; Huey Fremont et al., 2020), especially, it plays an important role in mediating the neural plasticity process of language function recovery in patients with aphasia after stroke (Di Pino et al., 2016). Animal studies have also shown that BDNF promotes long-term potentiation (LTP) through TrkB signaling (Lamb et al., 2015), which is considered to be essential for the intermittent memory process of the hippocampus (Zagrebelsky and Korte, 2014). Moreover, studies have found that BDNF can cross the blood–brain barrier through a high-volume saturated transport system. Animal studies have observed a positive correlation between the levels of BDNF in the brain and blood (Angelucci et al., 2011). Morichi et al. (2013) also found that changes in BDNF in peripheral blood are related to changes in cerebrospinal fluid (CSF) BDNF, and changes in BDNF at the peripheral level may reflect changes in BDNF in the brain. Therefore, some scholars believe that changes in BDNF at the peripheral level may reflect or at least partially reflect changes in brain BDNF (Fritsch et al., 2010). Winter et al. (2007) also found that the serum BDNF content increased significantly after high-intensity exercise, and the vocabulary learning speed was also significantly improved. They believed that the increase in the short-term learning success rate was related to the increase in the BDNF level.

Therefore, in this study, the peripheral blood of patients in the enrolled group was collected before and after low-frequency rTMS treatment, and the changes in peripheral serum BDNF content were measured, and then to explore the BDNF content changes in central nervous system. The results showed that the serum BDNF content of peripheral blood in the rTMS group increased significantly before and after treatment, and the serum BDNF content of the patients in the rTMS group was significantly higher than that of the patients in the S-rTMS group after treatment. This suggests that patients with non-fluent aphasia have a significant increase in serum BDNF levels in peripheral blood after low-frequency rTMS treatment. Many previous studies have confirmed that the increase in serum BDNF content in peripheral blood can reflect the changes in BDNF content in the patient's brain from the side. Therefore, we believe that after low-frequency rTMS stimulation, the content of BDNF in the brain of patients also increases to a certain extent, which may be one of the mechanisms of low-frequency rTMS.

In summary, our research results show that low-frequency rTMS combined with conventional speech training can significantly improve the language function of patients with non-fluent aphasia. In addition to directly changing the excitability of the cortex of the stimulated brain area, rTMS can inhibit the activation of different brain areas in the frontal and temporal lobes of the right cerebral hemisphere, and promotes the activation of different brain regions in the frontal and temporal lobes of the left dominant hemisphere, thereby improving the function of different brain regions and promoting changes in brain plasticity. It will also affect the transmission, expression and release of various cytokines and neurotransmitters, especially the expression and release of BDNF, which in turn promotes changes in brain plasticity. This may be one of the mechanisms by which rTMS promotes the improvement of the central nervous system, especially the language function of the brain.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of The Affiliated Hospital of Qingdao University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

QW and GB: conceptualization. LJ: methodology and formal analysis. YW: software. PM and XP: validation. SY and YZ: investigation. SH: resources and data curation. GB: writing—original draft preparation. GB, LJ, and QW: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

The authors wish to acknowledge Wenshuai, Qingdao University, for his help in interpreting the significance of the results of this study.

REFERENCES

Angelucci, F., Gelfo, F., De Bartolo, P., Caltagirone, C., and Petrosini, L. B. (2011). concentrations are decreased in serum and parietal cortex in immunotoxin 192 IgG-Saporin rat model of cholinergic degeneration. *Neurochem. Int.* 59, 1–4. doi: 10.1016/j.neuint.2011.04.010

Arheix-Parras, S., Barrios, C., Python, G., Cogné, M., Sibon, I., Engelhardt, M., et al. (2021). systematic review of repetitive transcranial magnetic stimulation in aphasia rehabilitation: Leads for future studies. *Neurosci. Biobehav. Rev.* 127, 212–241. doi: 10.1016/j.neubiorev.2021.04.008

Asadi, Y., Gorjipour, F., Behrouzifar, S., and Vakili, A. (2018). Irisin peptide protects brain against ischemic injury through reducing apoptosis and enhancing BDNF in a rodent model of stroke. *Neurochem. Res.* 43, 1549–1560. doi: 10.1007/s11064-018-2569-9

Breier, J. I., Juranek, J., Maher, L. M., Schmadeke, S., Men, D., and Papanicolaou, A. C. (2009). Behavioral and neurophysiologic response to therapy for chronic aphasia. *Arch. Phys. Med. Rehabil.* 90, 2026–2033. doi: 10.1016/j.apmr.2009.08.144

Bullmore, E., and Sporns, O. (2012). The economy of brain network organization. *Nat. Rev. Neurosci.* 13, 336–349. doi: 10.1038/nrn3214

Di Pino, G., Pellegrino, G., Capone, F., Assenza, G., Florio, L., Falato, E., et al. (2016). Val66Met BDNF polymorphism implies a different way to recover from stroke rather than a worse overall recoverability. *Neurorehabil. Neural. Repair.* 30, 3–8. doi: 10.1177/1545968315583721

Fahmy, E. M., and Elshebawy, H. M. (2021). Effect of high frequency transcranial magnetic stimulation on recovery of chronic post-stroke aphasia. *J. Stroke Cerebrovasc. Dis.* 30:105855. doi: 10.1016/j.jstrokecerebrovasdis.2021.105855

Fritsch, B., Reis, J., Martinowich, K., Schambra, H. M., Ji, Y., Cohen, L. G., et al. (2010). Direct current stimulation promotes BDNF-dependent synaptic plasticity: potential implications for motor learning. *Neuron* 66, 198–204. doi: 10.1016/j.neuron.2010.03.035

Haghighi, M., Mazdeh, M., Ranjbar, N., and Seifrabie, M. A. (2017). Further evidence of the positive influence of repetitive transcranial magnetic stimulation on speech and language in patients with aphasia after stroke: results from a double-blind intervention with sham condition. *Neuropsychobiology* 75, 185–192. doi: 10.1159/000486144

Harvey, D. Y., Mass, J. A., Shah-Basak, P. P., Wurzman, R., Faseyitan, O., Sacchetti, D. L., et al. (2019). Continuous theta burst stimulation over right pars triangularis facilitates naming abilities in chronic post-stroke aphasia by enhancing phonological access. *Brain Lang.* 192, 25–34. doi: 10.1016/j.bandl.2019.02.005

Huey Fremont, R., Manoochehri, M., Gazes, Y., Lee, S., Cosentino, S., et al. (2020). Effect of Functional BDNF and COMT polymorphismson symptoms and regional brain volume in frontotemporal dementia and corticobasal syndrome. *J. Neuropsych. Clin. Neurosci.* 32, 362–369. doi: 10.1176/appi.neuropsych.19100211

Lamb, Y. N., McKay, N. S., Thompson, C. S., Hamm, J. P., Waldie, K. E., and Kirk, I. J. (2015). Brain-derived neurotrophic factor Val66Met polymorphism, human memory, and synaptic neuroplasticity. *Wiley Interdiscip. Rev. Cogn. Sci.* 6, 97–108. doi: 10.1002/wcs.1334

Li, C. (2017). *Research on brain function plasticity in patients with motor aphasia after stroke based on resting state BOLD-f MRI*. Xinjiang: Shihezi University, 2–6.

Liu, Z. (2016). *Research on the plasticity of Broca's aphasia by BOLD-f MRI and DTI [D]*. Xinjiang: Shihezi University, 7–10.

Lopez-Romero, L. A., Riano-Carreno, D. M., Pachon-Poveda, M. Y., Mendoza-Sanchez, J. A., Leon-Vargas, Y. K., Moreno-Pabon, A., et al. (2019). Efficacy and safety of transcranial magnetic stimulation in patients with non-fluent aphasia, following an ischaemic stroke. A controlled, randomised and double-blind clinical trial. *Rev. Neurol.* 68, 241–249. doi: 10.33588/rn.6806.2018300

Menichelli, A., Furlanis, G., Sartori, A., Ridolfi, M., Naccarato, M., Caruso, P., et al. (2019). Thrombolysis' benefits on early post-stroke language recovery in aphasia patients. *J. Clin. Neurosci.* 70, 92–95. doi: 10.1016/j.jocn.2019.08.064

Morichi, S., Kashiwagi, Y., Takekuma, K., Hoshika, A., and Kawashima, H. (2013). Expressions of brain-derived neurotrophic factor (BDNF) in cerebrospinal fluid and plasma of children with meningitis and encephalitis/encephalopathy. *Int. J. Neurosci.* 123, 17–23. doi: 10.3109/00207454.2012.721829

Naeser, M. A., Ho, M. D., Martin, P. I., Hamblin, M. R., and Koo, B. B. (2020). Increased functional connectivity within intrinsic neural networks in chronic stroke following treatment with red/near-infrared transcranial photobiomodulation: case series with improved naming in aphasia. *Photobiomodul. Photomed. Laser Surg.* 38, 115–131. doi: 10.1089/photob.2019.4630

Neri, F., Romanella, S. M., Tomai Pitinca, M. L., Taddei, S., Monti, L., Benocci, S., et al. (2021). rTMS-induced language improvement and brain connectivity

changes in logopenic/phonological variant of Primary progressive Aphasia. *Clin. Neurophysiol.* 132, 2481–2484. doi: 10.1016/j.clinph.2021.07.017

Rossetti, A., Malfitano, C., Malloggi, C., Banco, E., Rota, V., and Tesio, L. (2019). Phonemic fluency improved after inhibitory transcranial magnetic stimulation in a case of chronic aphasia. *Int. J. Rehabil. Res.* 42, 92–95. doi: 10.1097/MRR.0000000000000322

Sollmann, N., Tanigawa, N., Ringel, F., Zimmer, C., Meyer, B., and Krieg, S. M. (2014). Language and its right-hemispheric distribution in healthy brains: an investigation by repetitive transcranial magnetic stimulation. *Neuroimage* 102, 776–788. doi: 10.1016/j.neuroimage.2014.09.002

Sreedharan, S., Arun, K. M., Sylaja, P. N., Kesavadas, C., and Sitaram, R. (2019). Functional connectivity of language regions of stroke patients with expressive aphasia during real-time functional magnetic resonance imaging based neurofeedback. *Brain Connect.* 9, 613–626. doi: 10.1089/brain.2019.0674

Szaflarski, J. P., Griffis, J., Vannest, J., Allendorfer, J. B., Nenert, R., Amara, A. W., et al. (2018). feasibility study of combined intermittent theta burst stimulation and modified constraint-induced aphasia therapy in chronic post-stroke aphasia. *Restor. Neurol. Neurosci.* 36, 503–518. doi: 10.3233/RNN-180812

Terao, Y., and Ugawa, Y. (2002). Basic mechanisms of TMS. *J. Clin. Neurophysiol.* 19, 322–343. doi: 10.1097/00004691-200208000-00006

Van den Heuvel, M. P., and Sporns, O. (2013). Network hubs in the human brain. *Trends Cogn. Sci.* 17, 683–696. doi: 10.1016/j.tics.2013.09.012

Winter, B., Breitenstein, C., Mooren, F. C., Voelker, K., Fobker, M., Lechtermann, A., et al. (2007). High impact running improves learning. *Neurobiol. Learn Mem.* 87, 597–609. doi: 10.1016/j.nlm.2006.11.003

Wise, R. J. (2003). Language systems in normal and aphasic human subjects: functional imaging studies and inferences from animal studies. *Br. Med. Bull.* 65, 95–119. doi: 10.1093/bmb/65.1.95

Yoon, T. H., Han, S. J., Yoon, T. S., Kim, J. S., and Yi, T. I. (2015). Therapeutic effect of repetitive magnetic stimulation combined with speech and language therapy in post-stroke non-fluent aphasia. *NeuroRehabilitation* 36, 107–114. doi: 10.3233/NRE-141198

Zagrebelsky, M., and Korte, M. (2014). Form follows function: BDNF and its involvement in sculpting the function and structure of synapses. *Neuropharmacology* 76, 628–638. doi: 10.1016/j.neuropharm.2013.05.029

Zhang, C., Xia, Y., Feng, T., Yu, K., Zhang, H., Sami, M. U., et al. (2021). Disrupted functional connectivity within and between resting-state networks in the subacute stage of post-stroke aphasia. *Front. Neurosci.* 15:746264. doi: 10.3389/fnins.2021.746264

Zhang, J., Zhong, D., Xiao, X., Yuan, L., Li, Y., Zheng, Y., et al. (2021). Effects of repetitive transcranial magnetic stimulation (rTMS) on aphasia in stroke patients: a systematic review and meta-analysis. *Clin. Rehabil.* 35, 1103–1116. doi: 10.1177/0269215521999554

Zhao, X., Li, Y., Tian, Q., Zhu, B., and Zhao, Z. (2019). Repetitive transcranial magnetic stimulation increases serum brain-derived neurotrophic factor and decreases interleukin-1β and tumor necrosis factor-α in elderly patients with refractory depression. *J. Int. Med. Res.* 47, 1848–1855. doi: 10.1177/0300060518817417

Zhu, D., Chang, J., Freeman, S., Tan, Z., Xiao, J., Gao, Y., et al. (2014). Changes of functional connectivity in the left frontoparietal network following aphasic stroke. *Front. Behav. Neurosci.* 8:167. doi: 10.3389/fnbeh.2014.00167

Zhu, Y., Bai, L., Liang, P., Kang, S., Gao, H., and Yang, H. (2017). Disrupted brain connectivity networks in acute ischemic stroke patients. *Brain Imag. Behav.* 11, 444–453. doi: 10.1007/s11682-016-9525-6

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 01 June 2022
doi: 10.3389/fnagi.2022.891433

Combined Multi-Atlas and Multi-Layer Perception for Alzheimer's Disease Classification

Xin Hong[1,2]*, *Kaifeng Huang*[1], *Jie Lin*[1], *Xiaoyan Ye*[3]*, *Guoxiang Wu*[4], *Longfei Chen*[5], *E. Chen*[6]* and *Siyu Zhao*[1]

[1] College of Computer Science and Technology, Huaqiao University, Xiamen, China, [2] Key Laboratory of Computer Vision and Machine Learning (Huaqiao University), Fujian Province University, Xiamen, China, [3] Fuzhou Comvee Network and Technology Co., Ltd, Fuzhou, China, [4] College of Foreign Languages, Huaqiao University, Quanzhou, China, [5] Department of Neurology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, China, [6] Department of Neurosurgery, Zhongshan Hospital Affiliated to Xiamen University, Xiamen, China

OPEN ACCESS

Edited by:
Yizhang Jiang,
Jiangnan University, China

Reviewed by:
Guodong Zhang,
Shenyang Aerospace University, China
Dehua Chen,
Donghua University, China
Wenhong Wei,
Dongguan University of Technology, China

***Correspondence:**
Xin Hong
10409035@qq.com;
xinhong@hqu.edu.cn
Xiaoyan Ye
yexy@comvee.cn
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: 07 March 2022
Accepted: 19 April 2022
Published: 01 June 2022

Citation:
Hong X, Huang K, Lin J, Ye X, Wu G, Chen L, Chen E and Zhao S (2022) Combined Multi-Atlas and Multi-Layer Perception for Alzheimer's Disease Classification. Front. Aging Neurosci. 14:891433. doi: 10.3389/fnagi.2022.891433

Alzheimer's disease (AD) is a progressive and irreversible neurodegenerative disease. To distinguish the stage of the disease, AD classification technology challenge has been proposed in Pattern Recognition and Computer Vision 2021 (PRCV 2021) which provides the gray volume and average cortical thickness data extracted in multiple atlases from magnetic resonance imaging (MRI). Traditional methods either train with convolutional neural network (CNN) by MRI data to adapt the spatial features of images or train with recurrent neural network (RNN) by temporal features to predict the next stage. However, the morphological features from the challenge have been extracted into discrete values. We present a multi-atlases multi-layer perceptron (MAMLP) approach to deal with the relationship between morphological features and the stage of the disease. The model consists of multiple multi-layer perceptron (MLP) modules, and morphological features extracted from different atlases will be classified by different MLP modules. The final vote of all classification results obtains the predicted disease stage. Firstly, to preserve the diversity of brain features, the most representative atlases are chosen from groups of similar atlases, and one atlas is selected in each group. Secondly, each atlas is fed into one MLP to fetch the score of the classification. Thirdly, to obtain more stable results, scores from different atlases are combined to vote the result of the classification. Based on this approach, we rank 10th among 373 teams in the challenge. The results of the experiment indicate as follows: (1) Group selection of atlas reduces the number of features required without reducing the accuracy of the model; (2) The MLP architecture achieves better performance than CNN and RNN networks in morphological features; and (3) Compared with other networks, the combination of multiple MLP networks has faster convergence of about 40% and makes the classification more stable.

Keywords: atlas, multi-layer perceptron, Alzheimer's disease, classification, PRCV competition

INTRODUCTION

Alzheimer's disease (AD) is a common neural degenerative disease, from which 60 to 70% of senile patients with dementia suffer (Jagust, 2013). A feature of AD is the damage induced by the irreversible and progressive cognitive function of human brains. It is continuously progressing when a normal-control (NC) gradually becomes a patient with AD. Mild cognitive impairment (MCI) is the early disease-developing stage (Reiman et al., 2010). Therefore, being able to correctly represent the disease-developing stage a patient is in helps in diagnosing and slowing the process of the disease. Over time, the condition of AD is often accompanied by brain atrophy. Recently, in Pattern Recognition and Computer Vision 2021 (PRCV 2021), the AD classification technology challenge[1] provided a dataset from multiple atlas partitions and extracted volume features. This dataset is used for three classification tasks of NC/MCI/AD. The data of each sample in the dataset consists of brain gray matter volume and average cortical thickness that are extracted from multiple atlases.

The AD classification frameworks directly analyze the patterns in neuroimaging data of AD/MCI/NC subjects. In addition, the classification framework is comprised of multi-components: feature extraction, feature selection, dimensionality downsampling, and feature-based classification. According to the PRCV 2021, the task of the challenge is to do the three classifications of patients. Over the past decade, the cortical thickness, voxel-wise, and hippocampal morphological features of sMRI were used to diagnose AD (Jagust, 2013). After jointly aligning whole-brain image data to associate each brain voxel, voxel features have extracted a vector with multiple scalar measurements. Gray matter voxels are used for input features and trained in the support vector machine (SVM) classifier to classify AD and NC categories (Klöppel et al., 2008). To improve the performance of the model, the researchers used a 3D CNN to make predictions about the stage of the disease that the AD patient was in based on MRI (Bron et al., 2015). In some work, researchers have also improved the accuracy of classification by pre-training or providing model complexity (Payan and Montana, 2015; Korolev et al., 2017). In the competition, most of the better-performing teams have optimized their methods based on the multi-layer perceptron (MLP) architecture. The adjustments on the network are, broadly, as follows: combining MLP with attention mechanisms, adjusting the depth of the MLP network, combining multiple networks for data processing, etc. For the processing of datasets, some teams filtered data based on the characteristics of the atlas or supplemented the data with interpolation.

Since comparative evaluations of these feature extraction techniques reveal several limitations for classifying AD, we present a multi-atlas multi-layer perceptron (MAMLP) approach to a one-dimensional long vector data extracted from multiple atlases. Compared to the CNN and rerrent neural network (RNN) methods, our method converges faster and has higher accuracy during the training process. A network composed of multiple MLP modules achieves higher accuracy in this task than a single MLP network. In addition, our method ranks the 10th in the competition.

RELATED WORK

Reliable diagnosis of AD ought to adapt to different datasets, such as MRI scans collected by several patient groups, to reduce differences in data distribution and bias against specific groups. The existing machine learning model has been applied to the detection of AD. According to existing studies, the cortical thickness, somatotopic and hippocampal morphological features extracted by sMRI can be used to diagnose AD (Jagust, 2013). After aligning whole brain image-feature data to associate each brain voxel in common, voxel features are extracted a vector with multi-scalar measurements. The coefficients of the series are calculated and normalized to eliminate the rotation translation effect and the features used to train the SVM-based classifier. Researchers applied the gray matter voxels as input features and trained the SVM classifier to classify AD and NC categories (Klöppel et al., 2008). In practical problems, there is often more than one factor affecting a thing, that is the dependent variable corresponds to more than one independent variable. For MRI data, we should also consider more image features. However, due to the limitations of extraction methods, the data inevitably have some biases and errors that need to be corrected by humans. And traditional machine learning methods are more demanding for data processing, and different processing methods may bring large differences in results.

The existing deep learning model has been applied to the classification of AD. 2D CNN was used to extract slice features from MRI scans. Deep learning aims to reduce the use of domain expert knowledge in designing and extracting the most appropriate discriminant features (Plis et al., 2014). In the AD classification task, the researchers used a model of 3D CNN to perform feature extraction of the complete MRI, which was then used for AD/NC classification (Bron et al., 2015), and some researchers have also used unsupervised auto-encoders to pre-train convolutional layers or a more complex network to improve the accuracy of classification (Payan and Montana, 2015; Korolev et al., 2017). In some studies, part of the CNN architecture was inspired by Hosseini-Asl et al. (2018), they provide a pre-trained 3DCNN network that learns to capture generic features of AD biomarkers and adapts to datasets from different domains. There are also studies using RNN to train an AD classifier (Velazquez et al., 2019). Cheng and Liu (2017) uses extracted inter-slice features to perform the final classification. Both CNN and RNN need a large number of training data and optimized structures to achieve reliable performance. These researches used CNN-based or RNN-based to extract essential features of MRI or acquired the dense representation of MRI to build a regression model for AD score prediction or to train a different classifier. Due to CNN's or RNN's excellent performance on image classification, more researches used several data modalities on different planes and clinical scores to build multi-channel CNN and increase the model prediction ability. Although these methods perform well

[1]Pattern Recognition and Computer Vision 2021 Alzheimer's disease classification technology challenge: https://competition.huaweicloud.com/information/1000041489/circumstance.

in image or text data, they may not be suitable for some discrete feature data, such as PRCV 2021 AD classification technology challenge dataset.

This paper uses AD classification methods based on deep learning for the PRCV 2021 AD classification technology challenge dataset, namely, SVM, RNN, CNN, region convolutional neural network (RCNN), and MLP. In order to solve the problem of the characteristics of the dataset itself and the small number of data samples, we used different MLPs to analyze the data from different atlases after screening. The advantage of this method is that it simplifies the structure of the network and prevents overfitting. At the same time, after atlas screening, some similar atlases are removed, which can reduce the negative impact of redundant data on the results. It is similar to the top-ranked methods, such as the use of multiple networks and atlas screening. In contrast to all these solutions, our method is carried out on the dataset. According to our model design and training method, the optimal model is obtained. Using the official scoring index of the competition, our model is better than other algorithms. However, due to the small number of samples in the data set, the results were somewhat unstable, and there was a gap between some optimization techniques that our team failed to surpass.

MATERIALS

Datasets

The dataset was provided by the PRCV 2021 AD classification technical challenge and contains 2,600 samples. **Table 1** shows the distribution and composition of the data. The age range of the samples was 32–91, with 1,982 samples concentrated between the age range of 60 and 80. The dataset contains the sample's brain gray matter volume and mean cortical thickness, which were extracted by the Computational Anatomy Tool12 (CAT12) based on multiple atlases.

The CAT12 software first aligns the MRI images and segments out the brain. Then, according to the different atlases, CAT12 segments the MRI and calculates the volume and cortical posteriority of the different regions. Finally, features of multiple atlases were combined to form a sequence of 28,169 one-dimensional features. These 28,169 eigenvalues are used as the feature data of this sample. **Table 2** shows the information on the templates. There are 13 types and 30 versions of templates used. The name in the table indicates the name of the template, while the version indicates the version used. Each template has a different region of interest (ROI), and based on ROI, the number and value of features extracted are different.

TABLE 1 | The distribution and composition of the data.

Class	Distribution	Subject total
Label	AD	671
	MCI	1,148
	NC	781
Age	Above 80	385
	60 – 80	1,983
	Under 60	232

DATA PREPROCESSING

This section introduces several novel contributions in data preprocessing. First, the atlases were filtered to reduce the dimensions when the dataset contains a small number of samples with high-dimensional morphological features. Second, the invalid value caused by the atlas mapping error was replaced by the average value or 0 when extracting the morphological feature from the brain atlas. Third, standardization was applied to adjust the data magnitude that is different between multiple atlases.

Based on the characteristics of the dataset, the data preprocessing methods, including atlas filtering, invalid value replacement, and data normalization, were established (**Figure 1**).

Atlas Filtering

The data in the PRCV 2021 AD classification technology challenge dataset combined 28,169 features extracted by 30 atlases. Among these atlases, some were similar to each other. For example, AAL1 to AAL2 to AAL3 was a process of gradual evolution and subdivision, which also had a similar relationship between Schaefer2018_1000 and Schaefer2018_100. Since a small sample with high-dimensional features caused over-fitting of the model, to reduce the feature redundancy of the template, we filtered out templates with similar functions and division basis and selected a template with the most detailed

TABLE 2 | The data summary of the atlases.

Name	Version	ROI number
AAL (Rolls et al., 2020)	AAL(1/2/3v1)	90/116/170
AICHA (Joliot et al., 2015)	AICHA_reordered	384
Brainnetome (Fan et al., 2016)	rBN_Atlas_246_1mm	246
Brodmann (Zilles and Amunts, 2010)	Brodmann	41
Gordon (Gordon et al., 2016)	Gordon	333
Hammersmith (Hammers et al., 2003)	Hammers-mith (83/95)	83/95
Harvard-Oxford (Desikan et al., 2006)	HarvardOxford	113
Jülich (Eickhoff et al., 2005)	Juelich-thr25	103
Melbourne	Tian_Subcortex (S1/S2/S3/S4_7T)	62/54/34/16
MIST (Urchs et al., 2019)	MIST (7/12/20/36/64/122/197/325/444)	7/12/20/36/64/122/197/325/444
Scheafer (Schaefer et al., 2018)	Schaefer 2018(100/200/400/600/800/1000)	100/200/400/600/800/1000
SUIT	Cerebellum-MNIflirt	28
Yeo (Thomas Yeo et al., 2011)	Yeo2011 (7/17)	7/17

"Version" represents the name of each version, where the different versions are indicated in "()" and "ROI Number" represents the number of ROIs for the different versions.

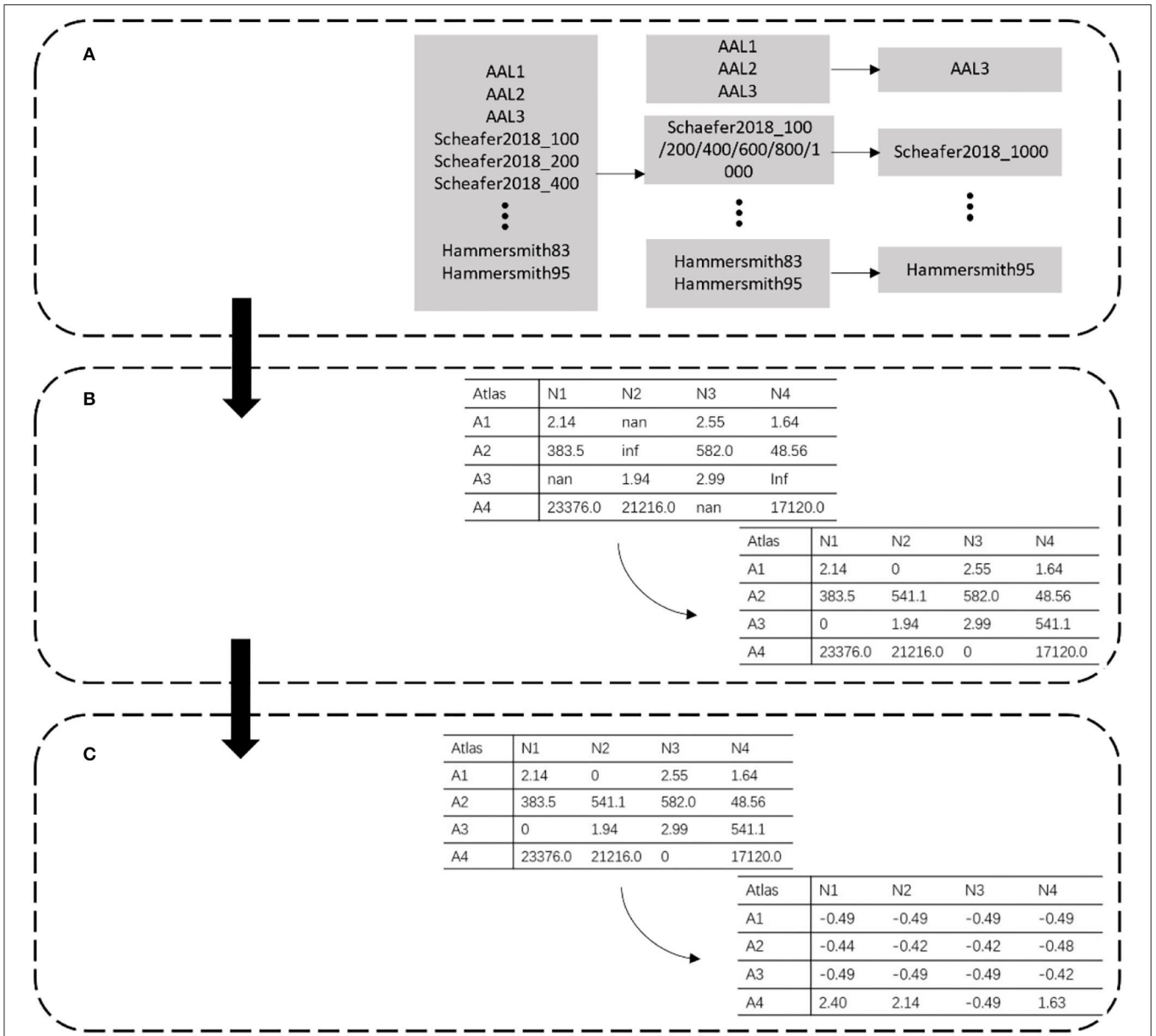


FIGURE 1 | Data preprocessing. Each row represents different partitions. Atlas is represented by A, and N in each column represents the extracted data number. **(A)** Atlas filtering. **(B)** Invalid value replacement. **(C)** Data standardization.

division among them (**Figure 1A**). For example, in the case of AAL templates, we kept the most detailed division of the AAL3v1 version as representative of this class of template. Meanwhile, we kept Schaefer2018_1000 as representative for the Schaefer2018 template.

Invalid Value Replacement

The morphological features are extracted from the MRI image by selecting a specific brain template with the CAT12 tool. During the extraction process, part of the data was lost due to the registration error of the template, which resulted in empty and infinite values. These invalid values directly led to the disappearance of the gradient in the model during the training process. As shown in **Figure 1B**, we dealt with these invalid values by replacing them. Empty and infinite values were replaced with 0 and the average value, respectively.

Data Standardization

The feature extracted from the different atlas had a magnitude difference. As shown in **Figure 1C**, the maximum data was >10,000, while the minimum data was <10. We standardized the data to adjust the values to the same magnitude. The mean and standard deviation of the whole dataset was calculated, and each data was divided into standard deviation from the mean.

The calculation of standardized data is as follows:

$$\hat{x}_{ij} = \frac{x_{ij} - mean}{std} \tag{1}$$

where i is the number of data and j is the number of the eigenvalues of the data i. Mean represents the average of the dataset, and std represents the standard deviation of the dataset. Equation (2) and Equation (3) show the calculation of mean and std, respectively.

$$mean = \frac{1}{ij}\sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij} \tag{2}$$

$$std = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij}}{ij}} \tag{3}$$

where the N and M are respectively, the size of the dataset and the length of each data.

After the data preprocessing, the length of data was reduced from 28,169 to 8,377. Invalid values in the data were removed by replacement. Finally, the data was standardized to reduce the gap in value.

METHODS

In this section, we present the MAMLP model using the challenge of dataset for AD prediction, specifically the one-dimensional long vector data extracted from multiple atlases. Then, we discussed the MAMLP architecture, which interlinks multiple MLP blocks with state connections, for modeling the differential information in the AD.

Further, this paper selects the data extracted from different atlases, constructs different small MLP networks according to different atlas for processing, and finally obtains the final prediction outputs combined with the results. Considering that constructing a huge MLP network often leads to overfitting due to insufficient samples of the dataset, this method not only avoided the overfitting caused by too small a sample size but also simplified the network to a certain extent and improves the efficiency of the algorithm. The structure of MAMLP is shown in **Figure 2**. We first separated the pre-processed data according to different atlases. The data from different atlases were input to different MLP network modules for analysis. Finally, the classification outputs of all MLP networks were combined to obtain the final result.

Mixed Layers MLP Modular

After the separation operation, the data of different atlases were input to different MLP networks for analysis. However, we observed that the number of ROIs between the various atlases was not consistent, and the number of different feature values was extracted based on different atlases. Therefore, a fixed MLP structure Was apparently more difficult to applied to all atlases.

To solve this problem, we designed a mixed-layer MLP network to facilitate the classification, and employed a two-layer or three-layer linear layer network to process the data according to the number of each atlas. As shown in **Figure 2**, a fully connected network containing three linear layers was used to process the data when the number of ROIs of the atlas was >100. Unlike the three linear layers network, if the number of atlases is <100, the number of linear layers is reduced to two. In the end, different network output classification results were based on original dataset from different atlases and were combined in the subsequent operation. The final classification result of the network can be expressed by the following equation:

$$O = \max\left(\sum_{j=1}^{N} SoftMax(M_j)\right) \tag{4}$$

where the M represents the output of the MAMLP subnetwork, j represents the jth subnetwork, and N represents the number of subnetworks.

Ativation Function and Loss Function

In the MLP network, superscript l is set to represent the data related to layer l, which consists of L layers. The input layer is marked as 0, the output layer is marked as l, and the subscript represents the matrix or a vector index. The deactivation value of layer L is equal to the activation value of the previous layer multiplied by the network weight matrix and adds the network deviation.

Equation (5) shows the calculation method for inactive value, where z^l represents the inactive value of Layer l, W^l represents the layer l network weight matrix, and b^l represents the layer l network bias. In addition, a^l represents the value of the l layer after the activation function, and the method of calculation is shown in Equation (6), where $h(z)$ denotes the activation function.

$$z^l = a^{l-1}W^l + b^l \tag{5}$$

$$a^l = h\left(z^l\right) = \max\left(0, z^l + N(0,1)\right) \tag{6}$$

Equation (5) facilitates the convergence of an end-to-end model training.

According to Equation (5) and Equation (6), z^l and a^l are calculated in order, and the output layer z^L is obtained. Loss function $C\left(a^L, y\right)$ is then calculated according to Equation (6), where y represents the label, and n^L represents the number of neurons in the output layer.

$$C\left(a^L, y\right) = -\sum_{i=1}^{n^L} y_i \, log \, a_i^L \tag{7}$$

The output $\hat{y}$ of the final network is the subscript with the highest probability in a^L. Equation (8) is the calculation method of $\hat{y}$.

$$\hat{y} = argmax_{i\in\{1,\cdots n^L\}} a^L \tag{8}$$

Therefore, standard MLPs are not equipped to deal with unreliable input data. We show in this section that the gain of MAMLP over those models increases in two important step with unreliable inputs: multi-step prediction and dealing with original data.

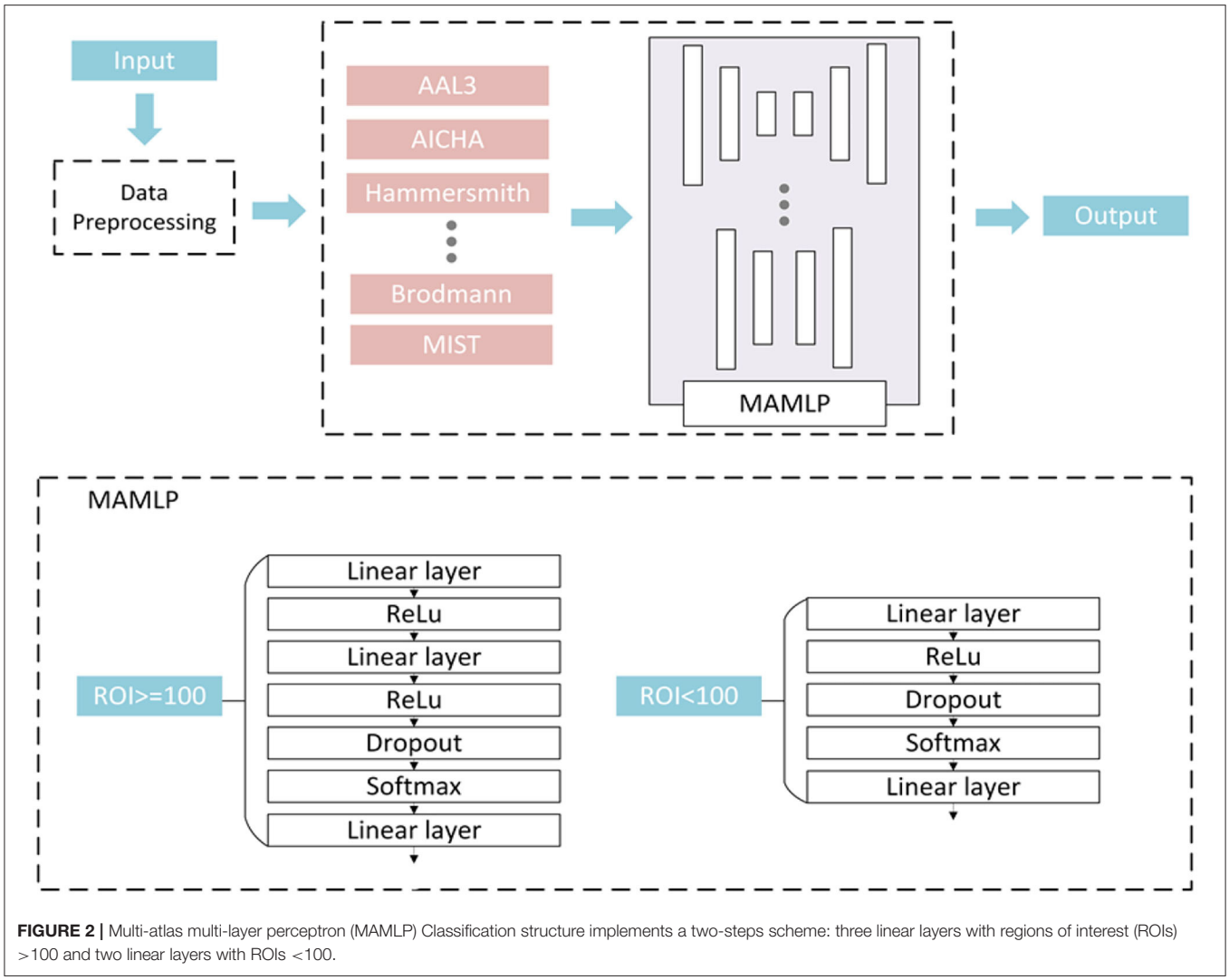


FIGURE 2 | Multi-atlas multi-layer perceptron (MAMLP) Classification structure implements a two-steps scheme: three linear layers with regions of interest (ROIs) >100 and two linear layers with ROIs <100.

Implementation

Our approach has two key components: the first is the filtering of atlases in data preprocessing, and the second is the analysis of the network structure using different fully connected networks for different atlases. In atlas filtering, we keep the most detailed atlases among similar atlases for division. The original dataset was processed using 30 atlases for MRI and 28,169 feature values were extracted. After processing, 13 atlases containing 8,377 feature values were finally retained. In the network structure, the data were processed using a hybrid network structure.

First, we separated the data from different atlases into 13 groups and fed them into different fully connected networks for analysis. Based on the number of ROIs of the atlases, data with a number >100 features are fed into a fully connected network with three linear layers for processing. Data with a number <100 features are fed into a fully connected network with two linear layers for processing. The structure of the fully connected network with three linear layers. The first linear layer was followed by a linear rectification function (ReLU) layer as the activation function. The second linear layer is followed by a dropout layer to prevent overfitting, while the last linear layer is followed by only a Softmax layer to obtain the final classification results. The fully connected network containing two linear layers removes the first linear layer and the ReLU layer compared to the network containing three linear layers. Finally, the results of each network are combined to obtain the final classification results.

Model Evaluation

In addition to using accuracy as the evaluation standard, we also introduce the F1 function as the evaluation index when evaluating the model. In statistics, the F1 function is used to simultaneously calculate the accuracy of unbalanced data classification problem under the consideration of the accuracy and recall of the model. The calculation formula is as follows:

$$F1 = 2\frac{Recall \times Precision}{Recall + Precision} \tag{9}$$

In the multi-classification problem, the F1 score of each category is usually calculated first and averaged to obtain the macro F1 score. The macro F1 score is then used to evaluate the performance of the model in our experiment. The calculation formula is as follows:

$$\text{macro F1score} = \frac{F1score_1 + F1score_2 + F1score_3}{3} \quad (10)$$

The Area Under Curve (AUC), as the evaluation index of binary classification standards, measures the ratio of true positive (recall) and false-positive classification. In multi-classification experiments, the macro F1 score is added as the evaluation index. Toward binary classification, AUC is also added to comprehensively evaluate the performance of models. The calculation formula is as follows:

$$\text{AUC} = \frac{\sum pred_{pos} > pred_{neg}}{\text{positiveNum} * \text{negativeNum}} \quad (11)$$

The denominator is the total number of combinations of positive and negative samples, while the numerator is the number of combinations where positive samples are greater than negative samples.

EXPERIMENTS AND DISCUSSION

To comprehensively evaluate the performance of the model, we set up several groups of experiments to compare and study the effects of the data dimension, network structure, and the number of atlases on the experimental results. Meanwhile, we further discuss the results of the competition and the advantages and disadvantages of our approach compared to other teams.

Parameter Setting of Experiment

The experimental environment of this paper was the PyTorch framework and NVIDIA–TITAN-XP GPU. During the training process, we adopted the following strategies: Cross-Entropy as the loss function; Stochastic Gradient Descent (SGD) as the optimizer; the learning rate is set to 0.001; the dropout layer in the network is set to 0.5. We divided the number of the training-set and test-set into 2,300:300, and 100 cases of each label were selected in the test-set. In the AD/NC/MCI experiment, four indicators were used for evaluation, including Accuracy, Precision, Recall, and F1score. AUC was used as an evaluation indicator in the binary classification experiment. The higher all the indicators, the better the effect of classification.

Comparative Experiments of Data Pre-processing

In the data pre-processing section, the following pre-processing operations are performed on the data: (1) Atlas Filtering for feature dimension reduction; (2) replacement of invalid values in the data; and (3) standardization of the data values. To demonstrate the effectiveness of these treatments, we conducted comparative experiments on data pre-processing.

Table 3 shows the impact of data pre-processing on the experiment. Compared with the unfiltered data and unstandardized data, the accuracy of the pre-processed data is improved greatly. These experiments were performed by default after the second pre-processing operation (invalid value replacement) because the model would have experienced gradient disappearance without this preprocessing. The results of the experiments show that (1) "Filtering Atlas" had an impact on the accuracy of the model, improving it by about five percent; and (2) "Numerical standardization" is significant. Without standardization, differences in extraction criteria between templates will make it difficult for the model to learn valuable information.

TABLE 3 | Effect of data preprocessing on the experiment.

Preprocessing	Accuracy	Precision	Recall	F1 score
Without preprocessing	0.25	0.15	0.11	0.13
Filtering atlas	0.51	0.47	0.43	0.45
Data standardization	0.64	0.63	0.64	0.64
Data standardization and filtering atlas	**0.67**	**0.68**	**0.67**	**0.68**

All experiments assume that the invalid values have been removed. Otherwise, the experiments cannot be performed. Bold values mean the best value in the comparative experiment.

Considering that a huge number of atlases are used in the data extraction process and that some atlases have high similarities, we filtered the models in the data preprocessing stage and selected one in the similar atlases. To further explore the accuracy of the model with the different number of atlases, we tried to keep more atlases or further removed them.

Figure 3 shows the experimental results with different numbers of atlases. The number of atlases after data preprocessing is 13. These results suggest that the classification accuracy was improved by removing similar atlases, but the classification accuracy showed a decreasing trend when atlases are further removed. These findings are understandable because using too many similar atlases causes the number of features per sample to exceed the sample size of the PRCV 2021 AD classification technology challenge dataset. A situation that over-fits the model while using too few atlases does not provide sufficient feature data. Therefore, choosing the appropriate number of atlases can further improve the classification accuracy of the model.

Different Methods Based on Different Data Dimensions

By splicing the data, the original one-dimensional data can be spliced into two- or three-dimensional data. Then, the convolution under the corresponding dimension can be used for data processing and analysis. We follow that these extracted data do not have image characteristics, such as color and form. Therefore, the method of using convolutional analysis after up dimensioning is considered to have poor performance for PRCV 2021 AD classification technology challenge dataset. We processed the data as two-dimensional and three-dimensional fake-image data and used classical CNN to process them.

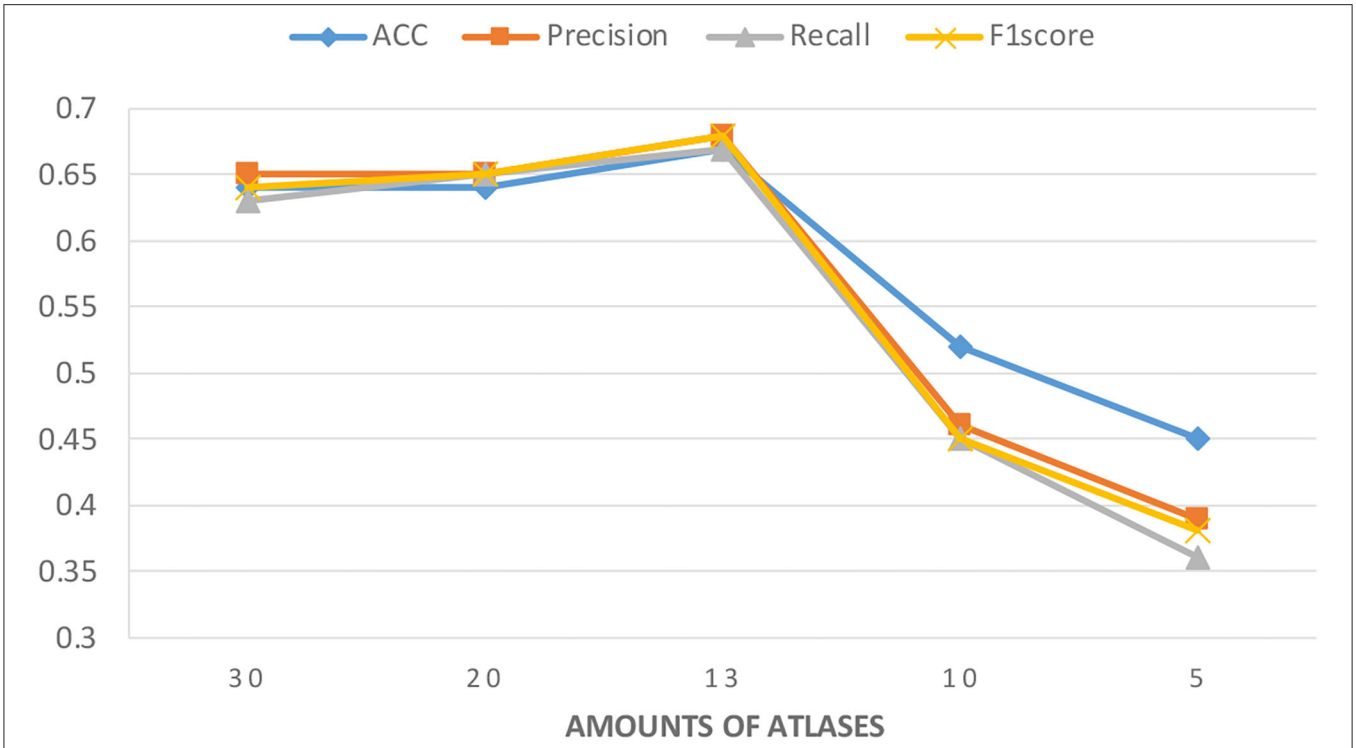


FIGURE 3 | Classification accuracy of the model under different amounts of atlases. The number of atlases after data preprocessing is 13.

In the experiment of different dimensions, the data are spliced as 168×168 two-dimensional data and 31×31×30 three-dimensional data. The data is then processed by invalid value replacement and standardization before training. On the two-dimensional data, visual geometry group (VGG) (Simonyan and Zisserman, 2014) and ResNet50 (He et al., 2016) are used to analyze the data, while the full-size diagnosis network (FDN) is used on the three-dimensional data (Li et al., 2019). For the comparison experiments, we use the same learning rate and batch size. The network structure is also the same as in the original paper, except that the FDN model uses a non-iterative version. The setup of these methods follows the original design of their papers. Meanwhile, we use an MLP network with 4 linear layers to compare with our method and evaluate the effectiveness of the method in four metrics, which are Accuracy, Precision, Recall, and F1score.

As shown in **Table 4**, our method obtained the best results in all four metrics. In addition, the method of raising the dimensionality does not effectively improve the classification accuracy. These results suggest that the method of using CNN for feature extraction on two-dimensional or three-dimensional data is not as effective as the method of using MLP on one-dimensional data. These findings are understandable because although the data has been improved on the dimension, it still does not have image features, such as color-feature or shape-feature. In addition, the CNN still cannot extract those disease-related features well. Compared with a single MLP network, since the data extracted from different atlas are not correlated, our method separates them and uses different networks for analysis, which can better prevent model overfitting and prevent mutual interference between different atlas data.

TABLE 4 | Data summary of different methods based on different data dimensions.

Data dimension	Models	Accuracy	Precision	Recall	F1 score
One-dimensional	MLP	0.66	0.66	0.65	0.66
	MAMLP (ours)	**0.67**	**0.68**	**0.67**	**0.68**
Two-dimensional	VGG	0.56	0.58	0.55	0.56
	ResNet50	0.64	0.66	0.61	0.62
Three-dimensional	FDN	0.55	0.53	0.54	0.55

The bold part represents the best result.

Different Methods Based on One-Dimensional Data

One-dimensional feature data in the PRCV 2021 AD classification technology challenge dataset comes from gray matter volume and mean cortical thickness components extracted from different atlases. Unlike MRI, the data in the dataset loses original image characteristics, such as color or shape. The methods which are used to process MRI on two-dimensional or three-dimensional had poor performance for this dataset. However, some methods for natural language processing

TABLE 5 | Data summary of in Alzheimer's disease (AD)/mild cognitive impairment (MCI)/normal control (NC) classification.

Methods	Accuracy	Precision	Recall	F1 score
CNN-1d (Kim, 2014)	0.55	0.52	0.52	0.52
RNN (Liu et al., 2015)	0.65	0.64	0.63	0.64
RCNN (Zhou et al., 2016)	0.62	0.62	0.62	0.63
MAMLP (ours)	**0.67**	**0.68**	**0.67**	**0.68**

The bold part represents the best result.

are often used to process one-dimensional feature data. Hence, we compared these methods with ours.

The procedure we followed can be briefly described as data pre-processing using different methods to analyze the data and four indicators to evaluate the model. We use three methods to compare with our method, including CNN-1d, RNN, and RCNN (Kim, 2014; Liu et al., 2016; Zhou et al., 2016).

As shown in **Table 5**, our method obtained the best results in all four metrics. The research we have done suggests that these natural language processing-related methods are not very good at extracting the relationship between features and disease stages compared to our methods. The CNN has advantages in performing two-dimensional image feature extraction, but does not work well for processing one-dimensional long vector data. Recurrent neural networks are mainly concerned with the temporal relationship between features and perform poorly in identifying the relationship between features and classification results. For PRCV 2021 AD classification technology challenge dataset, it has lost its original imaging features after atlas extraction, and the correlation between each feature is not obvious. As a result, CNN and RNN-related methods do not apply to this dataset compared to MLP.

To further measure the performance of the model, we take AUC as the evaluation standard and experiment on binary classification problems. Among them, the number of samples in each category in the classification problems of NC/MCI, NC/AD, and MCI/AD are 781/1,148, 781/671, and 1,148/671, respectively. We divide the train set and test set according to a ratio of 4:1. For the rest of the setup, it was kept consistent with the triple classification experiment.

Table 6 shows that the performance of the four methods in the three binary classification tasks. In the classification of NC/MCI and NC/AD, our model obtained the highest score. RNN model performs better in the classification of MCI/AD. In the experiments with dichotomous classification, the performance of the individual models was largely consistent with that of trichotomous classification, but in MCI/AD, the RNN performed much better. This phenomenon illustrates that our method is more sensitive to the differences between NC and AD/MCI and is more accurate in determining whether the disease is present.

Meanwhile, we compared the differences between the fixed MLP network and the hybrid MLP network, which is to verify whether this approach can improve the classification accuracy. As shown in **Table 7**, the mixed network structure exhibits a greater advantage in all metrics compared to the fixed one. This phenomenon is also easily explained by the fact that a small network is not suitable for large inputs when approaches use a fixed network structure and vice versa. If a fixed structure is used in all MLP sub-networks, the number of features per template should be fixed, which is difficult to achieve. Therefore, a mixed network structure is a more suitable method.

TABLE 6 | Data summary of different models in binary classification.

Methods	Accuracy	F1 score	AUC
(a) Data summary of NC/MCI classification			
CNN-1d (Kim, 2014)	0.65	0.75	0.60
RNN (Liu et al., 2015)	0.73	0.79	0.71
RCNN (Zhou et al., 2016)	0.73	0.80	0.69
MAMLP (ours)	**0.75**	**0.81**	**0.74**
(b) Data summary of NC/AD classification			
CNN-1d (Kim, 2014)	0.84	0.83	0.84
RNN (Liu et al., 2015)	0.86	0.84	0.85
RCNN (Zhou et al., 2016)	0.85	0.83	0.84
MAMLP (ours)	**0.89**	**0.89**	**0.90**
(c) Data summary of MCI/AD classification			
CNN-1d (Kim, 2014)	0.68	0.69	0.64
RNN (Liu et al., 2015)	**0.78**	**0.71**	**0.77**
RCNN (Zhou et al., 2016)	0.77	0.68	0.75
MAMLP (ours)	0.77	0.66	0.74

The bold part represents the best result.

TABLE 7 | Different numbers of linear layers on multi-layer perceptron (MLP) modules.

Number of linear layers	Accuracy	Precision	Recall	F1 score
Two	0.56	0.51	0.52	0.52
Three	0.59	0.58	0.57	0.58
Two and three mixed	**0.67**	**0.68**	**0.67**	**0.68**
Four	0.58	0.59	0.58	0.59

The bold part represents the best result.

In addition, we believe that there is no correlation between data from different atlases. Different from the original MLP network, referring to the Ortiz's method (Ortiz et al., 2016), the data is segmented according to different atlases and then sent into different MLP models for classification before the results are combined. In this way, we effectively reduce the complexity of the model and prevent the overfitting of the algorithm. **Figure 4** shows the change process of loss, accuracy, and f1score in the training process of different models. With the continuous improvement of training times, the value of loss continues to decline while the classification results of some models gradually deteriorate. It can be inferred that due to the small sample size and excessive training, the model has the phenomenon of overfitting, which is more obvious in the complex model. Compared with other models, our model

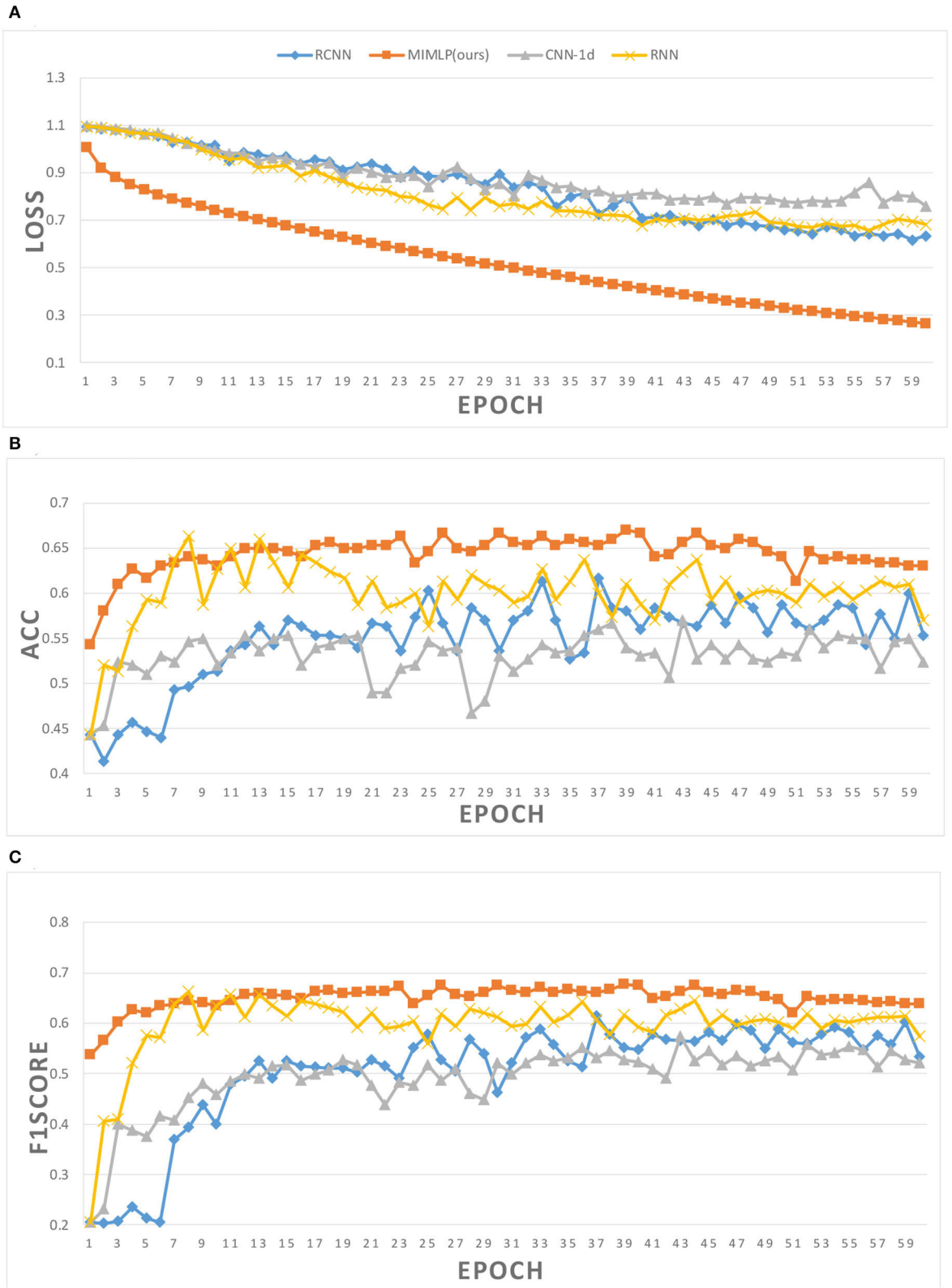


FIGURE 4 | Training details of different models. **(A)** Loss in training. **(B)** Accuracy in training. **(C)** F1score in training.

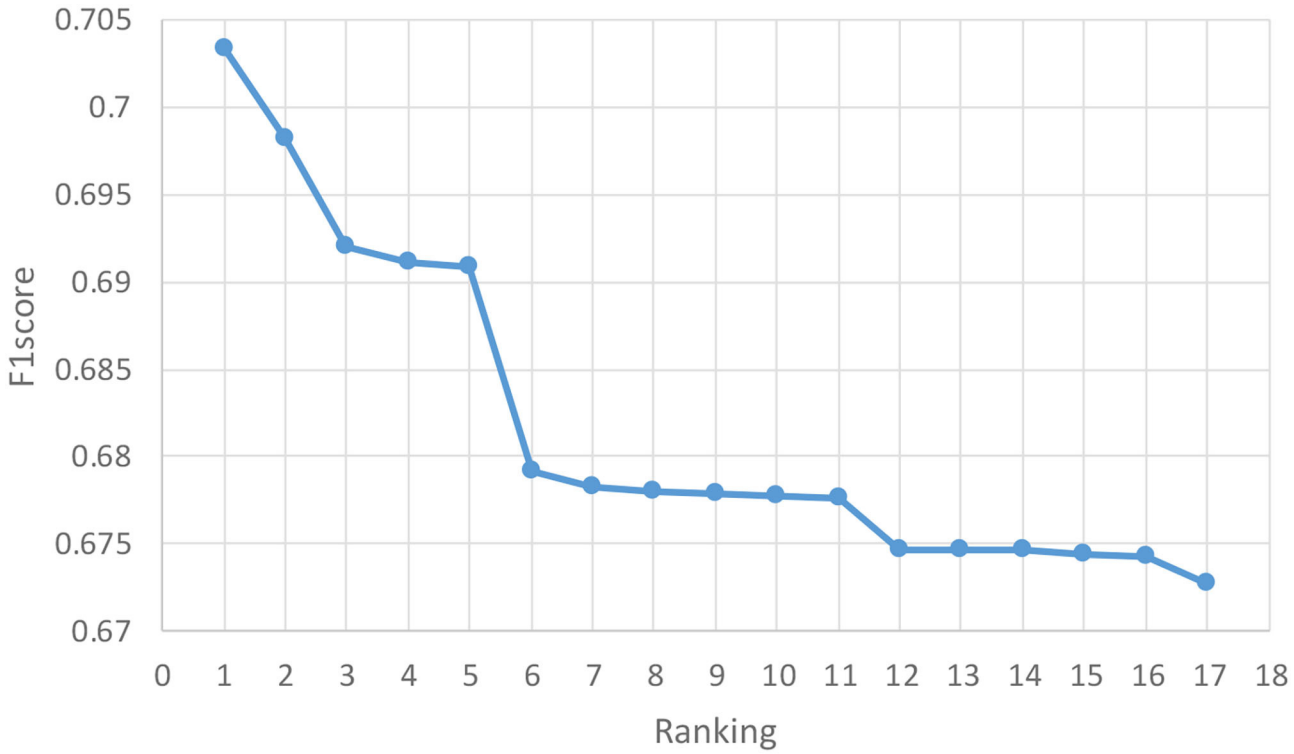


FIGURE 5 | Results of the competition. The horizontal axis represents the ranking of the competition and the vertical axis represents the score of the competition. The competition uses F1score as the final evaluation metric. The figure shows the 17 winning teams among 373 teams, among which our team gained the rank of 10th.

performs better in both the convergence speed of loss and the ability to prevent overfitting.

Discussion of PRCV2021

PRCV 2021 AD Classification Technical Challenge provides a dataset containing gray matter volumes and mean cortical thickness extracted from multiple atlases. Based on this dataset, PRCV 2021 proposes a triple classification task for AD. **Figure 5** shows the rankings and scores of all winning teams in the competition, among which our team ranks 10th. Most of the better performing teams in the competition have optimized their methods based on the MLP architecture. The adjustments on the network are as follows: combining MLP with attention mechanism, adjusting the depth of MLP network, and combining multiple networks for data processing, etc. For the processing of the dataset, some teams filtered the data based on the characteristics of the atlas or supplemented the data by interpolation.

In the competition, most of the teams used the MLP-based network and did various optimized operations. Among them, the best-performing method used a combination of MLP and attention and got the highest score of 0.7033. They added multiple attention modules to the network and connected outputs of different depths as input to the module. Compared with their method, we all used multiple different MLPs for training. The advantage of this is that it can effectively avoid the uncertainty of classification accuracy under a single model. However, their method adds an attention mechanism before obtaining the classification results so that the model can more accurately identify the characteristics related to the disease type and reduce the interference of other redundant data to solve the problem of overfitting.

There were also teams in the competition that used traditional machine learning algorithms, mainly random forests and SVM, and achieved good results. We think that traditional machine learning algorithms are also very applicable to this type of data. However, through post-competition experience sharing, we found that most of the machine learning teams focused their work on data processing and that most of the teams that won awards had a good approach to processing the dataset. Hence, in that task, the machine learning algorithms had higher requirements for data processing compared to deep learning related methods.

Similar to our method was that of the team that won fifth place. They also used different MLP networks to train data from different atlases. However, the difference is that our method removes some similar atlases before training, while their method selects the atlas based on the training results after training. After an analysis, we believe that their method is more appropriate because the correlation between the extracted results of the atlas and the disease should be judged by the model.

Compared to teams with similar scores to ours, our method still has a certain advantage. For example, the seventh-place

team uses a clever way to optimize. They train a large number of networks, and select the four with the best results to combine. Due to the small number of samples and larger number of feature values in the PRCV 2021 AD classification technology challenge dataset, most of the teams' methods suffer from overfitting problems. This is also evident in the training process, where the same model and parameters end up with a significant difference in classification accuracy. They take advantage of this feature to train a model that better fits the test set. Although this method has obtained good scores in the competition, its performance may not be good if the test set is re-divided. Compared with their method, our method is more versatile.

CONCLUSION

Against the dataset provided by the PRCV 2021 AD classification technology challenge, we propose a MAMLP model for Alzheimer's classification based on brain region data extracted by multi-atlas segmentation. The results of the experiment indicate that our model has better classification accuracy and generalization ability when targeting such datasets. Of course, our method is not optimal, as there are similarities in the ideas of the method compared to the teams ranked before us. For example, redundant data are removed by atlas selection and multiple networks are used for combination. The disadvantage is the lack of skill in training or the randomness caused by the small sample. An obvious limitation of this study is that the overfitting of the model due to the small sample has not been fully resolved. The next step is to use some small sample training methods to further improve the accuracy of the model. At the same time, compared with other teams' data processing methods, our method still has some gaps. In the face of high-dimensional data, dimensionality reduction is an important step, and if we can effectively remove some redundant data and duplicate data, we believe the classification effect of the model can become better.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: https://competition.huaweicloud.com/information/1000041489/circumstance.

AUTHOR CONTRIBUTIONS

XH supervised and managed the study, guided the experiment, and revised the manuscript. KH and JL performed the experiments, processed the data, wrote the manuscript, and were jointly responsible for its revision. XY, EC, and LC provided the technical support in the writing of the article. GW and SZ were responsible for the revision of the manuscript. All authors have read the manuscript and agreed to the published version.

FUNDING

This work was supported by Fuzhou science and technology planning Project: Development and application of intelligent management technology of Cambodian national chronic disease (No: 2020-DY-185). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

We thank the competition organizers, including Beijing University of Posts and Telecommunications and Huawei Cloud, for providing the platform and dataset for the study, and thank the National Key R&D Programme of China for providing funding support for the research.

REFERENCES

Bron, E. E., Smits, M., Van Der Flier, W. M., Vrenken, H., Barkhof, F., Scheltens, P., et al. (2015). Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge. *NeuroImage* 111, 562–579. doi: 10.1016/j.neuroimage.2015.01.048

Cheng, D., and Liu, M. (2017). "Combining convolutional and recurrent neural networks for Alzheimer's disease diagnosis using PET images", in: *2017 IEEE International Conference on Imaging Systems and Techniques (IST)* (Beijing: IEEE).

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., et al. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cereb. Cortex* 26, 3508–3526. doi: 10.1093/cercor/bhw157

Gordon, E. M., Laumann, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M., and Petersen, S. E. (2016). Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb. Cortex* 26, 288–303. doi: 10.1093/cercor/bhu239

Hammers, A., Allom, R., Koepp, M. J., Free, S. L., Myers, R., Lemieux, L., et al. (2003). Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum. Brain Mapp.* 19, 224–247. doi: 10.1002/hbm.10123

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition", in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition)*, (Las Vegas, NV: IEEE) 770–778.

Hosseini-Asl, E., Ghazal, M., Mahmoud, A., Aslantas, A., Shalaby, A. M., Casanova, M. F., et al. (2018). Alzheimer's disease diagnostics by a 3D deeply supervised adaptable convolutional network. *Front. Biosci.* 23, 584–596. doi: 10.2741/4606

Jagust, W. (2013). Vulnerable neural systems and the borderland of brain aging and neurodegeneration. *Neuron* 77, 219–234. doi: 10.1016/j.neuron.2013.01.002

Joliot, M., Jobard, G., Naveau, M., Delcroix, N., Petit, L., Zago, L., et al. (2015). AICHA: an atlas of intrinsic connectivity of homotopic areas. *J. Neurosci. Methods* 254, 46–59. doi: 10.1016/j.jneumeth.2015.07.013

Kim, Y. (2014). "Convolutional Neural Networks for Sentence Classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language*

Processing (Doha Qatar: Association for Computational Linguistics), 1746–1751.

Klöppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., et al. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain* 131, 681–689. doi: 10.1093/brain/awm319

Korolev, S., Safiullin, A., Belyaev, M., and Dodonova, Y. (2017). "Residual and plain convolutional neural networks for 3D brain MRI classification", in: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* (Melbourne: IEEE), 835–838.

Li, Q., Xing, X., Sun, Y., Xiao, B., Wei, H., Huo, Q., et al. (2019). "Novel iterative attention focusing strategy for joint pathology localization and prediction of MCI progression", in: *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Shenzhen: Springer), 307–315.

Liu, M., Zhang, D., Adeli, E., and Shen, D. (2015). Inherent structure-based multiview learning with multitemplate feature representation for Alzheimer's disease diagnosis. *IEEE. Trans. Biomed. Eng.* 63, 1473–1482. doi: 10.1109/TBME.2015.2496233

Liu, P., Qiu, X., and Huang, X. (2016). "Recurrent neural network for text classification with multi-task learning*", in: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*. (New York, NY: AAAI Press).

Ortiz, A., Munilla, J., Gorriz, J. M., and Ramirez, J. (2016). Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease. *Int. J. Neural Syst.* 26, 1650025. doi: 10.1142/S0129065716500258

Payan, A., and Montana, G. (2015). "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks," in *ICPRAM 2015 - 4th International Conference on Pattern Recognition Applications and Methods*, (Lisbon Portugal: Springer) Proceedings, Vol. 2, 2015.

Plis, S. M., Hjelm, D. R., Salakhutdinov, R., Allen, E. A., Bockholt, H. J., Long, J. D., et al. (2014). Deep learning for neuroimaging: a validation study. *Front. Neurosci.* 8, 229. doi: 10.3389/fnins.2014.00229

Reiman, E. M., Langbaum, J. B., and Tariot, P. N. (2010). Alzheimer's prevention initiative: a proposal to evaluate presymptomatic treatments as quickly as possible. *Biomark. Med.* 4, 3–14. doi: 10.2217/bmm.09.91

Rolls, E. T., Huang, C.-C., Lin, C.-P., Feng, J., and Joliot, M. (2020). Automated anatomical labelling atlas 3. *Neuroimage* 206, 116189. doi: 10.1016/j.neuroimage.2019.116189

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., et al. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* 28, 3095–3114. doi: 10.1093/cercor/bhx179

Simonyan, K., and Zisserman, A. (2014). "Very deep convolutional networks for large-scale image recognition," in *2015 ICLR International Conference on Learning Representations 2015* (San Diego, CA: arXiv.org), arXiv:1409.1556.

Thomas Yeo, B., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi: 10.1152/jn.00338.2011

Urchs, S., Armoza, J., Moreau, C., Benhajali, Y., St-Aubin, J., Orban, P., et al. (2019). MIST: a multi-resolution parcellation of functional brain networks. *MNI Open Res.* 1, 3. doi: 10.12688/mniopenres.12767.2

Velazquez, M., Anantharaman, R., Velazquez, S., and Lee, Y. (2019). "RNN-based Alzheimer's disease prediction from prodromal stage using diffusion tensor imaging", in: *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (San Diego, CA: IEEE), 1665–1672.

Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., et al. (2016). "Attention-based bidirectional long short-term memory networks for relation classification", in: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (Berlin: Association for Computational Linguistics), 207–212.

Zilles, K., and Amunts, K. (2010). Centenary of brodmann's map—conception and fate. *Nat. Rev. Neurosci.* 11, 139–145. doi: 10.1038/nrn2776

Conflict of Interest: XY was employed by Fuzhou Comvee Network and Technology Co., Ltd.

The remaining 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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 02 June 2022
doi: 10.3389/fnagi.2022.876202

Deep Learning Model for Prediction of Progressive Mild Cognitive Impairment to Alzheimer's Disease Using Structural MRI

Bing Yan Lim[1], Khin Wee Lai[1], Khairunnisa Haiskin[1], K. A. Saneera Hemantha Kulathilake[2], Zhi Chao Ong[3], Yan Chai Hum[4], Samiappan Dhanalakshmi[5], Xiang Wu[6] and Xiaowei Zuo[7]**

[1] Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia, [2] Department of Computing, Faculty of Applied Sciences, Rajarata University of Sri Lanka, Anuradhapura, Sri Lanka, [3] Department of Mechanical Engineering, Universiti Malaya, Kuala Lumpur, Malaysia, [4] Department of Mechatronics and Biomedical Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Petaling Jaya, Malaysia, [5] Department of Electronics and Communication Engineering, Faculty of Engineering and Technology, College of Engineering and Technology, SRM Institute of Science and Technology, Chennai, India, [6] School of Medical Information & Engineering, Xuzhou Medical University, Xuzhou, China, [7] Department of Psychiatry, The Affiliated Xuzhou Oriental Hospital of Xuzhou Medical University, Xuzhou, China

OPEN ACCESS

Edited by:
Sang-Bing Tsai, Wuyi University, China

Reviewed by:
Juan Yang, Suzhou University, China
Yufeng Yao, Changshu Institute of Technology, China

***Correspondence:**
Khin Wee Lai
lai.khinwee@um.edu.my
Xiaowei Zuo
2783369510@qq.com

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

Received: *15 February 2022*
Accepted: *11 May 2022*
Published: *02 June 2022*

Citation:
Lim BY, Lai KW, Haiskin K, Kulathilake KASH, Ong ZC, Hum YC, Dhanalakshmi S, Wu X and Zuo X (2022) Deep Learning Model for Prediction of Progressive Mild Cognitive Impairment to Alzheimer's Disease Using Structural MRI. Front. Aging Neurosci. 14:876202. doi: 10.3389/fnagi.2022.876202

Alzheimer's disease (AD) is an irreversible neurological disorder that affects the vast majority of dementia cases, leading patients to experience gradual memory loss and cognitive function decline. Despite the lack of a cure, early detection of Alzheimer's disease permits the provision of preventive medication to slow the disease's progression. The objective of this project is to develop a computer-aided method based on a deep learning model to distinguish Alzheimer's disease (AD) from cognitively normal and its early stage, mild cognitive impairment (MCI), by just using structural MRI (sMRI). To attain this purpose, we proposed a multiclass classification method based on 3D T1-weight brain sMRI images from the ADNI database. Axial brain images were extracted from 3D MRI and fed into the convolutional neural network (CNN) for multiclass classification. Three separate models were tested: a CNN built from scratch, VGG-16, and ResNet-50. As a feature extractor, the VGG-16 and ResNet-50 convolutional bases trained on the ImageNet dataset were employed. To achieve classification, a new densely connected classifier was implemented on top of the convolutional bases.

Keywords: Alzheimer's disease, deep learning, prediction, magnetic resonance imaging, mild cognitive impairment

INTRODUCTION

Alzheimer's disease (AD) is a progressive disease that causes neuronal loss and dementia in the elderly. Alzheimer's disease patients typically exhibit progressive memory loss at the outset, followed by cognitive decline and, eventually, loss of independence. It is predicted that by 2050, one out of every 85 people in the world will have AD (Brookmeyer et al., 2007). At the moment, there are approximately 90 million people who have been identified as having AD, and the number of diseased patients is expected to reach 300 million by 2050 (Zhu et al., 2015).

There are medications that can provide temporary moderate symptom relief or slow the progression of AD, and these treatments have been shown to help patients with AD by achieving

maximum cognitive function and maintaining independence for a period of time. However, there are currently no effective or safe drugs or therapies for curing Alzheimer's disease or altering the disease process in the brain (Tatiparti et al., 2020). The search for effective strategies to treat or prevent Alzheimer's disease remains one of the most difficult endeavors in medicine. As a result, it is critical to detect Alzheimer's disease in its early or prodromal stages so that patients can receive treatment before the disease progresses. Currently, the standard non-invasive clinical strategy for performing prognostic prediction for Alzheimer's disease is manual assessment *via* structural neuroimaging such as magnetic resonance imaging (MRI) or computed tomography (CT). Computer-aided methods based on artificial intelligence (AI) algorithms are currently being used to accomplish AD diagnostics (Wen et al., 2020).

In tandem with the rapid growth of AI, academics have been employing AI techniques such as deep learning to address complex problems in a variety of sectors, particularly medicine. Researchers have extended the use of multiple deep learning models to diagnose various stages of Alzheimer's disease. Current neuroimaging investigations that use computer-aided system studies have made substantial progress in classifying Alzheimer's disease (AD) and cognitively normal (CN) participants. Even though the binary classification of AD and CN participants performed admirably, it is not as useful as predicting the early-stage change of moderate cognitive impairment (MCI) to AD. The majority of research stopped at a binary categorization, without predicting whether a patient had MCI or the likelihood of converting to AD.

Detecting Alzheimer's disease in its prodromal stage, or anticipating its potential, is critical for its treatment, just as it is for other diseases. Treatments are successful if AD patients receive them as soon as feasible after being suspected of having AD biomarkers or symptoms. A 1-year delay in the progression of Alzheimer's disease can decrease the number of afflicted people by 10% (McKhann et al., 2011). According to the statistics, detecting Alzheimer's disease in its early stages is critical to reduce the number of patients worldwide.

Neurologists must manually study brain scans and undertake cognitive assessments during the diagnosis of Alzheimer's disease in order to make an accurate diagnosis of the symptoms and course of the disease. Because subtle changes in brain anatomy can be observed years before distinct biomarkers can be visualized by humans, it is realized that the human visual system is incapable of identifying subtle changes in underlying brain structure that may contain vital information about a patient's disease state, even when the analysis is performed by the experienced neurologists. As a result, an AI-based computer-aided system can assist neurologists in detecting complicated brain illnesses while reducing the potential for misdiagnosis. Moreover, it is expected to decrease the workload on medical professionals and reduce the frequency of patient visit and waiting time. Many recent studies (Basaia et al., 2019; Bi et al., 2020; Jiang et al., 2020; Guo et al., 2021; Hett et al., 2021; Mehdipour Ghazi et al., 2021; Deng et al., 2022) have been conducted to forecast early stages of Alzheimer's disease. The goal of this study is to build a computer-aided system based on a deep learning algorithm to evaluate the pathological brain structural changes in MRI data in order to forecast the early stages of Alzheimer's disease before it progresses to the severe stages. The contributions of this proposed study are as follows:

1. Performing novel preprocessing procedures on brain structural MRI used for training and testing the convolutional neural network.
2. Implementing CNN to perform multiclass classification (3-way) to classify cognitively normal (CN), MCI, and AD subjects.
3. Evaluating the performance by metrics such as accuracy, precision, recall, and F1-score.

The rest of this article is organized as follows: the following section discusses the materials and methods used in this study and the subsequent section elaborates the experimental results and discussion. The final section emphasizes the conclusion and future research directions.

MATERIALS AND METHODS

Dataset—Alzheimer's Disease Neuroimaging Initiative

The relevant data were retrieved from the database of the Alzheimer's Disease Neuroimaging Initiative (ADNI), which is available publicly upon approval from the ADNI. The ADNI database contains multiple collections of MRI images categorized by phase of the study, for example, ADNI1, ADNI2, ADNI-GO, and ADNI3 (as of August 2021). This study adopts all the sMRI data in the ADNI1 collection. A total of 819 subjects (229 CN, 398 with MCI, and 192 with AD) were enrolled at baseline. The CN class consists of healthy aging controls with no conversion within 3 years of follow-up visits from baseline. Subjects diagnosed with mild cognitive problems without losing their ability to carry out daily activities were retained in the MCI class. The AD class comprises patients identified as AD through diagnosis at baseline and exhibit no sign of reversion within 2 years of follow-up visits.

All the acquired sMRI were generated from scanners of various manufacturers, such as Philips, Siemens, and General Electric. On account of the various acquisition protocols, the dataset will undergo a preprocessing procedure. There is 1.2 mm spacing between two MRI scans, and the dimension of a voxel is $256 \times 256 \times 256$. In terms of resolution, there is only a slight difference found across the patients. The data used were restricted to the standard 1.5 T T1-weighted sMRI, which were acquired by the volumetric three-dimensional magnetization-prepared rapid gradient-echo (3DMPRAGE) protocol. Other data acquisition settings include 8-channel coil, TR = 650 ms, TE = minimum full, flip-angle = 8°, and FOV = 26 cm. Participants may have multiple scans at baseline and follow-up visits (after 1, 2, and 3 years).

The data used were restricted to the standard 1.5 T T1-weighted sMRI, which were acquired by the volumetric three-dimensional magnetization-prepared rapid gradient-echo (3DMPRAGE) protocol. Other data acquisition settings include 8-channel coil, TR = 650 ms, TE = minimum full, flip-angle = 8°,

and FOV = 26 cm. Participants may have multiple scans at baseline and follow-up visits (after 1, 2, and 3 years). It is important to note that not all participants appeared at every planned follow-up visit. Some participants were retained in the study without appearing at every follow-up meeting. There was also a significant decrease in follow-up visit rate after 2 years, indicating that fewer data were available over time. **Table 1** summarizes the demographic information for the 819 subjects with age ranges from 55 to 92 years, including 192 patients with AD, 398 subjects belonging to the MCI, and 192 who are cognitively normal and were included in the study. Based on **Table 1**, it can be seen that the CN group is more educated than the MCI and AD groups with mean education years of 16.0 ± 2.9 years, and the MCI group is the youngest among the three groups with a mean age of 74.7 ± 7.4 years.

TABLE 1 | Demographic of participants with MCI and AD and cognitive normal subjects from the study population.

Diagnostic type	Number of participants	Age	Gender (M/F)	Education (years)
CN	229	75.8 ± 5.0 (59.9–89.6)	119/110	16.0 ± 2.9 (6–20)
MCI	398	74.7 ± 7.4 (54.5–89.3)	257/141	15.7 ± 3.0 (4–20)
AD	192	75.3 ± 7.5 (55.1–90.9)	101/91	14.7 ± 3.1 (4–20)

Proposed Model

The process of the proposed approach is depicted in **Figure 1**. As a result, the acquired ADNI1 dataset is initially subjected to a number of preprocessing methods. The retrieved 2D images are then divided into training, validation, and testing sets. Three CNN models are evaluated: a CNN trained from scratch, VGG-16, and ResNet-50. The training data was supplemented before feeding the training data into the CNN models for training.

Preprocessing

Preprocessing was applied to each brain sMRI to normalize the data into desired form and format. The routine of preprocessing steps can be summarized into six different steps: (1) skull-stripping, (2) non-uniform intensity correction, (3) segmentation, (4) extraction of 2D image from 3D MRI volume, (5) pixel values normalization, and (6) data augmentation.

Skull Stripping

Skull stripping is the removal of the skull from a 3D brain MRI. For quantitative morphometric studies, the skull is the non-brain tissue that functions as noise, lowering CNN classification performance (Goceri and Songül, 2017). Aside from that, removing the skull from the brain enhances segmentation outcomes. To obtain solely the brain tissues, the skull section was stripped or deleted using the DeepBrain library. **Figures 2A,B** depicts the raw brain had its skull stripped together with intensity normalized using the DeepBrain library.

Bias Field Correction

Strong bias fields are known to cause voxel tissue type mislabeling, undermining the algorithm's accuracy, which is based on gray and white matter contrast (Gupta et al., 2019). To keep this impact to a minimum, the N4 bias field correction method was used in conjunction with the SimpleITK library for correcting low-frequency intensity presented non-uniformly in brain sMRI (Tustison et al., 2010). Following that, the intensity variation of the same brain tissue was deleted based on its location within the image. The bias-corrected brain displayed more consistent intensity in the white matter region (**Figure 2C**).

Tissue Segmentation

The hidden Markov random field (HMRF) tissue classifier was used to segment T1-weighted sMRI data that had previously been skull-stripped and bias field corrected (Zhang et al., 2001). The hidden Markov models were used to develop the HMRF idea. In contrast to hidden Markov, HMRF features an underlying Markov random field rather than a Markov chain.

The brain sMRI volumes were segmented into three different regions of GM, WM, and CSF using the HMRF tissue classifier from the DIPY library. These three main features were used to differentiate AD from MCI and CN. Alterations in WM and GM were commonly used for the analysis of AD progression (Klöppel et al., 2008). In ML approach studies, it would be laborious to perform tissue segmentation and feature extraction. Hence, automated segmentation is essential for a dataset with a large number of images. **Figure 2D** shows the plotting of the resulting segmentation with a clear separation between GM, WM, and CSF.

Extraction of 2D Images From the 3D Volume

The Matplotlib library was used to extract 2D slices or images from the segmented 3D MRI after the segmentation phase. More specifically, brain pictures in PNG format were recovered from the axial view of the 3D MRI slices ranging from the 160th to 170th slice. Slices in this range provide a wealth of information about the GM, WM, and CSF. For the three courses, a total of 2,387 brain scans were performed (CN, MCI, and AD). Good model performance is associated with selecting the best available slices containing relevant morphological information (Stoeckel and Fung, 2005). Given the preferable slice range, every interval of five slices (e.g., 160th, 165th, and 170th) of three brain images were extracted from the MRI volume of the AD and CN subjects, in which AD and MCI have 2,043 and 2,051 images, respectively. One scan was removed from the CN class due to file corruption. In addition, two brain images (160th and 165th) were extracted for the MCI class that yielded 2,044 images.

A padding private function was implemented to add padding to all final images, so that the output images have a uniform dimension of 271 × 271 pixels. Here, the images were saved in gray scale format and named according to their classes with a number suffix in an increasing sequence. After preprocessing, the data were all in the form of 2D images. This helps to substantially reduce the dataset size from 37 GB to 260 MB.

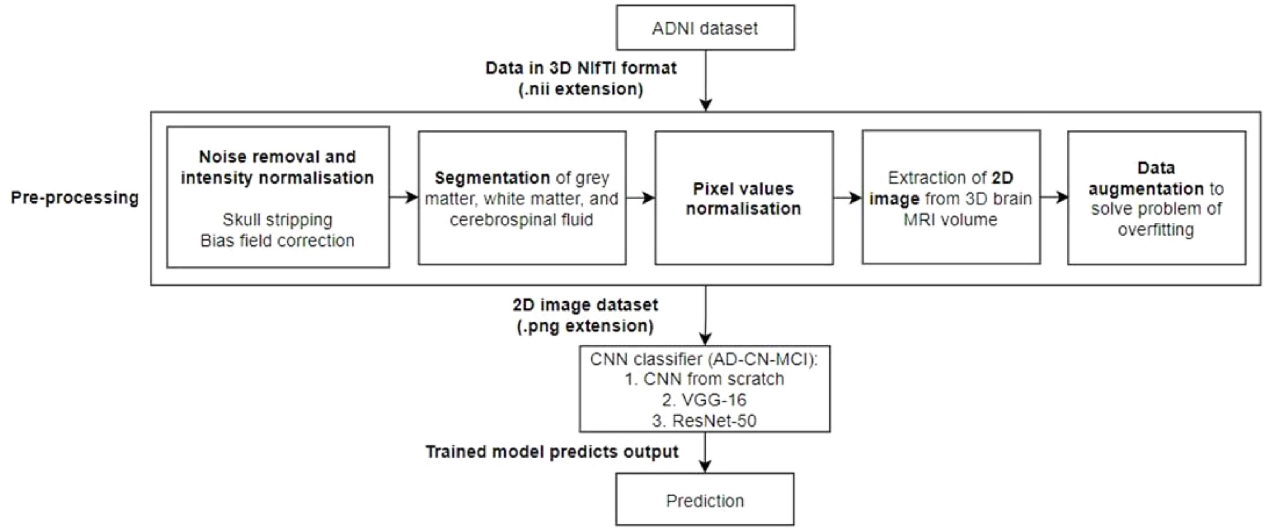


FIGURE 1 | Workflow of the proposed model. A preprocessing steps, which include noise removal and intensity normalization, segmentation, pixel value normalization, and 2D image extraction and data augmentation, will be performed followed by the classification by an AD-CN-MCI CNN classifiers.

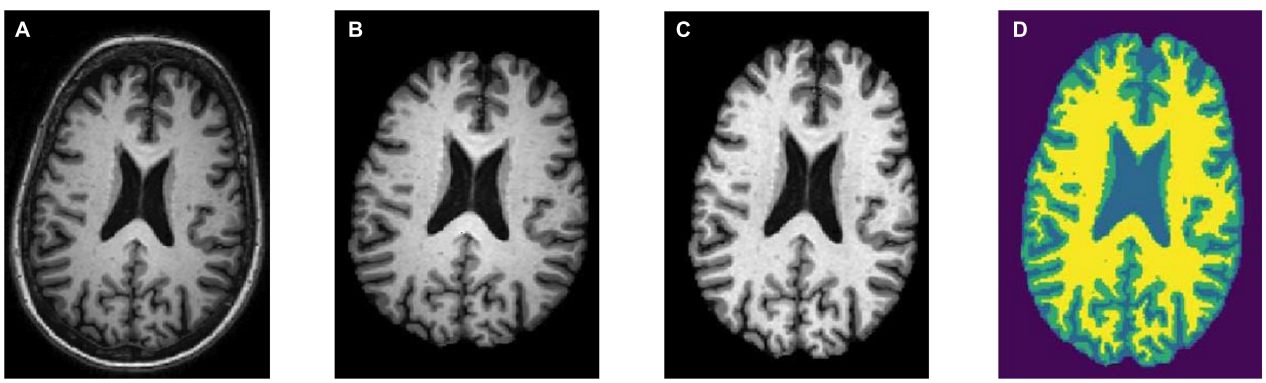


FIGURE 2 | Preprocessing steps: **(A)** raw bran MRI, **(B)** skull stripped MRI, **(C)** bias field corrected MRI, and **(D)** tissue segmented MRI (WM is denoted in yellow, GM is denoted in green, and CSF is denoted in light blue).

Pixel Values Normalization

As of this stage, every image data were in gray scale with pixel values ranging between 0 and 255 (8-bit). Before the training process, we normalize every image pixel value with a value between 0 and 1.

Data Augmentation

The process of data augmentation was performed to mitigate the general problem of the small dataset, which is overfitting during training, by applying various transformations on the images from the dataset. The transformations used were rotation of 15°, zoom range of 0.10°, height shift range of 0.10°, and width shift range of 0.10°.

Prediction Model

The CNN models used in this study will be described in detail here. To perform the 3-way classification task, three different CNN models were tested. The first model is a CNN that was trained from the ground up. Furthermore, the second and third models used the transfer learning technique. CNN models with pretrained ImageNet weights, such as VGG-16 and ResNet-50, were used instead of training a model from scratch. These models were trained to classify 1,000 different image classes using the ImageNet database, which contains over a million images.

Convolutional Neural Network From Scratch

Figure 3 depicts the 2D CNN architecture that was created from scratch. In a nutshell, the architecture consists of the following elements: five convolutional layers followed by ReLU activation; five max-pooling layers; two dropout layers; a flatten layer; a fully connected layer with 256 neurons followed by a dropout layer and a batch normalization layer; and, finally, an output layer with softmax activation that outputs the probability of prediction for each class.

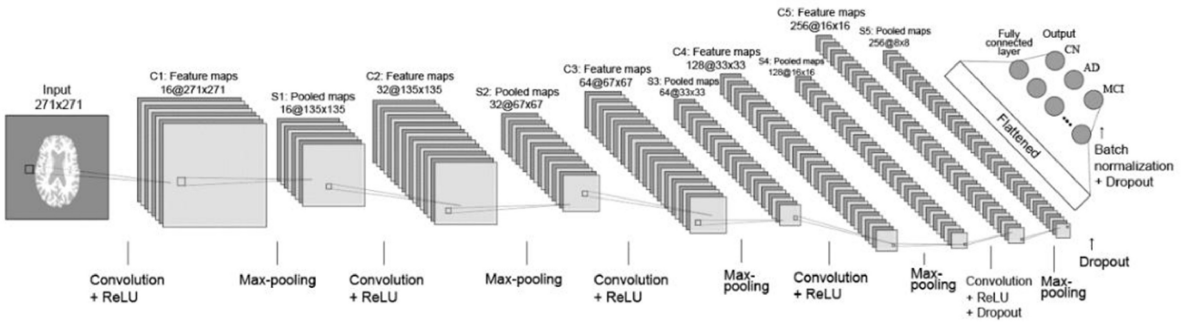


FIGURE 3 | Layout of CNN trained from scratch. Briefly, the architecture comprises the following: 5 convolutional layers followed by ReLU activation; 5 layers of max-pooling layers; 2 dropout layers; a flatten layer; a fully connected layer with 256 neurons followed by a dropout layer and a batch normalization layer; and ultimately an output layer with softmax activation, which outputs the probability of prediction for each class.

The CNN first layer was fed preprocessed axial view brain sMRI data. The second layer was a convolutional layer that performed convolution operations on input images and filtered the output to produce multiple feature maps. There were five convolution layers in total, each with 16-32-64-128-256 feature maps. All of the convolution filters had a size of 22, a stride of one, and "same" padding, which ensured that the output was the same size as the input. After each convolutional layer, a max-pooling layer with 22 regions was applied. The pooling layers functioned as down-sampling layers, resulting in the creation of multiple pooled maps. The final two pooling layers were followed by a dropout layer with a dropout rate of 0.5, which meant that 50% of the nodes in the layers would be dropped out to ensure regularization and prevent overfitting. The pooled feature maps were then flattened to a 1D vector and fed into the next fully connected layer with 256 neurons. A batch normalization layer was added before the dropout layer to improve the model's regularization even further. The final layer is the output layer with three nodes incorporating softmax activation function to determine the probabilities of each possible class of the classification task. Finally, a vector consisting of probabilities belonging to the AD, CN, and MCI classes was obtained as the final classification result.

VGG-16

In this study, the pretrained VGG-16 model was used in the form of a feature extractor (Simonyan and Zisserman, 2015). Also, VGG-16 with pretrained weights was used as a bootstrap feature extractor for feature extraction from the preprocessed brain sMRI images. The extracted features were then directed to a new classifier, which was trained from scratch.

It is important to note that the gray scale image dataset could not be directly fed to the VGG-16 model because it is a pretrained model with a fixed input configuration. VGG-16 requires RGB images with three channels as input. A gray scale image, on the contrary, has only one channel. The obvious solution is to iteratively repeat all of the image arrays in the dataset three times on a new dimension. As a result, the same image would appear in all three channels. This was accomplished by specifying the color mode as "RGB" in the Keras library's flow from directory method.

ResNet-50

The pretrained ResNet-50 model was used as a feature extractor, similar to VGG-16, and a new densely connected classifier was used for prediction (He et al., 2016). Deep neural network training is difficult because adding more layers causes the infamous vanishing gradient problem, also known as the exploding gradient problem. The main feature of ResNet is the design of residual connections. The residual block enabled ResNet to connect the previous layer to the current layer as well as the layer behind the previous layer. As a result, each layer can capture more than just the observations of the previous layer. Furthermore, the batch normalization layer is placed after each convolutional layer in ResNet. Batch normalizations normalize layer weights, allowing for faster training rates. This speeds up deep network training and reduces the vanishing gradient problem.

Parameters and Evaluation Metrics

Table 2 summarizes the best parameter combinations for training the three networks. In addition, the evaluation metrics were accuracy, precision, recall, and F1-score. Keras, an open-source high-level neural network API for building deep models, was used to build all of the deep learning models, with TensorFlow as the backend. Keras was chosen because it enables rapid prototyping and parallel computing with GPUs. In this study, training, validation, and testing routines were carried out on Google Colab in order to execute Python 3 codes for data preprocessing and the development of a CNN model. The GPU model would be assigned at random based on the availability on Google Colab. There was no published limit on the idle timeout period, RAM size, or disc size. Typically, a RAM size of around 13 GB and a disc size of around 70 GB would be allocated for GPU accelerated runtime.

In addition, to facilitate model training, two types of "callbacks" in Keras were implemented during training: Early Stop and ModelCheckpoint. Early Stop enabled the models to stop training if their performance did not improve after five epochs of monitoring validation loss. This is one method for preventing a model from overfitting. Next, ModelCheckpoint ensured that models always saved the best weights while training

to avoid loss of progression. Saving the weights is more efficient than saving the entire model's information because a large network like VGG-16 can take up at least 500 MB of memory.

RESULTS AND DISCUSSION

Training and Validation Performance

Table 3 reports the training and validation performance of the three different CNN models being experimented.

To avoid overfitting, all model training was completed with an early stop and a patience level of 15 epochs. The training and validation routines were halted when the validation loss began to deteriorate. The CNN trained from scratch finished training in 46 min, making it the quickest of the three models. Deep CNN, such as VGG-16 and ResNet-50, with multiple stacking layers, can be computationally expensive and take much longer to train than a shallow model trained from scratch. ResNet-50's longer training time can be attributed to a large number of trainable parameters. The ResNet-50 model, which has a frozen convolutional base and a swapped densely connected classifier, has 42.5 million trainable parameters.

The VGG-16 model, on the contrary, had an identical densely connected classifier and a frozen convolutional base with 8.4 million trainable parameters. Interestingly, despite having five times the number of trainable parameters as VGG-16, ResNet-50 spent only 21.33% more time on training. The inclusion of multiple batch normalization layers between the convolutional layer and the non-linear activation function may be the primary reason for this, allowing a higher learning rate to be used (He et al., 2016).

The loss function quantifies a model's performance in classifying input images from a dataset. The loss value indicates how well a model performs after each optimization epoch. The goal of training a deep learning network is to minimize the error calculated using the loss function while increasing testing accuracy. VGG-16 achieved a training loss value of 0.1511, while CNN from scratch and ResNet-50 achieved training loss values of 0.3102 and 0.2150, respectively. In terms of validation performance as measured by loss value, VGG-16 achieved the lowest loss value of 0.5263. ResNet-50 came in second with a loss value of 0.5901, and CNN from scratch came in third with a loss value of 0.7094.

The transfer learning method was tested for its ability to produce satisfactory results on small datasets, as seen in recent literature. Deep models with pretrained weights, such as VGG-16 and ResNet-50, were used for feature extraction instead of learning the convolutional bases from scratch. To improve the output classification scores, a new densely connected classifier trained from scratch was added to both models. Both VGG-16 and ResNet-50 outperform the CNN trained from scratch in this case. Despite the use of various regularization methods, such as dropout, batch normalization, and data augmentation, the overfitting problem persists in both models.

TABLE 2 | Hyperparameters of CNNs adopted in the experiments.

Parameter	CNN	VGG-16	ResNet50
Number of epochs	100	100	100
Batch size	512	256	256
Weight initializer	Xavier uniform	Xavier uniform	Xavier uniform
Optimizer	Adam	Adam	Adam
Adam parameters	β1 = 0.9, β2 = 0.999	β1 = 0.9, β2 = 0.999	β1 = 0.9, β2 = 0.999
Learning rate	10-4	10-5	10-4
Loss function	Categorical cross-entropy	Categorical cross-entropy	Categorical cross-entropy
Metrics	Accuracy	Accuracy	Accuracy
Data augmentation	Rotation, zoom, height shift, width shift, shear, horizontal flip	Rotation, zoom, height shift, width shift, shear, horizontal flip	Rotation, zoom, height shift, width shift, shear, horizontal flip

All the architectures adopted Xavier's uniform as the weight initializer and Adam as the optimizer.

TABLE 3 | Summary of training and validation performance.

Model	Training time (minutes)	Steps	Training		Validation	
			Accuracy	Loss	Accuracy	Loss
CNN	46	97	0.8755	0.3102	0.7270	0.7094
VGG-16	75	57	0.9492	0.1511	0.8066	0.5263
ResNet-50	91	56	0.9164	0.2150	0.7686	0.5901

Testing Performance and Discussion

After all of the models had been trained and validated, the 20% held out testing data were run on each and every model. The confusion matrix was used as a tool to assess model classification performance, along with a summary of prediction results. The number of correctly or incorrectly predicted predictions is summarized systematically in a table, with count values broken down by class. The confusion matrix is a table with three rows and three columns because it is a three-way classification task with three different classes. The predicted lab is represented by the rows (y-axis), and the predicted label is represented by the columns (x-axis). **Figure 4** depicts confusion matrices that describe each model's classification performance on test data. Using the seaborne library, each confusion matrix is visualized as a color-coded heat map. The darker cells for the diagonal elements can be seen in all of the plotted confusion matrices. This indicates that a large amount of data is correctly predicted according to its label. The off-diagonal elements with light shades, on the contrary, indicate model misclassifications.

The CNN predicted the MCI group with the highest accuracy and the CN group with the lowest accuracy when trained from scratch. It correctly classified 304 of 409 MCI images and 291 of 408 CN images. In contrast, the AD group has the highest classification accuracy in VGG-16 and ResNet-50, while the MCI group has the lowest classification accuracy. ResNet-50 classified 341 AD images out of 410 AD images predicted by VGG-16.

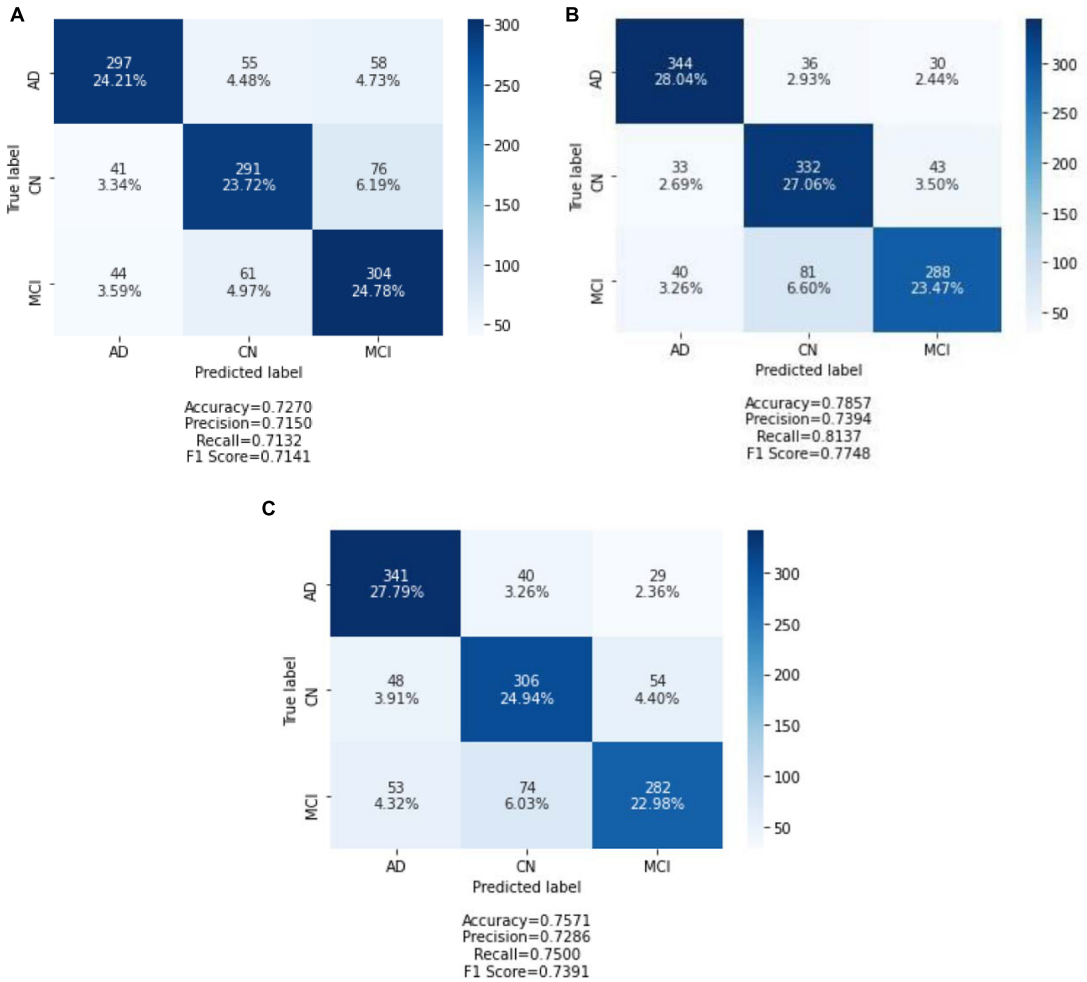


FIGURE 4 | Confusion matrix of three models on test data: **(A)** CNN from scratch; **(B)** VGG-16; and **(C)** ResNet-50. Each of the confusion matrices is visualized as a color-coded heat map using the seaborne library. It can be observed that all the plotted confusion matrices have darker cells for the diagonal elements. This indicates that more data are being predicted correctly to their respective label. Conversely, the off-diagonal elements with light shades indicate misclassifications done by the model.

In the MCI group, VGG-16 correctly predicted 288 images, and ResNet-50 correctly classified 282 of 409 AD images.

To further evaluate the classification model, classification metrics such as accuracy, precision, recall, and F1-score were calculated with the aid of the confusion matrices. For each classification model (CNN from scratch, VGG-16, and ResNet-50), the reported classification performance on test data is accuracy of 72.70, 78.57, and 75.71%, respectively; precision of 71.50, 73.94, and 72.86%, respectively; recall of 71.32, 81.37, and 75.00%, respectively; and F1-score of 71.41, 77.48, and 73.91%, respectively. Based on **Figure 5**, it is observed that VGG-16, which achieved the lowest loss value of 0.5263, performed the best on test data with an accuracy of 78.57%. The lowest testing accuracy of 72.70% is obtained using the CNN from scratch.

For further in-depth evaluation of performance on test data, the classification results for each class label are reported in **Table 4**. Similar to what was being analyzed using the confusion matrices, the AD group has the highest accuracy value for VGG-16 and ResNet-50. VGG-16 performed the greatest in predicting AD class with an accuracy of 83.90%, precision of 82.49%, recall of 83.90%, and F1-score of 83.19%. Interestingly, ResNet-50 has the lowest accuracy in predicting the MCI class. Overall, using VGG-16 improved the performance values for all three classes.

The AD group scored the highest accuracy value for VGG-16 and ResNet-50. VGG-16 performed the greatest in predicting AD class with an accuracy of 83.90%, precision of 82.49%, recall of 83.90%, and F1-score of 83.19%.

DISCUSSION

From the results obtained, the VGG-16 model outperformed the CNN trained from scratch and the ResNet-50 model. It has the best testing performance with an accuracy of 78.57%, precision of 73.94%, recall of 81.37%, and F1-score of 77.48%.

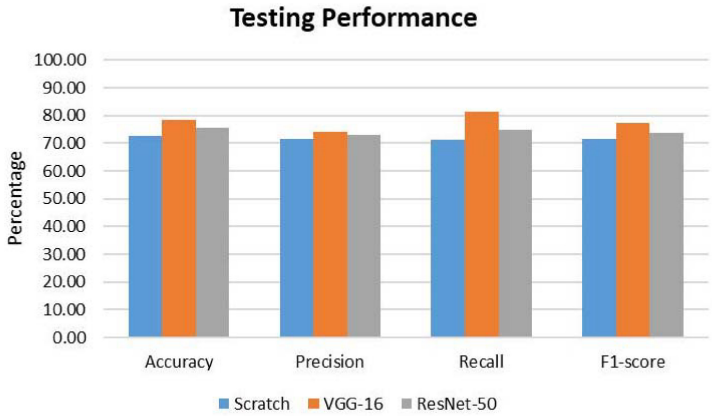


FIGURE 5 | Comparison of classification performance on test data. For all the metrics, VGG-16 ranks the highest.

Comparing its performance to other related works, VGG-16 has a performance below the average. Being trained on the ImageNet dataset, VGG-16 was able to extract representations using its convolutional base for learning the multiclass classification task. Despite the great performance on learning the representations, VGG-16 still encountered the typical overfitting problem due to the small dataset used. Several regularization methods were used, such as dropout, batch normalization, data augmentation, and early stopping. However, the signs of overfitting can still be noticed. This could be due to the high complexity of the classification task. The subtle discrepancies between the MCI and AD images require a large amount of data to learn the representation to classify them. With the small dataset being used in this project, the VGG-16 model could not learn the problem completely, hence the overfitting problem. Another possible reason could be that the dataset being used in this project has substantial differences as compared to the ImageNet dataset. The VGG-16 was pretrained on general images from the ImageNet without including medical images. Hence, the high-level features learned by the higher layers of the VGG-16 are not sufficient to differentiate the classes in this study.

Based on the results obtained, it is of importance to choose a proper training strategy for the model. Hence, the model is able to spend the least time training while trying to cover as many cases as possible. An adequate model capacity is essential for model generalization. Model depth should be kept as small as possible to prevent a model from overfitting on training data. The greater the depth, the more cases the model can memorize. As a consequence, the final system will perform worse on unseen data. Another possible reason behind inferior performance could be insufficient data augmentation. The data augmentation used is insufficient to generate diversity for the original dataset. An example of aggressive data augmentation can be seen in the study by Basaia et al. (2019). Apart from general augmentation transformations such as rotation, zooming, and scaling, the study implemented deformation, cropping, and flipping.

The strengths of this study are elaborated as follows. In general, most of the studies emphasized performing binary classification of different phenotypes of AD. In this study, three different classes (AD, CN, and MCI) are classified directly using a single classifier. This study is less common as most of the studies deal with the problem of multiple class labels by dividing the problem into several binary sub-problems. Moreover, tissue-segmented sMRI brain images were used, which substantially lower the requirement for computational costs in terms of power and time. Second, MRI images were segmented into GM, WM, and CSF for training and testing the model. Moreover, models were tested using an independent set of images held out from the dataset. In addition, the performance of popular deep transfer learning models such as VGG-16 and ResNet-50 was evaluated to study their performance on images not from the ImageNet domain.

TABLE 4 | Testing accuracy, precision, recall, and F1-score for all class label.

Model	Class label	Accuracy	Precision	Recall	F1-score
CNN from scratch	AD	0.7244	0.7775	0.7244	0.7500
	CN	0.7132	0.7150	0.7132	0.7141
	MCI	0.7433	0.6941	0.7433	0.7178
VGG-16	AD	**0.8390**	**0.8249**	**0.8390**	**0.8319**
	CN	0.8137	0.7394	0.8137	0.7748
	MCI	0.7042	0.7978	0.7042	0.7481
ResNet-50	AD	0.8317	0.7715	0.8317	0.8005
	CN	0.7500	0.7286	0.7500	0.7391
	MCI	0.6895	0.7726	0.6895	0.7287

Bold values are highest value.

CONCLUSION

In this study, we have conducted a series of experiments with different deep learning CNN architectures on preprocessed axial sMRI brain images retrieved from the ADNI database. To address the problem of classifying brain sMRI images of three distinct classes of AD, CN, and MCI, three different CNN models were

built, namely, a CNN from scratch, VGG-16, and ResNet-50. The VGG-16 model outperforms the other two models in testing. The results show that, despite being trained on general images from the ImageNet dataset, VGG-16 is capable of extracting relevant features for the classification task. Using the same dataset, the pretrained VGG-16 outperforms shallow CNN and classical machine learning algorithms. However, its performance is considered subpar when compared to other literature, which also employed deep learning techniques. Increasing the number of data for training is the main factor for improving classification performance. This project serves as a catalyst to motivate further study on computer-assisted AD diagnosis systems that can provide automated early diagnosis of AD and the detection of more phenotypes of AD.

For future studies, a list of improvements can be suggested. Effort should be devoted in attempting different pretrained CNN families such as AlexNet (Krizhevsky, 2014), Xception (Chollet, 2017), Inception (Szegedy et al., 2016), MobileNet (Howard et al., 2017), and other variants of VGG and ResNet as well as the more recent state-of-the-art network (Tan and Le, 2019, 2021) as base model for feature extraction. Furthermore, classification performance could be improved through fine-tuning. Unfreeze some layers, or even half of the model, for training with the classifier at a slower learning rate. Finally, a few different methodologies for improving classification performance to distinguish between AD, CN, and MCI may be investigated in the future. One method is to include multimodal data in the study. Multimodal research necessitates feature fusion to combine features from various modalities into a single feature vector. Another method for improving performance is to enrich the feature learning process by fusing low-dimensional features like clinical scores with the MRI features space.

DATA AVAILABILITY STATEMENT

The data were available from the database of the Alzheimer's Disease Neuroimaging Initiative (ADNI), which is available publicly upon approval from the ADNI.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This study was supported by the Universiti Malaya under Grant Nos. PV052-2019 and ACU UK under IF063-2021.

REFERENCES

Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., et al. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *Neuroimage Clin.* 21:101645. doi: 10.1016/j.nicl.2018.101645

Bi, X., Li, S., Xiao, B., Li, Y., Wang, G., and Ma, X. (2020). Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology. *Neurocomputing* 392, 296–304. doi: 10.1016/j.neucom.2018.11.111

Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement.* 3, 186–191. doi: 10.1016/j.jalz.2007.04.381

Chollet, F. (2017). "Xception: deep learning with depthwise separable convolutions," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (Honolulu, HI: IEEE), 1800–1807.

Deng, H., Zhang, Y., Li, R., Hu, C., Feng, Z., and Li, H. (2022). Combining residual attention mechanisms and generative adversarial networks for hippocampus segmentation. *Tsinghua Sci. Technol.* 27, 68–78. doi: 10.26599/TST.2020.9010056

Goceri, E., and Songül, C. (2017). "Automated detection and extraction of skull from MR head images: preliminary results," in *Proceedings of the 2017 International Conference on Computer Science and Engineering (UBMK)*, (Antalya: IEEE), 171–176. doi: 10.1109/UBMK.2017.8093370

Guo, X., Chen, K., Chen, Y., Xiong, C., Su, Y., Yao, L., et al. (2021). A computational Monte Carlo simulation strategy to determine the temporal ordering of abnormal age onset among biomarkers of Alzheimers disease. *IEEE/ACM Trans. Comput. Biol. Bioinform.* [Online ahead of print]. doi: 10.1109/TCBB.2021.3106939

Gupta, Y., Lee, K. H., Choi, K. Y., Lee, J. J., Kim, B. C., Kwon, G. R., et al. (2019). Early diagnosis of Alzheimer's disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of MRI T1 brain images. *PLoS One* 14:e0222446. doi: 10.1371/journal.pone.0222446

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (San Juan, PR: IEEE), 770–778. doi: 10.1109/CVPR.2016.90

Hett, K., Ta, V. T., Oguz, I., Manjón, J. V., and Coupé, P. (2021). Multi-scale graph-based grading for Alzheimer's disease prediction. *Med. Image Anal.* 67:101850. doi: 10.1016/j.media.2020.101850

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). MobileNets: efficient convolutional neural networks for mobile vision applications. *arxiv[Preprint]* Available online at: http://arxiv.org/abs/1704.04861 [accessed on April 17, 2017]

Jiang, J., Kang, L., Huang, J., and Zhang, T. (2020). Deep learning based mild cognitive impairment diagnosis using structure MR images. *Neurosci. Lett.* 730:134971. doi: 10.1016/j.neulet.2020.134971

Klöppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., et al. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain* 131, 681–689. doi: 10.1093/brain/awm319

Krizhevsky, A. (2014). One weird trick for parallelizing convolutional neural networks. *arxiv[Preprint]* Available online at: http://arxiv.org/abs/1404.5997 [accessed on April 23, 2014]

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. J., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005

Mehdipour Ghazi, M., Nielsen, M., Pai, A., Modat, M., Jorge Cardoso, M., Ourselin, S., et al. (2021). Robust parametric modeling of Alzheimer's disease progression. *Neuroimage* 225:117460. doi: 10.1016/j.neuroimage.2020.117460

Simonyan, K., and Zisserman, A. (2015). "Very deep convolutional networks for large-scale image recognition," in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, (San Diego, CA: ICLR), 1–14.

Stoeckel, J., and Fung, G. (2005). "SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information," in *Proceedings of the Fifth IEEE International Conference on Data Mining (ICDM'05)*, (Houston, TX: IEEE), 8. doi: 10.1109/ICDM.2005.141

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). "Rethinking the inception architecture for computer vision," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (Las Vegas, NV: IEEE), 2818–2826. doi: 10.1109/CVPR.2016.308

Tan, M., and Le, Q. V. (2019). "EfficientNet: rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning, ICML 2019 (International Machine Learning Society (IMLS))*, (Baltimore, MA: ICML), 10691–10700.

Tan, M., and Le, Q. V. (2021). EfficientNetV2: smaller models and faster training. *arxiv[Preprint]* Available online at: https://arxiv.org/abs/2104.00298v2 [accessed on April 1, 2021].

Tatiparti, K., Sau, S., Rauf, M. A., and Iyer, A. K. (2020). Smart treatment strategies for alleviating tauopathy and neuroinflammation to improve clinical outcome in Alzheimer's disease. *Drug Discov. Today* 25, 2110–2129. doi: 10.1016/j.drudis.2020.09.025

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., et al. (2010). N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. doi: 10.1109/TMI.2010.2046908

Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-González, J., Routier, A., Bottani, S., et al. (2020). Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation. *Med. Image Anal.* 63:101694. doi: 10.1016/j.media.2020.101694

Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45–57. doi: 10.1109/42.906424

Zhu, X., Suk, H.-I., Zhu, Y., Thung, K.-H., Wu, G., and Shen, D. (2015). Multi-view classification for identification of Alzheimer's Disease. *Mach. Learn. Med. Imaging* 9352, 255–262. doi: 10.1007/978-3-319-24888-2_31

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 10 June 2022
doi: 10.3389/fnagi.2022.918462

MPC-STANet: Alzheimer's Disease Recognition Method Based on Multiple Phantom Convolution and Spatial Transformation Attention Mechanism

Yujian Liu[1†], Kun Tang[1†], Weiwei Cai[2,3], Aibin Chen[1], Guoxiong Zhou[1], Liujun Li[4] and Runmin Liu[4,5]*

[1] College of Computer and Information Engineering, Central South University of Forestry and Technology, Changsha, China, [2] School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, China, [3] AiTech Artificial Intelligence Research Institute, Changsha, China, [4] Department of Civil, Architectural and Environmental Engineering, Missouri University of Science and Technology, Rolla, MO, United States, [5] College of Sports Engineering and Information Technology, Wuhan Sports University, Wuhan, China

OPEN ACCESS

Edited by:
Mohammad Khosravi,
Persian Gulf University, Iran

Reviewed by:
Xuebin Chen,
North China University of Science and Technology, China
Jianlei Kong,
Beijing Technology and Business University, China

***Correspondence:**
Aibin Chen
hotaibin@163.com

[†] These authors have contributed equally to this work

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

Received: 12 April 2022
Accepted: 04 May 2022
Published: 10 June 2022

Citation:
Liu Y, Tang K, Cai W, Chen A, Zhou G, Li L and Liu R (2022) MPC-STANet: Alzheimer's Disease Recognition Method Based on Multiple Phantom Convolution and Spatial Transformation Attention Mechanism. Front. Aging Neurosci. 14:918462. doi: 10.3389/fnagi.2022.918462

Alzheimer's disease (AD) is a progressive neurodegenerative disease with insidious and irreversible onset. The recognition of the disease stage of AD and the administration of effective interventional treatment are important to slow down and control the progression of the disease. However, due to the unbalanced distribution of the acquired data volume, the problem that the features change inconspicuously in different disease stages of AD, and the scattered and narrow areas of the feature areas (hippocampal region, medial temporal lobe, etc.), the effective recognition of AD remains a critical unmet need. Therefore, we first employ class-balancing operation using data expansion and Synthetic Minority Oversampling Technique (SMOTE) to avoid the AD MRI dataset being affected by classification imbalance in the training. Subsequently, a recognition network based on Multi-Phantom Convolution (MPC) and Space Conversion Attention Mechanism (MPC-STANet) with ResNet50 as the backbone network is proposed for the recognition of the disease stages of AD. In this study, we propose a Multi-Phantom Convolution in the way of convolution according to the channel direction and integrate it with the average pooling layer into two basic blocks of ResNet50: Conv Block and Identity Block to propose the Multi-Phantom Residual Block (MPRB) including Multi-Conv Block and Multi-Identity Block to better recognize the scattered and tiny disease features of Alzheimer's disease. Meanwhile, the weight coefficients are extracted from both vertical and horizontal directions using the Space Conversion Attention Mechanism (SCAM) to better recognize subtle structural changes in the AD MRI images. The experimental results show that our proposed method achieves an average recognition accuracy of 96.25%, F1 score of 95%, and mAP of 93%, and the number of parameters is only 1.69 M more than ResNet50.

Keywords: MPC-STANet, Multi-Phantom Convolution, Space Conversion Attention Mechanism, Synthetic Minority Over-sampling Technique, Alzheimer's disease recognition

INTRODUCTION

Alzheimer's disease (AD) is an insidious and slowly progressive neurodegenerative disease, which is mainly found in the elderly population over 60 years of age and is clinically manifested as amnesia, loss of mobility, language ability, etc. (Beitz, 2014; Andrieu et al., 2015). Alzheimer's disease has a long developmental cycle and is divided into five disease stages: Non-Demented, Very Mild Demented, Mild Demented, Moderate Demented, Severe Dementia. Very Mild Demented, where people with Mild Demented often have memory loss, and in severe cases, dementia; Mild Dementia, where people show a lack of memory, personality changes, disorientation, and difficulty performing daily tasks; Moderate Dementia, where patients experience significant personality changes and sleep disturbances, and already require additional care and support, which can be easily recognized by health care professionals. Severe Dementia, where patients with this condition already lack the ability to communicate, have difficulty completing the small tasks of life and require full-time treatment. Due to the long stage of Alzheimer's disease and the lack of obvious changes in the features of the early disease, it is difficult for patients themselves to realize this and it is difficult for doctors to make a correct judgment in time based on some of the small pathological features of patients in the early stages of the disease (the first four disease stages of Alzheimer's disease) (Wu and Swaab, 2005). When the symptoms of patients are obvious before they are diagnosed, Alzheimer's disease has already reached the late stage (the fifth disease stage: Severe Dementia). At this time, the patient has the problems of being unable to eat and incontinence and needs others to take care of their daily life, including eating or going to the toilet. A large number of nerves in patients have experienced irreversible death, and the reflex becomes abnormal, resulting in irreversible cognitive degeneration and dementia, which cannot achieve good therapeutic effects (Nelson et al., 2012). The use of deep learning research has little significance in recognizing severe dementia. Therefore, in this study, we only carry out diagnosis and recognition for the first four stages of Alzheimer's disease, which is of great significance for slowing and controlling the progress of the disease (Dubois et al., 2016).

The pathogenesis of Alzheimer's disease is complex, among which age is an important factor in the cause of this disease, and genetic factors, external trauma, education level, trace elements, etc., are also important reasons for the occurrence of this disease (De la Torre, 1999). The biological features of Alzheimer's disease include the formation of senile plaques due to the accumulation of β-amyloid (Aβ) in the cerebral cortex and the hippocampal region, neuronal cell reduction, and neurofibrillary tangles within neuronal cells, etc. (Zhao and Zhao, 2013). The brain structure of Alzheimer's disease patients is mainly characterized by brain atrophy, narrowing of the gyrus, enlargement of the sulcal gaps, and the degree of atrophy in the hippocampus region and medial temporal lobe atrophy compared to normal people. The observation of the brain structure of Alzheimer's patients is mainly through the Alzheimer's MRI medical images, which capture information about the relevant disease pattern of the patients through neuroimaging of the white matter area of the brain and assist doctors in judging the disease stage of Alzheimer's disease, while the Alzheimer's MRI medical images have the problems of difficulty in acquiring and the extremely unbalanced distribution of the acquired data volume (Chen and Glover, 2015). The manual recognition process of Alzheimer's disease is very complex. First, doctors need to ask the patient about recent living environment through psychological scales to assess whether his/her cognitive functions have deteriorated, then employ nuclear magnetic imaging to check whether the imaging structures of the brain of the patient have started to atrophy and change, and finally use electroencephalogram and long-term monitoring of the heartbeat to determine whether the patient is showing changes in cognitive functions and brain signals. Such a testing process relies on the professional knowledge of the physicians and clinical experience, but manual analysis of the medical image is time-consuming and laborious, and there is a risk of misdiagnosis. Therefore, if we can employ a computer to assist in diagnosis, we can improve the efficiency of doctors to a certain extent and also reduce the misdiagnosis and leakage caused by humans (Frisoni et al., 2010; Royce et al., 2019; Lu et al., 2020).

The stage recognition of Alzheimer's disease has been a popular research direction in the field of computer vision-aided diagnosis, and numerous studies have combined traditional machine learning methods to recognize this disease and achieved good recognition results (Negin et al., 2018; Wang et al., 2019). For example, Magnin et al. (2009). proposed and evaluated a novel automated method of whole-brain anatomical MRI based on support vector machine (SVM) classification to distinguish Alzheimer's disease (AD) patients from elderly control subjects, with a mean correct classification of 94.5% (mean specificity 96.6%; mean sensitivity 91.5%) for AD and control subjects. Lebedev et al. (2014) used a random forest classifier trained based on MRI measures of different structures for the diagnosis of Alzheimer's disease and achieve the best AD/HC sensitivity/specificity (88.6%/92.0%) results after combining with cortical thickness and volume measurements. However, although the above research methods were successfully applied to Alzheimer's disease classification and diagnosis, the extraction of effective features in Alzheimer's disease diagnosis often plays a more important role than the construction of classifiers, which requires manual selection of regions of interest before classification and a series of manual feature extraction steps with *a priori* knowledge, which is a tedious extraction process and has human factors interfering (Li et al., 2012; Sabuncu and Konukoglu, 2015). With the development of computer platforms, convolutional neural networks (CNNs) have been widely recognized for their good image recognition, and a large number of CNN-based Alzheimer's classification models have emerged. For example, Sarraf and Tofighi (2016) used convolutional neural networks to successfully classify functional MRI data from Alzheimer's brains with normal healthy brains, where the accuracy of the test data reach 96.85%. Ieracitano et al., 2019 proposed a data-driven approach to distinguish subjects with AD, MCI, and HC by acquiring electroencephalogram recordings and transforming the correlation spectra of 19 channels of electroencephalogram

traces into 2D grayscale images, and then classifying binary and multiple classes in 2D images using CNN models with 89.8 and 83.3% accuracy, respectively.

The above examples all show the application results of the field of deep learning in Alzheimer's disease well while demonstrating the better adaptability and data discrimination of the convolutional neural networks (CNN) within the field of Alzheimer's disease recognition. However, due to the complex structure of the human brain during Alzheimer's disease and the difficulty of detecting subtle structural changes in the brain during mild disease, and the fact that the aging process of normal people is accompanied by shrinkage of brain structures, Alzheimer's patients also suffer from shrinkage of brain areas, which poses many difficulties for research (Young et al., 2013; Lockhart and DeCarli, 2014). Only the correct determination of the changes in brain structure can effectively diagnose the different stages of Alzheimer's disease. Therefore, the main problems of this study are as follows: (1) the Alzheimer's MRI medical images acquired during the first four disease stages of Alzheimer's disease have the problem of unbalanced distribution in terms of data volume, which can affect the training effect of the model and make the classification results biased toward the class with more MRI images. (2) The brain structures in different disease stages of Alzheimer's disease produce subtle changes on MRI images, and the regions of interest (e.g., sulcal gaps, gyrus, hippocampal region, medial temporal lobe) account for a small proportion of the whole MRI image, complicating feature extraction. (3) The lack of distinctive features of Alzheimer's disease makes convolutional neural networks often accompanied by an increase in the number of convolutional layers to improve the ability of the neural network for feature extraction. However, when the number of layers of the neural network exceeds a certain threshold, there will be problems such as gradient disappearance and gradient explosion, making the neural network difficult to be trained, and the long-time training is not conducive to Alzheimer's disease prediction (Guo et al., 2017; Wu et al., 2018; Hu et al., 2021).

To deal with the problem of classification imbalance in the dataset, the most basic approach is either to directly copy the minority classes and add them to the sample set or to employ a certain percentage of the majority classes as the training set to obtain a relatively balanced dataset (López et al., 2013; Maxwell et al., 2018). However, this approach tends to lead to the problem of model overfitting, which makes the information learned by the model not generalized enough. To address this problem, Chawla et al. (2002) proposed the Synthetic Minority Oversampling Technique (SMOTE), which uses the similarity between the classes with fewer samples in the feature space to build synthetic new samples and add them to the minority classes. The SMOTE is a good solution to the problem that the information obtained by random oversampling is too special and not generalized enough. Therefore, we combine SMOTE with data expansion (flipping, adding random Gaussian noise, and contrast adjustment) for Alzheimer's disease to perform class-balancing preprocessing for better training results.

Given the little variation in the Alzheimer's MRI images in different disease stages and the small proportion of regions of interest, Toğaçar et al. (2021) used DeepDream, fuzzy color image enhancement, and super columnar techniques to process the Alzheimer's MRI dataset, input the processed MRI dataset into VGG-16 for feature extraction, and finally used Support Vector Machine (SVM) as a classifier. The recognition accuracy of using this method was 100% for MD and ND as well as 99.94% for VMD and MOD. Using the data enhancement algorithm on the Alzheimer's MRI dataset can enhance the features of each MRI image and suppress useless background information so that the deep learning model can better extract these features and achieve a high recognition rate. However, the use of data enhancement algorithms often has the problems of a cumbersome operation process and poor generalization ability, and the input of the deep learning model often requires a large amount of image data, which takes a long time to complete the feature enhancement operation. Thus, we improve the ability of the network to extract features by redesigning the structure of the deep neural network to avoid using more feature enhancement algorithms for the dataset. Therefore, we propose a recognition network of Alzheimer's disease based on Multi-Phantom Convolution and Space Conversion Attention Mechanism (MPC-STANet) with the residual network ResNet50 as the backbone network (He et al., 2016). Compared with VGG-16, ResNet50 has lower complexity and required parameters, and faster convergence speed. It has 50 training layers, which can extract more subtle features from Alzheimer's MRI images, with better classification accuracy. Moreover, the unique residual connection of ResNet50 breaks the symmetry of the neural network, improves the utilization of neurons in each layer, and multiple branches ensure that even if some layers degenerate, it will not affect the overall performance, which makes it widely used in the field of image recognition. Since the feature performance of the Alzheimer's MRI medical image is very different from that of an ordinary image, we structurally designed the convolutional layer of ResNet50 to better suit Alzheimer's disease: (1) To increase the feature extraction of small and scattered regions in the Alzheimer's MRI images, we employ dilated convolution (Yu and Koltun, 2015) instead of the original 7×7 convolution layer in STAGE1 to obtain a larger perceptual field without changing the number of parameters. (2) To extract more subtle pathological features of structures, we employ Multi-Phantom Convolution and the Space Conversion Attention Mechanism in the residual blocks to extract richer characterization information from patients' MRI images. Meanwhile, to avoid the redundancy of useless information, we add an average pooling layer to integrate space information in the shortcut branch of the residual block to improve the detection speed along with reducing the computation.

The contributions of this study are as follows.

(1) To solve the problem of classification imbalance in the Alzheimer's disease dataset, we increased the data volume in the minority classes using data expansion methods such as flipping, adding noise, and contrast adjustment (as depicted in **Figure 1**), and performed class-balancing operation using SMOTE (the results are displayed in **Table 1**). SMOTE performs a class-balancing operation by artificially

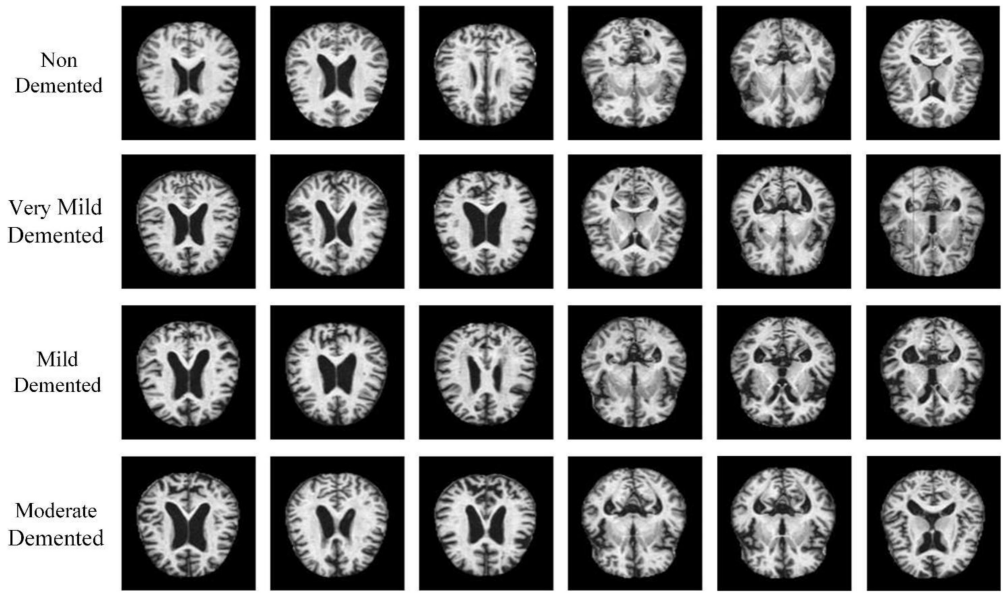


FIGURE 1 | Principles of Alzheimer's disease recognition.

TABLE 1 | The number of the four disease stages and their proportions.

Disease type	Original number	Percentage	Expanded number	Percentage
Non-Demented	3200	50%	3200	25%
Very Mild Demented	2240	35%	3200	25%
Mild Demented	896	14%	3200	25%
Moderate Demented	64	1%	3200	25%

TABLE 2 | The recognition accuracy of the original dataset and the preprocessed dataset in the three models.

Network model	Original data set	Preprocessed data set
ResNet50	76.9%	84.6%
ResNet50-SPAM	81.2%	89.4%
MPC-STANet	85.5%	96.2%

TABLE 3 | Comparison of accuracy and number of parameters of four networks.

Network model	Parameters	Accuracy
ResNet50	25.56M	84.6%
ResNet50-DC	25.56M	86.7%
MPC-STANet	27.25M	96.2%

TABLE 4 | Comparison of accuracy and number of parameters of three networks.

Network model	Parameters	Accuracy
ResNet50	25.56M	84.6%
ResNet50-MPRB	21.10M	89.5%
MPC-STANet	27.25M	96.2%

synthesizing new samples from the minority classes and adding them to the dataset for classification balance, which well solves the problem of model overfitting, as displayed in **Table 2**. In the MPC-STANet model, the recognition accuracy of the class-balancing processed dataset is improved by 10.7% compared to the unprocessed dataset.

(2) To extract the scattered and subtle pathological features of Alzheimer's disease, the MPC-STANet is proposed in this study. (a) We employ Dilated Convolution in STAGE 1 of the network to extract features from scattered pathological regions of Alzheimer's disease to obtain a larger range of feature information. As displayed in **Table 3**, the recognition accuracy of ResNet50 after using Dilated Convolution is improved by 2.1% compared with ResNet50. (b) To extract more subtle pathological features, we proposed Multi-Phantom Residual Block (including Multi-Conv Block and Multi-Identity Block) based on Multi-Phantom Convolution, average pooling layer, and

TABLE 5 | The influence of attention mechanisms on network accuracy.

Network model	Accuracy
ResNet50	84.6%
ResNet50-SE	85.9%
ResNet50-CMBA	87.8%
ResNet50- SCAM	90.1%
MPC-STANet	96.2%

TABLE 6 | Performance evaluations of each disease stages.

Network model	Recall	F1-score	Precision
Non-Demented	97%	96%	97%
Very Mild Demented	95%	94%	95%
Mild Demented	97%	97%	98%
Moderate Demented	95%	93%	94%

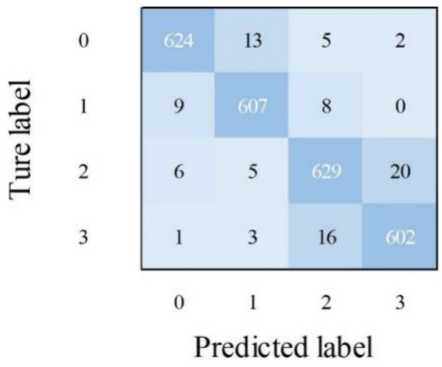


FIGURE 2 | Confusion matrix of the MPC-STANet.

Conv Block and Identity Block of ResNet50 to extract richer characterization information in the Alzheimer's MRI images, and the recognition accuracy is improved by 4.9% compared to ResNet50 and the model parameters decreased by 4.46M compared to ResNet50 (as displayed in **Table 4**). (c) Space Conversion Attention Mechanism is inserted between Multi-Phantom Convolution and 1 × 1 convolution, aiming to solve the problem of difficult recognition due to small differences between disease stages. Space Conversion Attention Mechanism preserves more important feature information (e.g., hippocampal region, brain gyrus, sulcal gaps, etc.) and discards redundant information (e.g., background) by assigning different weights in vertical and horizontal directions to enhance the extraction of tiny features, and the recognition accuracy is improved by 5.5% compared to ResNet50 (as displayed in **Table 5**).

(3) The recognition accuracy of the recognition methods proposed in this study for the first four stages of Alzheimer's disease in non-demented, very mild demented, Mild Demented and Moderate Demented are 97, 95, 98, and 94%, respectively. Other performance evaluations are shown in **Table 6** and the confusion matrix of the MPC-STANet is shown in **Figure 2**. In experiment 3.5, we tested the MPC-STANet and other networks in the same environment. The experimental results show that the Recall, F1-score, Precision, and mAP of the MPC-STANet proposed are 96, 95, 96, and 93%, respectively, which are higher than the other networks. The overall performance of the model is good, and the performance evaluations of other networks are shown in **Table 7**.

Therefore, we propose a method in this study for recognizing disease stages of Alzheimer's disease that combine class-balancing preprocessing and the MPC-STANet. The recognition principle is depicted in **Figure 3**. First, the minority classes are enhanced by flipping, adding noise and contrast adjustment, and then the class-balancing operation is achieved by SMOTE. Finally, the processed MRI dataset is input into the MPC-STANet for training and testing. To better extract the pathological features of Alzheimer's disease, Dilated Convolution, Multi-Phantom Residual Block (including Multi-Conv Block and Multi-Identity Block), and Space Conversion Attention Mechanism are incorporated in the MPC-STANet to achieve better recognition accuracy.

TABLE 7 | Evaluation indexes of the networks.

Network model	Recall	F1-score	Precision	mAP
ResNet50, He et al., 2016	83%	82%	85%	81%
VGG16, Toğaçar et al., 2021	80%	76%	77%	75%
U-Net, Hazarika et al., 2022	79%	75%	77%	73%
LeNet-5, Li et al., 2015	83%	82%	80%	75%
ADVIAN, Wang et al., 2019	84%	82%	85%	81%
MobileNet-SVM, Fei et al., 2022	90%	89%	89%	84%
DFNN, Huang et al., 2020	85%	82%	84%	81%
ResNet-STN, Sun et al., 2021	88%	89%	86%	83%
TReC, Xiao et al., 2021	91%	90%	92%	88%
Inception-v4, Bae et al., 2020	87%	90%	88%	85%
EfficientNetB0, Savaş, 2022	92%	94%	94%	92%
AlexNet, Hanmugam et al., 2022	77%	73%	75%	70%
GoogleNet, Hanmugam et al., 2022	84%	87%	86%	81%
MPC-STANet	96%	95%	96%	93%

MATERIALS AND METHODS

Data Acquisition

Dataset is an important part of the field of pattern recognition and data mining. Since the main motivation of this study is to design a deep learning framework for Alzheimer's disease classification, the adopted Alzheimer's MRI dataset was created by researcher Sarvesh Dubey (Kaggle) and was collected from multiple websites, hospitals, and public repositories. The dataset consists of 896 MRI Mild Dementia images, 64 MRI Moderate Dementia images, 3,200 MRI Non-Dementia images, and 2,240 MRI Very Mild Dementia images, and the distribution of the number of MRI images in different stages of Alzheimer's diseases is displayed in **Table 8**. All MRI images were preprocessed and resized to 128 × 128 pixels and saved in JPG format, and some of the images are depicted in **Figure 1**.

Class-Balancing Preprocessing Based on Data Expansion and SMOTE

As displayed in **Table 8**, the Alzheimer's MRI dataset acquired has the problem of unbalanced distribution in terms of data volume, which will lead to the imbalanced learning effect of the neural network model, and the problems of overfitting and under-fitting exist simultaneously. To address this problem, we can expand the dataset with the help of data expansion methods and Synthetic Minority Oversampling Technique (SMOTE) technique to balance the data volume of the first four disease stages to improve the accuracy of the neural network model.

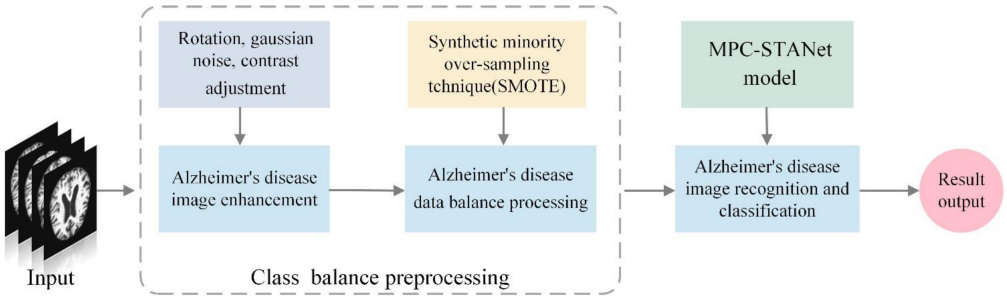


FIGURE 3 | Principles of Alzheimer's disease recognition.

TABLE 8 | Number distribution of Alzheimer's disease dataset.

Disease type	Original number	Percentage
Non-Demented (ND)	3200	50%
Very Mild Demented (VMD)	2240	35%
Mild Demented (MD)	896	14%
Moderate Demented (MOD)	64	1%

Data Expansion

Training the neural network model with more datasets allows it to learn more effective feature points to improve the recognition accuracy of the model, prevent overfitting, etc. We use MATLAB 2020b to flip the image, add random Gaussian noise, contrast adjustment, and other data expansion methods to expand the minority classes to improve the training effect of the neural network model. Also, the data expansion method is an important way to balance the data volume of different classes and the result images are shown in **Figure 4**.

Synthetic Minority Oversampling Technique

Table 8 displays the number of images in the dataset for the first four disease stages of Alzheimer's disease, which indicates that the classes of the dataset are unbalanced. If a class-unbalanced dataset is used for prediction, the predictions tend to yield conclusions that are also biased, that is, the classification results will be biased toward the majority class. To address this problem, we apply synthetic minority oversampling technique (SMOTE) to this dataset, which addresses the classification imbalance in the dataset by randomly replicating the classes with fewer samples in the dataset to match the classes with more samples. We oversample the classes with fewer samples using the seeds of 42 random number generators, and **Table 1** displays the distribution of the Alzheimer's disease dataset after using the data expansion and SMOTE.

Suppose the number of the minority classes samples are $T$and set a sampling ratio to determine the magnification N according to the sample imbalance ratio so that the sample can be expanded by N times after sampling. The algorithm steps of the SMOTE are as follows:

Step1: Consider a sample $x \in \{1, ..., T\}$ in the minority class, calculate its distance to all samples in the minority classes based on the Euclidean distance, and select K the nearest neighbors.

$$|x| = \sqrt{x_1^2 + x_2^2 + x_3^2 + ... + x_n^2} \tag{1}$$

Step2: Randomly select a sample B from the K nearest neighbors and combine it with the original sample to synthesize a new sample according to the following formula.

$$x_{new} = a + rand(0, 1) \times |a - b| \tag{2}$$

Step3: Repeat Step 2 and Step 3 N times.

Step4: Repeat the above steps for T samples of the minority classes.

ResNet50 Backbone

ResNet50 constructs the deep network model as a shallow network model and an additional layer of self-mapping connects the trained shallow structure with the additional layer of self-mapping through residual units, transmits the input across layers through a shortcut, and then adds the output after convolution to achieve the effect of fully training the underlying network. ResNet50 has 6 STAGE (STAGE1∼ STAGE6), containing 49 convolutional layers and 1 fully connected layer. Among them, the 49 convolutional layers consist of two basic blocks. As shown in **Figure 5**, one is Identity Block, which has the same dimension of input and output, so it can be concatenated with more than one for deepening the network layers; the other basic block is Conv Block, which has an inconsistent dimension of input and output, so it cannot be concatenated consecutively and its role is to change the dimension of the feature vector.

The actual Alzheimer's disease needs to be judged by looking at the pathological features such as the degree of enlargement of the sulcal gaps, and the degree of atrophy in the hippocampus region and medial temporal lobe. However, due to the small variation in MRI image features during the four disease stages: mild dementia (MID), moderate dementia (MOD), non-dementia (ND), and very mild dementia (VMD), more subtle pathological features need to be extracted to better discriminate. To extract more subtle

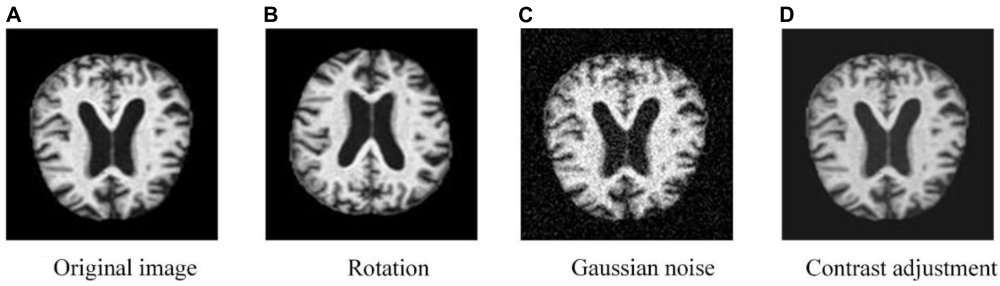


FIGURE 4 | MRI image processed by augmentation methods.

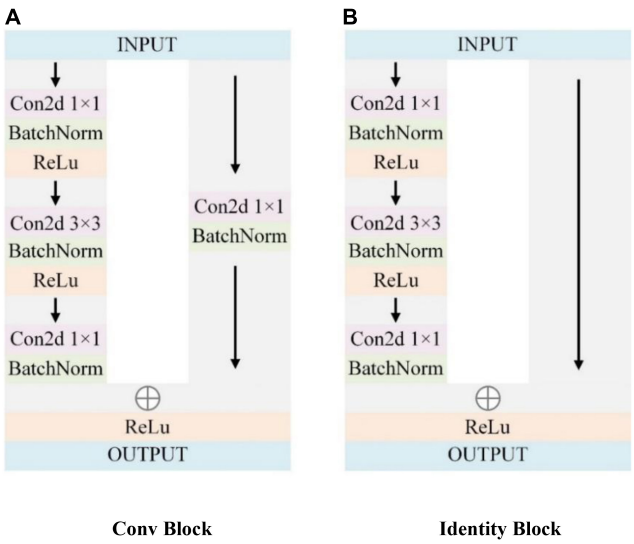


FIGURE 5 | ResNet residual block.

pathological features, this study employs ResNet50, the winner of the ImageNet large-scale visual recognition competition in 2015, as the basic network. However, the properties of the Alzheimer's MRI images are very different from those of ordinary images, and designing according to the pathological characteristics of Alzheimer's disease can effectively improve the accuracy of the model. Therefore, we propose a recognition network of Alzheimer's disease based on Multi-Phantom Convolution and Space Conversion Attention Mechanism (MPC-STANet), which is improved based on ResNet50.

Recognition Network of Alzheimer's Disease Based on Multi-Phantom Convolution and Space Conversion Attention Mechanism

The MPC-STANet is upgraded based on ResNet50, and the network structure is depicted in **Figure 6**.

The feature extraction network of ResNet50 consists of 7 ×7 convolution and 3 ×3 maximum pooling layer (STAGE1), convolutional residual extraction network composed of Conv Block and Identity Block (STAGE2˜STAGE5), average pooling layer, and fully connected layer (STAGE6). The MPC-STANet proposed in this study is based on ResNet50, using the Dilated Convolution (DC) instead of 7 ×7 convolution of STAGE1; changing the two basic blocks of Conv Block and Identity Block using Multi-Parallel Convolution (MPC) and averaging pooling layer, proposing Multi-Conv Block and Multi-Identity Block; adding the Space Conversion Attention Mechanism (SCAM) between the convolution blocks. This study improves the network structure, and more details will be provided in the following chapters.

Dilated Convolution

The pathological feature points of different disease stages of Alzheimer's disease are obscure and scattered. To classify disease stages based on the Alzheimer's MRI medical images more accurately, more effective subtle pathological features need to be extracted. ResNet50 uses a 7 × 7 convolution with a large perceptual field in STAGE1, which is sufficient for extracting features from common and ordinary images in the ImageNet database, but it is difficult to adequately

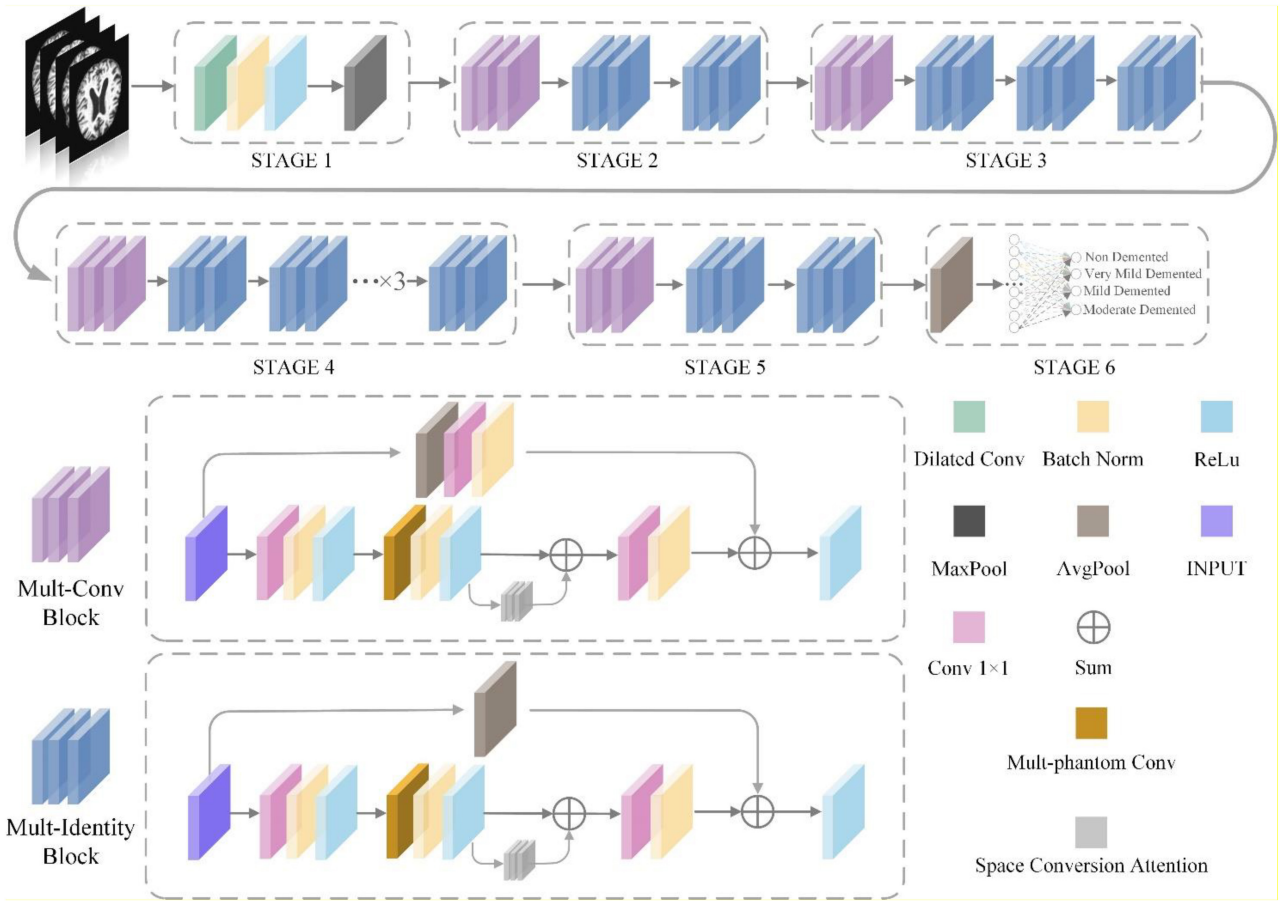


FIGURE 6 | The t overall architecture of the MPC-STANe.

consider the subtle pathological features of MRI. Therefore, to reduce the information loss during the extraction process and improve the recognition ability of the model, we employ Dilated Convolution (DC) to replace the 7 × 7 convolution in STAGE1. Dilated Convolution increases the perceptual field while maintaining the size of the feature map unchanged and does not cause problems such as information loss.

Dilated Convolution expands the perceptual field size of ordinary convolution by setting different dilation rates (r). Among them, r determines the interval size of the holes injected in the convolution. If r is too small, the range of the perceptual field is limited, and if r is too large, the features in the perceptual field lose some relevance. Dilated Convolution can be regarded as inserting a zero value of *r-1* into the convolution kernel during ordinary convolution. For ordinary convolution, the convolution kernel of 3 × 3 is calculated on the feature map, and the perceptual field of the new feature point is three, as depicted in **Figure 7A**. For the dilation convolution with dilation rate $r = 2$, one zero value is inserted between the 3 × 3 convolution kernels to obtain its perceptual field of five, as depicted in **Figure 7B**, which results in the equivalent of two ordinary 3 × 3 convolutions with only one computation.

Assuming that Dilated Convolution kernel is $k \times k$, and the dilated rate is r, then the actual convolution kernel is:

$$K = k + (k-1) \times (r-1) \tag{3}$$

After Dilated Convolution process, the relationship between the size of the input and output feature maps is as follows:

$$W_2 = \frac{W_1 + 2p - r \times (k-1) - 1}{s} + 1 \tag{4}$$

Among them, W_1 and W_2 represent the size of the input and output feature maps, respectively, s and p represent the step-size and the patch.

Multi-Conv Block and Multi-Identity Block Based on Multi-Phantom Convolution

ResNet50 residual block mainly consists of a linear branch (one 1 × 1 convolution layer and two 3 × 3 convolution layers) and a shortcut branch with 1 × 1 convolution, where the linear branch is used to extract feature information in the feature map and generate the output feature matrix; the shortcut branch uses 1 × 1 convolution to increase the number of channels and match the number of channels of the linear branch, which is used to avoid

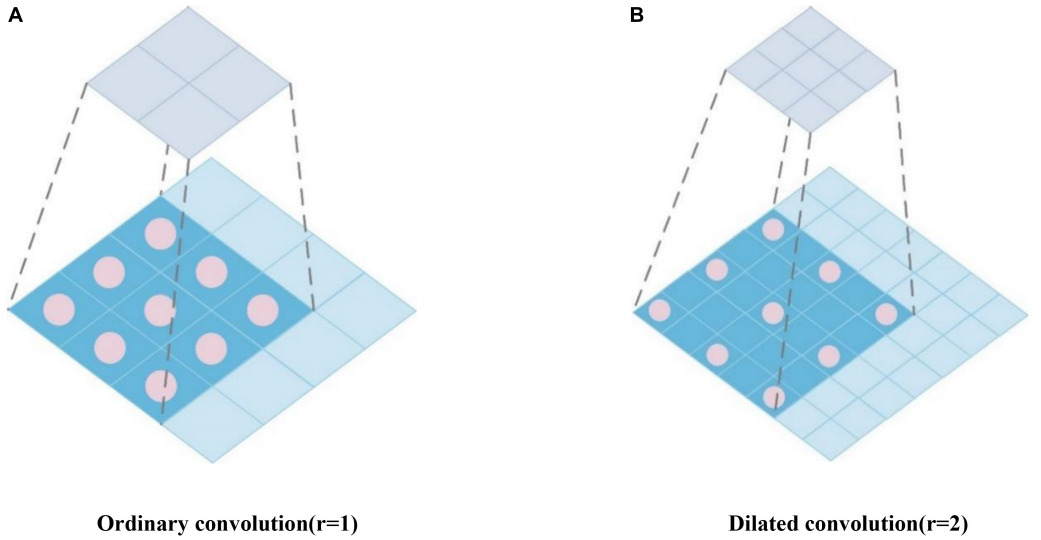


FIGURE 7 | Perceptual field of ordinary convolution and extended convolution.

the problems of gradient disappearance and gradient explosion caused by network depth. Finally, the output feature matrixes of the two branches are summed to obtain the feature map of residual block output, and then the feature map is put into the Relu activation function to enhance the non-linearization of the model. However, due to the variety of structural changes in MRI images in different stages of Alzheimer's disease such as the changes in the cerebral cortex, especially in the temporal and parietal regions, the features that appear tend to show only subtle differences; whereas the structure of the hippocampal region of suffering from Alzheimer's disease is significantly changed in different stages.

To address this problem, we propose the Multi-Phantom Convolution (MPC) by borrowing the convolution by channel direction in the Inception network (Szegedy et al., 2016), and incorporating MPC into the residual block to propose Multi-Phantom Residual Block (MPRB), which has two blocks: Multi-Conv Block and Multi-Identity Block, to extract the features of more abundant characterization information in patients' MRI. MPRB divides the feature matrix map output from 1×1 convolution into 4 parts of feature maps equally according to the channel direction, and then the feature maps of the different parts are extracted by different convolution and pooling operations for multi-scale feature extraction, and finally concatenated according to the channel direction. MPRB can extract more subtle pathological features; meanwhile, the MPRB reduces the training parameters and speeds up the convergence of the model when dividing the input feature maps and parallel convolution operations. In addition, it should be noted that the pathological features account for small regions of the whole MRI image and the proportion of information to be acquired is small. o avoid the redundancy of useless information, we add a 2×2 average pooling layer to integrate spatial information in the shortcut branch of MPRB, and the structure of MPRB is depicted in **Figure 8**. The average pooling layer has no parameters and does not change the global number of parameters while preventing overfitting at this layer.

The specific implementation process of Multi-Phantom Residual Block:

Step1: The input feature matrix is successively passed through the 1×1 convolution layer, batch norm layer, and Relu activation function, and takes the result as the input of step 2; the input feature matrix is successively passed through the average pool layer, 1×1 convolution layer and batch norm layer as the output of the shortcut branch.

Step2: The feature matrix of the linear branch is divided into four parts according to the channel direction.

Step3: The feature map I: passing through the 1×1 convolution layer; the feature map II: passing through the 1×1 convolution layer and 3×3 convolution layer successively; the feature map III: passing through the 1×1 convolution layer and 5×5 convolution layer successively; the feature map IV: passing through the maximum pool layer and 1×1 convolution layer successively.

Step4: The four feature maps are concatenated according to the channel direction and passed through the batch norm layer and Relu activate function.

Step5: The feature map output from step 4 with the feature map output through the Space Conversion Attention Mechanism (SCAM) are summed by pixels, and the sum result successively passes through the 1×1 convolution layer and batch norm layer.

Step6: Sum the linear branch (the output of step 5) with the shortcut branch, and pass through the Relu activate function.

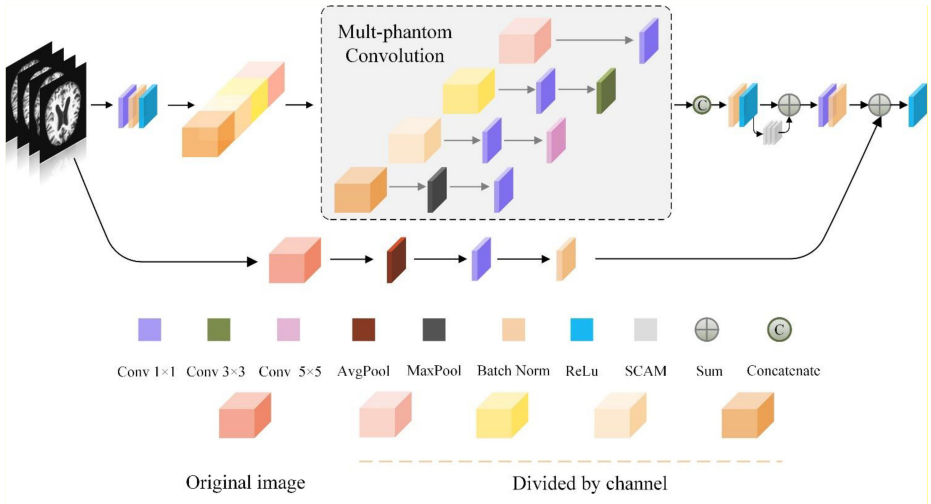


FIGURE 8 | The structure of multi-phantom residual block.

Space Conversion Attention Mechanism

The visual attention mechanism is a brain signal processing mechanism that is unique to human vision. By quickly scanning the global image, human vision obtains the target region to be focused on, which is generally called the focus of attention, and then devotes more attention resources to this region to obtain more detailed information about the target to be focused on, while suppressing other useless information. The core goal of adding an attention mechanism to the network is essentially similar to the human selective visual attention mechanism, which also selects the more critical information for the current task goal from a multitude of information and ignores other redundant information to successfully improve the expressive power of the network.

The more popular attention mechanisms include SENet (Hu et al., 2018), CBAM (Woo et al., 2018), and Non-local Neural Networks (Wang et al., 2018), and many pieces of research combined with the recognition of attention mechanisms have achieved good recognition results, for example, Huang et al. (2020). proposed a brain tumor diagnosis system based on a differential feature neural network (DFNN), which mainly consists of an innovative differential feature map (DFM) block and a squeeze-and-excitation (SE) block. The experimental results indicated that the average accuracy of DFNN in classifying the brain as abnormal and normal on two databases was 99.2 and 98%, respectively. Xiao et al. (2021) proposed an early diagnosis method for pathological brain called the TReC, which imported the CBAM convolutional channel attention mechanism into the pre-trained ResNet residual block and replaced the fully connected layer with a new FC layer. The experimental results indicated an accuracy of 100% in the two-class classification task and an accuracy of 97.44% in the multi-class classification task. Sun et al. (2021) proposed a recognition method of the residual network (ResNet) combining space transformation network (STN) and non-local attention mechanism (non-local attention) to consider the long-range correlation in feature space, and successfully applied the method to the early diagnosis of Alzheimer's disease, with the recognition accuracy of up to 97.1%, macroscopic accuracy of up to 95.5%, macroscopic recall of up to 95.3%, and macroscopic F1 value of up to 95.4%.

The space dimension of the image refers to the height (H) and width (W) of the image, and C represents the feature channel of the image. The space attention mechanism pays attention to the importance of the space location features of the image, generating space attention coefficients for the output feature maps, and enhancing or suppressing different space location features according to the feature weights. The traditional space attention mechanism tends to focus on weight assignment in only one direction, which inevitably leads to the loss of important information in the image. For the Alzheimer's MRI images, it is important to observe space changes in different disease stages, such as small changes in the cerebral cortex and structural changes in the hippocampal region, which are important for determining the stage of Alzheimer's disease. Therefore, we propose a Space Conversion Attention Mechanism (SCAM) that assigns weights based on both vertical and horizontal directions, and the specific structure is depicted in **Figure 9**.

The Space Conversion Attention Mechanism is composed of three parts:

(a) The horizontal spatial attention mechanism is used to generate horizontally oriented weight coefficients for each row of features; the vertical spatial attention mechanism is used to generate vertically oriented weight coefficients for each column of features.

$$c_i = \sum_{j=1}^{n} \frac{exp(e_{i,j})}{\sum_{k=1}^{n} exp(e_{ik})} h_j \tag{5}$$

Among them, $e_{i,j}$ represents the weight coefficients assigned by the horizontal or vertical attention mechanism, pixel j represents the sequence feature, i represents the temporal features

at a certain moment, and h_j represents the hidden layer information of the feature sequence j. $c_{\mathrm{I}} = \{c_1, c_2 \ldots c_{i-1}, c_i\}$ represents the weight coefficients of the vertical attention mechanism in the feature sequence; $c_{\Pi} = \{c_1, c_2 \ldots c_{i-1}, c_i\}$ represents the weight coefficients of the horizontal attention mechanism in the feature sequence.

(b) To further expand the difference between the weight coefficients, we add the horizontal and vertical weight coefficients (addition strategy). for example, the small weight coefficient maybe 0.1 + 0.3 and the large weight coefficient maybe 0.8 + 0.9. In contrast, the difference between the summed weight coefficients is more obvious.

$$add = c_{\mathrm{I}} + c_{\Pi} \tag{6}$$

(c) To select the more interesting regions, we match the horizontal and vertical weight coefficients to find the maximum value (maximum strategy), which is used to complement the results of the second part of the weight coefficients [e.g., max(0.3, 0.8)].

$$max = max(c_{\mathrm{I}}, c_{\Pi}) \tag{7}$$

Finally, we concatenate the weight coefficients calculated by the above strategy with c_{I} and c_{Π} through formula (6), and the concatenated results are passed through 1×1 convolution and sigmoid function to make the dimension of input and output consistent.

$$SPA = concatenate([c_{\mathrm{I}}, c_{\Pi}, add, max]) \tag{8}$$

$$Weight = \sigma(F_h(SPA)) \tag{9}$$

Among them, *SPA* represents the Space Conversion Attention Mechanism, F_h represents 1×1 convolution, σ represents the sigmoid function, and *Weight* represents the feature weights.

RESULTS AND ANALYSIS

Experimental Environment and Settings

All of the trains and tests in this work are carried out on the same hardware and software platform. The hardware environment is Windows (64bit) operating system, Intel Core i7-9700U CPU, and 2080Ti GPU. The software programming environment for data expansion is MATLAB 2020b; The software programming environment for the MPC-STANet is Python 3.8.12, Pytorch 1.8.2, and CUDA 10.2. Considering the memory size of the GPU and the time of the experiment, we set the *Batchsize* to 32 for training and 8 for testing. The learning rate *Ir* was set to 10^{-3}, and the *epochs* was set to 140. The Adam optimizer and Cross-Entropy Loss were used during training, and the incremental gradient descent was used as the training method. After class-balancing preprocessing, there were 12,600 MRI images in Alzheimer's disease dataset. In this experiment, according to the ratio of 7:2:1, we divided the dataset into a training set, test set, and validation set for training and testing the MPC-STANet.

Effectiveness Experiment of the Module

Effectiveness Experiment of Preprocessing

To verify whether the training with a class-balanced preprocessed dataset can improve the performance of the model and improve the recognition accuracy, we input the original dataset and the preprocessed dataset into ResNet50, ResNet50-SPAM, and the MPC-STANet, respectively, for experiments. **Table 2** displays the recognition accuracy of the original dataset and the preprocessed dataset in the three networks. The results show that the recognition accuracy of the three networks in the preprocessed dataset is significantly higher than that of the original dataset. This is because the dataset is expanded by flipping, adding noise, and contrast adjustment, which increases the diversity of the dataset and avoids network coverage. The SMOTE algorithm is used to make the samples achieve class balance, to avoid the information learned during training to tend to the disease majority class. As a result, following preprocessing, the accuracy of the dataset has increased in all three models.

Effectiveness Experiment of Dilated Convolution

We used Dilated Convolution (DC) in STAGE 1 of the MPC-STANet. To verify the effect of the DC on classification performance, we conducted experiments on ResNet50, ResNet50-DC, and the MPC-STANet under the same test environment. **Table 3** displays that using DC in ResNet50 can improve the recognition accuracy without changing the model parameters.

Effectiveness Experiment of Multiple-Phantom Residual Block Based on Multiple-Phantom Convolution

To verify the effect of the Multiple-Phantom Residual Block (MPRB) on model accuracy and parameter, we trained and tested ResNet50, ResNet50-MPRB, and the MPC-STANet using the same dataset. As displayed in **Table 4**, the experimental results show that ResNet50-MPRB with multiple-phantom residual blocks can greatly improve the accuracy of the network and reduce the number of parameters of the model.

Effectiveness Experiment of Space Conversion Attention Mechanism

To more intuitively understand the improvement of the network accuracy by Space Conversion Attention Mechanism (SCAM), we trained and tested ResNet50, ResNet50-SE, ResNet50-CMBA, ResNet50-SCAM, and the MPC-STANet, respectively, on the preprocessed dataset. **Table 5** displays the accuracy of the networks with different attentions on the test set. The experimental results show that after using the attention mechanism, ResNet50-SE, ResNet50-CMBA, and ResNet50-SCAM improve 1.3, 3.2, and 5.5%, respectively, in terms of accuracy compared to ResNet50. The SCAM outperforms the other attention mechanisms

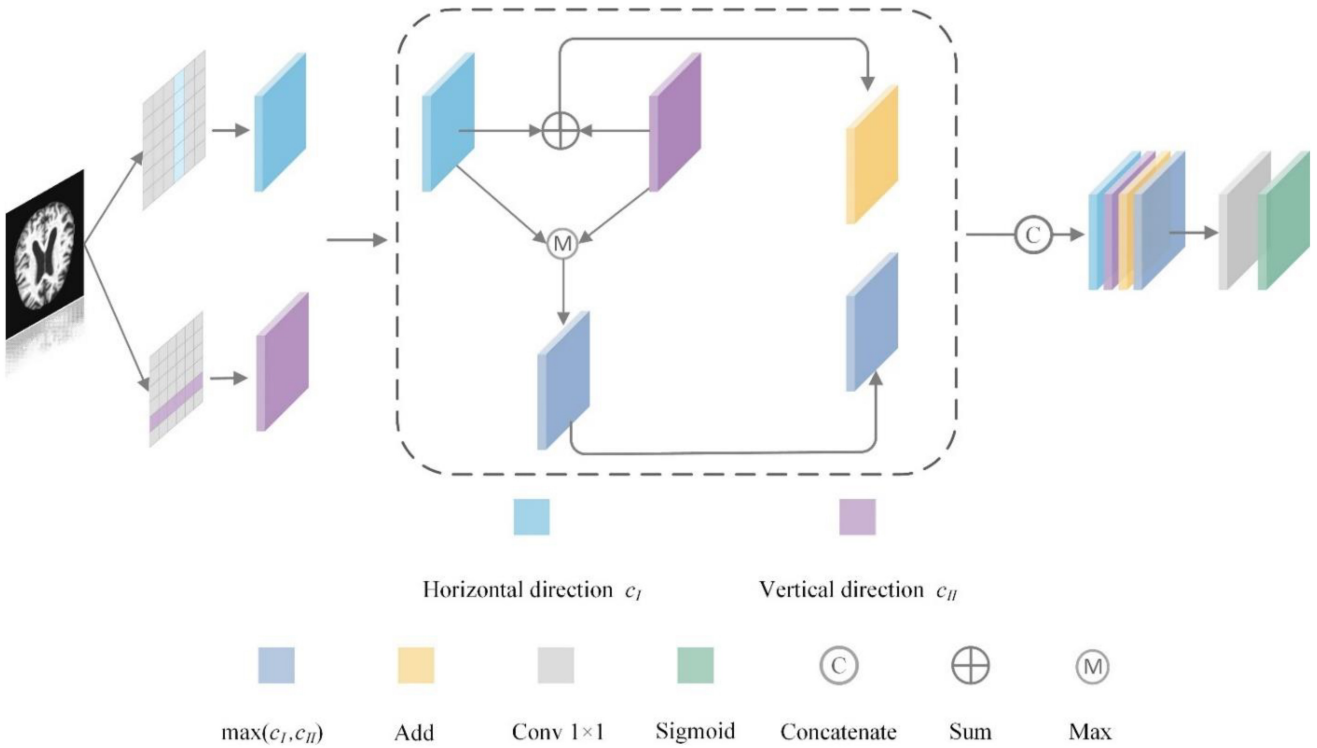


FIGURE 9 | Space conversion attention mechanism structure.

TABLE 9 | Comparison of recognition accuracy and parameters of different networks.

Network model	Parameters	Accuracy
ResNet50	25.56M	84.6%
ResNet50-DC	25.56M	86.7%
ResNet50-MPRB	21.10M	89.5%
ResNet50-SCAM	31.17M	90.1%
ResNet50-DC-MPRB	21.10M	93.3%
ResNet50-DC-SCAM	31.17M	93.8%
ResNet50-MPRB-SCAM	27.25M	94.6%
MPC-STANet	27.25M	96.2%

in improving accuracy by considering the weight feature relationship in both horizontal and vertical directions. The accuracy of the MPC-STANet proposed in this study is 96.2%, which indicates that the Alzheimer's MRI images features are deeply extracted, and the network is effective in recognizing.

Table Ablation Experiment

To fully validate the effectiveness of the method proposed in this study, we employed the same dataset and experimental environment, and only changed the parts that needed to be compared in each experiment. In this experiment, ResNet50 is selected as the backbone network, and one or more of the three methods, Dilated Conv (DC), Multi-Phantom Residual Block (MPRB), and Space Conversion Attention Mechanism (SCAM), are added to compare the effects of different schemes on model parameters and recognizing accuracy. The comparing results are displayed in **Table 9**.

Based on the accuracy of the network, the accuracy of the MPC-STANet was higher than other networks, reaching 96.2%. When SCAM was applied to ResNet50, its accuracy improved by 5.5% compared to the original ResNet50. Similarly, ResNet50 using Dilated Conv or MPRB methods improved by 2.1 and 4.9%, respectively, compared to the original ResNet50. The preceding evidence indicates that all three methods are effective for increasing accuracy. And the solution of DC paired with MPRB or SCAM has the largest improvement in accuracy with 8.7 and 9.2%, respectively.

Based on the number of parameters of the network, the network with Dilated Conv is the same in terms of the number of parameters as the network that keeps a single variable, which is consistent with the principle that Dilated Convolution does not change the number of parameters. In terms of the number of parameters, ResNet50-MPRB is 4.46M less than ResNet50, demonstrating that the MPRB method aids in network compression.

Overall Evaluation of the MPC-STANet

In the same environment, the MPC-STANet has a more stable learning process and higher recognition accuracy than its

backbone network ResNet50. The performance of the MPC-STANet on the four disease stages are depicted in **Figure 2**. In the confusion matrix, the numbers 0, 1, 2, and 3 represent the four disease stages, Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented, respectively. We tested the MPC-STANet using a total of 2,560 MRI test sets and displayed the test results in the confusion matrix. The diagonal of the confusion matrix is the number of correctly predicted images with a total of 2,462 MRI images. The overall recognition rate of the MPC-STANet is 96.2%. **Table 6** displays the recognition rate of the four disease stages in the MPC-STANet. It can be seen that the highest precision of Mild Demented reached 98%, and that of Moderate Demented was only 94%.

Comparison With Other Networks

We employ four indexes, Recall, F1-score, Precision, and mAP to evaluate the performance of the MPC-STANet. The results are displayed in **Table 7**, The performance indexes of the MPC-STANet all surpass 90%, higher than those of other networks, indicating that this network is more advantageous in recognizing Alzheimer's disease than other networks, and is better for classifying the disease stage.

DISCUSSION

In this study, we construct the MPC-STANet capable of discriminating the first four disease stages of Alzheimer's disease and use the Alzheimer's MRI images created by researcher Sarvesh Dubey as the dataset. We employ data expansion and SMOTE to perform class-balancing preprocessing of the dataset, and then input the preprocessed dataset into the MPC-STANet for recognition, and its average recognition accuracy reaches 96.25%. The experiments show that the combination of class-balancing preprocessing and MPC-STANet for the recognition of the first four disease stages of Alzheimer's disease is effective and does not require operations such as numerous feature enhancement preprocessing or manual feature extraction, but the following explorations are needed: (1) The researcher who provided the Alzheimer's MRI dataset did not provide any statistical information about patients and did not account for this condition, which raises doubts about our recommended approach. Therefore, Alzheimer's disease datasets with detailed statistics need to be further considered in future explorations to be more convincing. (2) In the Space Conversion Attention Mechanism, we employ the maximum strategy to match the vertical and horizontal weight coefficients to select the regions of interest. However, we tend to ignore the data in the small value regions using the maximum value strategy, resulting in data loss. Therefore, it is worth thinking about considering both maximum and minimum values. (3) The actual data volume of Alzheimer's disease collected in this study is not enough. In the future, the Alzheimer's MRI image data should be further enriched to improve the generalization ability of the model.

CONCLUSION

To address the problems of classification imbalance of the Alzheimer's MRI datasets, small structural changes during different disease stages, small proportion of feature regions to the whole MRI image, and scattered features, we propose a novel method for recognizing different disease stages of Alzheimer's disease based on class-balancing preprocessing and Multi-Phantom Convolution and Space Conversion Attention Mechanism recognition network (MPC-STANet). First, we perform class-balancing preprocessing on the Alzheimer's MRI datasets using data expansion methods such as flipping, adding noise and contrast adjustment, and SMOTE. Then, we propose the MPC-STANet with ResNet50 as the backbone network. In the MPC-STANet, Dilated Convolution is used to increase the perceptual field of the network to recognize scattered feature regions, and Space Conversion Attention Mechanism is used to enhance feature extraction of subtle changes in the MRI Alzheimer's image. Based on Multi-Phantom Convolution, Multi-Phantom Residual Block (including Multi-Conv Block and Multi-Identity Block) is proposed to extract subtle brain feature points. For the recognition of different disease stages of Alzheimer's disease, the proposed MPC-STANet has higher recognition accuracy and a smaller number of parameters compared with the ResNet50 backbone network. The experimental results indicate that the recognition accuracy of the MPC-STANet is 96.2% and the number of parameters is only 1.69M higher than that of ResNet50.

Based on the detection of the disease stages of Alzheimer's disease has been a hot research topic in the field of computer vision-aided diagnosis, The MPC-STANet can be used for disease stage recognition after acquiring the Alzheimer's MRI dataset, which is significant for doctors to distinguish the disease and take corresponding treatment. Future research in this study will focus on how the network can handle complex structural brain features, how to enhance the extraction ability for subtle and scattered features, and how to handle datasets that are not preprocessed. In addition, we need to consider how to further optimize the structure of the network model to facilitate a more effective recognition of Alzheimer's disease and delay the deterioration of this disease promptly.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

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.

AUTHOR CONTRIBUTIONS

YL: methodology, writing-original draft preparation, conceptualization, and data curation. KT: software, data acquisition, and investigation. GZ: validation and project administration. AC: supervision and funding acquisition. WC: model guidance. RL: formal analysis and resources. LL: visualization and writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Changsha Municipal Natural Science Foundation (Grant No. kq2014160); in part by the National Natural Science Foundation in China (Grant No. 61703441); in part by the key projects of the Department of Education Hunan Province (Grant No. 19A511); and in part by the Hunan Key Laboratory of Intelligent Logistics Technology (2019TP1015).

REFERENCES

Andrieu, S., Coley, N., Lovestone, S., Aisen, P. S., and Vellas, B. (2015). Prevention of sporadic Alzheimer's disease: lessons learned from clinical trials and future directions. *Lancet Neurol.* 14, 926–944. doi: 10.1016/S1474-4422(15)00153-2

Bae, J. B., Lee, S., Jung, W., Park, S., Kim, W., Oh, H., et al. (2020). Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. *Sci. Rep.* 10:22252. doi: 10.1038/s41598-020-79243-9

Beitz, J. M. (2014). Parkinson's disease: a review. *Front. Biosci.* 6, 65–74. doi: 10.4103/0028-3886.226451

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16, 321–357. doi: 10.1186/1756-0381-6-16

Chen, J. E., and Glover, G. H. (2015). Functional magnetic resonance imaging methods. *Neuropsychol. Rev.* 25, 289–313.

De la Torre, J. C. (1999). Critical threshold cerebral hypoperfusion causes Alzheimer's disease? *Acta Neuropathol.* 98, 424–436. doi: 10.1007/s004010051044

Dubois, B., Padovani, A., Scheltens, P., Rossi, A., and Dell'Agnello, G. (2016). Timely diagnosis for Alzheimer's disease: a literature review on benefits and challenges. *J. Alzheimers Dis.* 49, 617–631. doi: 10.3233/JAD-150692

Fei, Z., Yang, E., Yu, L., Li, X., Zhou, H., Zhou, W., et al. (2022). A Novel deep neural network-based emotion analysis system for automatic detection of mild cognitive impairment in the elderly. *Neurocomputing* 468, 306–316.

Frisoni, G. B., Fox, N. C., Jack, C. R. Jr., Scheltens, P., and Thompson, P. M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nat. Rev. Neurol.* 6, 67–77.

Guo, L., Li, N., Jia, F., Lei, Y., and Lin, J. (2017). A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* 240, 98–109.

Hanmugam, J. V., Duraisamy, B., Simon, B. C., and Bhaskaran, P. (2022). Alzheimer's disease classification using pre-trained deep networks. *Biomed. Signal Process. Control* 71:103217.

Hazarika, R. A., Maji, A. K., Syiem, R., Sur, S. N., and Kandar, D. (2022). Hippocampus Segmentation Using U-Net Convolutional Network from Brain Magnetic Resonance Imaging (MRI). *J. Digit. Imaging*. [Epub ahead of print]. doi: 10.1007/s10278-022-00613-y

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (Las Vegas, NV: IEEE), 770–778.

Hu, H., Wei, Y., and Zhou, Y. (2021). Product-harm crisis intelligent warning system design based on fine-grained sentiment analysis of automobile complaints. *Complex Intell. Syst.* 28, 1–8. doi: 10.1007/s40747-021-00306-z

Hu, J., Shen, L., and Sun, G. (2018). "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (Salt Lake City, UT: IEEE), 7132–7141.

Huang, Z., Xu, H., Su, S., Wang, T., Luo, Y., Zhao, X., et al. (2020). A computer-aided diagnosis system for brain magnetic resonance imaging images using a novel differential feature neural network. *Comput. Biol. Med.* 121:103818. doi: 10.1016/j.compbiomed.2020.103818

Ieracitano, C., Mammone, N., Bramanti, A., Hussain, A., and Morabito, F. C. (2019). A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings. *Neurocomputing* 323, 96–107.

Kaggle (2022). *Alzheimer Detection and Classification*. San Francisco: Kaggle.

Lebedev, A. V., Westman, E., Van Westen, G. J., Kramberger, M. G., Lundervold, A., and Aarsland, D. (2014). Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage* 6, 115–125. doi: 10.1016/j.nicl.2014.08.023

Li, J., Luong, M. T., Jurafsky, D., and Hovy, E. (2015). When are tree structures necessary for deep learning of representations? *arXiv* [Preprint]. doi: 10.48550/arXiv.1503.00185

Li, Y., Wang, Y., Wu, G., Shi, F., Zhou, L., Lin, W., et al. (2012). Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. *Neurobiol. Aging* 33, e15–e427. doi: 10.1016/j.neurobiolaging.2010.11.008

Lockhart, S. N., and DeCarli, C. (2014). Structural imaging measures of brain aging. *Neuropsychol. Rev.* 24, 271–289. doi: 10.1007/s11065-014-9268-3

López, V., Fernández, A., García, S., Plade, V., and Herrera, F. (2013). An insight into classification with imbalanced data: empirical results and current trends on using data intrinsic characteristics. *Inf. Sci.* 250, 113–141.

Lu, L., Zhang, J., Xie, Y., Gao, F., Xu, S., Wu, X., et al. (2020). Wearable health devices in health care: narrative systematic review. *JMIR mHealth uHealth* 8:e18907. doi: 10.2196/18907

Magnin, B., Mesrob, L., Kinkingnéhun, S., élégrini-Issac, M., Colliot, O., Sarazin, M., et al. (2009). Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology* 51, 73–83. doi: 10.1007/s00234-008-0463-x

Maxwell, A. E., Warner, T. A., and Fang, F. (2018). Implementation of machine-learning classification in remote sensing: an applied review. *Int. J. Remote Sens.* 39, 2784–2817.

Negin, F., Rodriguez, P., Koperski, M., Kerboua, A., Gonzàlez, J., Bourgeois, J., et al. (2018). PRAXIS: towards automatic cognitive assessment using gesture recognition. *Expert Syst. Appl.* 106, 21–35.

Nelson, P. T., Alafuzoff, I., Bigio, E. H., Bouras, C., Braak, H., Cairns, N. J., et al. (2012). Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature. *J. Neuropathol. Exp. Neurol.* 71, 362–381. doi: 10.1097/NEN.0b013e31825018f7

Royce, C. S., Hayes, M. M., and Schwartzstein, R. M. (2019). Teaching critical thinking: a case for instruction in cognitive biases to reduce diagnostic errors and improve patient safety. *Acad. Med.* 94, 187–194. doi: 10.1097/ACM.0000000000002518

Sabuncu, M. R., and Konukoglu, E. (2015). Clinical prediction from structural brain MRI scans: a large-scale empirical study. *Neuroinformatics* 13, 31–46. doi: 10.1007/s12021-014-9238-1

Sarraf, S., and Tofighi, G. (2016). Classification of alzheimer's disease using fmri data and deep learning convolutional neural networks. *arXiv* [Preprint]. doi: 10.48550/arXiv.1603.08631

Savaş, S. (2022). Detecting the Stages of Alzheimer's Disease with Pre-trained Deep Learning Architectures. *Arab. J. Sci. Eng.* 47, 2201–2218.

Sun, H., Wang, A., Wang, W., and Liu, C. (2021). An improved deep residual network prediction model for the early diagnosis of alzheimer's disease. *Sensors* 21:4182. doi: 10.3390/s21124182

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, J., et al. (2016). "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (Las Vegas, NV: IEEE), 2818–2826.

Toğaçar, M., Cömert, Z., and Ergen, B. (2021). Enhancing of dataset using DeepDream, fuzzy color image enhancement and hypercolumn techniques to detection of the Alzheimer's disease stages by deep learning model. *Neural Comput. Appl.* 33, 9877–9889.

Wang, S., Tang, C., Sun, J., and Zhang, Y. (2019). Cerebral micro-bleeding detection based on densely connected neural network. *Front. Neurosci.* 13:422. doi: 10.3389/fnins.2019.00422

Wang, X., Girshick, R., Gupta, A., and He, H. (2018). "Non-local neural networks," in . *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (Salt Lake City, UT: IEEE), 7794–7803. doi: 10.3390/s21206873

Woo, S., Park, J., Lee, J. Y., and Kweon, I. S. (2018). "Cbam: Convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, (Cham: Springer), 3–19. doi: 10.1371/journal.pone.0264551

Wu, Y., Deng, L., Li, G., Zhu, J., and Shi, L. (2018). Spatio-temporal backpropagation for training high-performance spiking neural networks. *Front. Neurosci.* 12:331. doi: 10.3389/fnins.2018.00331

Wu, Y. H., and Swaab, D. F. (2005). The human pineal gland and melatonin in aging and Alzheimer's disease. *J. Pineal Res.* 38, 145–152. doi: 10.1111/j.1600-079X.2004.00196.x

Xiao, Y., Yin, H., Wang, S. H., and Zhang, Y. D. (2021). TReC: transferred ResNet and CBAM for Detecting Brain Diseases. *Front. Neuroinform.* 15:781551. doi: 10.3389/fninf.2021.781551

Young, J., Modat, M., Cardoso, M. J., Mendelson, A., Cash, D., Ourselin, S., et al. (2013). Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. *NeuroImage* 2, 735–745. doi: 10.1016/j.nicl.2013.05.004

Yu, F., and Koltun, V. (2015). Multi-scale context aggregation by dilated convolutions. *arXiv* [Preprint]. doi: 10.48550/arXiv.1511.07122

Zhao, Y., and Zhao, B. (2013). Oxidative stress and the pathogenesis of Alzheimer's disease. *Oxidative Med. Cell. Longev.* 2013: 316523.

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 24 June 2022
doi: 10.3389/fnagi.2022.943436

Discriminant Subspace Low-Rank Representation Algorithm for Electroencephalography-Based Alzheimer's Disease Recognition

*Tusheng Tang[1], Hui Li[1], Guohua Zhou[2,3], Xiaoqing Gu[3] and Jing Xue[4]**

[1] School of Computer Science and Information Engineering, Changzhou Institute of Technology, Changzhou, China, [2] School of Information Engineering, Changzhou Institute of Industry Technology, Changzhou, China, [3] School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, China, [4] Department of Nephrology, The Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi, China

Alzheimer's disease (AD) is a chronic progressive neurodegenerative disease that often occurs in the elderly. Electroencephalography (EEG) signals have a strong correlation with neuropsychological test results and brain structural changes. It has become an effective aid in the early diagnosis of AD by exploiting abnormal brain activity. Because the original EEG has the characteristics of weak amplitude, strong background noise and randomness, the research on intelligent AD recognition based on machine learning is still in the exploratory stage. This paper proposes the discriminant subspace low-rank representation (DSLRR) algorithm for EEG-based AD and mild cognitive impairment (MCI) recognition. The subspace learning and low-rank representation are flexibly integrated into a feature representation model. On the one hand, based on the low-rank representation, the graph discriminant embedding is introduced to constrain the representation coefficients, so that the robust representation coefficients can preserve the local manifold structure of the EEG data. On the other hand, the least squares regression, principle component analysis, and global graph embedding are introduced into the subspace learning, to make the model more discriminative. The objective function of DSLRR is solved by the inexact augmented Lagrange multiplier method. The experimental results show that the DSLRR algorithm has good classification performance, which is helpful for in-depth research on AD and MCI recognition.

Keywords: electroencephalography, Alzheimer's disease, low-rank representation, subspace learning, classification

OPEN ACCESS

Edited by:
Mohammad Khosravi, Persian Gulf University, Iran

Reviewed by:
Ximing Xia, Nanjing Institute of Technology (NJIT), China
Runmin Liu, Wuhan Sports University, China

****Correspondence:***
Jing Xue
xuejing@njmu.edu.cn

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

Received: *13 May 2022*
Accepted: *06 June 2022*
Published: *24 June 2022*

Citation:
Tang T, Li H, Zhou G, Gu X and Xue J (2022) Discriminant Subspace Low-Rank Representation Algorithm for Electroencephalography-Based Alzheimer's Disease Recognition. Front. Aging Neurosci. 14:943436. doi: 10.3389/fnagi.2022.943436

INTRODUCTION

Alzheimer's disease (AD) is a disease characterized by memory loss, slow and gradual changes in brain function, and the manifestations of intellectual loss (Zhang et al., 2021). With the advancement of global aging, AD has now become a major public health problem affecting the world. The existing treatment of AD can only temporarily help relieve memory and cognition, but

not a cure. To obtain disease-controlling treatments, it is an urgent need to classify the course of AD for early diagnosis. And especially, the National Institutes of Health revised the clinical diagnostic criteria for AD, characterizing research guidelines for early diagnosis and treatment (Cummings, 2021). The progression of AD is mainly divided into three stages. The first is the early clinical stage with no symptoms; the second is the intermediate stage with mild cognitive impairment (MCI); and the final stage with dementia symptoms (Mirzaei and Adeli, 2022).

More researchers are studying methods that can sensitively and conveniently monitor AD, involving cognitive neuropsychological detection, biochemical detection, neuroimaging detection, and so on. In recent years, electroencephalography (EEG) has become an important tool for studying human brain activity (Ghorbanian et al., 2015). Noninvasive EEG imaging methods are directly related to neural local field potentials and have a high temporal resolution. The millisecond-level temporal resolution and direct electrophysiological information provided by EEG can accurately reflect cognitive behaviors related to human neural activity. Therefore, more studies are beginning to use EEG for the diagnosis and prediction of early AD. For example, EEG spectral studies have revealed that EEG diffuse slow waves are a major feature of AD. EEG studies of AD patients have shown that the reduced power in the alpha (8–15 Hz) band and the increased power in the delta (0.5–4 Hz) band are significant features of AD (Fröhlich et al., 2021). The increase in power in the theta (4–8 Hz) band and the decrease in power in the beta (15–30 Hz) band also indicate that they can be useful for detecting MCI to AD transitions (Maturana-Candelas et al., 2020). Recently, machine learning technology has been widely used in the analysis of brain imaging data, which has greatly promoted the development of cognitive neuroscience. Most of the research revolves around feature extraction and classifier optimization. In terms of feature extraction, Wen et al. (2020) first converted the EEG signals into multispectral images and then used a deep convolutional neural network learning model for EEG classification. Similarly, Ieracitano et al. (2019a) drew the power spectral density of the EEG into the form of a spectrogram, and converted the EEG signal classification into a CNN-based image classification problem. Ieracitano et al. (2019b) spliced the continuous wavelet transform features and bispectral features of EEG signals to achieve the fusion of the two types of features. The advantage of this algorithm is that the fused features can obtain higher accuracy than only using one type of feature. The disadvantage is that the correlation between features is not considered enough. At the same time, the dimension of fusion features is greatly increased, which is easy causing the over-fitting problem.

In terms of classification algorithms, Miltiadous et al. (2021) compared six classification algorithms for EEG analysis for frontotemporal dementia in AD and verified the effectiveness of these algorithms. This study provided solutions for the early diagnosis of frontotemporal dementia. Anuradha and Jamal (2021) detected the progression of AD by detecting abnormal behavior in EEG. The authors used a feed-forward artificial neural network as a classifier to perform EEG feature analysis on abnormal and normal subjects and obtained a classification accuracy of 94.4%. Ge et al. (2020) exploited the robust biomarkers in EEG, combined linear discriminant analysis as a classifier, and proposed a systematic identification framework based on signal processing and computer-aided techniques for the detection of AD. Araujo et al. (2022) developed an intelligent system that can distinguish various stages of AD through EEG signals. The system used wavelet packet to extract multi-band features of EEG signals and used multiple machine learning methods as classification models.

Electroencephalography signals can reflect the functional state of the brain and the activity of brain physiological structures. The difficulties in classifying EEG signals using machine learning algorithms are as follows: first, the amplitude of the EEG signals is usually around 50 μv. The EEG signals are very weak, and their background noise is usually very strong. Second, EEG signals have strong randomness. In the process of acquisition, EEG signals will not only be stimulated by the outside world but also produce interference signals due to their own blinking and other actions. Therefore, it is still a challenging task to use machine learning methods to identify AD based on EEG signals. To solve this problem, the researchers usually reduce the dimension of EEG high-dimensional data and extract a small amount of the most valuable compact information, which not only saves storage space and processing time but also enables learning a robust model (Lei et al., 2021). Subspace learning and low-rank representation can well achieve this goal. Subspace learning is a well-known dimension reduction method in machine learning. Its main goal is to adopt appropriate strategies to map high-dimensional original data into the low-dimensional subspace to reduce the data dimension. Low-rank representation (LRR) can effectively separate the noise in the EEG signals to restore clean data and obtain accurate subspace segmentation of data.

Inspired by the strong theory of subspace learning and low-rank representations, this paper proposes an EEG-based discriminant subspace low-rank representation learning algorithm (DSLRR) for AD recognition. On the one hand, based on the low-rank representation, DSLRR utilizes the supervised information and local manifold information by least squares regression (LSR) and graph discriminant embedding. On the other hand, DSLRR introduces principal component analysis (PCA) and global preserved constraints into the subspace of learning. The algorithm optimization adopts a strategy of alternating parameter updates using the inexact augmented Lagrange multiplier method. Our contribution is as follows: (1) The DSLRR algorithm combines subspace learning and low-rank representation in a flexible manner. (2) By introducing global graph embedding and PCA term, the data projection can preserve the global structure information of EEG data in the discriminant subspace. (3) The learned low-rank representation coefficient can effectively avoid the negative effects of the original data's redundant features and noise information. (4) By introducing LSR and graph discriminant

embedding, the learned low-rank representation coefficient can explicitly contain the intrinsic local manifold structure and discriminant information of EEG data. The experiments on four EEG datasets verify that the DSLRR algorithm can be effectively used for the recognition of AD, MCI, and healthy control (HC).

BACKGROUND

Electroencephalography Dataset for Alzheimer's Disease and Mild Cognitive Impairment Recognition

The EEG data were obtained from 109 participants recruited at the IRCCS Centro Neurolesi Bonino-Pulejo in Italy, including 23 HC, 49 AD, and 37 MCI (Fiscon et al., 2018). The age of men and women and the proportion of genders are shown in **Figure 1**. The EEG data collection time was from 2012 to 2013. The scalp electrode position was determined using the international 10–20 system, and EEG data from 19 electrodes were collected. The sampling frequency was 256 or 1,024, and the acquisition time of EEG signals was 300 s. To reduce the effect of the artifact, the EEG signals from 60 to 240 s were selected, and the adopted normalized sampling frequency was 256 Hz. Feature extraction adopted the fast Fourier transform, which divided 180 s of data into six epochs of 30 s, and extracted 16 Fourier coefficients. Therefore, 304 features (19 electrodes × 16 Fourier coefficients) were available for each sample.

Subspace Learning

We have a labeled dataset with n samples $Y = \left[y_1, ..., y_n\right] \in R^{d\times n}$, where y_i represents the ith training sample, and its class label matrix is $\bar{Y} = \left[\bar{y}_1, ..., \bar{y}_n\right] \in R^{C\times n}$. The dimension of the sample is d, and n samples are divided into C classes. When the dimensionality of the original EEG data is high, the data computational and storage costs will be very large. Thus, a common solution is to project the high-dimensional data into a low-dimensional space (Lei et al., 2021). Let $Q \in R^{d\times C}$ be the projection matrix, the projection data can be represented as $V = [v_1, ..., v_n] \in R^{C\times n}$ in the label space, where $V = Q^T Y$.

Generally speaking, the premise of manifold subspace learning is that the data exists in high-dimensional space in the form of manifold embedding from low-dimensional space data. The key point of manifold learning is to ensure that low-dimensional data can reflect the inherent structural information contained in high-dimensional space (Zhang et al., 2020). As a commonly used manifold learning algorithm, locality-preserving projection (LPP) preserves the local neighbor relationship of the data by using an adjacency graph and affinity matrix (Weng and Shen, 2008). The LPP algorithm consists of three steps. Step 1 is to construct an adjacency graph. For example, we construct an adjacency graph using the k-nearest neighbor algorithm. The nearest neighbors of each point connected to it are known as neighbor nodes. Step 2 is to assign weights to each edge. In the adjacency graph, the affinity matrix represents the similarity between sample points, which can generally be calculated using the two-value method, cosine distance or Gaussian kernel function. For example, the affinity matrix **E** constructed by the two-value method can be defined as follows:

$$E_{\mathrm{ij}} = \begin{cases} 1, & \text{if } y_i \in N_k(y_j) \text{ or } y_j \in N_k(y_i) \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where $N_k(y_i)$ represents the k nearest neighbor nodes of y_i .

Low-Rank Representation

Low-rank representation aims to exploit the sparsity of matrix singular values to model high-dimensional data in multi subspace (Li et al., 2017; Jiang et al., 2021). Given a dataset **Y**, the LRR algorithm regards the input data itself as a dictionary and uses the basis in the dictionary to linearly represent the sample points, while minimizing its rank. The optimization problem of LRR can be described as follows:

$$\begin{aligned} &\min_{L} \operatorname{rank}(L), \\ &s.t.\ Y = YL, \end{aligned} \tag{2}$$

where $L \in R^{n\times n}$ is the representation coefficients of **Y**, which reflects the global correlation between the original data samples. In theory, the coefficient matrix **L** obtained by the LRR should be a block diagonal matrix. That is to say, each block corresponds to a subspace, the number of blocks represents the number of data subspaces, and the size of the block corresponds to the dimension of the subspace.

Eq. (2) is not a convex optimization problem due to its discrete. Using the nuclear norm instead of rank(L), Eq. (2) can be transformed into the convex optimization problem as:

$$\begin{aligned} &\min_{L} ||L||_* , \\ &s.t.\ Y = YL, \end{aligned} \tag{3}$$

where $||||_*$ is the nuclear norm.

Considering the noise or sparse error in **Y**, LRR enhances the model's robustness by improving the correlation between the individual columns of **L**, and the problem of LRR can be written as:

$$\begin{aligned} &\min_{L,S} ||L||_* + \theta\, ||S||_1 , \\ &s.t.\ Y = YL + S, \end{aligned} \tag{4}$$

where $S \in R^{d\times n}$ is sparse component of **Y**. θ is the regularization parameter.

Obviously, LRR decomposes the data **Y** into low-rank representation YL and sparse representation **S**. The former component YL generally represents the main features contained in **Y**, and the latter generally represents the redundant features and noise information contained in **Y**. In the clean data scenario, S represents the reconstruction error. Therefore, **L** can accurately indicate the subspace segmentation of **Y**, which ensures the robustness of the learned model. However, LRR ignores the role of local structure information in data and does not exploit the supervised information in the training data. Therefore, LRR cannot reflect the intra-class identity and inter-class dissimilarity in low-rank representation.

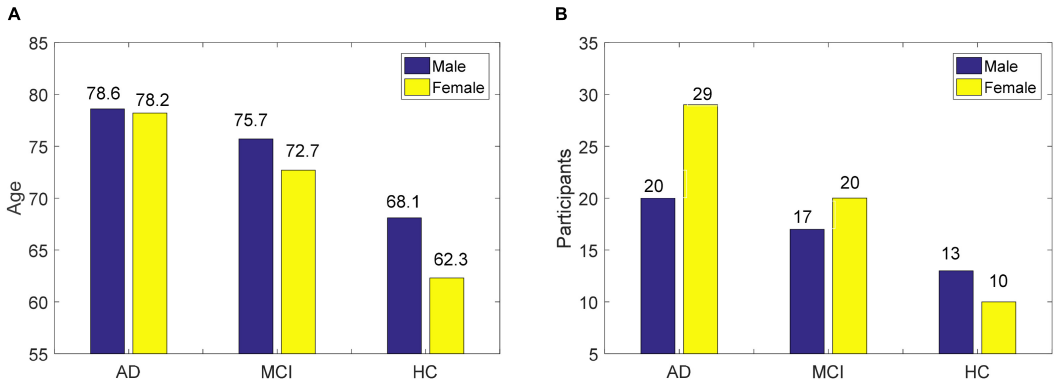


FIGURE 1 | The basic information of EEG data used in this study, **(A)** age of men and women, and **(B)** proportion of gender.

DISCRIMINANT SUBSPACE LOW-RANK REPRESENTATION ALGORITHM

Objective Function

Discriminant Margin Term on Representation Coefficients

To learn the discriminant low-rank representations, we introduce graph discriminant embedding (Huang et al., 2018) into our algorithm, which combines supervised information to define intra-class and inter-class graph affinity matrices. We think if two EEG samples are closer in the original space, their representation coefficients will be close to each other. The compactness between samples of the same class and the separability between samples of different classes is the important knowledge in discriminant low-rank representations. To this end, we define affinity matrices E^{com} and E^{sep} to represent the similar relationship between intra-class and inter-class, respectively:

$$E_{\text{i,j}}^{\text{com}} = \begin{cases} e^{\frac{-||y_i - y_j||^2}{t}}, & \text{if } y_i \in \hat{N}_k(y_i) \text{ or } y_j \in \hat{N}_k(y_i), \bar{y}_j = \bar{y}_i \\ e^{\frac{-t}{||y_i - y_j||^2}}, & \text{if } y_i \notin \hat{N}_k(y_i) \text{ or } y_j \notin \hat{N}_k(y_i), \bar{y}_j = \bar{y}_i \\ 0, & \text{if } \bar{y}_j \neq \bar{y}_i \end{cases} \tag{5}$$

$$E_{\text{i,j}}^{\text{sep}} = \begin{cases} e^{\frac{-||y_i - y_j||^2}{t}}, & \text{if } y_i \in \tilde{N}_k(y_i) \text{ or } y_j \in \tilde{N}_k(y_i), \bar{y}_j \neq \bar{y}_i \\ e^{\frac{-t}{||y_i - y_j||^2}}, & \text{if } y_i \notin \tilde{N}_k(y_i) \text{ or } y_j \notin \tilde{N}_k(y_i), \bar{y}_j \neq \bar{y}_i \\ 0, & \text{if } \bar{y}_j = \bar{y}_i \end{cases} \tag{6}$$

where $\hat{N}_k()$ and $\tilde{N}_k()$ represent the k-nearest neighbor samples of intra-class and inter-class, respectively. The parameter t ($t > 0$) is the weight parameter used to adjust the correlation between two samples. We set $t = 1$ in this study.

Then we define the discriminant margin term $\varsigma_1(L)$ on representation coefficients:

$$\begin{aligned} \varsigma_1(L) &= \sum_{i=1}^{n} \varsigma_i(L_i) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} (\left|\left|L_i - L_j\right|\right|^2 E_{\text{i,j}}^{\text{com}} - \left|\left|L_i - L_j\right|\right|^2 E_{\text{i,j}}^{\text{sep}}) \\ &= \text{Tr}(L^T U L) \end{aligned} \tag{7}$$

where $U = E^{\text{com}} - E^{\text{sep}} + \varepsilon I$, ε is a very small positive. Eq. (7) represents the intra-class compactness and the inter-class dissimilarity in representation coefficients. Its essence is to excavate the local structural information representation coefficients. In addition, Eq. (7) can avoid the influence of the redundant information and noise of the original data.

Global Structure Term on Projection

We adopt the affinity matrix $\mathbf{E}$ to represent the correlation between two samples using supervised information. The element e_{ij} in $\mathbf{E}$ is computed as:

$$e_{ij} = \begin{cases} 1, & \text{if } y_i \text{ and } y_j \text{ are of the same class} \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

To preserve the global discriminant information of the original data in the subspace, we introduce the global structure term on projection:

$$\begin{aligned} \varsigma_2(Q) &= \tfrac{1}{2} \sum_{i,j} e_{\text{ij}} \left|\left|Q^T y_i - Q^T y_j\right|\right|_2^2 - \beta \text{Tr}(Q^T Y Y^T Q) \\ &= \text{Tr}(Q^T Y E Y^T Q) - \beta \text{Tr}(Q^T Y Y^T Q) \\ &= \text{Tr}(Q^T Y (E - \beta I) Y^T Q) \end{aligned} \tag{9}$$

where β is the regularization parameter.

The first factor $\sum_{i,j} e_{ij} \left|\left|Q^T y_i - Q^T y_j\right|\right|_2^2$ in Eq. (9) is the global preserved component on projection. Obviously, when this component reaches the minimum, the distance of samples of the same class will be as close as possible in the projection subspace. The second component $\text{Tr}(Q^T Y Y^T Q)$ in Eq. (9) is the PCA component on projection. Its goal is to ensure that the projecting data in the low-dimensional subspace can depict the inherent structure information contained in the original space.

Least Squares Regression Term

As an effective supervised learning method, LSR learns the linear projection that transforms the sample to the label space, and obtains the regression vector as the data representation in the label space (Zhao et al., 2022). Therefore, we try to find a projection matrix with the help of LSR in the low-rank representation. Different from the traditional projection method on the original data, the DSLRR algorithm only uses clean data

representation to learn the projection matrix in the low-rank representation framework, which can not be affected by the redundant information of EEG data. This idea can be obtained as:

$$\begin{aligned} &\varsigma_3(Q,L,S) = ||L||_* + \theta\,||S||_1 + \gamma\left|\left|V-\bar{Y}\right|\right|_F^2 + \eta\,||V||_F^2, \\ &s.t.\ Y = YL + S, \\ &\quad V = Q^T YL, \\ &\quad 1_n^T L = 1_n^T. \end{aligned} \tag{10}$$

where γ and η are regularization parameters.

Equation (10) tries to minimize the least squares loss between the regression results and the corresponding regression target. In addition, in the low-rank representation framework, the compact representation of the data can be learned through subspace projection.

The Objective Function

We integrate Eqs (7), (9), and (10) into a learning model, and obtain the objective function of the DSLRR algorithm:

$$\begin{aligned} &\min \varsigma_1(L) + \varsigma_2(Q) + \varsigma_3(Q,L,S) \\ &= \min_{Q,L,S} \mu \mathrm{Tr}(L^T UL) + \alpha \mathrm{Tr}(Q^T Y(E-\beta I)Y^T Q) + ||L||_* + \\ &\theta\,||S||_1 + \gamma\left|\left|V-\bar{Y}\right|\right|_F^2 + \eta\,||V||_F^2, \\ &\quad s.t.\ Y = YL + S, \\ &\qquad V = Q^T YL, \\ &\qquad 1_n^T L = 1_n^T. \end{aligned} \tag{11}$$

where α and μ are regularization parameters.

From Eq. (11), we can see that the DSLRR algorithm combines subspace learning and low-rank representation into a learning model. Based on low-rank representation learning, the compact and discriminant low-rank representation can be reinforced by graph discriminant embedding. Based on subspace learning, the discriminant projection can be obtained by LSR, global structure preserved, and PCA technologies.

Optimization

There are three unsolved parameters {**Q**, **L**, **S**} in Eq. (11). To make Eq. (11) separable, the relaxation matrix Λ is introduced to represent L. Substitute the constraint $V = Q^T YL$ into Eq. (11), Eq. (11) can be re-written as:

$$\begin{aligned} &\min_{Q,L,S} ||\Lambda||_* + \theta\,||S||_1 + \gamma\left|\left|Q^T YL - \bar{Y}\right|\right|_F^2 + \eta\left|\left|Q^T YL\right|\right|_F^2 + \\ &\mu \mathrm{Tr}(L^T UL) + \alpha \mathrm{Tr}(Q^T Y(E-\beta I)Y^T Q), \\ &s.t.\ Y = YL + S, \\ &\quad 1_n^T L = 1_n^T, \\ &\quad L = \Lambda. \end{aligned} \tag{12}$$

We optimize three parameters by the inexact augmented Lagrange multiplier algorithm in an iterative optimization strategy (Kang et al., 2015). Eq. (12) has the following form:

$$\begin{aligned} &\min_{Q,L,S,\Lambda} ||\Lambda||_* + \theta\,||S||_1 + \gamma\left|\left|Q^T YL - \bar{Y}\right|\right|_F^2 + \eta\left|\left|Q^T YL\right|\right|_F^2 + \\ &\mu \mathrm{Tr}(L^T UL) + \alpha \mathrm{Tr}(Q^T Y(E-\beta I)Y^T Q) + \mathrm{Tr}(\tau_a^T(Y - YL - S)) \\ &+\mathrm{Tr}(\tau_b^T(\Lambda - L)) + \mathrm{Tr}(\tau_c^T(1_n^T L - 1_n^T)) + \tfrac{\delta}{2}(||\text{Y-YL} - S||_F^2 + \\ &||\Lambda - L||_F^2 + \left|\left|1_n^T L - 1_n^T\right|\right|_F^2), \end{aligned} \tag{13}$$

where δ is a trade-off parameter. The matrices $\tau_a \in R^{d\times n}$, $\tau_b \in R^{d\times n}$, $\tau_c \in R^{d\times n}$, and $\tau_d \in R^{d\times n}$ are the Lagrange multipliers.

1) Optimize **Q**, while fixing the other parameters. Eq. (13) can be written as:

$$\min_Q \gamma\left|\left|Q^T YL - \bar{Y}\right|\right|_F^2 + \eta\left|\left|Q^T YL\right|\right|_F^2 + \alpha \mathrm{Tr}(Q^T Y(E-\beta I)Y^T Q). \tag{14}$$

We can get the closed-solution of **Q** as:

$$Q = [Y((\alpha(E-\beta I)) + (\gamma+\eta)LL^T)Y^T]^{-1} YL\bar{Y}^T. \tag{15}$$

2) Optimize Λ , while fixing the other parameters. Eq. (13) can be written as:

$$\begin{aligned} &\min_\Lambda\ ||\Lambda||_* + \mathrm{Tr}(\tau_c^T(L-\Lambda)) + \tfrac{\delta}{2}||L-\Lambda||_F^2 \\ &= \min_\Lambda\ \tfrac{1}{\delta}||\Lambda||_* + \tfrac{1}{2}\left|\left|\Lambda - (L + \tfrac{1}{\delta}\tau_c)\right|\right|_F^2. \end{aligned} \tag{16}$$

We use the singular value thresholding operator (Cai et al., 2010; Li et al., 2017) to solve Eq. (16). We employ the singular value decomposition algorithm on $L + \frac{1}{\delta}\tau_c$ as $L + \frac{1}{\delta}\tau_c = \mathrm{H}\Sigma\Delta$, where H is the diagonal matrix with its element being a group of singular values $\{\Theta_k\}, 1 \le k \le p$, p is the rank. The matrix Λ can be computed by $\Lambda = \mathrm{H}\Omega_{(1/\delta)}\Sigma\Delta$, in which $\Omega_{(1/\delta)}\Sigma = diag(\{\Theta_k - \frac{1}{\delta}\}_+)$, where "+" means the positive part.

3) Optimize **L**, while fixing the other parameters. Eq. (13) can be written as:

$$\begin{aligned} &\min_L \gamma\left|\left|Q^T YL - \bar{Y}\right|\right|_F^2 + \eta\left|\left|Q^T YL\right|\right|_F^2 + \mu \mathrm{Tr}(L^T UL) + \\ &\mathrm{Tr}(\tau_a^T(Y - YL - S)) + \mathrm{Tr}(\tau_b^T(\Lambda - L)) + \mathrm{Tr}(\tau_c^T(1_n^T L - 1_n^T)) + \\ &\tfrac{\delta}{2}(||\text{Y-YL} - S||_F^2 + ||\Lambda - L||_F^2 + \left|\left|1_n^T L - 1_n^T\right|\right|_F^2), \end{aligned} \tag{17}$$

Let the first derivative of **L** in Eq. (16) be zero, we have,

$$\Theta_a L = \Theta_b + \Theta_c, \tag{18}$$

$$\Theta_a = 2U + 2(\gamma+\eta)Y^T QQ^T Y + \delta(Y^T Y + 1_n 1_n^T + I_n), \tag{19}$$

$$\Theta_b = 2\gamma Y^T Q\bar{Y} + \delta(Y^T \text{Y-Y}^T S + \Lambda + 1_n 1_n^T), \tag{20}$$

$$\Theta_c = Y^T \tau_a - \tau_b - 1_n \tau_c, \tag{21}$$

We can get the closed-solution of **L** as:

$$L = \Theta_a^{-1}(\Theta_b + \Theta_c). \tag{22}$$

4) Optimize **S**, while fixing the other parameters. Eq. (13) can be written as:

$$\min_{S} \frac{\theta}{\delta} ||S||_{2,1} + \frac{1}{2} \left|\left| S - (Y - YL + \frac{1}{\delta}\tau_a) \right|\right|_F^2, \quad (23)$$

According to the theory of (Liu et al., 2013), we can obtain the **S** by

$$S(:, i) = \begin{cases} \frac{||\tau_i|| - \theta}{||\tau_i||} ||\tau_i||, & if \ \frac{\theta}{\delta} < ||\tau_i|| \\ 0, & \text{otherwise} \end{cases} \quad (24)$$

where τ_i is the ith column vector of the matrix τ_a .

Testing

Given test EEG data Y_{test} , we first compute its low-rank representation L_{test} using Eq. (11), while setting parameters γ =0, α =0, and μ =0. Second, we construct the new training set YL and test set $Y_{\text{test}}L_{\text{test}}$. Third, we use the training set YL to train a classifier and build a classifier to predict the label of $Y_{\text{test}}L_{\text{test}}$. In this study, we used nearest neighbor (NN) algorithm as the classifier. The whole training and testing procedure for EEG data recognition are summarized in **Algorithm 1**.

Algorithm 1: DSLRR algorithm for EEG data recognition.

Input: The training EEG data Y and its label $\tilde{Y}$, the testing EEG data Y_{test} ;

Output: the class label of Y_{test} ;

// Construct the new training data on Y

Calculate matrices **E** by Eq. (8), E^{com} by Eq. (5) and E^{sep} by Eq. (6);

Repeat:

  Optimize **Q** using Eq. (15) withΛ , **L**, and **S** fixed;

  OptimizeΛ using Eq. (16) with **Q**, **L**, and **S** fixed;

  Optimize **L** using Eq. (22) withΛ , **Q**, and **S** fixed;

  Optimize **S** using Eq. (24) withΛ , **L**, and **Q** fixed;

Obtain the new testing data YL;

Until Eq. (13) convergence

// Construct the new testing data on Y_{test}

Repeat:

  Optimize **Q** using Eq. (15) withΛ , **L**, and **S** fixed, while setting γ =0, α =0, and μ =0;

  OptimizeΛ using Eq. (16) with **Q**, **L**, and **S** fixed, while setting γ =0, α =0, and μ =0;

  Optimize **L** using Eq. (22) withΛ , **Q**, and **S** fixed, while setting γ =0, α =0, and μ =0;

  Optimize **S** using Eq. (24) withΛ , **L**, and **Q** fixed, while setting γ =0, α =0, and μ =0;

Until Eq. (13) convergence

Obtain the new test data $Y_{\text{test}}L_{\text{test}}$;

// Train a classifier and predict the class label

Train a classifier using training data YL (such as NN classifier, support vector machine);

Test and output the class label of $Y_{\text{test}}L_{\text{test}}$ using the trained classifier.

EXPERIMENT

Experimental Settings

To verify the effectiveness of the DSLRR algorithm, we compared the DSLRR algorithm with the SPCA (Jiang, 2011), LRR (Liu et al., 2013), LRDLSR (Chen and Yang, 2014), JSLC (Lu et al., 2021), and NRLRL (Gao et al., 2021) in the experiment. The LRR algorithm is the baseline algorithm of the DSLRR algorithm. SPCA and JSLC algorithms are subspace learning algorithms. LRDLSR and NRLRL are low-rank representation algorithms. For SPCA, the weight parameters are set in the covariance mixture, α is set inversely proportional to the sample size, and η is searched in $[2^{-5}, 2^{-4}, ..., 2^5]$. For LRR, the parameter λ is searched in $[1, 4, ..., 30]/\sqrt{d}$, where d is the data dimension. For LRDLSR, the parameters α and β are searched in $[10^{-4}, 10^{-3}, ..., 1]$, and the parameters γ and λ are set to be 0.01. For JSLC, subspace dimension and the size of the dictionary are searched in [50, 100,. . . , 300]. The regularization parameters are searched in [0.5, 1,. . . , 5]. For NRLRL, the size of the dictionary is searched in [50, 100,. . . , 300], and λ ,γ , and η are searched in $[2^{-4}, 2^{-3}, ..., 2^4]$. For DSLRR, all regularization parameters are searched in$[2^{-4}, 2^{-3}, ..., 2^4]$, and k-nearest neighbors in $\hat{N}_k()$ and $\tilde{N}_k()$ are searched in [1,. . . , 11].

Due to the limited training EEG samples, we expand the EEG data with the data augmentation strategy. The number of EEG samples in HC, MCI, and AD is 69, 74, and 98, respectively. In this section, the experiments are conducted on four EEG datasets for AD and MCI recognition, namely, (1) HC & AD, (2) HC & MCI, (3) HC & (MCI+AD), and (4) MCI & AD. The ratio of the two classes of samples is 1:1. We randomly select 50 samples in each class for model training, and the rest samples are used for testing. We perform our experiments 10 times and record the classification performance in terms of accuracy, sensitivity, specificity, precision, F-measure, G-mean, and Jaccard. All experiments are conducted by MATLAB on a Windows machine.

Classification Results

The classification results in four EEG datasets are reported in **Tables 1–4**, where the best results are highlighted in bold. These four data sets are binary classification problems. According to the results in **Tables 1–4**, we can see that:

(1) Alzheimer's disease is a population suffering from AD, which has shown clinical symptoms. The EEG signal differentiation between AD and healthy people is the most significant, and the difference between EEG features is more obvious. Therefore, the classification performance in the dataset of AD and HC is high. Although the symptoms of MCI are not as significant as those of AD, there is a certain probability of AD. The difference between the EEG features and those of healthy people is also significant, and the difference between EEG features is also obvious, so the classification performance in the dataset of MCI and HC is also high. In addition, AD and MCI are mixed into one class in the third dataset of HC and (AD+MCI), which

TABLE 1 | Classification results of the comparison algorithms in HC and AD dataset.

	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean	Jaccard
SPCA	92.54	92.11	92.98	93.51	92.46	92.35	86.33
	±2.68	±1.58	±2.09	±2.22	±2.69	±2.75	±3.03
LRR	88.16	92.11	84.21	85.29	88.45	87.99	79.55
	±2.56	±2.06	±1.53	±2.24	±1.79	±1.53	±2.51
LRDLSR	93.42	**99.87**	86.84	88.42	93.84	93.18	88.42
	±2.78	±2.03	±2.72	±2.86	±2.78	±2.77	±2.78
JSLC	94.68	96.84	92.63	93.10	94.87	94.66	90.82
	±1.93	±1.66	±1.80	±2.59	±3.55	±2.01	±1.49
NRLRL	96.05	92.11	99.70	99.74	95.71	95.88	92.11
	±3.03	±1.08	±1.36	±2.08	±2.99	±2.78	±3.46
Our algorithm	**97.74**	95.49	**99.81**	**99.79**	**97.55**	**97.65**	**95.49**
	±2.68	±2.14	±1.57	±1.91	±3.19	±1.64	±1.72

The bold values mean the best performance results.

TABLE 2 | Classification results of the comparison algorithms in MCI and AD dataset.

	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean	Jaccard
SPCA	91.12	91.71	90.53	91.71	91.01	90.70	84.19
	±3.23	±3.24	±2.62	±2.92	±3.00	±1.77	±2.54
LRR	87.89	80.42	95.37	94.77	86.23	87.15	76.85
	±1.99	±3.20	±2.39	±2.95	±2.61	±1.79	±2.52
LRDLSR	92.32	84.32	**99.89**	**99.90**	90.46	91.93	84.32
	±1.39	±2.09	±2.98	±2.78	±2.98	±2.91	±2.55
JSLC	92.98	88.42	97.54	97.70	92.04	92.46	86.41
	±2.49	±2.38	±2.44	±2.55	±2.60	±2.08	±1.78
NRLRL	94.74	89.47	99.53	99.47	94.29	94.51	89.47
	±2.13	±3.34	±2.33	±1.69	±2.22	±3.29	±3.88
Our algorithm	**95.61**	**94.74**	96.49	96.67	**95.43**	**95.48**	**91.40**
	±1.83	±2.09	±1.75	±2.35	±2.19	±2.04	±2.67

The bold values mean the best performance results.

TABLE 3 | Classification results of the comparison algorithms in HC and (MCI+AD) dataset.

	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean	Jaccard
SPCA	92.11	84.21	99.25	99.13	91.43	91.77	84.21
	±2.62	±1.39	±2.98	±2.53	±2.37	±2.46	±2.12
LRR	89.47	84.21	94.74	94.12	88.89	89.32	80.00
	±3.48	±3.08	±1.94	±3.43	±2.92	±2.34	±2.28
LRDLSR	94.74	94.74	94.74	95.16	94.62	94.56	89.90
	±3.08	±3.17	±3.63	±1.25	±2.38	±2.44	±1.37
JSLC	95.61	94.74	96.49	96.37	95.44	95.55	91.67
	±2.18	±1.91	±2.60	±2.46	±1.57	±1.61	±2.67
NRLRL	96.26	92.22	**99.43**	**99.35**	95.86	95.99	92.26
	±2.78	±3.57	±1.62	±3.38	±3.78	±1.89	±3.38
Our algorithm	**98.42**	**99.34**	96.84	97.00	**98.46**	**98.40**	**97.00**
	±2.50	±1.91	±2.65	±2.86	±1.44	±1.29	±2.49

The bold values mean the best performance results.

is significantly distinguishable from healthy EEG signals. Therefore, its classification performance is expectable. The classification accuracy of DSLRR algorithm in AD and HC is 97.74%. The classification accuracy of the DSLRR algorithm in MCI & and HC is 95.61%. The classification accuracy of the DSLRR algorithm in HC and MCI+AD is 98.42%. The classification accuracy of these three datasets is above 97.26%. The experimental results illustrate that DSLRR can better identify MCI and AD from HC.

(2) Compared with the first three datasets, the difference between EEG features between MCI and AD is relatively low. Therefore, the classification performance of each algorithm decreases to a certain extent in the MCI & AD dataset. However, we can see that the DSLRR algorithm still

TABLE 4 | Classification results of the comparison algorithms in HC and MCI dataset.

	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean	Jaccard
SPCA	86.84	**98.76**	73.68	79.17	91.01	90.70	84.19
	±3.20	±3.13	±3.36	±3.56	±2.75	±1.75	±3.09
LRR	84.21	84.21	84.21	84.21	86.23	87.15	76.85
	±2.64	±3.22	±2.36	±3.15	±3.05	±3.00	±2.57
LRDLSR	88.60	89.47	87.72	88.12	90.46	91.93	84.32
	±2.18	±3.20	±2.97	±2.64	±3.08	±3.03	±3.06
JSLC	89.47	98.32	78.95	82.61	92.04	92.46	86.41
	±2.77	±3.01	±3.38	±2.17	±3.05	±1.58	±3.03
NRLRL	90.79	92.11	89.47	90.08	94.29	94.51	89.47
	±2.68	±3.49	±1.69	±3.00	±2.39	±2.71	±1.83
Our algorithm	**93.42**	94.74	**92.11**	**92.46**	**95.43**	**95.48**	**91.40**
	±2.39	±1.93	±2.25	±1.69	±2.12	±2.06	±2.04

The bold values mean the best performance results.

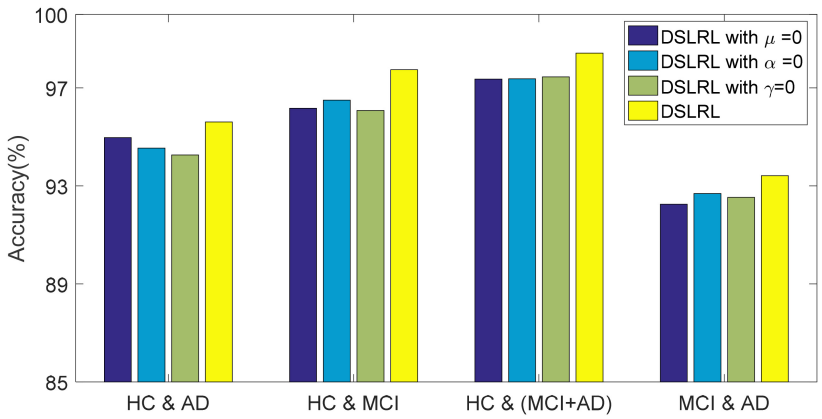


FIGURE 2 | Classification accuracy of DSLRR with ablation experiment in four EEG datasets.

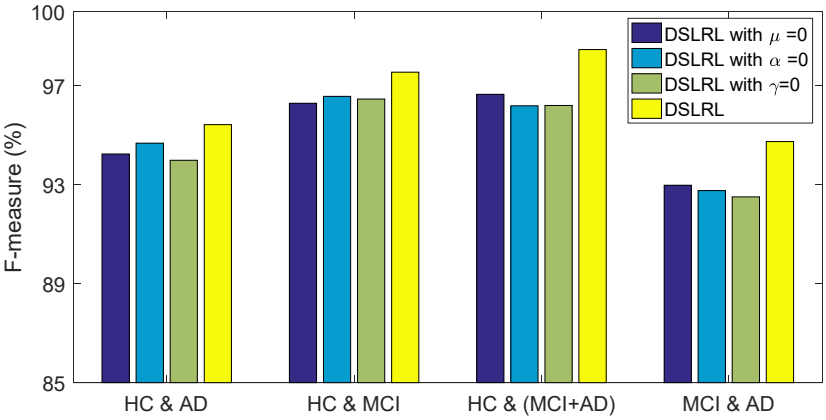


FIGURE 3 | F-measure of DSLRR with ablation experiment in four EEG datasets.

achieves the best values of accuracy, F-measure, G-mean, and Jaccard. On the one hand, through the joint learning of subspace and low-rank representation, the DSLRR algorithm can learn the robust and discriminant projection subspace. On the other hand, by making full use of Laplace manifold and LSR technologies, the DSLRR algorithm can exploit the structure knowledge and manifold structure information of EEG signals. Furthermore, the sum of the

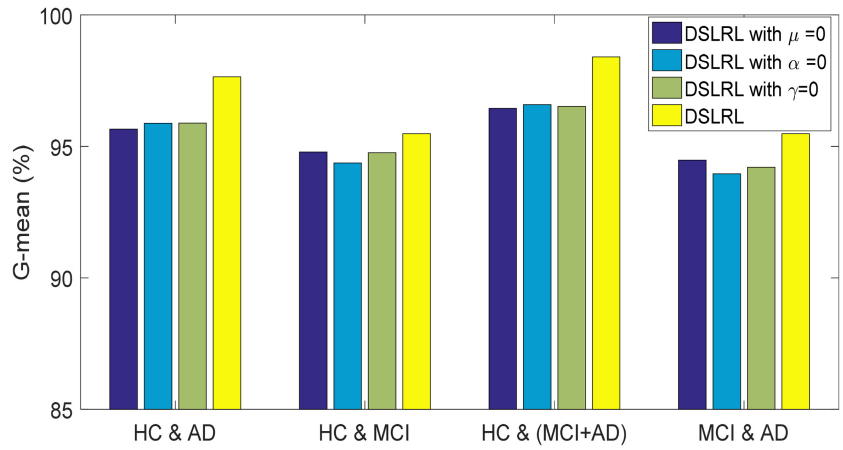


FIGURE 4 | G-mean of DSLRR with ablation experiment in four EEG datasets.

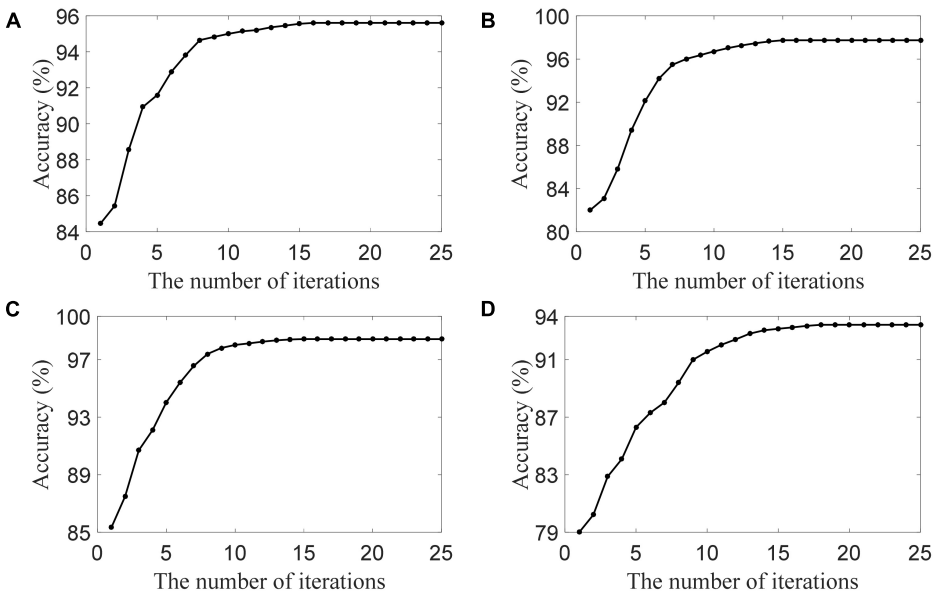


FIGURE 5 | The convergence of the DSLRR algorithm in four datasets, **(A)** HC and AD, **(B)** HC and MCI, **(C)** HC and (MCI+AD), and **(D)** MCI and AD.

columns of each low-rank coefficient matrix $\mathbf{L}$ of 1 has a positive effect on the classification.

(3) The LRR algorithm can describe the correlation of data, and the coefficient matrix is low rank. However, this algorithm doesn't consider the local structural characteristics of the data, and often cannot effectively exploit the discriminant information in the data. In this case, the LRR algorithm is not directly applicable to the EEG classification for AD recognition. The JSLC algorithm achieves good results in four datasets. JSLC is a low-rank representation model based on dictionary learning, which integrates discriminant information of samples into dictionary learning, and can also eliminate the influence of noise information on the classification model. This result shows that joint learning of low-rank representation and subspace learning is an effective means to solve EEG classification. The NRLRL algorithm conducts low-rank learning in the original data space. Its classification performance is lower than DSLRR in four datasets, which further shows that more data dimensions may not improve model performance. Due to the redundant information and noise in EEG data, it is effective to obtain the compact and discriminant feature representation through subspace learning and low-rank representation.

Ablation Experiment

The DSLRR algorithm integrates discriminant margin term, global structure term, and LSR term on the basis of the LRR algorithm. To verify the role of these terms, we performed ablation experiments on four EEG datasets. For discriminant

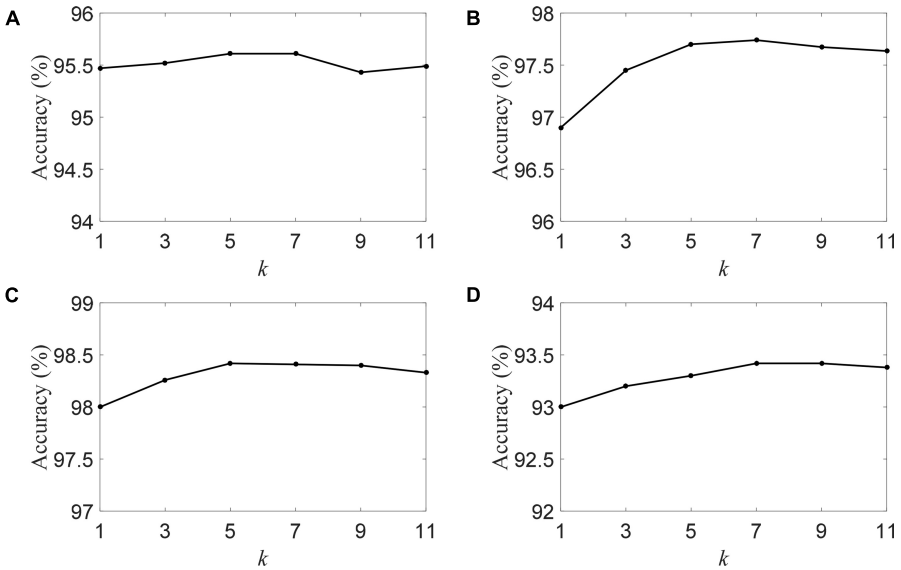


FIGURE 6 | The accuracy of the DSLRR algorithm with different k in four datasets, **(A)** HC and AD, **(B)** HC and MCI, **(C)** HC and (MCI+AD), and **(D)** MCI and AD.

margin term, its purpose is to use supervised information to establish graph embedding, to improve the distinguishing ability of the model. To verify its effect, we remove this item from Eq. (11), that is, set the parameter $\mu = 0$. For global structure terms, their purpose is to preserve the structure information of data in subspace. To verify its effect, we remove this item from Eq. (11) by setting the parameter $\alpha = 0$. For the LSR term, its purpose is to use the least square constraint to utilize the discriminant information in the data. Similarly, to verify its effect, we remove this item from Eq. (11), that is, set the parameter $\gamma = 0$. The classification accuracy, F-measure, and G-means of DSLRR with an ablation experiment in four EEG datasets are shown in **Figures 2**–**4**, respectively. From the results in **Figure 2**, we can see that if any one of three terms is removed from Eq. (11), the classification accuracy in the four EEG datasets has decreased to varying degrees. This is because each term has a corresponding contribution to the EEG classification task, which also illustrates the necessity of the coexistence of these three terms from another perspective. The results in **Figures 3**, **4** show that this conclusion is well verified. Therefore, the lack of any term will degrade the performance of the DSLRR algorithm.

Parameter Analysis

To show the convergence of the DSLRR algorithm, we plot its convergence curve in **Figure 5**. As shown in **Figure 5**, the DSLRR algorithm converges quickly in several iterations across four EEG datasets. The results show that the DSLRR algorithm is acceptable in the running time, which shows that the DSLRR algorithm has high practical worthiness.

We plot the classification accuracy of the DSLRR algorithm with different k-nearest neighbors in **Figure 6**. **Figure 6** visually shows that the classification is mildly sensitive to k. The DSLRR algorithm can achieve good classification accuracy when the parameter k is in the range of [5, 7, 9]. When k is <5 or k is greater than 9, the classification accuracy is slightly lower. Therefore, we can fix $k = 7$ in the experiment.

CONCLUSION

With the emergence of global aging, the prediction and diagnosis of AD have attracted extensive attention. In recent years, EEG technology has been developed and has become an important means to detect abnormal brain activity in patients with AD. To realize the early diagnosis of AD, we propose the DSLRR learning algorithm. The DSLRR algorithm inherits the advantages of low-rank representation, removes redundant information and noise, and improves the discriminant ability of low-rank representation through graph discriminant embedding. Meanwhile, based on subspace learning, the DSLRR algorithm introduces LSR and global structure preserving constraints to further improve the discriminative ability of the model. Extensive experimental results on real EEG data verify the effectiveness of the DSLRR algorithm.

In the future, we will continue to explore our work in the following aspects. First, the DSLRR algorithm is essentially a linear learning method. The brain is a nonlinear system with the ability of self-adaptation and self-regulation. Under some internal or external stimuli, the regulation and application functions of biological tissue will inevitably affect the electrophysiological signals, so that neurons have chaotic discharge phenomena, which present nonlinear characteristics. This makes the DSLRR algorithm unable to exert its performance in complex EEG data. To this end, we consider introducing a nonlinear learning model to improve the stability and accuracy of the DSLRR algorithm, so that it can be better suitable for various complex application scenarios. Second, the DSLRR algorithm is suitable for EEG classification using single-feature information. At present, the

technologies of feature processing and feature exaction are more mature, and the obtained feature information is correspondingly more diverse. In the next stage, we will extend the proposed algorithm to multi-feature scenarios to form a richer AD recognition system. Third, with the popularization of EEG acquisition equipment, using the existing labeled samples to analyze the unlabeled samples in multiple domains is a difficult problem in EEG-based AD recognition. We will use transfer learning technology to extend our algorithm in the future, to further enhance the generalization of the algorithm.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The EEG dataset analyzed in this study can be found in: https://github.com/tsyoshihara/Alzheimer-s-Classification-EEG.

AUTHOR CONTRIBUTIONS

TT, XG, and JX conceived and developed the model and wrote the manuscript. HL and GZ ran the experiment and analyzed the results. All authors read, edited, and approved the manuscript.

FUNDING

This work was supported in part by the Science and Technology Project of Changzhou city under grant CE20215032.

REFERENCES

Anuradha, G., and Jamal, D. N. (2021). Classification of dementia in EEG with a two-layered feed forward artificial neural network. *Eng. Technol. Appl. Sci. Res.* 11, 7135–7139. doi: 10.48084/etasr.4112

Araujo, T., Teixeira, J. P., and Rodrigues, P. M. (2022). Smart-data-driven system for Alzheimer disease detection through electroencephalographic signals. *Bioengineering (Basel)* 9:141. doi: 10.3390/bioengineering9040141

Cai, J., Candèsm, E. J., and Shen, Z. (2010). A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* 20, 1956–1982. doi: 10.1137/080738970

Chen, J., and Yang, J. (2014). Robust subspace segmentation via low-rank representation. *IEEE Trans. Cybern.* 44, 1432–1445. doi: 10.1109/TCYB.2013.2286106

Cummings, J. (2021). The role of neuropsychiatric symptoms in research diagnostic criteria for neurodegenerative diseases. *Am. J. Geriatr. Psychiatry* 29, 375–383. doi: 10.1016/j.jagp.2020.07.011

Fiscon, G., Weitschek, E., Cialini, A., Felici, G., Bertolazzi, P., Salvo, S. D., et al. (2018). Combining EEG signal processing with supervised methods for alzheimer's patients classification. *BMC Med. Inform. Decis. Mak.* 18:35. doi: 10.1186/s12911-018-0613-y

Fröhlich, S., Kutz, D. F., Müller, K., and Voelcker-Rehage, C. (2021). Characteristics of resting state EEG power in 80+-year-olds of different cognitive status. *Front. Aging Neurosci.* 13:675689. doi: 10.3389/fnagi.2021.675689

Gao, M., Liu, R., and Mao, J. (2021). Noise robustness low-rank learning algorithm for electroencephalogram signal classification. *Front. Neurosci.* 15:797378. doi: 10.3389/fnins.2021.797378

Ge, Q., Lin, Z. C., Gao, Y. X., and Zhang, J. X. (2020). A robust discriminant framework based on functional biomarkers of EEG and its potential for diagnosis of Alzheimer's disease. *Healthcare* 8:476. doi: 10.3390/healthcare8040476

Ghorbanian, P., Ramakrishnan, S., and Ashrafiuon, H. (2015). Stochastic non-linear oscillator models of EEG: the Alzheimer's disease case. *Front. Comput. Neurosci.* 9:48. doi: 10.3389/fncom.2015.00048

Huang, P., Li, T., Cao, G., and Geng, Y. (2018).) Feature extraction based on graph discriminant embedding and its applications to face recognition. *Soft Comput.* 23, 7015–7028. doi: 10.1007/s00500-018-3340-5

Ieracitano, C., Mammone, N., Bramanti, A., Hussaind, A., and Morabitoa, F. C. (2019a). A convolutional neural network approach for classification of dementia stages based on 2Dspectral representation of EEG recordings. *Neurocomputing* 323, 96–107. doi: 10.1016/j.neucom.2018.09.071

Ieracitano, C., Mammone, N., Hussain, A., and Francesco, M. (2019b). A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia. *Neural Netw.* 123, 176–190. doi: 10.1016/j.neunet.2019.12.006

Jiang, X. (2011). Linear subspace learning-based dimensionality reduction. *IEEE Signal Process. Mag.* 28, 16–26. doi: 10.1109/MSP.2010.939041

Jiang, Y., Gu, X., Wu, D., Hang, W., Xue, J., Qiu, S., et al. (2021). A novel negative-transfer-resistant fuzzy clustering model with a shared cross domain transfer latent space and its application to brain CT image segmentation. *IEEE ACM Trans. Comput. Biol. Bioinform.* 18, 40–52. doi: 10.1109/TCBB.2019.2963873

Kang, M., Kang, M., and Jung, M. (2015). Inexact accelerated augmented Lagrangian methods. *Comput. Optim. Appl.* 62, 373–404. doi: 10.1007/s10589-015-9742-8

Lei, W., Ma, Z., Liu, S., and Lin, Y. (2021). EEG mental recognition based on RKHS learning and source dictionary regularized RKHS subspace learning. *IEEE Access* 9, 150545–150559. doi: 10.1109/ACCESS.2021.3124028

Li, P., Yu, J., Wang, M., Zhang, L., Cai, D., and Li, X. (2017). Constrained low-rank learning using least squares-based regularization. *IEEE Trans. Cybern.* 47, 4250–4262. doi: 10.1109/TCYB.2016.2623638

Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., and Ma, Y. (2013). Robust recovery of subspace structures by low-rank representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 171–184. doi: 10.1109/TPAMI.2012.88

Lu, J., Zhou, G., Zhu, J., and Xue, L. (2021). Joint subspace and low-rank coding method for makeup face recognition. *Math. Prob. Eng.* 2021, 1–8. doi: 10.1155/2021/9914452

Maturana-Candelas, A., Gómez, C., Poza, J., Ruiz-Gómez, S. J., and Hornero, R. (2020). Inter-band bispectral analysis of EEG background activity to characterize Alzheimer's disease continuum. *Front. Comput. Neurosci.* 14:70. doi: 10.3389/fncom.2020.00070

Miltiadous, A., Tzimourta, K. D., Giannakeas, N., Tsipouras, M. G., Afrantou, T., Ioannidis, P., et al. (2021). A1zheimer's disease and frontotempora1 dementia a robust c1assification method of EEG signa1s and a comparison of va1idation methods. *Diagnostics* 11:1437. doi: 10.3390/diagnostics11081437

Mirzaei, G., and Adeli, H. (2022). Machine learning techniques for diagnosis of alzheimer disease, mild cognitive disorder, and other types of dementia. *Biomed. Signal Process. Control* 72:103293. doi: 10.1016/j.bspc.2021.103293

Wen, D., Li, P., Li, X., Wei, Z., Zhou, Y., Pei, H., et al. (2020). The feature extraction of resting-state EEG signal from amnestic mild cognitive impairment with type 2 diabetes mellitus based on feature-fusion multispectral image method. *Neural Netw.* 4, 373–382. doi: 10.1016/j.neunet.2020.01.025

Weng, B., and Shen, J. (2008). Classification of multivariate time series using locality preserving projections. *Knowl. Based Syst.* 21, 581–587. doi: 10.1016/j.knosys.2008.03.027

Zhang, Y., Chung, F., and Wang, S. (2020). Clustering by transmission learning from data density to label manifold with statistical

diffusion. *Knowl. Based Syst.* 193:105330. doi: 10.1016/j.knosys.2019.105330

Zhang, Y., Wang, S., Xia, K., Jiang, Y., and Qian, P. (2021). Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion. *Inf. Fusion* 66, 170–183. doi: org/10.1016/j.inffus.2020.09.002

Zhao, S., Wu, J., Zhang, B., and Fei, L. (2022). Low-rank inter-class sparsity based semi-flexible target least squares regression for feature representation. *Pattern Recognit.* 123:108346. doi: 10.1016/j.patcog.2021.108346

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 24 June 2022
doi: 10.3389/fnagi.2022.924113

Evaluation of Feature Selection for Alzheimer's Disease Diagnosis

Feng Gu[1,2†], Songhua Ma[3,4†], Xiude Wang[1,2], Jian Zhao[5], Ying Yu[5] and Xinjian Song[6,7]*

[1] Department of Medical Image, Affiliated Nantong Rehabilitation Hospital of Nantong University, Nantong, China, [2] Department of Medical Image, The Second People's Hospital of Nantong, Nantong, China, [3] Department of Neurology, Affiliated Nantong Rehabilitation Hospital of Nantong University, Nantong, China, [4] Department of Neurology, The Second People's Hospital of Nantong, Nantong, China, [5] Nantong Center for Disease Control and Prevention, Nantong, China, [6] Department of Rehabilitation Medicine, Affiliated Nantong Rehabilitation Hospital of Nantong University, Nantong, China, [7] Department of Rehabilitation Medicine, The Second People's Hospital of Nantong, Nantong, China

Accurate recognition of patients with Alzheimer's disease (AD) or mild cognitive impairment (MCI) is important for the subsequent treatment and rehabilitation. Recently, with the fast development of artificial intelligence (AI), AI-assisted diagnosis has been widely used. Feature selection as a key component is very important in AI-assisted diagnosis. So far, many feature selection methods have been developed. However, few studies consider the stability of a feature selection method. Therefore, in this study, we introduce a frequency-based criterion to evaluate the stability of feature selection and design a pipeline to select feature selection methods considering both stability and discriminability. There are two main contributions of this study: (1) It designs a bootstrap sampling-based workflow to simulate real-world scenario of feature selection. (2) It develops a decision graph to determine the optimal combination of supervised and unsupervised feature selection both considering feature stability and discriminability. Experimental results on the ADNI dataset have demonstrated the feasibility of our method.

Keywords: artificial intelligence, Alzheimer's disease, feature selection, stability, discriminability

OPEN ACCESS

Edited by:
Sang-Bing Tsai,
Wuyi University, China

Reviewed by:
Hongru Zhao,
Soochow University, China
Kaifa Zhao,
Hong Kong Polytechnic University, Hong Kong SAR, China

***Correspondence:**
Xinjian Song
xj_song@yahoo.com

† These authors have contributed equally to this work

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

Received: *20 April 2022*
Accepted: *06 June 2022*
Published: *24 June 2022*

Citation:
Gu F, Ma S, Wang X, Zhao J, Yu Y and Song X (2022) Evaluation of Feature Selection for Alzheimer's Disease Diagnosis. Front. Aging Neurosci. 14:924113. doi: 10.3389/fnagi.2022.924113

INTRODUCTION

Alzheimer's disease (AD) (Xiao-Cong et al., 2018; Hou et al., 2020; Mishra and Li, 2020; Subasi, 2020; He et al., 2022) is a degenerative disease of the central nervous system, which is clinically manifested as progressive memory impairment, cognitive dysfunction, language dysfunction, and personality change, etc. AD has a serious impact on the lives of patients, but also brings a heavy economic burden to patients' families. At present, the research progress of AD is slow, and the disease factors cannot be accurately determined. It is usually found at an advanced stage, and even treatment will not produce a better therapeutic effect. Therefore, the early diagnosis of AD is very critical, which can effectively inhibit the development of the disease, and even avoid the occurrence of clinical symptoms by taking timely treatment. Mild Cognitive Impairment (MCI) is considered as an intermediate state between health and AD. In patients with MCI, the probability of progressing to AD is about 10–15% (He et al., 2022). Therefore, if patients with MCI can be effectively identified and actively intervened, it is of great significance for the control of AD.

With the rapid development of artificial intelligence (Jiang et al., 2020; Xia et al., 2020; Zhang et al., 2020a,b, 2021a,b, 2022), intelligent models are widely used in MCI or AD recognition.

Kloppel et al. (2008) input gray matter features of brain images of AD patients into linear support vector machines (SVM), so as to apply the trained SVM to clinical studies. Ashburner and Friston (2000) applied morphometric methods to the diagnosis of AD, which spatially normalized high-resolution images of all subjects into the same stereotactic space. Then, gray matter was separated from the spatially normalized images and data smoothing was performed on them. Voxel parameter test statistics were performed on the two groups of smoothed gray images to improve the uneven intensity of the brain artifact images. Hinrichs et al. (2009) also proposed an AD recognition framework based on the smoothness of three-dimensional image coordinate space. It directly integrates the spatial relations of voxels into the learning framework and does not require image preprocessing information of other modes, thus automatically classifying subjects according to structural or functional imaging features. In addition, MCI was associated with changes in cortical morphology, such as cortical thickness, sulcus depth, surface area, gray matter volume, and mean curvature in different brain regions. These features have been shown to have a specific neuropathological and genetic basis. However, most methods have focused on univariate prediction models, and cortical features are usually isolated. Therefore, Li et al. (2014) used a multivariate approach to study the abnormalities of multiple cortical features in patients with mild cognitive impairment, and identified subtle patterns of changes in cortical anatomical structure through a classification model. Liu et al. (2013) used non-linear global data structure to map multivariable MRI data such as regional brain volume and cortical thickness into a low-dimensional local linear space through local linear embedding method, and trained a disease classifier by embedding brain features to predict whether MCI would be transformed into AD in the future. Möller et al. (2016) took the voxel values extracted from the voxel data as the original feature data, and proposed a feature selection method to apply to the original feature vector, so as to reduce the dimension of the original feature vector to a low-dimensional space and carry out the next classification task. From the above-mentioned studies, we can summarize the general process of MCI/AD recognition based on intelligent model, as shown in **Figure 1**. From **Figure 1**, it can be found that the general process of MCI/AD recognition contains four components, preprocessing, feature extraction, feature selection, and prediction. Preprocessing aims to process the original images including registration, standardizing and smoothing. Feature extraction aims to extract original features from the images after preprocessing. Feature selection aims to select discriminant features from the original feature set. Prediction aims to build a classification model to recognize MCI or AD patients. In the phase of prediction, based on the selected features, a prediction model is established for MCI/AD recognition.

From **Figure 1**, it can be found that feature selection is a key phase in the process of MCI/AD recognition. The goal of feature selection is to select discriminant features with low relevance between each other and high relevance to the outcome. In recent 2 years, some excellent feature selection work has emerged in the field of medical images. For example, Demir and Akbulut (2022) proposed a new residual- convolutional neural network to extract deep features from MRI images. Mainenti et al. (2022) proposed a radiomics-based pipeline to enhance MRI-based risk stratification in patients with endometrial cancer. Although previous studies have achieved great success in feature selection, feature discriminability is often the first important factor and feature stability is always omitted. In this study, first of all, feature stability, variance, and pairwise correlation were analyzed. Then, the least absolute shrinkage and selection operator (LASSO) and recursive feature elimination (RFE) were employed to search for the optimal feature set (Mainenti et al., 2022).

In this study, we focus on feature selection because few studies consider both the stability and performance of feature selection so far, which are two key factors for the classification phase. The main contributions cover two aspects. The first one is that we introduce a frequency-based criterion to evaluate the stability of a feature selection method. The second is that we propose a bootstrap-based flow chart and a decision graph to select the best combination of supervised and unsupervised feature selection methods. The following sections are organized as follows. Section "Data and Methods" presents the data we used and the methods we proposed. Section "Results" reports the experimental results, section "Discussion" discusses the experimental results and the last section concludes the whole study.

DATA AND METHODS

Data

In this study, we select 103 patients with MRI and PET from the Alzheimer's Disease Neuroimaging Initiative (ADNI) as our datasets. ADNI is a 5-year public partnership sponsored by several institutes, companies, and non-profit organizations (Zhang et al., 2021b). Owning to the original images cannot be directly used for our study, we set up a data preprocessing pipeline, which contains three main steps. Firstly, each subject in ADNI contains 96 PET images. Statistical parametric mapping (SPM) (Muzik et al., 2000) is used to fuse these PET images to construct a 3-D one which has brain spatial information and the feature information between tissue structures are also retained. In addition, motion correction is performed due to head motion. Secondly, the MRI image and PET image of each subject are registered, and affinely aligned. In the third step, the average template data generated is used to spatially normalize all PET images to the standard MNI space. PET images are also smoothed (8 mm Gaussian) to avoid the influences caused by noises. The AAL (automated anatomical atlas) (Rolls et al., 2020) which is available as a toolbox[1] for SPM is used as a template to extract original features from PET images. Based on AAL, the brain is segmented into 116 regions, and we select 90 regions from the cerebrum for feature extraction. To be specific, firstly, the PET images are resampled to the same size as the AAL template so that each region is in correspondence spatially. The size of AAL template is 61 $\times$73 $\times$61. Then we extract average intensity

[1]http://www.gin.cnrs.fr/AAL

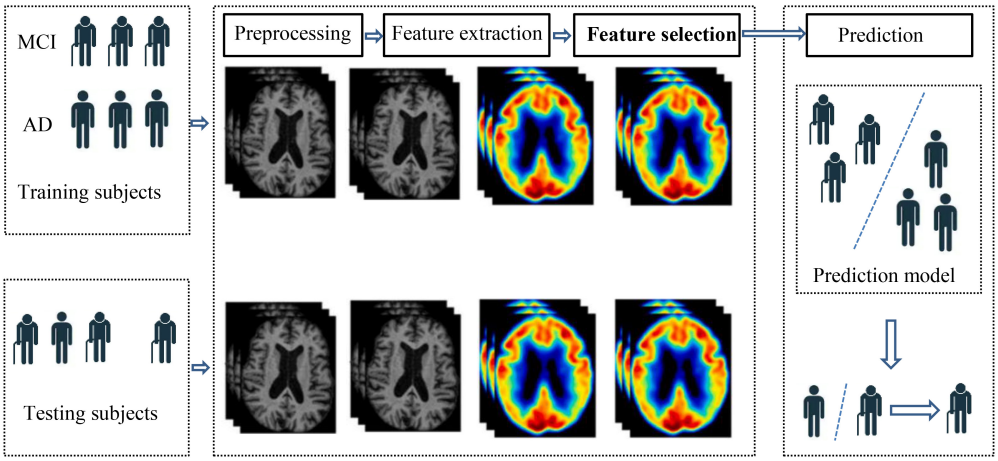


FIGURE 1 | General process of MCI/AD recognition.

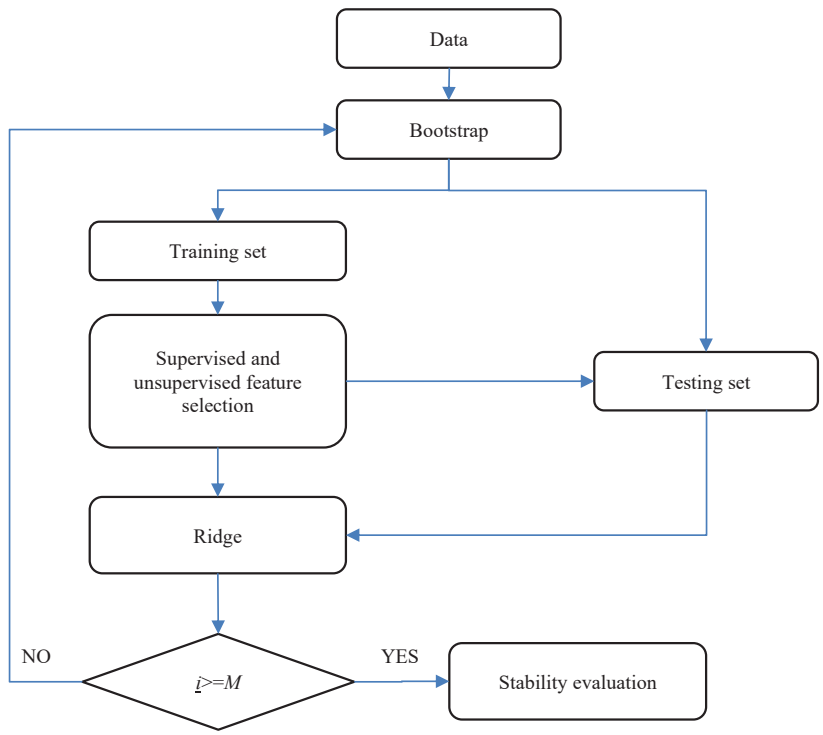


FIGURE 2 | Flow chart for stability evaluation.

values from all regions of PET images as original features for our proposed classification model.

Methods

Stability Evaluation Metrics

In this study, we use a frequency-based criterion to measure the stability of a feature select method (Nogueira et al., 2017). For clarity, suppose we have a feature selection method Φ and a d-dimensional dataset X. The feature selection method is performed on the d-dimensional dataset X to select discriminant features. The feature selection process is repeated M times by a bootstrap strategy. Then we can define a binary matrix Z, as shown in (1) to indicate the feature selection results of M tries,

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1d} \\ z_{21} & z_{22} & \cdots & z_{2d} \\ \cdots & \cdots & \cdots & \cdots \\ z_{M1} & z_{M2} & \cdots & z_{Md} \end{bmatrix} \quad (1)$$

In Z, each row represents one try of feature selection. In each row, $z_{ij} = 1(i = 1, 2, ..., M, j = 1, 2, ..., d)$ represents that the j-th feature is selected in the i-th try; otherwise, the j-th feature is not selected. Based on the binary matrix Z, the stability of feature selection method Φ in terms of the frequency-based criterion can be defined as:

$$Stability(Z) = 1 - \frac{\frac{1}{d}\sum_{f=1}^{d}\left[\frac{M}{M-1}(\frac{1}{M}\sum_{i=1}^{M} z_{if})(1 - \frac{1}{M}\sum_{i=1}^{M} z_{if})\right]}{\frac{\frac{1}{M}\sum_{i=1}^{M}\sum_{f=1}^{d} z_{if}}{d}(1 - \frac{\frac{1}{M}\sum_{i=1}^{M}\sum_{f=1}^{d} z_{if}}{d})} \quad (2)$$

From (2), we can see that $Stability(Z)$ ranges from 0 to 1, the greater the value, the better the stability.

Stability Evaluation Workflow

In this study, we use a supervised feature selection method to reduce features irrelative to the outcome, and an unsupervised feature selection method to reduce redundant features. To evaluate the stability of feature selection, a bootstrap sampling-based flow chart is established, which is shown in **Figure 2**. Firstly, the AD dataset is split into the training set (70%) and the testing set (30%) by bootstrap sampling. Then supervised and unsupervised feature selection is performed on the training set to select discriminant features. The testing set is updated with the selected features. Finally, a Ridge regression model is trained based on the selected features. The bootstrap sampling is repeated M times so that the matrix **Z** in (1) can be obtained. Based on **Z**, we can use (2) to evaluate the stability of the supervised and unsupervised feature selection methods we used.

Decision Graph for Feature Selection

In Li et al. (2017), a feature selection package was shared which contains 33 different kinds of supervised and unsupervised feature selection methods. In this study, we aim to choose a best supervised and unsupervised combination from this package for AD diagnosis. First of all, we set up an initial exclusion criterion to select a part of supervised and unsupervised feature select methods from the package provided by Li et al. (2017). The exclusion criterion states: (1) if prediction performance in terms of AUC of a feature selection method is lower than 0.5, the method is excluded. (2) If the running time of one try of a feature selection method is more than 30 min, the method is excluded. These exclusion criteria are defined for two reasons. The first is that if the prediction performance of the feature selection method is lower than 0.5, it indicates that the prediction performance of the method is close to the randomness level. Second, if the running time of a feature selection method exceeds 30 min, it will exceed the normal tolerance range when the training set size is not large. With the exclusion criterion, we finally select F score (denoted as S1:), T Score (denoted as S2), ReliefF (denoted as S3), and Fish Score (denoted as S4) as supervised feature selection methods, and Lap_score (denoted as U1), spectral feature selection (SPEC, denoted as U2), Monte Carlo feature selection (MCFS, denoted as U3), non-negative discriminative feature selection (NDFS, denoted as U4), unsupervised discriminative feature selection (UDFS, denoted as U5), and Person_score (denoted as U6) as unsupervised feature selection methods. Therefore, we have 24 combinations, i.e., S1U1, S1U2,…, S4U6, as shown in **Table 1**. Secondly, as we stated before that both performance and stability are important for Alzheimer's disease diagnosis.

TABLE 1 | All combinations of supervised and unsupervised feature selection methods.

Combination name	Name of supervised method	Name of unsupervised method
S1U1	F score	Lap_score
S1U2		SPEC
S1U3		MCFS
S1U4		NDFS
S1U5		UDFS
S1U6		Person score
S2U1	T score	Lap_score
S2U2		SPEC
S2U3		MCFS
S2U4		NDFS
S2U5		UDFS
S2U6		Person score
S3U1	ReliefF	Lap_score
S3U2		SPEC
S3U3		MCFS
S3U4		NDFS
S3U5		UDFS
S3U6		Person score
S4U1	Fish score	Lap_score
S4U2		SPEC
S4U3		MCFS
S4U4		NDFS
S4U5		UDFS
S4U6		Person score

Based on **Figure 2**, we can generate the matrix **Z.** Thus, we can use (2) to evaluate the stability of the supervised and unsupervised feature selection methods we used. Therefore, we design a decision graph, as shown in **Figure 3**, to determine the best combination of the supervised and unsupervised feature selection methods.

RESULTS

The decision graph of all combinations for MRI features is shown in **Figure 4**. It is observed that the combination S2U6 wins the best in terms of *AUC*Stability*, which means that the combination of T Score (supervised feature selection method) and Person Score (unsupervised feature selection method) performs better than other combinations in terms of both AUC and stability. Therefore, the supervised feature selection method *T Score* and the unsupervised feature selection method *Person Score* will be selected as the feature selection methods for modeling.

The decision graph of all combinations for PET features is shown in **Figure 5**. Similar to **Figure 4**, it is observed that the combination S1U1 and S4U6 wins the best. Therefore, the combination *F score + Lap score* or the combination *Fish Score + Person Score* will be selected for the following phase of modeling.

From **Figures 4**, **5**, it can be found that this is no combinations that always perform best. Our method is case-dependent, which means that it provides decision support for users.

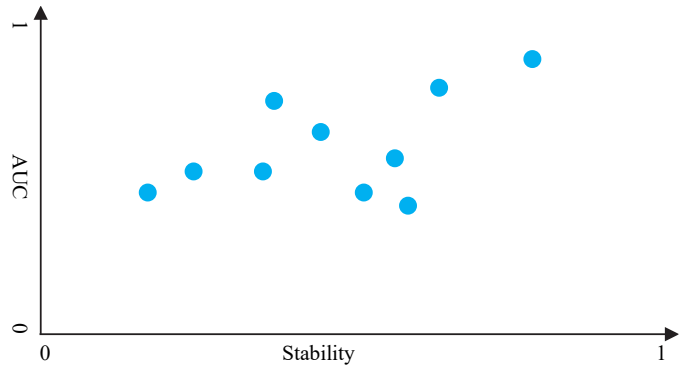


FIGURE 3 | Decision graph for feature selection.

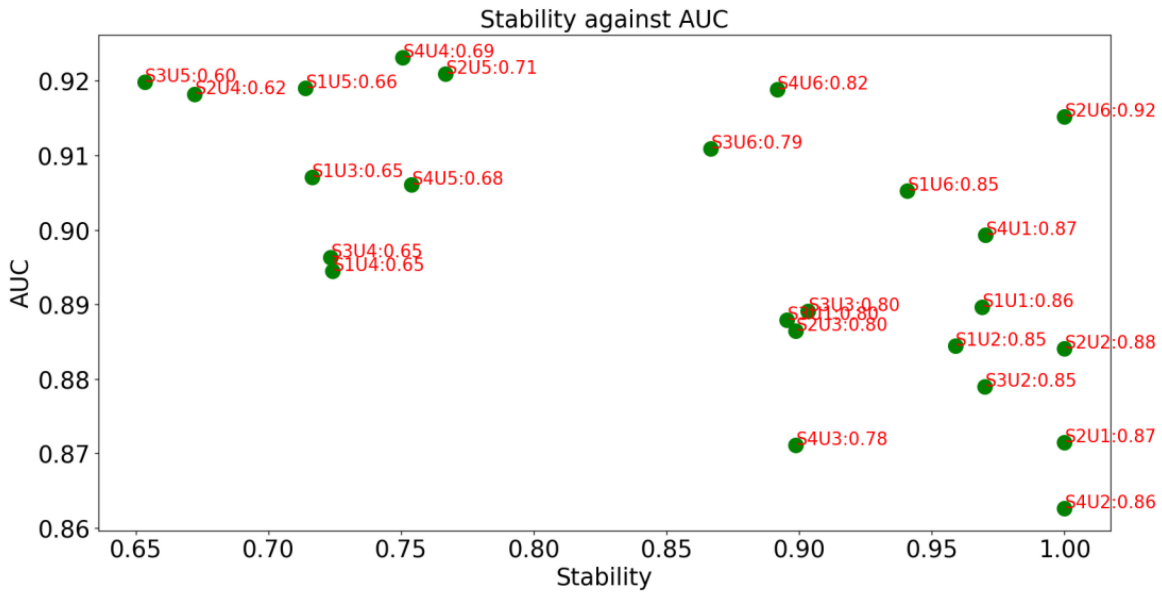


FIGURE 4 | Decision graph for MRI features.

DISCUSSION

In this study, we have 103 subjects, for both MRI and PET, the feature dimension of each subject is 93, which is near to the number of subjects. When classification models are applied to the high-dimensional data, a critical issue is known as the curse of dimensionality, which refers to the phenomenon that data becomes sparse in high-dimensional space may occur (Li et al., 2017). Therefore, feature selection plays a very significant role in the recognition of AD or MCI. So far, many feature selection methods have been successfully applied in the field of medical image-based diagnosis. For example, in Salvatore et al. (2015), employed PCA (principle component analysis) to select discriminant features from the density maps of WM (white matter) and GM (gray matter) as input of SVM for AD recognition. In Liu et al. (2013), employed LLE (local linear embedding) as the unsupervised feature reduction method to reduce features from the space of multivariate regional brain volume and cortical thickness MRI to a locally low-dimensional linear space while maintaining the global non-linear data structure. Then, the reduced brain features in the low-dimensional space were used to train the prediction model. Unlike Liu et al. (2013) and Salvatore et al. (2015) in Beheshti et al. (2015) proposed a filter-based supervised feature reduction method containing three main steps. First of all, feature extraction was carried out by using the voxel clusters that are detected by the voxel-based morphometric (VBM) on sMRI and the voxel values as the volume of interest (VOI). Secondly, the probability distribution function of the VOI was employed to represent the statistical information of the respective high-dimensional structural MRI samples. Thirdly, the final selected features were employed to train a SVM classifier to perform the AD recognition task. In Nir et al. (2015) extracted DTI-based features and proposed a tractography-based model to recognize AD and MCI. First of all, the authors used tractography and clustering techniques to locate and organize fibers into 18

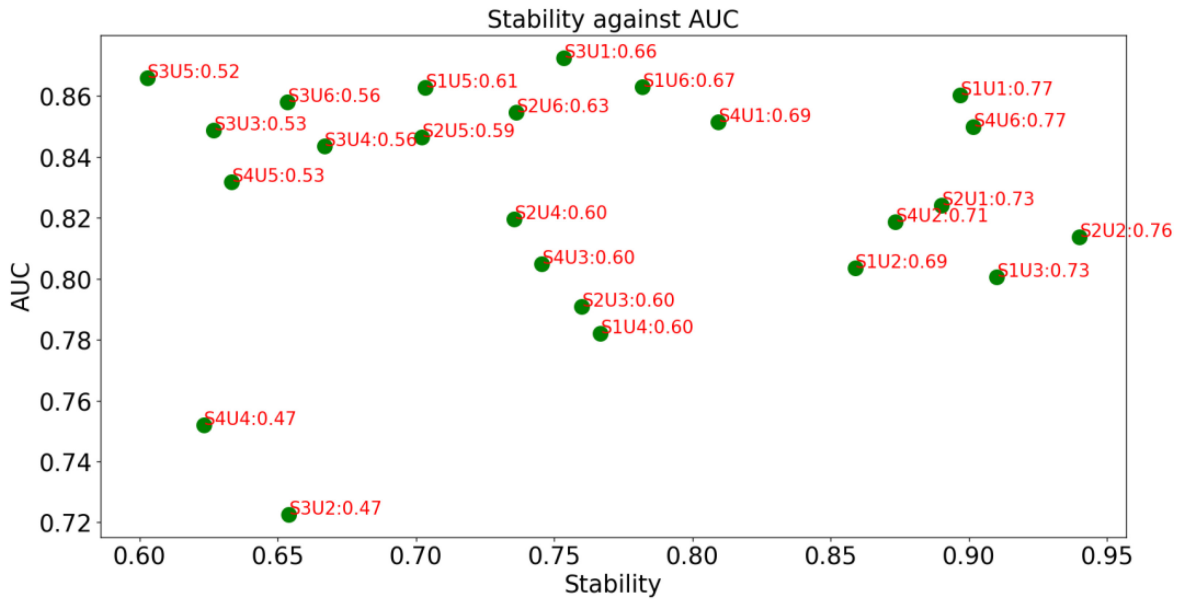


FIGURE 5 | Decision graph for PET features.

fiber bundles. Secondly, the authors computed density maps to quantify the number of fibers passing through each voxel and used the shortest path graph search to reduce the fiber bundles based on maximum density path (MDP) so that the fiber bundles can be expressed in a compact and low-dimensional space. Thirdly, the diffusivity measures of fractional anisotropy (FA) and MD computed along all the registered across subjects (MDPs) were selected as the features to train an SVM classifier. Feature selection methods in this category can be characterized as making use of the global or local statistical information. In De Martino et al. (2008) employed multivariate feature selection to select features to model functional MRI spatial patterns. To be specific, the authors employed RFE combined with an SVM classifier (REF-SVM) to reduce the irrelevant voxels recursively. Similarly, in Wee et al. (2011), based on DTI images, Wee et al. proposed a framework for MCI recognition. In this framework, the original features come from the anatomical regions, and REF-SVM was also used to reduce the original feature set.

Although different kinds of feature selection (reduction) methods have been widely used for AD and MCI recognition, an important thing that is not fully considered is the stability of the feature selection methods. In practice, we expect that the selected feature selection method can maintain robustness when training data changes slightly. Therefore, in this study, we introduce a frequency-based criterion to evaluate the stability and design a pipeline to select feature selection methods considering both stability and discriminability. Experimental results shown in **Figures 4**, **5** indicate that the proposed pipeline works well and can help us to determine the best combination of feature selection methods. That is to say, the proposed criterion *AUC*Stability* can find the optimal combination of supervised and unsupervised feature selection methods.

CONCLUSION

In this study, we introduce a frequency-based criterion to evaluate the stability of feature selection and design a pipeline to select feature selection methods considering both stability and discriminability.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The data is available on http://adni.loni.usc.edu/about/.

AUTHOR CONTRIBUTIONS

FG and SM contributed to the writing and experiments. XW, JZ, and YY contributed to the data collection and preprocessing. XS supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Project of Nantong Health Commission (MB2020045) and the Science and Technology Project of Nantong City (MS22021027).

ACKNOWLEDGMENTS

We thank the reviewers whose comments and suggestions helped improve this manuscript.

REFERENCES

Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage* 11, 805–821.

Beheshti, I., Demirel, H., and Alzheimer's Disease Neuroimaging Initiative. (2015). Probability distribution functionbased classification of structural mri for the detection of alzheimer's disease. *Comput. Biol. Med.* 64, 208–216. doi: 10.1016/j.compbiomed.2015.07.006

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., and Formisano, E. (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fmri spatial patterns. *Neuroimage* 43, 44–58. doi: 10.1016/j.neuroimage.2008.06.037

Demir, F., and Akbulut, Y. (2022). A new deep technique using R-CNN model and L1NSR feature selection for brain MRI classification. *Biomed. Signal Process. Control* 75:103625.

He, B., Bukhari, S., Fox, E., Abid, A., Shen, J., Kawas, C., et al. (2022). AI-enabled in silico immunohistochemical characterization for Alzheimer's disease. *Cell Rep. Methods* 2:100191. doi: 10.1016/j.crmeth.2022.100191

Hinrichs, C., Singh, V., Mukherjee, L., Xu, G., Chung, M. K., and Johnson, S. C. (2009). Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. *Neuroimage* 48, 138–149. doi: 10.1016/j.neuroimage.2009.05.056

Hou, K., Zhao, J., Wang, H., Li, B., Li, K., Shi, X., et al. (2020). Chiral gold nanoparticles enantioselectively rescue memory deficits in a mouse model of Alzheimer's disease. *Nat. Commun.* 11:4790. doi: 10.1038/s41467-020-18525-2

Jiang, Y., Zhang, Y., Lin, C., Wu, D., and Lin, C. T. (2020). EEG-based driver drowsiness estimation using an online multi-view and transfer TSK fuzzy system. *IEEE Trans. Intell. Transport. Syst.* 22, 1752–1764.

Kloppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., et al. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain* 131, 681–689. doi: 10.1093/brain/awm319

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., et al. (2017). Feature selection: a data perspective. *ACM Comput. Surv.* 50, 1–45.

Li, S., Yuan, X., Pu, F., Li, D., Fan, Y., Wu, L., et al. (2014). Abnormal changes of multidimensional surface features using multivariate pattern classification in amnestic mild cognitive impairment patients. *J. Neurosci.* 34, 10541–10553. doi: 10.1523/JNEUROSCI.4356-13.2014

Liu, X., Tosun, D., Weiner, M. W., Schuff, N., and Alzheimer's Disease Neuroimaging Initiative. (2013). Locally linear embedding (lle) for mri based Alzheimer's disease classification. *Neuroimage* 83, 148–157. doi: 10.1016/j.neuroimage.2013.06.033

Mainenti, P. P., Stanzione, A., Cuocolo, R., Del Grosso, R., Danzi, R., Romeo, V., et al. (2022). MRI radiomics: a machine learning approach for the risk stratification of endometrial cancer patients. *Eur. J. Radiol.* 149:110226. doi: 10.1016/j.ejrad.2022.110226

Mishra, R., and Li, B. (2020). The application of artificial intelligence in the genetic study of Alzheimer's disease. *Aging Dis.* 11:1567. doi: 10.14336/AD.2020.0312

Möller, C., Pijnenburg, Y., Wm, V. D. F., Versteeg, A., Tijms, B., De Munck, J., et al. (2016). Alzheimer disease and behavioral variant frontotemporal dementia: automatic classification based on cortical atrophy for single-subject diagnosis. *Radiology* 279, 838–848. doi: 10.1148/radiol.2015150220

Muzik, O., Chugani, D. C., Juhász, C., Shen, C., and Chugani, H. T. (2000). Statistical parametric mapping: assessment of application in children. *Neuroimage* 12, 538–549. doi: 10.1006/nimg.2000.0651

Nir, T. M., Villalon-Reina, J. E., Prasad, G., Jahanshad, N., Joshi, S. H., Toga, A. W., et al. (2015). Diffusion weighted imaging-based maximum density path analysis and classification of Alzheimer's disease. *Neurobiol. Aging* 36, S132–S140. doi: 10.1016/j.neurobiolaging.2014.05.037

Nogueira, S., Sechidis, K., and Brown, G. (2017). On the stability of feature selection algorithms. *J. Mach. Learn. Res.* 18, 6345–6398.

Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., and Joliot, M. (2020). Automated anatomical labelling atlas 3. *Neuroimage* 206:116189. doi: 10.1016/j.neuroimage.2019.116189

Salvatore, C., Cerasa, A., Battista, P., Gilardi, M. C., Quattrone, A., and Castiglioni, I. (2015). Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front. Neurosci.* 9:307. doi: 10.3389/fnins.2015.00307

Subasi, A. (2020). Use of artificial intelligence in Alzheimer's disease detection. *Artif. Intell. Precis. Health* 2020, 257–278.

Wee, C.-Y., Yap, P.-T., Li, W., Denny, K., Browndyke, J. N., Potter, G. G., et al. (2011). Enriched white matter connectivity networks for accurate identification of mci patients. *Neuroimage* 54, 1812–1822. doi: 10.1016/j.neuroimage.2010.10.026

Xia, K., Zhang, Y., Jiang, Y., Qian, P., Dong, J., Yin, H., et al. (2020). TSK fuzzy system for multi-view data discovery underlying label relaxation and cross-rule & cross-view sparsity regularizations. *IEEE Trans. Indust. Inform.* 17, 3282–3291.

Xiao-Cong, P. A. N. G., De, K. A. N. G., Jian-Song, F. A. N. G., Ying, Z. H. A. O., Lv-Jie, X. U., Wen-Wen, L. I. A. N., et al. (2018). Network pharmacology-based analysis of Chinese herbal Naodesheng formula for application to Alzheimer's disease. *Chin. J. Nat. Med.* 16, 53–62. doi: 10.1016/S1875-5364(18)30029-3

Zhang, Y., Chung, F. L., and Wang, S. (2020a). Clustering by transmission learning from data density to label manifold with statistical diffusion. *Knowl. Based Syst.* 193:105330.

Zhang, Y., Lam, S., Yu, T., Teng, X., Zhang, J., Lee, F. K. H., et al. (2022). Integration of an imbalance framework with novel high-generalizable classifiers for radiomics-based distant metastases prediction of advanced nasopharyngeal carcinoma. *Knowl. Based Syst.* 235:107649.

Zhang, Y., Wang, G., Chung, F. L., and Wang, S. (2021a). Support vector machines with the known feature-evolution priors. *Knowl. Based Syst.* 223:107048.

Zhang, Y., Wang, S., Xia, K., Jiang, Y., Qian, P., and Alzheimer's Disease Neuroimaging Initiative. (2021b). Alzheimer's disease multiclass diagnosis *via* multimodal neuroimaging embedding feature selection and fusion. *Inform. Fusion* 66, 170–183.

Zhang, Y., Zhou, Z., Bai, H., Liu, W., and Wang, L. (2020b). Seizure classification from EEG signals using an online selective transfer TSK fuzzy classifier with joint distribution adaption and manifold regularization. *Front. Neurosci.* 14:496. doi: 10.3389/fnins.2020.00496

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 05 July 2022
doi: 10.3389/fnagi.2022.948704

CNNG: A Convolutional Neural Networks With Gated Recurrent Units for Autism Spectrum Disorder Classification

Wenjing Jiang[1,2†], Shuaiqi Liu[1,2†], Hong Zhang[1,2], Xiuming Sun[3], Shui-Hua Wang[4], Jie Zhao[1,2] and Jingwen Yan[5]*

[1] College of Electronic and Information Engineering, Hebei University, Baoding, China, [2] Machine Vision Technological Innovation Center of Hebei, Baoding, China, [3] School of Mathematics and Information Science, Zhangjiakou University, Zhangjiakou, China, [4] School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China, [5] School of Engineering, Shantou University, Shantou, China

OPEN ACCESS

Edited by:
Yu-Dong Zhang,
University of Leicester,
United Kingdom

Reviewed by:
Geng Peng,
Shijiazhuang Tiedao University, China
Vipin Tyagi,
Indian Institute of Technology
Kharagpur, India

***Correspondence:**
Xiuming Sun
sunxiuming@zjku.edu.cn

[†] These authors have contributed equally to this work

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

Received: 20 May 2022
Accepted: 16 June 2022
Published: 05 July 2022

Citation:
Jiang W, Liu S, Zhang H, Sun X, Wang S-H, Zhao J and Yan J (2022) CNNG: A Convolutional Neural Networks With Gated Recurrent Units for Autism Spectrum Disorder Classification. Front. Aging Neurosci. 14:948704. doi: 10.3389/fnagi.2022.948704

As a neurodevelopmental disorder, autism spectrum disorder (ASD) severely affects the living conditions of patients and their families. Early diagnosis of ASD can enable the disease to be effectively intervened in the early stage of development. In this paper, we present an ASD classification network defined as CNNG by combining of convolutional neural network (CNN) and gate recurrent unit (GRU). First, CNNG extracts the 3D spatial features of functional magnetic resonance imaging (fMRI) data by using the convolutional layer of the 3D CNN. Second, CNNG extracts the temporal features by using the GRU and finally classifies them by using the Sigmoid function. The performance of CNNG was validated on the international public data—autism brain imaging data exchange (ABIDE) dataset. According to the experiments, CNNG can be highly effective in extracting the spatio-temporal features of fMRI and achieving a classification accuracy of 72.46%.

Keywords: ASD classification, CNNG, CNN, spatio-temporal features, ABIDE

INTRODUCTION

The neurodegenerative diseases such as autism spectrum disorder (ASD) have received increasing attention in recent years. ASD, also referred to as autism, is a common neurodevelopmental cognitive disorder in children, mostly related to genetic factors. Due to the unclear etiology of autism, lack of specific drug treatment and life-long incurable, the patient's family needs to bear heavy psychological and economic pressure for a long time. ASD is characterized by complexity and heterogeneity. ASD mainly relies on the doctor's diagnosis on the foundation of the Diagnostic and Statistical Manual of Mental Disorders. It is not only time-consuming but also highly subjective, which can easily lead to misdiagnosis. Therefore, the development of a fully automatic ASD diagnostic technology will alleviate the burden on doctors and be helpful to detect symptoms and obtain early intervention and treatment in childhood.

With the development of medical imaging, many functional neuroimaging techniques have been proposed to use in brain research, such as Electroencephalogram (EEG), magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and so on (Laxer, 1997; Wu et al., 2001; Holdsworth and Bammer, 2008). fMRI has the advantages of non-invasiveness and

high temporal and spatial resolution. fMRI can enable people to more intuitively understand the physiological and pathological functional activities of the brain. Therefore, fMRI is widely used in clinical and basic research in many fields such as neuroscience, cognitive science and psychology (Heuvel and Pol, 2010; Liu S. et al., 2019; Liu S. et al., 2020; Vakamudi et al., 2020). The fMRI which takes blood oxygenation level dependent (BOLD) signal imaging as the fundamental principle can be divided into task state and resting state in brain research. Task-fMRI means that fMRI data is collected by subjects under the specified task, such as staring at a certain color of a certain mark or moving a finger for a period of time. As a method of acquiring brain signals with the high spatial and temporal resolution, resting-state fMRI (rs-fMRI) requires subjects in a state of complete relaxation without accepting any specified or strenuous tasks. The acquisition method is simple and fast and is suitable for ASD patients, so it is widely used in ASD classification. As with most classification studies of neurological disorders, the data used in this paper were resting-state fMRI. Due to the lack of subtype data of ASD in the current public datasets, ASD classification studies are mainly aimed at dichotomizing ASD and typical controls (TC). We also aim to distinguish ASD and TC.

In recent years, with the advancements in computer technology and machine learning, artificial intelligence has been broadly applied in different industrial fields. Scholars are committed to using machine learning to process and analyze medical data. The processing based on medical data has received more and more attention from researchers. Brain neuroimaging has also gradually provided a new way for the classification research of brain neurological diseases. The study of fMRI-based ASD classification can be divided into two directions in terms of model composition: traditional machine learning and deep learning.

Scholars from various countries have proposed different ASD classification and identification methods with traditional machine learning. The main steps include feature extraction and classification. In 2015, Plitt et al. (2015) used three groups of regions of interest to generate three independent fMRI time-course correlation matrices for subjects. Then, the generated feature matrix is used for classification by combining linear kernel-based support vector machine (SVM), and the classification accuracy was 73.89% in 178 subjects. In 2020, Wang et al. (2020) put forward a multi-site adaption framework via low-rank representation decomposition to address the differences between multiple sites. The key idea is to establish a common low-rank representation for data from multiple sites. One site can be treated as the target domain and the rest as the source domain. So, each site can be mapped into a common space by using the low-rank representation. It can reduce the distribution difference of data at different sites by using the data of the target domain to linearly represent the data of the source domain. Finally, the proposed algorithm used a linear kernel-based SVM classifier for ASD classification, and its classification accuracy is 71.88% in 468 subjects. In 2020, Zhao et al. (2020) extracted the time-invariant features in the low-order or high-order dynamic functional connectivity network of fMRI data by using central moment. By integrating the traditional functional connectivity network, the low-order dynamic functional connectivity network and features were extracted from the high-order dynamic functional connectivity network, and a linear kernel-based SVM classifier was used to obtain up to 83.00% accuracy in 45 ASD patients and 47 TCs. In the same year, Karampasi et al. (2020) used the time series extracted by the CC200 atlas, demographic information, texture and divergence features of the BOLD signal as manually extracted features. Then, five feature selection algorithms such as recursive feature elimination with correlation bias reduction, local learning, infinite feature selection, minimum redundancy maximum correlation and Laplace score were used for feature selection. Finally, SVM based on linear kernel and Gaussian kernel, K-nearest neighbor classifier, linear discriminant analysis and random forest were used for ASD classification. Among them, the linear kernel-based SVM classifier achieved the highest classification accuracy of 72.5% among 871 subjects. Sun et al. (2021) first investigated the statistical differences among six resting-state networks. Then, they analyzed subjects with independent component analysis and applied an image-based meta-analysis to explore the consistency of spatial patterns across different sites. Finally, using these patterns as features, the results were predicted by an SVM classifier based on the Gaussian radial basis sum function. The six resting-state networks achieved classification accuracies of 66.10, 53.20, 59.70, 50.00, 75.80, and 88.70% in 295 subjects, respectively.

The process of feature selection in traditional machine learning algorithms is often accompanied by a certain degree of subjectivity. With the rapid progress of computer technology, classification algorithms based on deep learning have gained popularity. Deep learning-based methods can learn optimal classification strategies directly from raw data by using hierarchies of varying complexity. Compared with traditional machine learning methods, it has stronger classification and recognition capabilities. In 2018, Heinsfeld et al. (2018) used the CC200 functional atlas to segment the brain into 200 regions of interest (ROI) and calculated the Pearson correlation coefficient between each ROI to generate a functional connectivity matrix. Then, by removing the upper triangular and diagonal parts of the functional connectivity matrix, the remaining parts were spread into a one-dimensional vector to be used as classification features. Finally, two stacked denoising self-encode network with Softmax activation function was used for ASD classification, which obtained an accuracy of 70% in 1,035 subjects. In 2018, Xiao et al. (2018) divided the dataset of each subject into 30 independent components. Then, 20 key components were selected based on the maximum energy criterion for all bands. The array of 84 key features for all subjects was reshaped into a 3,400*84-dimensional key feature matrix. After performing normalization, the feature matrix was fed into a stacked autoencoder and the subjects were classified by using a Softmax classifier, which obtained an average classification accuracy of 87.21% in 84 subjects. In 2019, Rathore et al. (2019) obtained a classification accuracy of 69.2% in 1,035 subjects by using a simple 3-layer neural network with functional correlation and its topological features of EEG signals. In 2020, Thomas et al. (2020) trained a full 3D CNN containing only the average pooling layer and two convolutional layers, and the classification accuracy achieved 66% on 1,162 subjects.

Niu et al. (2020) introduced a multichannel deep attention neural network for ASD classification, whose classification accuracy achieved 73.2% in 809 subjects. Li et al. (2020b) put forward an ASD classification algorithm on the basis of interpretable graph neural networks. In this algorithm, each graph convolution block contains a nodal convolution layer and a nodal pooling layer. This algorithm segmented brain images into 84 ROIs by using Desikan-Killiany mapping and constructed a functional connectivity matrix by using Pearson correlation coefficients. The functional connectivity matrix was fed into the proposed graph neural network for ASD classification, which obtained 79.7% classification accuracy in 118 subjects.

In 2019, Khosla et al. (2019) extracted ROI time series features by different atlases and further proposed an integrated learning strategy based on 3D CNN. The new network used the full-resolution 3D spatial structure of rs-fMRI data to fit a non-linear prediction model and obtained a classification result of 72.8%. Li et al. (2020a) presented an ASD classification algorithm by combining attention, long and short-term memory recurrent neural networks and self-encoder networks. This algorithm utilized functional connectivity as a feature and achieved 71.3% inter-site classification accuracy.

Functional magnetic resonance imaging images are an arrangement of a series of three-dimensional images obtained in a time series with a large number of data voxels. Most current methods used atlases to segment the brain into multiple regions of interest and construct a functional connectivity matrix as features. Then feature extraction methods were used to select some of the optimal features to input into a classifier for ASD classification. These algorithms did not fully exploit the spatio-temporal information of the source images. And they destroyed the temporal and spatial correlation of the original data. Therefore, we design an ASD classification algorithm based on 3D CNN and GRU. The representative high-level features of 3D images at each time point are gradually extracted by 3D convolutional neural networks. Then, the above spatial features at each time point are fed into GRU in series to analyze their temporal correlation information. Finally, a fully connected layer with a Sigmoid activation function is used to predict the category.

The main contributions of this paper are: (1) We combine the strengths of 3D CNN and GRU to construct a CNNG network. The CNNG network performs well in extracting the spatio-temporal features of fMRI data and hence obtains better ASD classification performance. (2) CNNG adopts intercepting time dimension, scaling brain image size as well as regularization and Dropout to prevent the overfitting phenomenon during model training. (3) We select the data of 871 subjects in the commonly used ABIDE database as the experimental data so that the trained model has better generalization ability for the diagnosis of ASD.

THE PROPOSED ALGORITHM

Some studies have been conducted for ASD classification by using CNN (Li et al., 2018; El-Gazzar et al., 2019). Because of the complexity and high dimensionality of fMRI images, only a few studies are using intact brain images directly as input data. Researchers have devoted themselves to reducing the input dimensionality by downscaling four-dimensional images into two-dimensional images or segmenting brain regions to construct functional connectivity matrices. And then the CNN networks or brain functional networks are constructed for classification. However, the above methods severely neglect the spatio-temporal information in fMRI data. Because the original fMRI data has high spatial and temporal dimensions, it will cause a serious overfitting phenomenon when the original fMRI data is the direct input of the network. Therefore, in this paper, the fMRI data are spatially reduced and intercepted with fixed temporal dimensions. For better results, we construct a deep learning classification model based on spatio-temporal features by combining it with a 3D convolutional neural network and gated recurrent unit, called the CNNG model. The CNNG model uses multiple 3D CNN networks with shared weights to extract the spatial features of brain images at each time point, and then uses the GRU to resolve the temporal information. The structure of the CNNG model is presented in **Figure 1**.

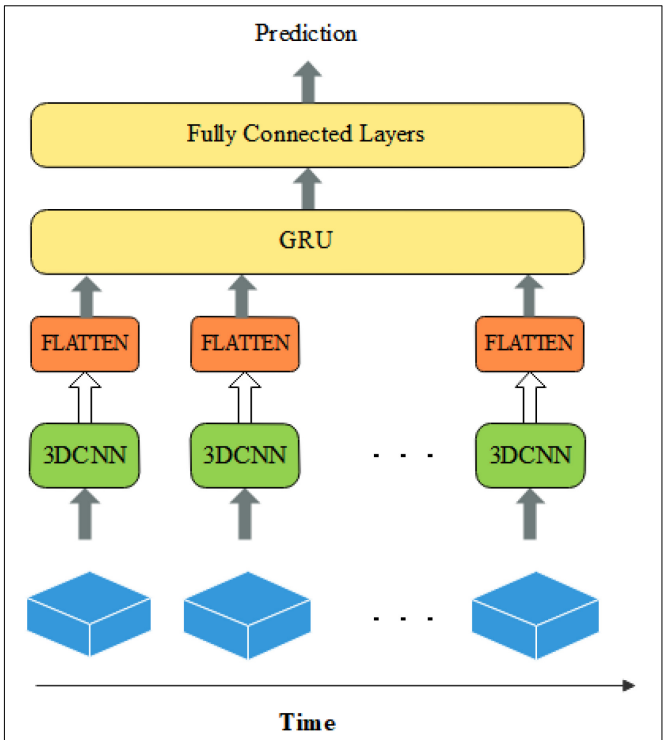


FIGURE 1 | The structure of the CNNG model.

Figure 1 illustrates that the 3D spatial structure of fMRI at each time point is sent to 3D CNN for spatial feature extraction. The extracted spatial features are separately flattened and sent to GRU for temporal feature extraction. The last layer is the fully connected layers (FC) with the Sigmoid activation function, which predicted classification results. Each node of the layer in FC is attached to all nodes of the previous layer, and the features extracted in the previous layer are combined to output the prediction probability. Each part is described in detail below.

Three-Dimensional Convolutional Neural Network

A convolutional neural network is a deep feed-forward neural network with local connectivity features and weight sharing. 3D convolution extends 2D convolution to 3D and extracts features of 3D data by 3D kernel convolution. Assuming that element $k_{ij}^{x_0 y_0 z_0}$ is the value at the position (x_0, y_0, z_0) of the j-th feature map of the i-th layer, then the three-dimensional convolution can be expressed as:

$$k_{ij}^{x_0 y_0 z_0} = \partial\left(b_{ij} + \sum_{c}\sum_{p=0}^{P_i-1}\sum_{q=0}^{Q_i-1}\sum_{r=0}^{R_i-1} w_{ijc}^{pqr} k_{(i-1)c}^{(x_0+p)(y_0+q)(z_0+r)}\right) \quad (1)$$

where ∂ is the activation function. Pi, Q_i, and R_i are the dimensional magnitudes of the three directions, respectively. w_{ijc}^{pqr} is the value of the convolution kernel connecting the c-th feature map of the $i-1$-th layer with the j-th feature map of the i-th layer at the position (p, q, r). b_{ij} is bias.

Medical images contain two-dimensional, three-dimensional and four-dimensional images, etc. 3D convolution can extract spatial features of 3D images, which is increasingly used in medical image analysis. The fMRI contains data from 3D brain space images, so the 3D CNN is suitable for the 3D spatial feature extraction of fMRI. In CNN, large convolutional kernels can be replaced with repeated small convolutional kernels. The different sizes of the convolutional kernels bring the different sizes of the perceptual field. So, it is often used to replace one layer of large convolutional kernels with multiple layers of small convolutional kernels to reduce the number of parameters and computation while maintaining the same perceptual field. For example, it is very common to replace one layer of 5×5 convolutional kernels with two layers of 3×3 kernels, and to replace one layer of 7×7 kernels with three layers of 3×3 kernels.

The structure of the 3D CNN model used in this paper is presented in **Figure 2**. The input size of 3D CNN is $28 \times 28 \times 28$, and it contains three convolutional layers. Each convolutional layer has 8 convolution kernels with the size of $3 \times 3 \times 3$, and they are all connected with ReLU layers. The fourth layer is the maximum pooling layer with a step size of 2 and a kernel size of $2 \times 2 \times 2$. The main purpose of the maximum pooling layer is to reduce the image size, prevent overfitting and reduce the running time. To extract more advanced features, three sets of repeated convolutional and pooling layers are added after the pooling layer. And the size of each convolutional kernel is $3 \times 3 \times 3$. The number of filters in each convolutional kernel is 16, 32, and 64. The size of the pooling kernel after each convolutional layer is $2 \times 2 \times 2$.

Gated Circulation Unit

After extracting the fMRI spatial features by using 3D CNN, we use GRU to process the spatial features arranged along the time dimension after flattening. GRU is an improved version of long short-term memory (LSTM) presented by Cho et al. (2014), in which many ideas are borrowed from LSTM. LSTM has three inputs and three outputs, while the GRU has two inputs and two outputs (Xin et al., 2021; Liu S. et al., 2022). GRU can accelerate the training and enhance the network performance because of fewer parameters. The structure of GRU neurons is shown in **Figure 3**.

Let be the input of the GRU, and c_t is the output of the GRU. As can be seen from **Figure 3**, the expression of GRU is also slightly different from that of LSTM, with the following equation:

$$z_t = \sigma\left(W_z \cdot \left[h_{t-1}, x_t\right]\right) \quad (2)$$

$$r_t = \sigma\left(W_r \cdot \left[h_{t-1}, x_t\right]\right) \quad (3)$$

$$\tilde{h}_t = tanh\left(W \cdot \left[r_t * h_{t-1}, x_t\right]\right) \quad (4)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (5)$$

where z_t denotes update gate. r_t denotes reset gate. $\tilde{h}_t$ denotes hidden unit. h_t is the current moment output. W_z, W_r, and W denote weights. *tanh* is the activation function.

In terms of operation, the GRU and LSTM work in a similar way. But the GRU unit uses a hidden state to combine the forgetting and input gates into a single update gate. It controls both how much information needs to be forgotten from the hidden layer of the previous moment and how much memory information from the hidden layer of the current moment is added. There is also a new "gate" in the GRU called the reset gate, which controls whether the computation of $\tilde{h}_t$ depends on state h_{t-1} at the previous moment. When, $r_t = 0$, $\tilde{h}_t$ is only related to the current input x_t and has nothing to do with the history state. When, $r_t = 1$, $\tilde{h}_t$ is related to x_t and h_{t-1}. The advantage of GRU over LSTM is that there is less internal "gating" and fewer parameters than LSTM. GRU can achieve equivalent levels of performance, and it is easier to train, which can greatly increase training efficiency. Therefore, we use GRU for feature extraction in the time dimension to obtain better ASD classification results.

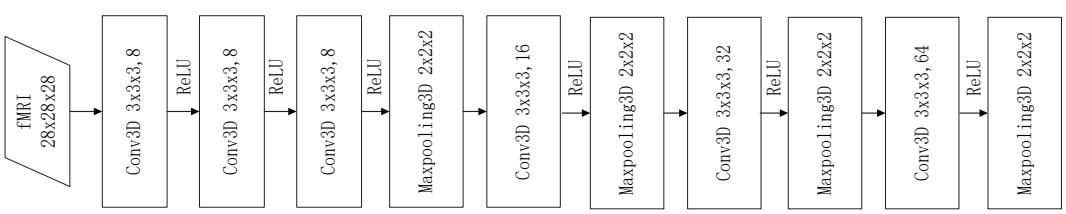


FIGURE 2 | The structure of single-frame convolutional neural network (CNN).

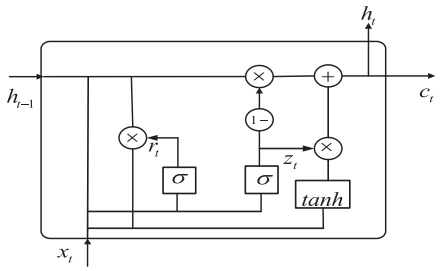


FIGURE 3 | The structure of gate recurrent unit (GRU).

Model Training

The proposed model extracts the spatial features before the temporal analysis. The specific parameter settings of the single-frame CNN model are given in **Table 1**. We first use two repeated three-dimensional convolutions with the size of $3 \times 3 \times 3$ to extract the low-level features. Then, we use repeated pooling and convolution to extract the high-level features. And the repeated two-layer convolution is replaced by a single convolutional layer with a kernel size of $3 \times 3 \times 3$, which is reduce the number of parameters. The extracted spatial features are flattened and input into a GRU with 32 neurons. Finally, the predicted values are output by a fully connected layer with a Sigmoid activation function.

The adaptive moment estimation (Adam) optimization algorithm is used for model optimization in the network training. The loss function is the cross-entropy loss function. The input batch size is set to 1, and the learning rate is 0.00001. Dropout means that in the training process of the network, neural network units are randomly discarded from the network according to a certain probability. To avoid overfitting of the proposed model, the values of dropout and recurrent_dropout of the parameters in GRU are set to 0.3. The two-parameter regularization is carried out in the Dense layer with the parameter 0.00001.

DATA PREPROCESSING

The rs-fMRI data used in this paper are from the international publicly available Autism Brain Imaging Data Sharing Project dataset. The dataset brings together 1,112 subjects (539 ASD patients and 573 TCs) from 17 sites around worldwide, including: California Institute of Technology (Caltech), Carnegie Mellon University (CMU), Kennedy Krieger Institute (KKI), Ludwig Maximilians University Munich (MaxMun), New York University Langone Medical Center (NYU), Olin Institute of Living at Hartford Hospital (Olin), Oregon Health and Oregon Health and Science University (OHSU), San Diego State University (SDSU), Social Brain Lab (SBL), Stanford University (Stanford), Trinity Centre for Health Sciences (Trinity), University of California, Los Angeles (UCLA), University of Leuven (Leuven), University of Michigan (UM), University of Pittsburgh School of Medicine (Pitt), University of Utah School of Medicine (USM), and Yale Child Study Center (Yale). The corresponding sites and the sizes of samples are shown in **Table 2**, and all the data can be downloaded from the official website from ABIDE I (2022). The database includes rs-fMRI, structural MRI, and extensive phenotypic information for each subject. In this paper, the subjects with missing partial information were excluded. A final dataset of 871 subjects, including 403 ASD patients and 468 TCs, was obtained by removing the samples with incomplete brain coverage, high motion peaks, ghosting and other scanner artifacts.

During fMRI acquisition, a lot of noise is generated, so preprocessing is required before use. The preprocessing method used in this paper is the configurable pipeline for the analysis of connectomes (CPAC), and the specific processing steps are as follows:

(1) Time slice correction. There is a time difference in the acquisition of fMRI images. To ensure the accuracy of the images, 3dTshift of functional neuroimaging analysis was used to correct the time slices.

(2) Head movement correction. When collecting data, it is impossible to guarantee that the subject does not move at all.

TABLE 1 | Structure of single-frame 3D convolutional neural network (CNN) model.

Layer	Type	Output size	Filter	Core size
1	Conv3D	28× 28× 28	8	3× 3× 3
2	Conv3D	28× 28× 28	8	3× 3× 3
3	Conv3D	28× 28× 28	8	3× 3× 3
4	MaxPooling3D	14× 14× 14	8	2× 2× 2
5	Conv3D	14× 14× 14	16	3× 3× 3
6	MaxPooling3D	7× 7× 7	16	2× 2× 2
7	Conv3D	7× 7× 7	32	3× 3× 3
8	MaxPooling3D	4× 4× 4	32	2× 2× 2
9	Conv3D	4× 4× 4	64	3× 3× 3
10	MaxPooling3D	2× 2× 2	64	2× 2× 2

TABLE 2 | Names of the 17 sites and their sample sizes.

Serial number	Sites	ASD	TC	Total subjects
1	Caltech	19	19	38
2	CMU	14	13	27
3	KKI	22	33	55
4	Leuven	29	35	64
5	MaxMun	24	33	57
6	NYU	79	105	184
7	OHSU	13	15	28
8	Olin	20	16	36
9	Pitt	30	27	57
10	SBL	15	15	30
11	SDSU	14	22	36
12	Stanford	20	20	40
13	Trinity	24	25	49
14	UCLA	62	47	109
15	UM	68	77	145
16	USM	58	43	101
17	Yale	28	28	56

Some slight movements can lead to huge data differences, so head movement correction is needed.

(3) Alignment. The skewed functional or structural image is adjusted to the vicinity of the spatial standard position, so that the subsequent processing algorithm can quickly find the optimal value and ensure a higher quality alignment.

(4) Numerical normalization. The 4D fMRI images were globally normalized with the global mean value equal to 1,000.

(5) Interference signal regression. The Friston 24-parameter model regression was used to eliminate the head movement effect of the functional image after alignment. To reduce the effect of respiration and heartbeat, the regression was done. Regression is also used to remove low-frequency drift generated by the long machine operation (Wang, 2020).

(6) Filtering. To reduce the influence of noise such as breathing and heartbeat and remove the low-frequency drift and the high-frequency noise, the low-frequency signal in the range of 0.01–0.1 Hz is selected. This frequency band can reflect the individual's spontaneous neural activity and has certain biological significance (Lu et al., 2007).

(7) Spatial normalization. In general, the size of the human brain varies. In order to unify the standard, the image space is normalized to the template space of the Montreal Neurological Institute with a resolution of $3 \times 3 \times 3\ mm^3$.

In the ABIDE dataset, the dimensions of the 3D spatial brain images were consistent for each site. While the temporal dimensions varied, the OHSU site had the least temporal dimension of 78, and the CMU site had the highest temporal dimension of 316. Since the model requires a fixed input size, the fMRI data is preprocessed before being fed into the network. Specifically, the fMRI of the first ten time points was removed in the time dimension, and 32 consecutive frames of 3D brain images were taken from the eleventh frame. Spatially, the spatial dimension of each image (61, 73, 61) is downsampled to (28, 28, 28). After the processing of temporal and spatial dimensions, the size of the obtained fMRI data is (28, 28, 28, 32). This ensures the same model input and preserves the spatio-temporal characteristics of fMRI. The selection of the time input size is discussed in detail in the experimental results analysis section.

EXPERIMENTAL RESULTS AND ANALYSIS

The experiments in this paper are based on the Tensenflow 1.0 platform. The environment is the Ubuntu18.4 operating system. The hardware is a server with 32G memory, Intel(R) Xeon(R) CPU E5-2667 processor and NVIDIA Tesla K40c.

As we all know, in the field of deep learning, it is very important to divide the training set and test set reasonably. For the traditional machine learning stage (the size of data set is less than 10,000), the general allocation ratio is that the ratio of training set to test set is 7:3 or 8:2. Try to keep the distribution of training set and test set consistent. For verifying the validity of CNNG and retaining as much training data as possible, the data is categorized into a training set and a test set in a ratio of 8:2.

In binary classification studies, Accuracy, Sensitivity and Specificity are commonly used indicators, which can be expressed as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{7}$$

$$Specificity = \frac{TN}{TN + FP} \tag{8}$$

In this experiment, the label for ASD patients is "1," and the label for TC is "0." The above equation True Positive (TP) indicates the number of samples with label "1" predicted to the number of samples with label "1." False Positive (FP) indicates the number of samples that predict a label of "0" to a label of n"1" True Negative (TN) indicates the number of samples with the label "0" predicted to the number of samples with the label "0." False Negative (FN) indicates the number of samples with the label "1" predicted as the label "0." TP+FP+TN+FN is the total number of samples. TP+FN is the total number of samples with the true label "1". TP+FP is the total number of samples with the prediction label "1," including both correct and incorrect predictions. FP+TN represents the total number of samples with the true label "0." TN+FN represents the total number of samples with the prediction label "0," including both correct and incorrect predictions.

It can be seen from the above description that the sensitivity reflects the ability of ASD patients to be correctly distinguished. The higher the sensitivity means the higher the probability that a patient with ASD will be correctly diagnosed. The specificity reflects the effect of TC subjects being correctly classified. The accuracy reflects the overall classification ability. The higher the accuracy, the greater the value for practical medical diagnosis applications.

Ablation Experiments

Effects of Different Convolution Kernel Sizes

For the purpose of obtaining the optimal model, we select the number of convolution layers and the size of the convolution kernel by comparison experiments. First, the number of convolution layers and the kernel size before the first pooling layer are determined. The experiments were conducted by using convolution kernels with a size of $5 \times 5 \times 5$ and $7 \times 7 \times 7$ as well as replacing them with repeated small convolution kernels. The result is listed in **Table 3**. From **Table 3**, it is clear that the repeated small convolutional kernels have better classification results than the corresponding large convolutional kernels. For example, the superposition of three convolution kernels with a size of $3 \times 3 \times 3$ achieves a maximum accuracy of 72.46%. It is about 2% higher than the accuracy of the model with the corresponding convolutional kernel size of $7 \times 7 \times 7$. Similarly, when two convolution kernels with a size of $3 \times 3 \times 3$ are superimposed, the accuracy rate is 70.04%. While the classification accuracy of the model with a convolution kernel size of $5 \times 5 \times 5$ is only 67.63%. The above results show that the superposition of small convolutional kernels has the advantages of a small number of parameters and the low computational complexity. It is clear

that the classification effect of only one convolution layer with a convolution kernel size of $3 \times 3 \times 3$ is not ideal. It may be due to the fact that the receptive field is too small to effectively extract the features.

Comparison of Long Short-Term Memory Module and Gate Recurrent Unit Module

Since LSTM and GRU are the variations of RNN, both are widely used in temporal information extraction. The classification results based on temporal feature extraction selection by LSTM and GRU, respectively, are shown in **Table 4**. When GRU is replaced by LSTM in the proposed method, the classification accuracy is 68.60%, the sensitivity is 56.60%, and the specificity is 81.19%. The accuracy is significantly lower compared to GRU, so we use GRU for temporal feature extraction.

Effects of Different Time Dimensions

The selection of the number of fMRI time points has an important influence on the model training. In the time dimension, 8, 16, 32, and 48 frames of fMRI images are used for experiments in this paper. **Table 5** presents the classification effects.

According to **Table 5**, the classification accuracy improves when the temporal dimension increases. However, it starts to decrease when the temporal dimension is 48. Specifically, when the time dimension is taken as eight, the classification accuracy is low. This may be mainly due to the short time resulting in the short feature vector extracted by the tandem CNN, which cannot extract the temporal features effectively. And when the time dimension is 48, the number of parameters and computation of model training increases, which may easily lead to the phenomenon of overfitting. So, in the proposed algorithm, we finally choose the data of 32-time points, which can archive the best classification effect.

Effects of Different Numbers of Gate Recurrent Units

The selection of GRU has experimented in the previous section, and the number of GRU units also determines the performance of CNNG. So, we set the number of GRU units to 16, 32 and 48 for experimental analysis. From **Table 6**, the accuracy, sensitivity and specificity are lower when the number of GRU units is too small or too lager. This is because when the number of GRU units is less than 32, the proposed model is limited by the number of units and is not sufficient to fully express the information contained in the temporal dimension of the fMRI data. And when the number of units increases to 48, the classification performance shows different degrees of degradation due to overfitting because the parameters of the units are too redundant. Therefore, the model performance is optimal when the number of GRU units is taken as 32.

Comparison With Traditional Machine Learning Algorithms

For verifying the validity of CNNG, we compare it with the ASD classification algorithm by using traditional machine learning. The comparison algorithms are: (1) ASD classification algorithm by using graph Fourier transform (GFT) and support vector machine proposed by Brahim and Farrugia (2020), which is abbreviated as RBF-SVC; (2) An ASD classification algorithm based on functional connectivity networks and recursive-clustering elimination support vector machine proposed by Chaitra et al. (2020), which is abbreviated as RCE-SVM; (3) A hybrid ASD classification algorithm by combining different brain segmentation definitions, functional connectivity matrix construction methods and feature extraction methods proposed by Graa and Silva (2021), which is abbreviated as HFR; (4) The ASD classification algorithm proposed by Abraham et al. (2016) based on CC400 brain atlas and support vector machine, which is abbreviated as C-SVC; (5) The ASD classification algorithm on the basis of functional connectivity and ridge regression classifier proposed by Yang et al. (2019), which is abbreviated as FCR.

TABLE 3 | Classification performance of different convolutional kernel sizes.

Kernel sizes (number of layers)	Accuracy	Sensitivity	Specificity
7 x 7x 7 (1)	70.53%	64.15%	77.23%
5 x 5 x 5 (1)	67.63%	62.37%	72.64%
3 x 3 x 3 (3)	**72.46%**	**65.35%**	**79.25**%
3 x 3 x 3 (2)	70.04%	65.09%	75.25%
3 x 3 x 3 (1)	62.32%	59.80%	66.04%

The bold values in this table represent the optimal values of accuracy, specificity and sensitivity.

TABLE 4 | Performance of different temporal feature extraction modules.

Time feature extraction module	Accuracy	Sensitivity	Specificity
LSTM	68.60%	56.60%	**81.19%**
GRU	**72.46%**	**65.35%**	79.25%

The bold values in this table represent the optimal values of accuracy, specificity and sensitivity.

TABLE 5 | Classification performance of different time interceptions.

Time dimension	Accuracy	Sensitivity	Specificity
8	63.74%	59.33%	67.32%
16	69.08%	63.46%	72.12%
32	**72.46%**	**65.35%**	**79.25%**
48	69.57%	62.37%	76.42%

The bold values in this table represent the optimal values of accuracy, specificity and sensitivity.

TABLE 6 | Classification performance with different numbers of gate recurrent unit (GRU) units.

Number of GRU units	Accuracy	Sensitivity	Specificity
16	71.50%	64.15%	77.28%
32	**72.46%**	**65.35%**	**79.25%**
48	71.01%	62.38%	74.26%

The bold values in this table represent the optimal values of accuracy, specificity and sensitivity.

The test results of the compared algorithm used in this paper are all from the test results of the code offered by the author in the corresponding reference. The test set used in this paper comes from 17 different sites, so the final metrics obtained are the average accuracy, average sensitivity and specificity. **Table 7** presents the performance of CNNG and the comparison algorithm on the test set.

As presented in **Table 7**, the accuracy, sensitivity and specificity of the CNNG model in 871 subjects reached 72.46, 71.35, and 79.25%, respectively. The accuracy is obviously higher than other traditional machine learning algorithms. When classifying ASD, many traditional machine learning algorithms need to divide the brain into multiple regions of interest, which is treated as a node for subsequent feature selection or calculation. This process obviously loses fMRI spatial information of the data. After the original image is preprocessed, the CNNG model directly extracts and classifies features through the model, which fully exploits the spatiotemporal information of fMRI data, thereby extracting more discriminative features and further enhancing the classification capability of the algorithm. In addition, the manual features extracted by the fixed computational feature algorithm are sensitive to noise, scanning equipment and parameters, and make a big influence on the overall classification capability of traditional machine learning algorithms.

TABLE 7 | Classification performance of traditional machine learning algorithms and CNNG.

Classification	Accuracy	Sensitivity	Specificity
RBF-SVC	66.70%	62.35%	72.35%
RCE-SVM	67.30%	64.5%	70.10%
HFR	71.10%	67.00%	75.00%
C-SVC	67.00%	53.20%	78.30%
FCR	71.98%	70.89%	71.53%
CNNG	**72.46%**	**71.35%**	**79.25%**

The bold values in this table represent the optimal values of accuracy, specificity and sensitivity.

TABLE 8 | Classification performance of deep learning algorithms and CNNG.

Classification	Accuracy	Sensitivity	Specificity
CNN-MLP	70.22%	62.35%	72.35%
SVC	71.10%	67.00%	75.00%
DiagNet	70.30%	68.03%	72.20%
HI-GCN	67.20%	65.90%	68.40%
GAT	68.02%	74.06%	62.26%
CNNG	**72.46%**	**74.35%**	**79.25%**

The bold values in this table represent the optimal values of accuracy, specificity and sensitivity.

Comparison With Deep Learning Algorithms

We also carry out a comparison between CNNG and the deep learning-based ASD classification algorithm. The comparison algorithms are: (1) The ASD classification algorithm based on convolutional neural network and multilayer perceptron presented by Sherkatghanad et al. (2020), which is abbreviated as CNN-MLP; (2) The ASD classification algorithm based on functional connection network, extreme random tree and support vector machine proposed by Liu Y. et al. (2020), which is abbreviated as SVC; (3) The ASD classification algorithm based on joint representation learning deep multimodal model proposed by Eslami et al. (2019), which is abbreviated as DiagNet; (4) The ASD classification algorithm based on the hierarchical graph convolutional neural network framework introduced by Hao et al. (2020), which is abbreviated as HI-GCN; (5) The ASD classification algorithm based on graph attention network proposed by Hu et al. (2021), which is abbreviated as GAT. The test results of the comparison algorithm used in this paper are all from the test results of the code offered by the author in the corresponding reference. The test set used in this paper comes from 17 different sites, so the final metrics obtained are the average accuracy, average sensitivity and specificity. **Table 8** presents the results of CNNG and the comparison algorithm on the test set.

Table 8 shows that the proposed algorithm obtains an accuracy of 72.46% in the experiment of 871 subjects, which is 5.26% higher than that of HI-GCN, 4.44% higher than that of GAT, and 2.44% higher than that of CNN. It is also a certain improvement compared to SVC and DiagNet. The proposed algorithm also obtains a specificity of 79.25% and a sensitivity of 74.35%. All the results reveal that the overall performance of CNNG is superior to the other deep learning algorithms, which suggests that directly extracting spatio-temporal features from 4D fMRI data for classification has better results for ASD classification than just by using 2D or 3D fMRI data or functional

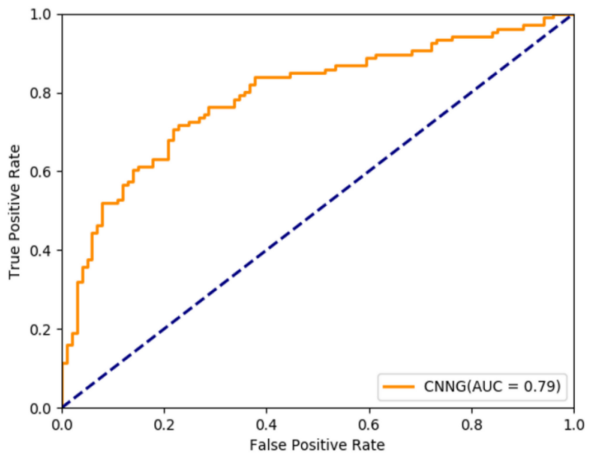


FIGURE 4 | Receiver operating characteristic (ROC) curve of CNNG model.

connectivity. For further evaluating the performance of CNNG, the receiver operating characteristic (ROC) curve and the area under the curve (AUC) values are plotted in **Figure 4**. In **Figure 4**, the horizontal coordinate represents the false positive rate (FPR) and the vertical coordinate represents the true positive rate (TPR). The ROC curve reflects the trend of the TPR and the FPR. The closer the area is to 1, the stronger the recognition ability is. Among the above proposed deep learning algorithms, the values of AUC for CNN-MLP, DiagNet, HI-GCN, and GAT are 0.7486, 0.764, 0.745, and 0.7358, respectively. As shown in **Figure 4**, the AUC value of the proposed algorithm is 0.79, which is 5.42% higher compared to GAT. This is an improvement of 4.5% compared to HI-GCN. Compared to CNN-MLP, the ACU of CNNG is improved by 4.14%. There is also a magnitude improvement compared to DiagNet. These data indicate that CNNG performs well for classification.

CONCLUSION

In this paper, we put forward a deep learning model—CNNG, which can fully exploit the spatio-temporal information in fMRI data to avoid excessive dimensionality reduction and missing information caused by using the manual features for classification. CNNG is mainly composed of 3D CNN and GRU. In CNNG, spatial feature extraction is extracted by using 3D CNN, and then GRU is used to analyze temporal information. The validity of the CNNG model is proved by comparing it with the algorithms based on traditional machine learning and the algorithms based on deep learning. The experimental results indicate that CNNG performs better than other algorithms for ASD classification. CNNG can extract fMRI data features from the perspective of spatio-temporal convolution, which has some clinical value for the early diagnosis of ASD. At present, the sensitivity of the proposed algorithm does not obtain a large improvement, and next, we will optimize the algorithm for better ASD classification.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: Autism Brain Imaging Data Exchange (ABIDE) dataset.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Autism Brain Imaging Data Exchange (ABIDE) dataset. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HZ performed the computer simulations. SL, XS, and JZ analyzed the data. SL and WJ wrote the original draft. XS and S-HW revised and edited the manuscript. JY polished the manuscript. All authors confirmed the submitted version.

FUNDING

This work was supported in part by National Natural Science Foundation of China under Grant 62172139, Natural Science Foundation of Hebei Province under Grants F2022201055 and F2020201025, Science Research Project of Hebei Province under Grant BJ2020030, Natural Science Interdisciplinary Research Program of Hebei University under Grant DXK202102, Open Project Program of the National Laboratory of Pattern Recognition (NLPR) under Grant 202200007, Open Foundation of Guangdong Key Laboratory of Digital Signal and Image Processing Technology under Grant 2020G DDSIPL-04, Project Funded by China Postdoctoral under Grant 271355, and Research Project of Hebei University Intelligent Financial Application Technology R&D Center (XGZJ2022022), and Open Foundation of Guangdong Key Laboratory of Digital Signal and Image Processing Technology under Grant 2020G DDSIPL-04.

ACKNOWLEDGMENTS

We thank the support provided by the High-Performance Computing Center of Hebei University.

REFERENCES

ABIDE I (2016). *Autism Brain Imaging Data Exchange I*. Available online at: http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html (accessed June 24, 2016).

Abraham, A., Milham, M., Martino, A. D., Craddock, R. C., Samaras, D., Thirion, B., et al. (2016). Deriving reproducible biomarkers from multi-site resting-state data: an Autism-based example. *NeuroImage* 147, 736–745. doi: 10.1016/j.neuroimage.2016.10.045

Brahim, A, and Farrugia, N. (2020). Graph Fourier transform of fMRI temporal signals based on an averaged structural connectome for the classification of neuroimaging. *Artif. Intell. Med.* 106:101870. doi: 10.1016/j.artmed.2020.101870

Chaitra, N., Vijaya, P. A., and Deshpande, G. (2020). Diagnostic prediction of autism spectrum disorder using complex network measures in a machine learning framework. *Biomed. Signal Process. Control* 62:102099. doi: 10.1016/j.bspc.2020.102099

Cho, K., Merrienboer, B. V., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., et al. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Comput. Sci.* 10, 1–11. doi: 10.3115/v1/D14-1179

El-Gazzar, A., Quaak, M., Cerliani, L., Bloem, P., Wingen, G. V., and Thomas, R. M. (2019). "A hybrid 3DCNN and 3DC-LSTM based model for 4D spatio-temporal fMRI data: an ABIDE autism classification study," in *International Workshop on Machine Learning in Clinical Neuroimaging*, (Shenzhen), 95–102. doi: 10.48550/arXiv.2002.05981

Eslami, T., Mirjalili, V., Fong, A., Laird, A., and Saeed, F. (2019). ASD-DiagNet: a hybrid learning approach for detection of autism spectrum disorder using fMRI data. *Front. Neuroinformatics* 13:70. doi: 10.3389/fninf.2019.00070

Graa, M., and Silva, M. (2021). Impact of machine learning pipeline choices in autism prediction from functional connectivity data. *Int. J. Neural Syst.* 31, 2150009–2150009. doi: 10.1142/S012906572150009X

Hao, J. A., Peng, C., Myx, A., Jy, A., and Oz, C. (2020). Hi-GCN: a hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction. *Comput. Biol. Med.* 127, 104096–104096. doi: 10.1016/j.compbiomed.2020.104096

Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., and Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage Clin.* 17, 16–23. doi: 10.1016/j.nicl.2017.08.017

Heuvel, M., and Pol, H. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* 20, 519–534. doi: 10.1016/j.euroneuro.2010.03.008

Holdsworth, S. J., and Bammer, R. (2008). Magnetic resonance imaging techniques: fMRI, DWI, and PWI. *Semin. Neurol.* 28, 395–406. doi: 10.1055/s-0028-1083697

Hu, J., Cao, L., Li, T., Dong, S., and Li, P. (2021). GAT-LI: a graph attention network based learning and interpreting method for functional brain network classification. *BMC Bioinformatics* 22:379. doi: 10.1186/s12859-021-04295-1

Karampasi, A., Kakkos, I., Miloulis, S. T., Zorzos, I., and Dimitrakopoulos, G. N. (2020). "A machine learning fMRI approach in the diagnosis of autism," in *2020 IEEE International Conference on Big Data (Big Data)*, (Atlanta, GA: IEEE), 3628–3631. doi: 10.1109/BigData50022.2020.9378453

Khosla, M., Jamison, K., Kuceyeski, A., and Sabuncu, M. R. (2019). Ensemble learning with 3D convolutional neural networks for functional connectome-based prediction. *NeuroImage* 199, 651–662. doi: 10.1016/j.neuroimage.2019.06.012

Laxer, K. D. (1997). Clinical applications of magnetic resonance spectroscopy. *Epilepsia* 38, S13–S17. doi: 10.1111/j.1528-1157.1997.tb04534.x

Li, X., Dvornek, N. C., Papademetris, X., Zhuang, J., and Duncan, J. S. (2018). "2-channel convolutional 3D deep neural network (2CC3D) for fMRI analysis: ASD classification and feature learning," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, (Washington, DC: IEEE), 1252–1255. doi: 10.1109/ISBI.2018.8363798

Li, X., Zhou, Y., Dvornek, N. C., Zhang, M., Zhuang, J., Ventola, P., et al. (2020b). "Pooling regularized graph neural network for fMRI biomarker analysis," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, (Lima), 625–635. doi: 10.1007/978-3-030-59728-3_61

Li, X., Gu, Y., Dvornek, N., Staib, L. H., and Duncan, J. S. (2020a). Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results. *Med. Image Anal.* 65, 101765–101765. doi: 10.1016/j.media.2020.101765

Liu, S., Yin, L., Miao, S., Ma, J., and Hu, S. (2020). Multimodal medical image fusion using rolling guidance filter with CNN and nuclear norm minimization. *Curr. Med. Imaging* 16, 1243–1258. doi: 10.2174/1573405616999200817103920

Liu, S., Zhao, C., An, Y., Li, P., and Zhang, Y. (2019). Diffusion tensor imaging denoising based on Riemannian geometric framework and sparse Bayesian learning. *J. Med. Imaging Health Inform.* 9, 1993–2003. doi: 10.1166/jmihi.2019.2832

Liu, S., Zhao, L., Zhao, J., Li, B., and Wang, S. H. (2022). Attention deficit/hyperactivity disorder classification based on deep spatio-temporal features of functional magnetic resonance imaging. *Biomed. Signal Process. Control* 71:103239. doi: 10.1016/j.bspc.2021.103239

Liu, Y., Xu, L., Li, J., Yu, J., and Yu, X. (2020). Attentional connectivity-based prediction of autism using heterogeneous rs-fMRI data from CC200 atlas. *Exp. Neurobiol.* 29, 27–37. doi: 10.5607/en.2020.29.1.27

Lu, H., Zuo, Y., Gu, H., Waltz, J. A., Zhan, W., Scholl, C. A., et al. (2007). Synchronized delta oscillations correlate with the resting-state functional MRI signal. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18265–18269. doi: 10.1073/pnas.0705791104

Niu, K., Guo, J., Pan, Y., Gao, X., Peng, X., Li, N., et al. (2020). Multichannel deep attention neural networks for the classification of autism spectrum disorder using neuroimaging and personal characteristic data. *Complexity* 2020, 1–9. doi: 10.1155/2020/1357853

Plitt, M., Barnes, K. A., and Martin, A. (2015). Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *NeuroImage Clin.* 7, 359–366. doi: 10.1016/j.nicl.2014.12.013

Rathore, A., Palande, S., Anderson, J. S., Zielinski, B. A., Fletcher, P. T., and Wang, B. (2019). "Autism classification using topological features and deep learning: a cautionary tale," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, (Istanbul), 736–744. doi: 10.1007/978-3-030-32248-9_82

Sherkatghanad, Z., Akhondzadeh, M., Salari, S., Zomorodi-Moghadam, M., and Salari, V. (2020). Automated detection of autism spectrum disorder using a convolutional neural network. *Front. Neurosci.* 13:1325. doi: 10.3389/fnins.2019.01325

Sun, J. W., Fan, R., Wang, Q., Wang, Q. Q., and Ma, H. B. (2021). Identify abnormal functional connectivity of resting state networks in autism spectrum disorder and apply to machine learning-based classification. *Brain Res.* 1757, 147299–147299. doi: 10.1016/j.brainres.2021.147299

Thomas, R. M., Gallo, S., Cerliani, L., Zhutovsky, P., and Wingen, G. V. (2020). Classifying autism spectrum disorder using the temporal statistics of resting-state functional MRI data with 3d convolutional neural networks. *Front. Psychiatry* 11:440. doi: 10.3389/fpsyt.2020.00440

Vakamudi, K., Posse, S., Jung, R., Cushnyr, B., and Chohan, M. O. (2020). Real-time presurgical resting-state fMRI in patients with brain tumors: quality control and comparison with task-fMRI and intraoperative mapping. *Hum. Brain Mapp.* 41, 797–814. doi: 10.1002/hbm.24840

Wang, C. (2020). *Research on Intelligent Algorithm for ASD Auxiliary Diagnosis based on Resting-State fMRI brain Functional Connectivity*. Ph.D. thesis. Nanchang: Nanchang University.

Wang, M., Zhang, D., Huang, J., Yap, P. T., and Liu, M. (2020). Identifying autism spectrum disorder with multi-site fMRI via low-rank domain adaptation. *IEEE Trans. Med. Imaging* 39, 644–655. doi: 10.1109/TMI.2019.2933160

Wu, O., Koroshetz, W. J., Ostergaard, L., Buonanno, F. S., Copen, W. A., Gonzalez, R. G., et al. (2001). Predicting tissue outcome in acute human cerebral ischemia using combined diffusion- and perfusion-weighted MR imaging. *Stroke* 32, 933–942. doi: 10.1161/01.STR.32.4.933

Xiao, Z., Wang, C., Jia, N., and Wu, J. (2018). SAE-based classification of school-aged children with autism spectrum disorders using functional magnetic resonance imaging. *Multimed. Tools Appl.* 77, 22809–22820. doi: 10.1007/s11042-018-5625-1

Xin, Q., Hu, S., Liu, S., Zhao, L., and Wang, S. (2021). WTRPNet: an explainable graph feature convolutional neural network for epileptic EEG classification. *ACM Trans. Multimed. Comput. Commun. Appl.* 17, 1–18. doi: 10.1145/3460522

Yang, X., Islam, M. S., and Khaled, A. (2019). "Functional connectivity magnetic resonance imaging classification of autism spectrum disorder using the multisite ABIDE dataset," in *IEEE EMBS International Conference on Biomedical & Health Informatics. Mathematics & Computer Science*, (Magnolia: IEEE), 1–4. doi: 10.1109/BHI.2019.8834653

Zhao, F., Chen, Z., Rekik, I., Lee, S. W., and Shen, D. (2020). Diagnosis of autism spectrum disorder using central-moment features from low- and high-order dynamic resting-state functional connectivity networks. *Front. Neurosci.* 14:258. doi: 10.3389/fnins.2020.00258

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 13 July 2022
doi: 10.3389/fnagi.2022.908143

Soft Attention Based DenseNet Model for Parkinson's Disease Classification Using SPECT Images

Mahima Thakur[1], Harisudha Kuresan[1], Samiappan Dhanalakshmi[1], Khin Wee Lai[2]* and Xiang Wu[3]**

[1] Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Chennai, India, [2] Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia, [3] School of Medical Information Engineering, Xuzhou Medical University, Xuzhou, China

OPEN ACCESS

Edited by:
Sang-Bing Tsai,
Wuyi University, China

Reviewed by:
Xiongtao Zhang,
Huzhou University, China
Rostom Mabrouk,
Bishop's University, Canada
Bing Li,
Affiliated Cancer Hospital of Zhengzhou University, China

***Correspondence:**
Samiappan Dhanalakshmi
dhanalas@srmist.edu.in
Khin Wee Lai
lai.khinwee@um.edu.my
Xiang Wu
wuxiang@xzhmu.edu.cn

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

Received: 30 March 2022
Accepted: 05 May 2022
Published: 13 July 2022

Citation:
Thakur M, Kuresan H, Dhanalakshmi S, Lai KW and Wu X (2022) Soft Attention Based DenseNet Model for Parkinson's Disease Classification Using SPECT Images. Front. Aging Neurosci. 14:908143. doi: 10.3389/fnagi.2022.908143

Objective: Deep learning algorithms have long been involved in the diagnosis of severe neurological disorders that interfere with patients' everyday tasks, such as Parkinson's disease (PD). The most effective imaging modality for detecting the condition is DaTscan, a variety of single-photon emission computerized tomography (SPECT) imaging method. The goal is to create a convolutional neural network that can specifically identify the region of interest following feature extraction.

Methods: The study comprised a total of 1,390 DaTscan imaging groups with PD and normal classes. The architecture of DenseNet-121 is leveraged with a soft-attention block added before the final classification layer. For visually analyzing the region of interest (ROI) from the images after classification, Soft Attention Maps and feature map representation are used.

Outcomes: The model obtains an overall accuracy of 99.2% and AUC-ROC score 99%. A sensitivity of 99.2%, specificity of 99.4% and f1-score of 99.1% is achieved that surpasses all prior research findings. Soft-attention map and feature map representation aid in highlighting the ROI, with a specific attention on the putamen and caudate regions.

Conclusion: With the deep learning framework adopted, DaTscan images reveal the putamen and caudate areas of the brain, which aid in the distinguishing of normal and PD cohorts with high accuracy and sensitivity.

Keywords: neural networks, Parkinson's disease (PD), DenseNet architecture, region of convergence (ROC), area under the curve

INTRODUCTION

Parkinson's disease (PD) is recognized as a chronic neurodegenerative condition of the central nervous system that primarily affects older adults (Pahuja et al., 2019) by Pereira et al. Researchers recognize the lack of dopaminergic neurons as the major cause of PD (Prediger et al., 2014). In the etiology of PD, oxidative stress is becoming a key factor in dopaminergic neuron degeneration (Zhou et al., 2009; Blesa et al., 2015). Loss of dopaminergic neurons is observed in substantia nigra of the mid-brain and later in loss of dopamine transporters in the striatum (Porritt et al., 2005). The striatum is the most significant component of the brain's basal ganglia region, which produces

is responsible for releasing the dopamine neurons in the midbrain. The disease's progressive nature may be attributed to the gradual deterioration in the striatum with age (Shahed and Jankovic, 2007).

Early PD is defined as the time frame before the start of severe motor symptoms and before the beginning of significant neurological impairment; yet, there is a scarcity of evidence that indicates the true potential of early therapy in terms of clinical and financial results (Zhou et al., 2009). More clarification is needed to study the true effect of early intervention on these outcomes. Future research should examine the impact of new diagnostic tools like genetic biomarkers on a wide range of medical issues (Porritt et al., 2005; Zhou et al., 2009).

Non-motor symptoms of PD result include anosmia (which affects the olfactory system), fatigue, disturbed sleep cycle, and fluctuations in bodyweight, disorders involving in temperament and cognitive aptitude, coronary artery disorders, bladder and bowel incontinence and digestive tract disorders. While the motor symptoms include resting tremor, rigidity, impaired body balance, slowing and freezing down of body movements or bradykinesia. As a result of such motor symptoms, the affected individual suffers from micrographia, dystonia and overall struggle in daily life activities. One may be able to find significant traits that are not normally employed in the clinical diagnosis of PD using machine learning algorithms, and depend on these alternative measures to diagnose PD in preclinical stages or atypical forms.

A person's gait and movement patterns are closely scrutinized during a medical examination. Parkinson's disease is characterized by bradykinesia (slow, tiny movements). Rigidity, or the quality of being rigid. During a medical exam, we passively move the patient's joints to discover this. The arms, legs, and neck are often rigid in those with Parkinson's disease. Restless tremors. While individuals aren't paying attention or are preoccupied, these tremors come out, therefore a good opportunity to notice this is when someone is walking. To diagnose this illness, there are no biomarkers (tests or assessments). We don't need to do imaging or laboratory testing unless we feel that there is a different reason for the patient's symptoms. Movement indicators are just a portion of the picture for Parkinson's disease, since it affects every region of the body.

Changes in memory and cognition, difficulty sleeping, emotional symptoms including worry and sadness, or even hallucinations, are among the most often reported symptoms. Patients with Parkinson's disease are more likely to have issues with their autonomic nervous system. Controls such as heart rate, digestion, breathing, pupillary response and urine and sexual desire are all under the control of this system, which is mostly unconsciously active (Marine et al., 2019).

Neuroimaging technique in the recent past, particularly SPECT (single-photon emission computed tomography), have presented promising potential because of their sensitivity and specificity in diagnosing early PD. SPECT is proved to be more accessible to clinicians, being less expensive (Lauretani et al., 2015; Jinjin et al., 2019). The SPECT method of imaging avails 123I-FP-CIT, i.e., 123I-Ioupane. This radioligand binds the dopamine transporters in the striatum and termed as SPECT DaTSCAN Dopamine transporter levels in the brain may be seen using the DaTSCAN procedure, a form of Single-Photon Emission Computed Tomography (SPECT) (Harisudha et al., 2021). Traditionally, a standardized analysis and detection of such subject images are carried out by specialized technicians and radiologists. Notably, smaller putamen and caudate regions (the dopamine transporters) are observed in the case of PD patients, mainly because of the steady deficiency of dopaminergic neurons (Shahed and Jankovic, 2007; Mohammed et al., 2020).

In the healthcare industry, machine learning techniques are becoming more prevalent. Machine learning allows an algorithms to learn and extract meaningful representations from data in a semi-automated way, as the term indicates. Machine learning models have been used to diagnose Parkinson's disease using a variety of data modalities, such as handwriting trends, gait patterns, and neuroimaging methodologies (Dhanalakshmi and Venkatesh, 2016; Matesan et al., 2018; Oláh et al., 2021).

Patients with PD who are diagnosed and treated early have decreased chance of progression and perhaps cheaper long-term care expenditures. Computer-Aided Diagnosis models that sufficiently make use of Artificial Intelligence (AI) techniques, particularly Deep Learning (DL) methods in the recent past; have suitably specialized as a reliable diagnostic tool (Matesan et al., 2018; Oláh et al., 2021). With the advancement in central processing unit (CPU) and graphics processing unit (GPU), better availability of reliable databases with ease of access in online platforms, and rapid improvisation of learning algorithms (Rumman et al., 2019; Bevilacqua et al., 2020).

Parkinson's disease is distinct from other disorders in various ways, including how well it responds to levodopa. PD may be differentiated using a variety of neuroimaging methods, according to current scientific research. An imaging study using positron emission tomography (PET) has revealed a possible mechanism for the lack of response to PD treatment, as the study was also used in the preservation of dopamine receptors in PSP (Brooks et al., 1992). MRI with a high field strength (1.5 T) and a heavy T2 weighting, [18F]-fluorodopa positron emission tomography, [11C] raclopride imaging of dopamine D2 receptors, and single photon emission computed tomography of striatal dopamine reuptake sites are all possible imaging investigations (Warren et al., 2007).

MRI is the best structural imaging method that does not use ionizing radiation when compared to nuclear imaging. In the early stages of Parkinson's disease, the vast majority of routine MRI methods failed to detect disease-specific abnormalities. Brain parenchyma sonography, a commonly used diagnostic tool for Parkinson's disease (Brooks et al., 1992), recently revealed abnormal hyper echogenicity in both PD and essential tremor (Warren et al., 2007). As a recent research found that 77% of levodopa patients first reacted favorably to the medicine, levodopa has become an essential treatment for Parkinson's disease (PD) (Brooks et al., 1992). Levodopa has been argued by physicians to be detrimental to prognosis since it is not definite of Parkinson's disease (Warren et al., 2007). To distinguish Parkinson's disease from other Parkinsonian illnesses, Apo morphine injections have been tried subcutaneously, however, they are ineffective and contribute very little to the diagnosis of PD (Brooks et al., 1992).

Despite promising pre-clinical data, many previously proposed medicines have failed clinical trials, underscoring the need of a well-thought-out study plan. Recent advances in our knowledge of the pathogenic processes and anatomical bases of Parkinson's disease (PD) symptoms have opened up new therapeutic options, and it now seems likely that approaches to treating the disease will change considerably in the years ahead. Crediting the recent success of deep learning in medical image classification, this study relies on a similar motive, to detect the disease as early as possible, making optimum use of the convolutional neural network (CNN), a DL based architectural topology (Tagare et al., 2017; Rui et al., 2020).

In the case of making precise judgments based on large datasets, deep neural networks are clearly an asset. The methodology implemented in this research are Deep learning algorithms in which DenseNet 121 performed well when compared to other techniques. All previous layers provide extra input to DenseNet layers, which in turn provide their own feature maps to all following layers. Instead of adding the activations generated by one layer to the activations generated by subsequent levels, the activations are simply concatenated together. As the layers build upon one another, they share a "collective wisdom." In order to maintain some kind of global state, the original inputs and activations from prior levels are retained at each layer (or, to be more accurate, between blocks of layers). A smaller number of parameters for a given depth is the result of this approach, which facilitates the reuse of existing features. Since dense networks can handle smaller datasets, they're especially well-suited to them. Because there are no duplicate feature maps to train when DenseNets are connected in this manner, they need less parameters than a similar classic CNN. Some ResNets versions have also shown that many layers contribute little and may be eliminated. As a result, ResNets have a large number of parameters since each layer has its own weights to learn. A tiny number of new feature maps is all that DenseNets layers do, since they are relatively narrow (e.g., 12 filters). There was also an issue with training in extremely deep networks, due to the flow of information and gradients stated above. Because the gradients from the loss function and the original input picture can be accessed directly by each layer in DenseNets, this problem is alleviated (Ahlrichs and Lawo, 2013; Minja et al., 2019; Latha et al., 2020).

RELATED WORKS

Rumman et al. (2019) proposed an image processing and Artificial Neural Network (ANN) based approach to find the domain of putamen and caudate as the region of interest from SPECT images for detecting PD in its early stage. The region values of the putamen and caudate were then fetched to the ANN classifier for recognition.

Wolfswinkel et al. (2021) developed a convolutional neural network called DaTNet-3 to differentiate and classify normal and PD subjects that underwent the DaTSCAN procedure. They collected the imaging data from Parkinson's Progression Marker Initiative (PPMI) and a hospital-based dataset. Wenzel et al. (2019) explored variable image characteristics at different camera settings using FP-CIT SPECT to train the InceptionV3 CNN model for automated classification. Three image settings: unsmoothed, smoothed, and combination of smooth and unsmoothed were fed into the neural network.

Dhanalakshmi et al. (2019) introduced the role of isosurface to extract and collect only the most relevant features from complex 3D DaTSCAN images. This method was further utilized to implement CNN architectures such as LeNet and ALexNet for PD classification. Martínez-Murcia et al. (2017) performed an exhaustive analysis of DaTSCAN images implementing a voxel-based logistic lasso model. The model helped to define the regional voxels in the caudate, putamen, and globus pallidus area for an informed classification of control and PD categories. Additionally, another ML technique called logistic component analysis was utilized for judging feature differences within the same population or groups.

Ortiz et al. (2019) utilized Alexnet architecture and introduced an image normalization layer to capture the region of interest from SPECT images. The model helps achieve high classification accuracy for classifying PD and control groups. Adams et al. (Subhrajit et al., 2018) performed a quantitative analysis of DAT SPECT imaging by combining the baseline score of DAT image scans with UDPRS_III (motor function scores) as base input parameters. These features, which included motor function and DAT scan scores, were then provided as input to the CNN model for prediction and classification for PD.

Oliveira et al. (Mohammed et al., 2020) assessed certain features that contribute to dopaminergic degeneration for PD using [123I] FP-CIT SPECT brain scans. A total of seven features were calculated and employed for the assessment using ML classifiers that include Support Vector Machine (SVM), K-Nearest Neighbours (KNN), and Logistic Regression (LR). Most accurate results were obtained using the SVM classifier with the seven features.

Martínez-Murcia et al. (Oliveira et al., 2018) developed and proposed a 3-D CNN system for a fast feature diagnosis of PD using SPECT imaging modality. Activation maps were constructed and visualized for practical feature understanding of the network. Shiiba et al. (2020) studied and assessed shape features of SPECT images combined with semi-quantitative parameters to feed into Machine Learning classifiers. The semi-quantitative features and the shape features were used to extract and study the region of interest in classifying PD.

Pianpanit et al. (2021) investigated different model interpretation methods using SPECT images with deep learning model approaches. Techniques like SHAP (Shapley Additive explanations) and guided backpropagation were explored for their attributes and their performance compared for distinguishing between normal and PD subjects.

It is critical to diagnose PD at an early stage, given that the severity of PD and its many phases are essential in determining when to intervene. Many researchers have suggested a model predict and diagnose the disease. Deep learning for image analysis has yielded some of the most impressive progress in recent years (Adams et al., 2018). Various deep learning and machine learning approaches have been used to predict PD in multiple

studies. Imaging modalities including MRI (Magnetic Resonance Imaging) and SPECT for detecting PD have raised high interest in research studies.

Sivaranjini et al. (Adams et al., 2018) attempted to classify PD and healthy control images from MRI modality by following the architecture of the transfer learning model, AlexNet. This model successfully recognizes the structural differences in normal and PD subjects and yields optimal results. Chakraborty et al. (Magesh et al., 2020) carried out data pre-processing for 3T T1-Weighted MRI brain scans and designed a 3D CNN model for extracting complex patterns in the brain images of normal and PD cohorts.

All studies that employed accuracy in model assessment obtained a diagnostic accuracy above chance values for each research. It's possible certain data types may not be generalizable enough to forecast how effectively they may help us discriminate between Parkinson's disease and other Parkinsonian illnesses, but the application of machine learning to many various kinds of data led to great diagnostic accuracy in PD. Data splitting procedures and cross validation were not described despite the great diagnostic accuracy and performance reported in many investigations. When 2D slices are derived from 3D volumes in data modalities like 3D MRI scans, several slices may be created for a single patient. Data leakage and an overestimation of model performance may occur if the same subject is used in the training, validation, and testing sets. This compromises the repeatability of reported findings.

In the proposed work the following approaches are taken as objectives for this study:

1. Taking the PPMI repository's DaTSCAN SPECT images as the adequate dataset for usage in the study.
2. Comparing healthy Vs. PD cohorts contour edge imaging technique.
3. Analyze and train the images on DenseNet-121, CNN topologies with a soft attention block in addition to it, influenced by the works of Martínez-Murcia et al. (2017), Oliveira et al. (2018), and Wolfswinkel et al. (2021).
4. Compare the results with other pre-trained CNN models.
5. Visualize the successful model results using Soft Attention Map and Feature Map representation.
6. Analyze the results using statistical metrics.

The paper is structured as follows: The section 3 explains about the materials and methods followed by section 4 deals with the methodology in detail along with the results and discussion. The session 5 concludes the research work.

MATERIALS AND METHODS

Dataset

The data for this research was acquired from the PPMI public repository, which is a multimodal, prolonged study of radiomic feature observations, neuroimaging, and biological markers in PD patients and healthy controls (HC).

Various industries' scientists, researchers, sponsors, and study populations have continued to work together to build this substantial searchable archive to make PD research and therapies easier and more effective by finding progression biomarkers.

Data Preprocessing

A SPECT scan generates a volumetric image of the basal ganglia. Typically, a collection of axial view planes is then created for clinical evaluation. These picture sequences have been anonymized and exported in PNG format. All images were created using a single slice that is most typical of the basal ganglia's anatomical location.

To begin, the image is first pre-processed and the ROI, the putamen and caudate area, is segmented. The caudate and putamen area segmentation areas are computed and given as features to the Deep learning algorithms. With the help of the training data, the DL algorithms are trained and the prediction model may be utilized to distinguish PD patients from healthy ones. The deep learning algorithms implemented in this research work are DenseNet 121, Xception, ResNet 50, MobileNet V2, Inception ResNet V2, ResNet 152V2, EfficientNet B1. The **Figure 1** depicts the suggested procedure using SPECT pictures from the PPMI database.

Contour Edge Detection

Edge detection is a technique for detecting the borders of objects in images. It detects brightness disparities in image processing, computer vision, and machine vision fields. Edge detection is used to extract images and data. Edge detection are essential in computer vision since they involve identifying and classifying objects in images.

In Parkinson's disease, nigrostriatal loss is often disproportionate, with greater degradation observed in the putamen relative to the caudate nucleus. Corresponding with Parkinson's disease are aberrant appearances such as symmetrical loss of uptake in both putamen and total loss in absorption in the caudate and putamen despite usual functioning. **Figure 1** uses canny-edge detection to highlight the putamen and caudate region in normal and healthy cohorts. Finding a closed form and drawing the object's border is the primary goal of contour detection as shown in **Figure 1**. It is also possible to employ contour detection to estimate an object's form based on such attributes as its aspect ratio, length, and solidity. The images are accustomed to working with grayscale images, the first step is to transform the image to gray. To approximate contours, a simple threshold is utilized. Smoother contours may be achieved by using the OPEN and CLOSE technique. A list of contours is obtained and the final contours are sketched on the color.

Data Augmentation

Figures 2, **3** depict the data augmentation that was performed on both the subtypes, Health Control and PD patients. Data augmentation was adopted to correct the balance of the dataset due to its moderate size and the small number of HC participants engaged in the screening procedure. The description of the eight different type of augmentations is mentioned in the **Table 1**. The eight augmentation techniques were so chosen

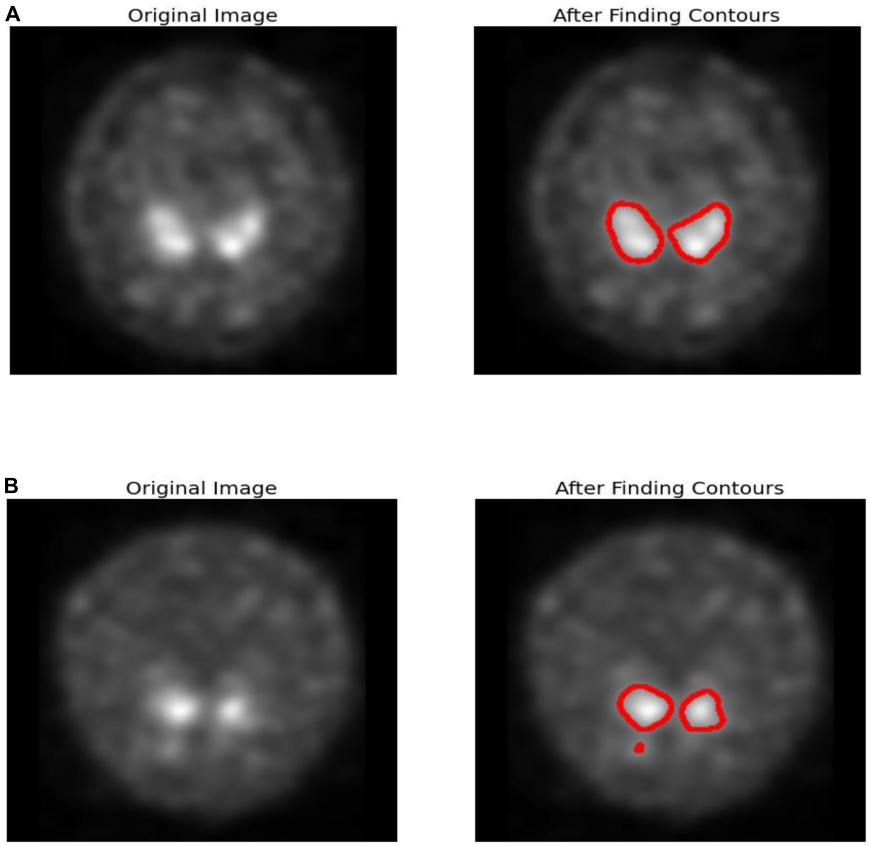


FIGURE 1 | Edge detection **(A)** healthy subject **(B)** PD subject.

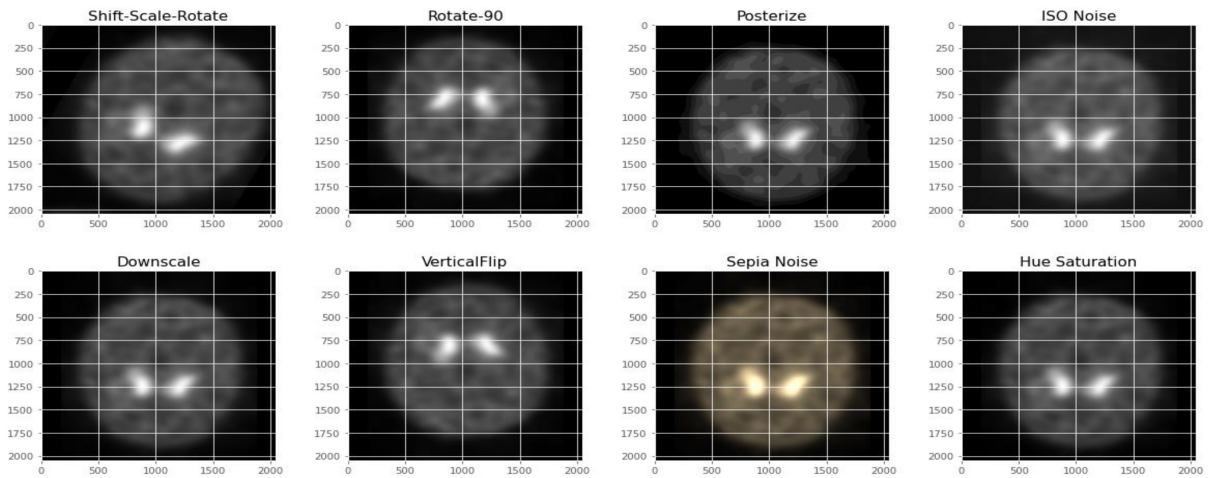


FIGURE 2 | Data augmentation for healthy control participants.

manually that the model extracts different spatial representations throughout the dataset in order to ensure that the final models function effectively in the event of several limitations like over fitting conditions. After data augmentation, 1,840 images of HC participants and 2,002 images of PD patients are generated. The imaging data selected for this study includes 1,390 DaTscan SPECT images, which are split into two classes: PD (with 1,160 images) and Healthy Controls (with 230 images), as shown in **Table 2**. The PPMI imaging support validated the diagnosis of PD by confirming that the screening DaT-SPECT (123I FP-CIT) is

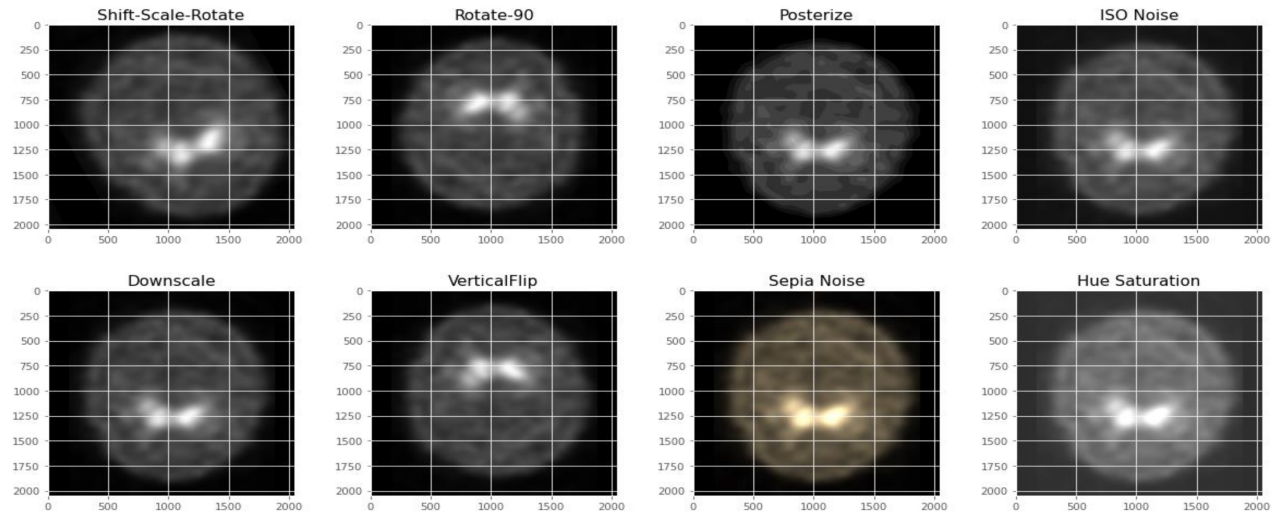


FIGURE 3 | Data augmentation for PD participants.

TABLE 1 | Image augmentation types and its description.

Augmentation type	Description
Shift Scale Rotate	Shifts along the x/y axis, scale (zoom in/out) and rotates on a random value
Rotate-90	Rotate by 90 degrees.
Posterize	Reduces the number of bits for each color channel
ISO Noise	Random sensor noise or Gaussian noise.
Downscale	Reduces the overall resolution of the image
Vertical Flip	Flip the input vertically around the x-axis.
Sepia Noise	Sepia filter is added randomly
Hue Saturation	Adds hue saturation to the image

TABLE 2 | Dataset of DaTscan SPECT images.

Study class	No. of subjects	Average age	
		Male	Female
PD	1160	63.68 ±	62.35 ±
CN	230	61.87 ±	59.26 ±

associated with a DaT deficiency. The network is divided into two segments during the training phase. The augmentation network uses two images of the same kind as the input image and outputs a layer with the same resolution as the input image. This layer is used to create an "enhanced" version of the original picture. Finally, the enhanced picture is sent into a second network, which is called a classification network. An end-of-network drop in classification accuracy is due to a loss in cross entropy on class sigmoid. End-to-end, an addition loss is calculated to control how well augmented images match their input counterparts. As a result, the total loss is the sum of these two losses.

The overall block diagram of diagnosis of PD using SPECT images is shown in **Figure 4**. The SPECT image is pre-processed, augmented and classified using various Deep learning algorithms. The performance metrics are compared for with and without augmented images. The metrics are accuracy, sensitivity, specificity, precision and F1 score. The SPECT quantitation of a given picture feature is affected by a wide range of physical parameters, but three stand out: attenuation, scatter and detector response (or finite spatial resolution limited by the collimator). As the image feature size falls, detector response, or limited spatial resolution, becomes more critical in SPECT quantification.

The detected activity concentration drops with the volume of features smaller than nearly twice the detector's spatial resolution. This is because the SPECT image's count values are dispersed across a broader area than the emission source itself. As a result, the actual concentration is lowered. If a source is big enough, the dispersion of counts away from that source is counterbalanced to a greater extent than with smaller sources. Linear deconvolution filtering, such as Wiener or Metz filters, may be used to adjust detector response. The detector response will become blurry if the filter gain is greater than unity at low spatial frequencies. High-frequency picture noise may be controlled by "rolling-off" the filter to zero gain. The need for a model to display what location it is attending to while making a decision/prediction in deep learning. Various attention mechanisms have been introduced in the past years. With the development of automated pattern learning mechanisms, particularly models that can be trained to focus on specific regions, it is now possible to focus on critical areas for attention.

METHODOLOGY

DenseNet Architecture

The DenseNet (densely connected convolutional network) is recognized for having convolutional neural network architecture that is state-of-art, when validated for classification using the popular ImageNet dataset. Huang et al. validated the technique of using direct connections in a feed-forward manner from each

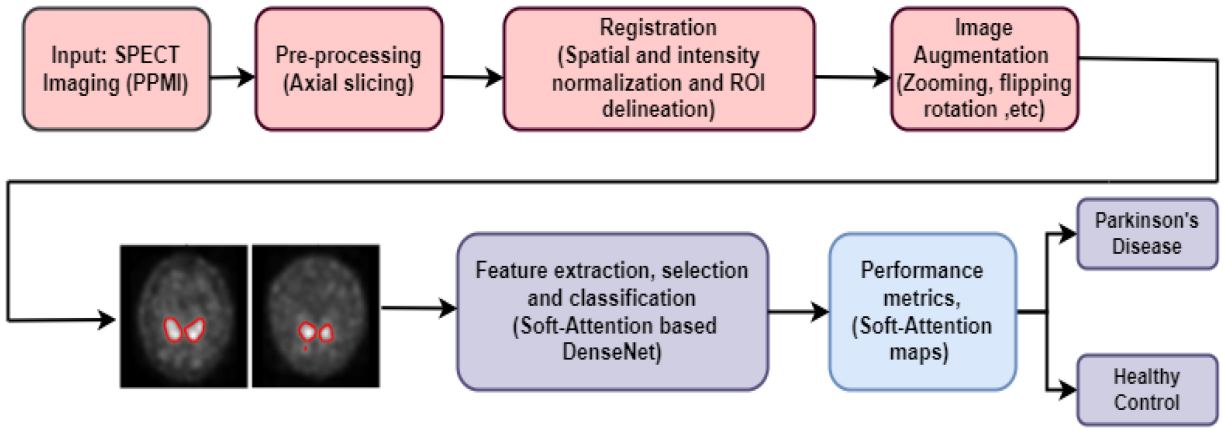


FIGURE 4 | Overall block diagram of the proposed work.

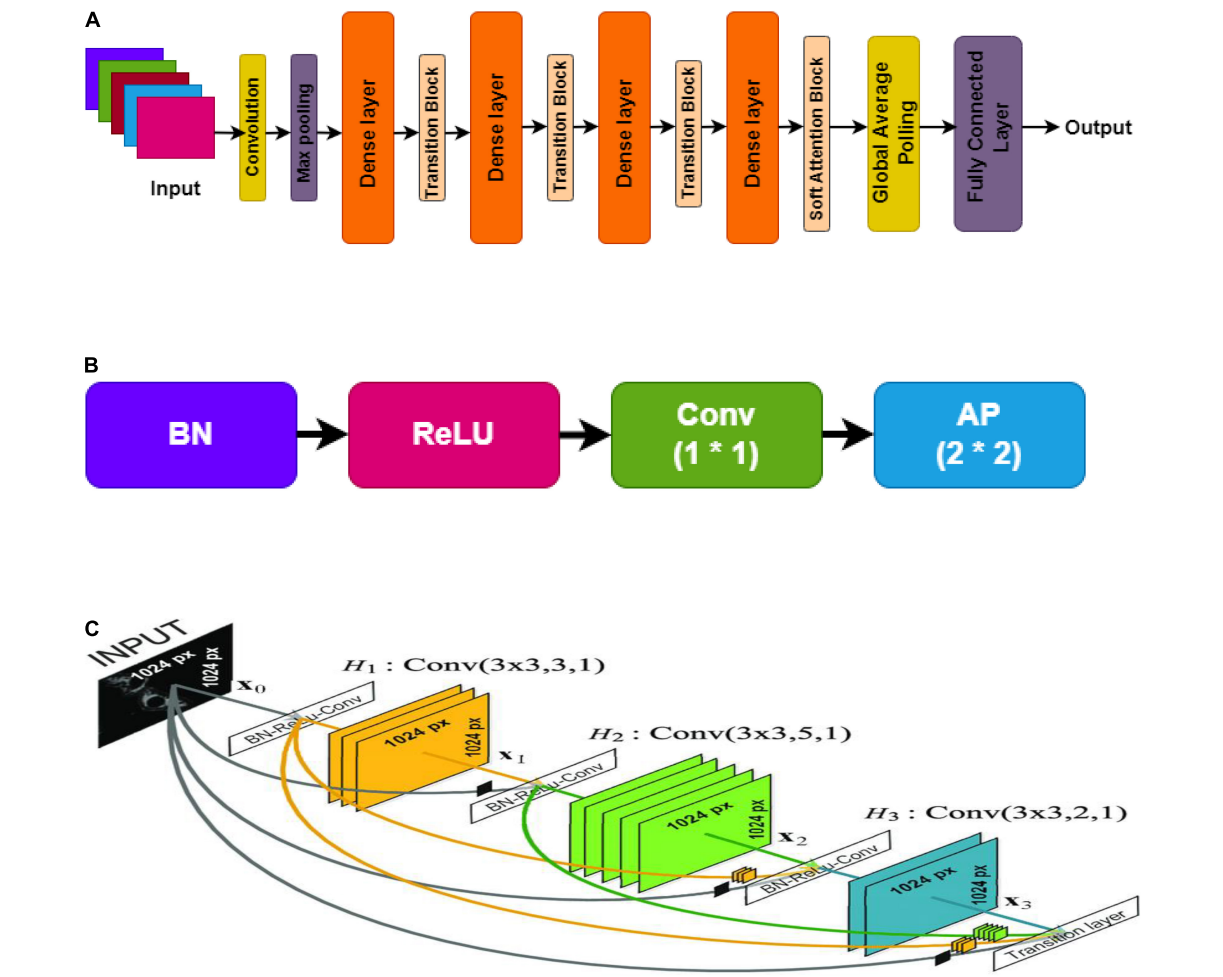


FIGURE 5 | DenseNet, **(A)** overall architecture; **(B)** transition layer; **(C)** dense layer.

layer to every other layer. Every layer in the model architecture takes the target input and concatenation of the preceding layers' feature maps. It performs non-linear operations such as batch normalization, ReLU, and convolution or pooling. The resultant feature maps of each layer are provided as inputs to the succeeding connected layers after the non-linear function's

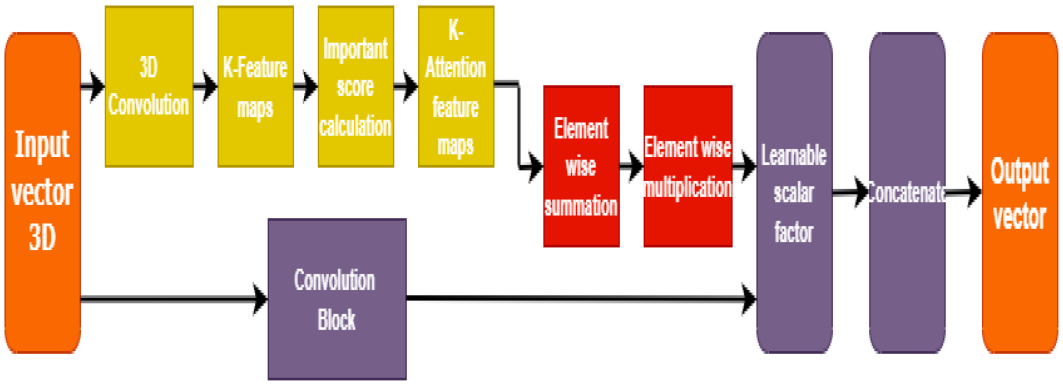


FIGURE 6 | Soft attention mechanism.

TABLE 3 | DenseNet121 architecture with Soft Attention Block for classification of PD and HC cases.

Layers	Output shape	Kernel size and details
Convolution 2D	112 × 112	7 × 7 conv, stride 2 (Rui et al., 2019)
Max Pooling 2D	56 × 56	3 × 3 max−pool,stride 2
Dense Block (Pahuja et al., 2019)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (Pahuja et al., 2019)	56 × 56	1 × 1 conv
	28 × 28	2 × 2 average pool,stride 2
Dense Block (Prediger et al., 2014)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (Prediger et al., 2014)	28 × 28	1 × 1 conv
	14 × 14	2 × 2 average pool,stride 2
Dense Block (Blesa et al., 2015)	14 × 14	$\begin{bmatrix} 1 \times 1\text{conv} \\ 3 \times 3\text{conv} \end{bmatrix} \times 24$
Transition Layer (Blesa et al., 2015)	14 × 14	1 × 1 conv
	7 × 7	2 × 2 average pool,stride 2
Dense Block (Zhou et al., 2009)	7 × 7	$\begin{bmatrix} 1 \times 1\text{conv} \\ 3 \times 3\text{conv} \end{bmatrix} \times 16$
Soft Attention Block	7 × 7	Soft Attention × 1
Classification	1 × 1	7 × 7 global average pool
Layer	2	Fully Connected Dense Layer, Softmax

computation. If the size of the feature maps changes, the concatenation procedure is unsuccessful. Hence, the need for pooling operation is crucial when the size of the feature maps varies (Chakraborty et al., 2020).

The architecture is organized into distinct blocks, i.e., Dense blocks (densely connected) to assist in the pooling process. The layers between dense blocks are transition layers that conduct the tasks of convolution, batch normalization, and pooling. On an average note, each function generates K unique feature maps, a hyper-parameter known as the growth rate which determines the number of feature maps each layer delivers to the network (Sivaranjini and Sujatha, 2020). Once updated, the feature maps may be viewed throughout the network. Unlike other traditional CNN models, this also waives the need to reproduce one layer to another.

Each layer in the network reproduces k feature maps and causes many parameter inputs as shown in **Figure 5**. As a solution, to limit the number of input feature maps to 4k, a [1 × 1] size of convolution was employed in the bottleneck layer. Thus, minimizing the amount of feature mappings at transition layers is another optimal feature of DenseNet. The number of feature maps in a dense block with n feature maps results in θn, later in the transition layer, lying in the factor range of $0 < \theta \leq 1$, known as the compression factor (Gao et al., 2019). DenseNet's design provides many advantages in addition to network compactness: it overcomes the vanishing gradient problem, optimizes feature transfer, and minimizes the rate of parameters. DenseNet121 network architecture was utilized in this paper.

Soft Attention Block

SPECT DaTSCAN method helps in distinguishing between Parkinson's and control subjects by helping to visualize the basal ganglia region. The dopamine transporters: putamen and caudate regions are reported to get smaller in size due to the loss of dopaminergic neurons in the case of PD patients. Soft attention can be a useful idea to detect the region in the image where minor to significant distortion is found, which is considered an abnormality and needs further analysis.

Soft attention takes the robust approach of promoting the most relevant input (in this case, pixels in an image) while still allowing a subset of the other information to contribute to the model's decision-making as shown in **Figure 6**. It is taking inspiration from the good works of Martínez-Murcia et al. (2017) and Dhanalakshmi et al. (2019), a soft attention block that utilizes 3D-convolution to attend and identify the essential features responsible for classification.

This way, the high-level features are first extracted, and the resultant convolution of 'K' kernels generates a feature map (having K attention heads). This feature map is further

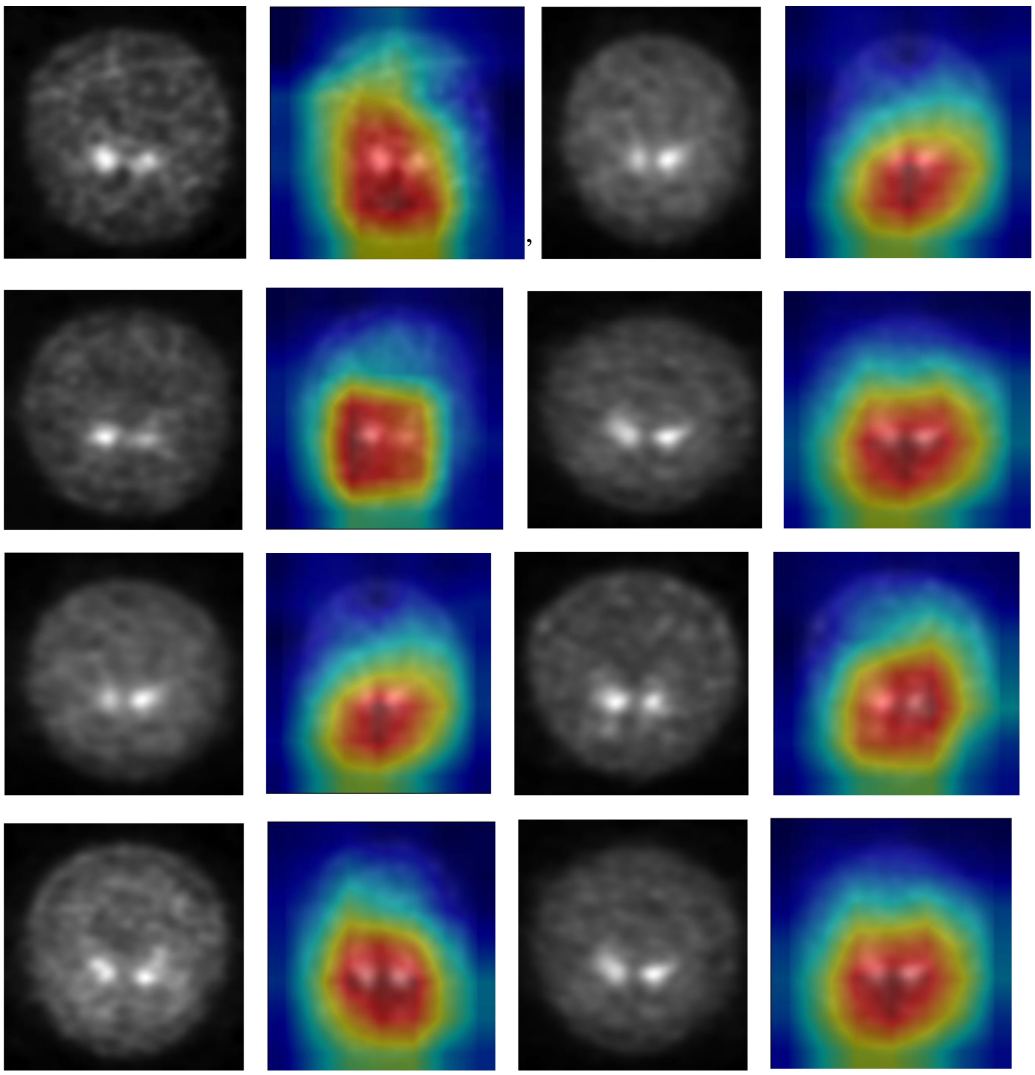


FIGURE 7 | Soft attention map visualisation of PD patients with the ROI highlighted.

normalized to calculate the importance score (or *soft attention* score) based on the appropriate location from the extracted feature map. A 3-D tensor $f^x \in \mathbb{R}^{h^x \times w^x \times d^x}$ having a 3-D kernel of size $H \in 3\ x\ 3\ x\ n^x$ is input to the convolutional layer that results in a feature map$f_{3d} \in \mathbb{R}^{h^x \times w^x \times 1}$.

Such K kernels, work as attention points to generate $f_{3d} \in \mathbb{R}^{h^x \times w^x \times K}$. These feature maps are further normalized and averaged to be retained for calculating the importance score (for dominant locations) or the soft attention score. The equation for soft attention score denoted by "*S*" is given below:

$$S = \sum_{k=1}^{K} \frac{\exp\left(f_{3d_{ij}}\right)}{\sum_{i=1}^{w^x} \sum_{j=1}^{h^x} \exp\left(f_{3d_{ij}}\right)}; \text{where } f_{3d} = H\left(f^x\right) \quad (1)$$

The resultant f^x tensor is thereby multiplied with the soft attention score S, so the value becomesf_s^x. A learnable scalar is assigned to compute the weights, which in this case is y with a value of 0.01. The total weight is calculated by the equation given below:

$$\alpha_y = f^x + y f_s^x \quad (2)$$

The finalized soft attention layer helps the model decide the specific locations of the feature map that has important attributes on the whole.

Figure 5 shows the soft attention maps.

Convolution and pooling are the foundation of DenseNet. In order to get to the classification layer, there are four more dense blocks followed by transition layers. After that comes a dense block followed by yet another transition layer. The DenseNet121 architecture with Soft Attention Block for classification of PD and HC cases is shown in **Table 3**.

The stride is 2 and the first convolution block comprises 64 filters of size 7 × 7. After that, there's a MaxPool layer with 3 × 3 max pooling and a stride of 2. ReLU activation and the real

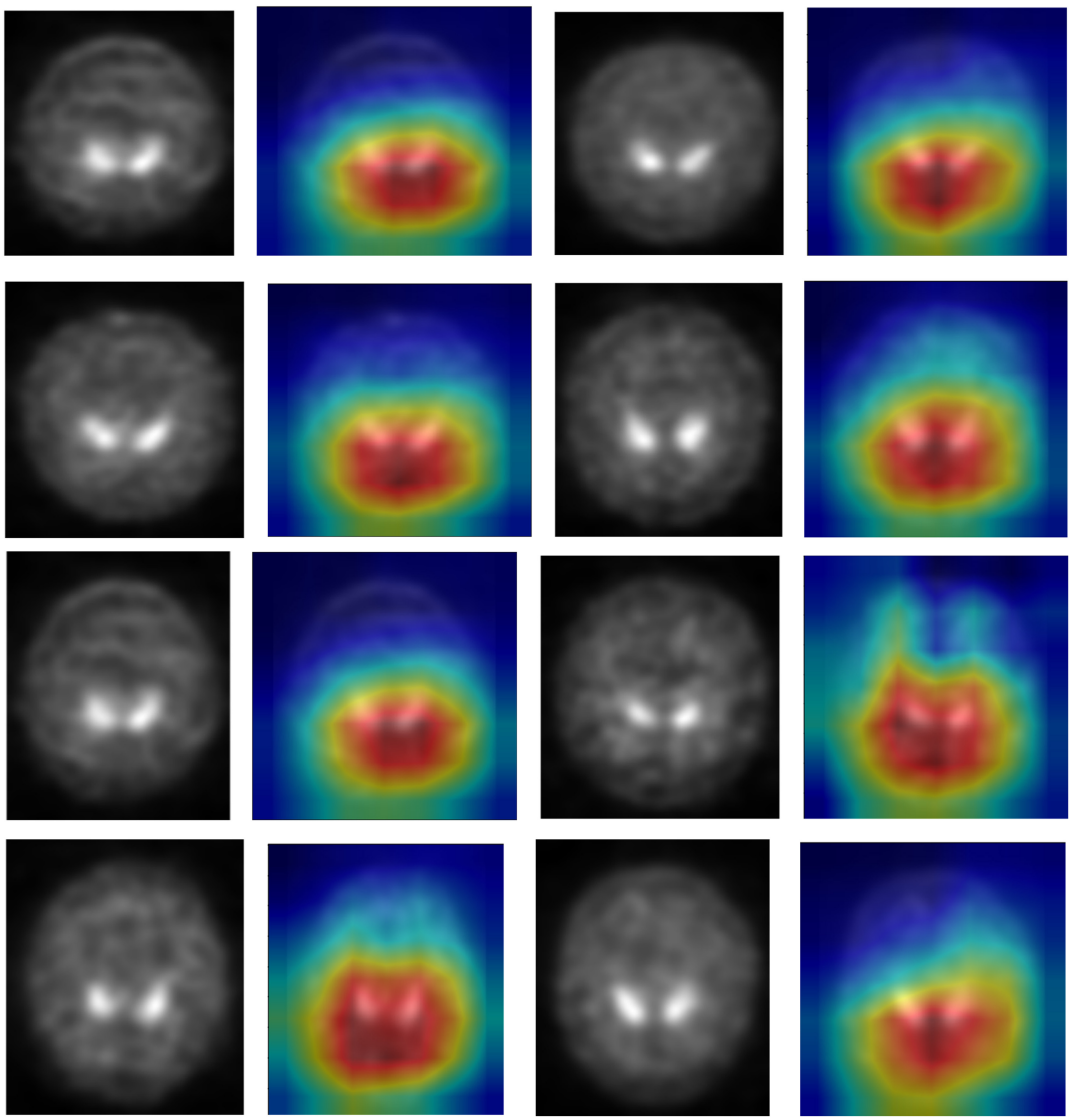


FIGURE 8 | Soft attention map visualisation of HC patients with the ROI highlighted.

Conv2D layer follow BatchNormalization in every convolutional block. In **Table 3**, convolutions with 1 × 1 and 3 × 3 kernel sizes are used in each dense block. This is repeated six times in dense block 1, twelve times in dense block 2, twenty-four times in dense block 3, and ultimately sixteen times in dense block 4. Each 1 × 1 convolution has four times the number of filters in dense block. As a result, 4 filters are employed, yet only 3 of those filters are ever used. In addition, the input tensor and the output tensor must be joined.

The number of channels in the transition layer is to be reduced to half of the current channels. An average pool layer with a stride of two is used in conjunction with a 1 × 1 convolutional layer. bn rl conv already has a kernel size of 1 × 1, therefore we don't need to declare it again.

Half of the channels in the transition layers must be removed. To figure out how many channels there are, we need to acquire half of the input tensor x. As a result, we may utilize Keras backend (K) to produce a tuple with the dimension of x when given a tensor x. For our purposes, we simply need to know how many filters there are in that form. So [−1] is added. This number of filters may be divided by two to reach the desired result.

The dense blocks and transition layers have now been defined. The thick blocks and transition layers must now be stacked one on top of the other. Since the repetitions are 6,12,24,16 we build a "for loop" to go through them. In this way, the loop is executed four times, each time with a different number from the range of 6, 12, 24, or 16. The dense blocks and transition layers are now complete.

There is a final output layer, then Global Average Pooling. Following Dense Block 4, there is no transition layer between Dense Blocks 3 and 4, but it goes straight into the Classification Layer after Dense Block 4. Global Average Pooling is used on the

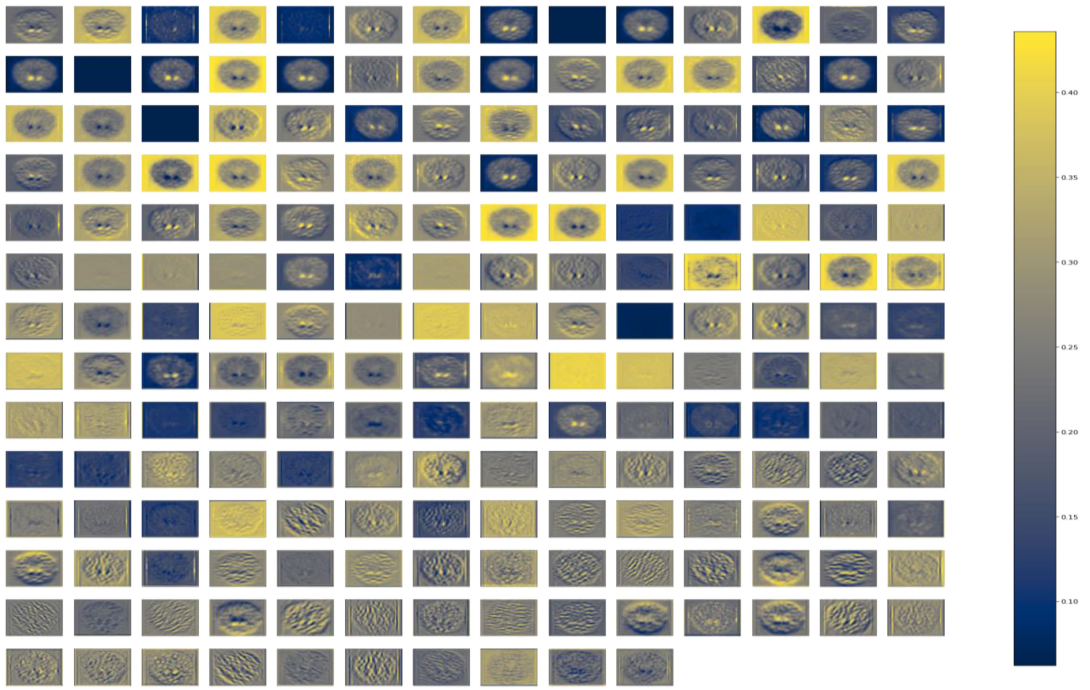


FIGURE 9 | Dense Block of DenseNet121 with batch normalization operation.

connection 'd,' not the one on 'x,' as was previously stated. To eliminate the for loop from the above code and stack the levels one after the other without a transition layer is another option.

RESULTS AND DISCUSSION

Visual Assessment

DenseNet is compared with the other deep earning algorithms such as ResNet, Inception ResNet, Xception, MobileNet, and EfficientNet V2. The vanishing gradient issue was solved by introducing the idea of residual connections in ResNet V2. Inception ResNet, a ResNet version that employs several size kernels inside the same layer, is utilized since it is difficult to select on a ResNet kernel size. For example, Xception proposed the notion of depth-wise separable convolution in order to minimize the number of parameters without compromising performance. There are now between 100 and 1,000 less parameters since MobileNet has included point wise convolution in addition to depth wise convolution. The Soft Map visualization for PD and healthy controls are shown in **Figures 7**, **8**.In soft attention, instead of utilizing the image x as an input, we use weighted image characteristics compensated for attention in soft attention. The areas of the image that get the most attention seem brighter. The weighted characteristics of the DL algorithms, as well as the PD and normal it predicted, are shown in the image above. The low weight of the feature map multiplied by the soft focus discredits places that aren't significant. As a result, regions with high levels of attention retain their original worth while those with low levels of attention approach zero (become dark in the visualization). With the "PD and normal patient," the attention module creates a new feature map with all areas darkened except the region of interest area.

In order to know the working of the model, and how it defines the ROI through its various layers, feature maps visualization may be useful. **Figure 8** shows layers inside the dense block, including batch normalization and ReLU activation, and how they influence in the overall classification work of the model. These feature maps highlighted help in conclusively deciding the putamen and caudate regions taking the major role in predicting the desired class.

The **Figure 9** represents the DenseNet 121 algorithm implementing Batch normalization operation. Normalizing network activations over a mini-batch of a certain size is what batch normalization is all about. It is possible to normalize a mini-batch of data by computing the mean and variance for each characteristic. To get the feature's standard deviation, remove the mean and divide the feature by the mini-batch standard deviation. Batch normalization enhances the model's training speed by smoothing the loss function and improving the model's parameters. Poorly initialized neural networks are addressed by batch normalization. Pre-processing may be done at every level of the network, according to this interpretation. At the start of training, it compels the activations in a network to take on a unit Gaussian distribution. One of the most often used techniques for training deep neural networks (DNNs) is Batch Normalization (BatchNorm). The gradients are more predictable and steady as a result of this smoothness, making it easier to train.

Internal covariate shift is no longer an issue. This ensures that each layer's input is spread around a common mean and standard

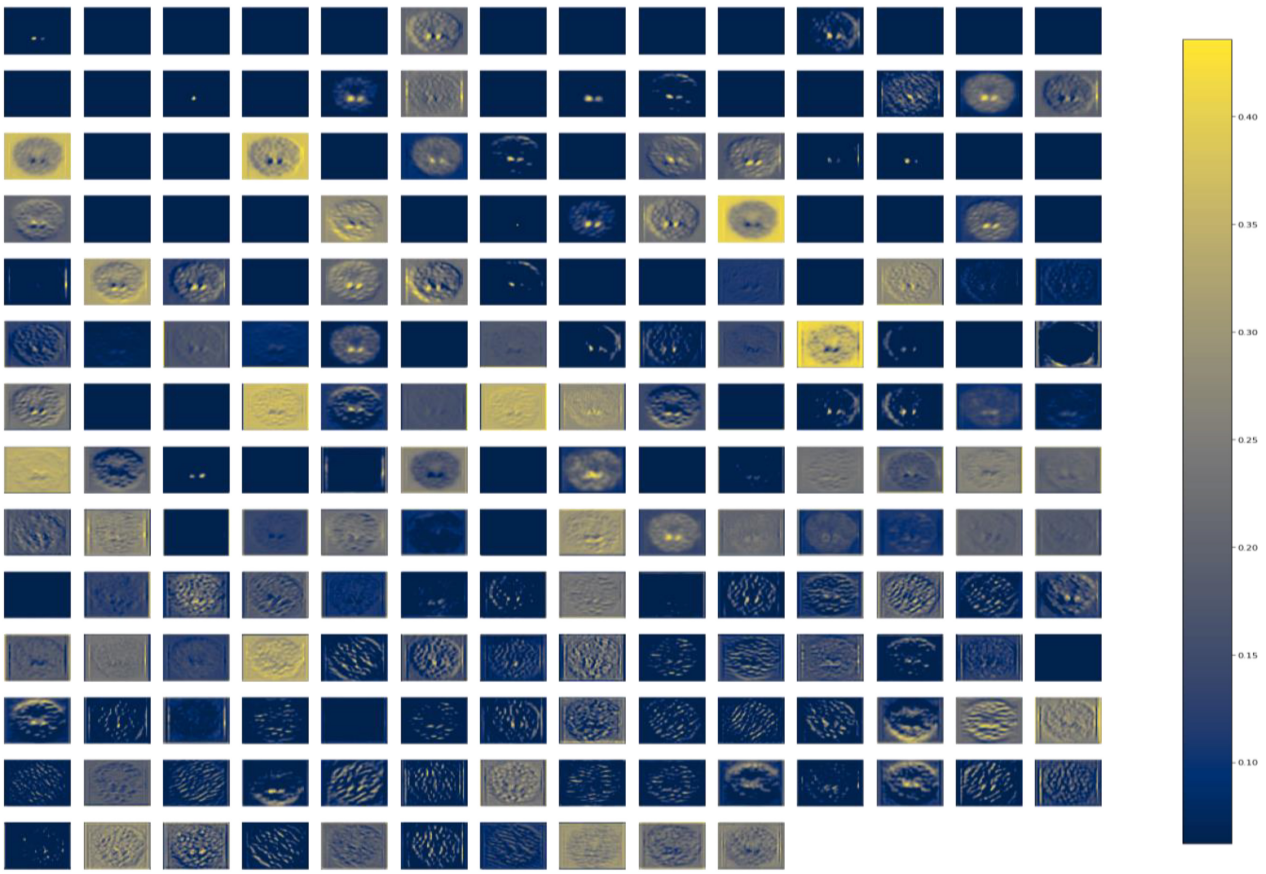


FIGURE 10 | Dense Block of DenseNet121 with ReLU activation.

deviation. Assume for a moment that we're training an image classification model that sorts photos into one of two groups: PD or NPD. If we just have photographs of PD, these images will also be distributed in a certain manner. The model's settings will be updated as a result of using these photographs. if we get a fresh batch of photos from people who are not diagnosed with PD. As a result of this change, the distribution of these new photos will be somewhat altered. Using these fresh photos, the model will adjust its parameters. As a result, the distribution of the concealed activation will shift as well. It's called an internal covariate shift, and it's a change in the concealed activation. Batch normalization enhances the model's training speed by smoothing the loss function and improving the model's parameters.

Figure 10 represents the DenseNet 121 algorithm implementing ReLU activation function. It is possible to increase the learning pace of deep neural networks by using ReLU activation functions in the hidden layers. Deep neural networks now employ the rectified linear unit as their typical activation function. Using ReLU activation function, the vanishing gradient issue is avoided. This is the reason why the deep neural network's learning speed can be improved by activating ReLU. As a result of avoiding the need to do exponential and division computations, employing rectified linear units speeds up computations significantly. Squeezing values from 0 to the maximum imparts sparsity into the hidden units, another ReLU feature. ReLUs may readily overfit when compared to sigmoid functions, although the dropout approach has been used to mitigate this problem, and deep neural networks with corrected networks have shown enhanced performance. Because of its simplicity and dependability, the ReLU and its derivatives have been included into several deep learning systems.

Quantitative Assessment

The **Figure 11A** shows the validation accuracy plot having an accuracy of 99.2% and the validation loss is shown in **Figure 11B**. An AUC of 99% is achieved for DenseNet 121 architecture. Classification methods rely on the AUC-ROC statistic to gauge their effectiveness is shown in **Figure 11C**.

Figure 12 shows the performance metrics for DenseNet 121 implemented for with and without augmentation. The accuracy with augmented images is better than without augmented images. The AUC-ROC measure gives us a good idea of a model's ability to differentiate between different classes. The more AUC a model has, the better it is judged to be. For every conceivable cut-off for a test or combination of tests, AUC-ROC curves are widely used to illustrate the relationship and trade-off between sensitivity and specificity. The accuracy for with and without implementing soft attention Map visualization is 95% and 99.62% is achieved.

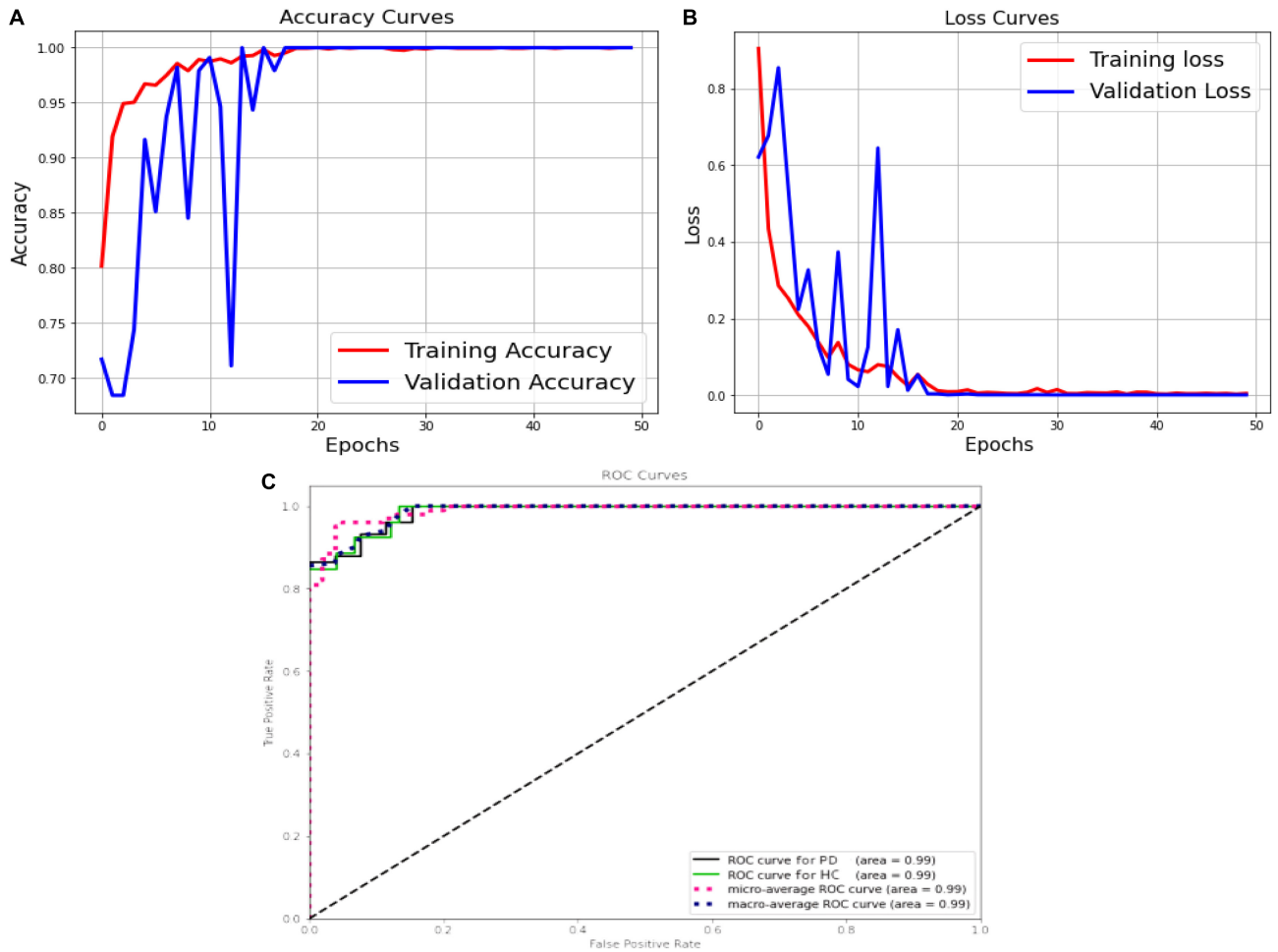


FIGURE 11 | DenseNet, **(A)** accuracy; **(B)** loss curve; **(C)** ROC curve.

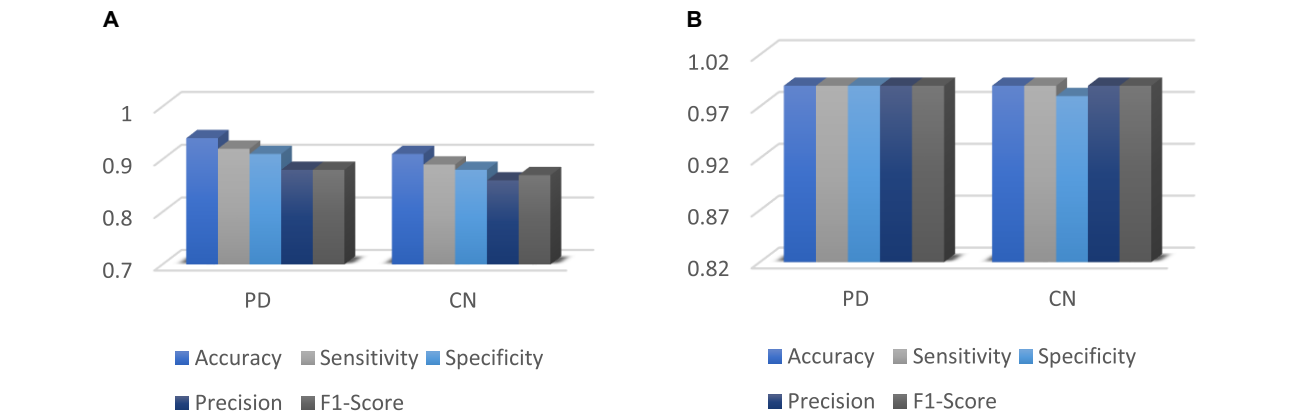


FIGURE 12 | (A) Without augmented images; **(B)** with augmented images.

The ROC curve's area under the curve provides an indication of the test's value in answering the underlying issue. At different threshold values, AUC—ROC curves may also be used as a performance evaluation. Using the AUC-ROC to assess the performance of a classification model is vital. When a model's accuracy is improved *via* the use of this test, its value and correctness are both increased. In classification issues, the true positive rate and the predictive value of a predictive model may be

TABLE 4 | Overall accuracy of DenseNet 121 for PD and control subject.

	Without augmented images					With augmented images				
Class	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**F1-Score**	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**F1-Score**
PD	94%	92%	91%	88%	88%	99%	99%	99%	99%	99%
CN	91%	89%	88%	86%	87%	99%	99%	98%	99%	99%
Overall Score	92.5%	90.5%	89.5%	87%	87.5%	99%	99%	98.5%	99%	99%

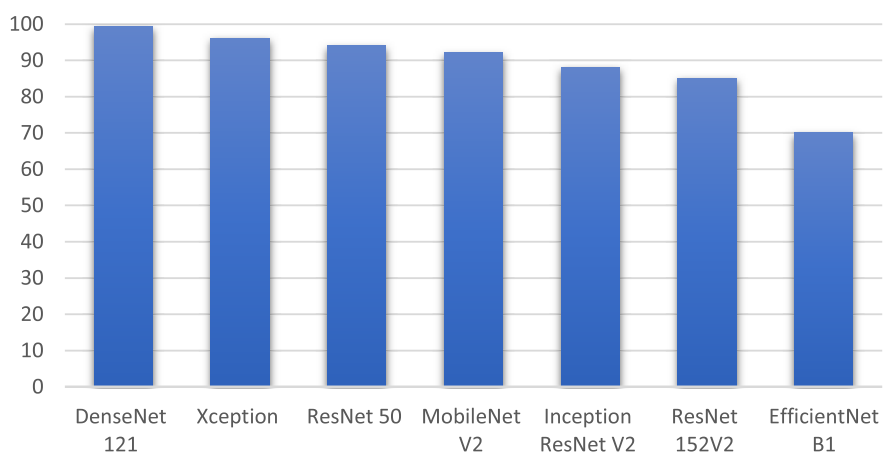


FIGURE 13 | Comparison of DenseNet using soft attention with other deep learning algorithms.

TABLE 5 | Comparison of various deep learning models.

Deep learning algorithms	**Computational time**
Xception	1.5 min
ResNet 50	2.25 min
MobileNet V2	1.7 min
Inception ResNet V2	1.75 min
ResNet 152V2	3.3 min
EfficientNet B1	3.2 min
DenseNet 121	2.15 min

summarized using this technique, which helps us understand the trade-off between the two. **Table 4** compares the augmented and non-augmented on the scales of accuracy, sensitivity, specificity, precision and accuracy.

DenseNet 121 gave an improvement of 3.2, 5.2, 7.2, 11.2, 14.2, and 29.2% in accuracy when compared to other deep learning algorithms such as Xception, ResNet 50, MobileNet V2, Inception ResNet V2, ResNet 152V2, and EfficientNet B1 as shown if **Figure 13**. DenseNet 121 solves the vanishing-gradient issue and encouraging feature reuse. DenseNet also decrease the number of parameters which yields an increase in accuracy.

In **Table 5** the comparison of various deep learning techniques based on the computational time is listed. DenseNet 121 has the computational time of 2.15 min with an accuracy of 99.2%.

Many researchers have expanded their horizons by employing numerous deep learning frameworks to detect PD from normal and other disease categories, which might be used in future analysis and examinations. Magesh et al. (Adams et al., 2018) used transfer learning (with VGG16 as the leading model architecture) for classifying PD from normal groups. Local Interpretable Model-Agnostic Explainer (LIME) was selected as an illustratable method to find the region of interest to analyze PD and normal group DaTscan images. LIME proved to be vital substitute for explainable-AI based diagnosis to be used instead of Grad-CAM and saliency mapping representation.

Chien et al. (2020) reflected the focus on putamen and caudate region (from SPECT images) and demonstrated the use of artificial neural network for detecting PD and parkinsonism caused by other disorders. Sensitivity and specificity of 81.8% and 88.6% were achieved, though these research classes could be further investigated for study.

Nazari et al. (2022) tested layer-wise relevance propagation (LRP) based CNN for classification of normal and reduced patients. The study achieved a sensitivity and specificity of 92.8% and 98.7% respectively. Relevance maps were plotted which could be further investigated for clarity.

Chen et al. (2021) put focus on striatum scanning and implemented a CNN model based on attenuation correction. Monte-carlo based simulation results were drawn for a clearer visual assessment based on voxel-wise, patch-wise and image-wise imaging methods. Although computationally expensive, this strategy showed promise as a substitute in clinical scenario.

Tufail et al. (2021) attempted at developing 3D-CNN to extract attributes of Alzheimer's Disease (AD), PD and normal classes from both PET and SPECT images. This multi-classification based experiment helped in establishing the relationship between AD and PD patients. This study reflects that 3D CNN models at relatively cheaper cost in computational levels could be thus developed for voxel based understanding of 3D SPECT images with explainable-AI based techniques implemented.

Leung et al. (2021) developed an approach by evaluating based on mean absolute error (MAE) and mean absolute percentage error (MAPE) for outcome prediction. The study was based on a three staged ensemble method to reveal spatiotemporal attributes, to demonstrate the connection between imaging and non-imaging information, for predictions based on motor outcomes.

The limitations of this study are, the selected features have been specifically tailored to the diagnosis of Parkinson's disease. In people with Parkinson's disease (PD), the illness advances in a predictable manner: First, the putamen on the side of the patient's clinical symptoms begins to decline in DaT concentration, and subsequently the caudate. Striatum on DaTSCAN loses its comma-shape and becomes dot-shaped or vanishes completely when this occurs. It is possible to "force" a pre-defined region into the form of an exclamation point, resulting in a semi-quantitative metric that is still high but solely represents caudate binding, and therefore does not account for putamen dysfunction. When training, the ReLU may become unstable, resulting in the death of certain gradients. This is a serious drawback. So some neurons die and the weight updates don't activate in subsequent data points, preventing learning since dead neurons offer zero activation.

CONCLUSION

This work demonstrates that significant clinical examination performance may well be attained utilizing deep learning for SPECT scan interpretation and analyses. In order to accurately diagnose PD, it may be necessary to use DaTscan imaging to evaluate pathophysiological changes. Although our method allows to describe and visualize normal and PD cohorts relatively explicitly, of DaTscan SPECT images, using soft attention maps, it cannot be used for clinically analyzing the motor outcomes from SPECT images. Instead of using Gradient-weighted Class Activation Mapping (Grad-CAM), soft attention mapping is used which is cost-effective.

This is possible even with a modest number of participants by exploiting the strength of huge pre-trained neural networks through the transfer learning process along with manual addition of soft-attention block, as was done with DenseNet architecture in this study. The necessity for an end-to-end 3D CNN architecture should also be noted for future study. There were five CNN models employed in comparison with our CNN model: DenseNet 121, Xception 50, Resnet 50, Mobilenet V2, Inception ResNet V2, and EfficientNet B1. An AUC of 99% and an accuracy of 99.2 % are achieved in this system, compared to previously suggested methods. Further DenseNet-121 with the soft attention block retains features with low level of complexity.

This was a semi-automatic diagnostic process, not an entirely automated diagnosis monitoring system. Utilizing the complete scan volume rather than just a single slice may prove to be a rewarding topic of future study. If the transfer learning process is to be employed, this would need the usage of a properly pre-trained 3-D convolutional neural network. Though the simulation findings and study are intriguing, they can be corroborated with much bigger datasets. It is imperative to deal with the requirement for an end-to-end 3-D CNN model that can retrieve relevant features from the 3-D SPECT image data itself for improved clearer outcomes in future. This mandates that we continue to make progress on our research in deep learning and explainable-AI methods on Parkinsonism and related disorders.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

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.

AUTHOR CONTRIBUTIONS

HK was responsible for ideation, technique and design, data gathering and visualization, formal analysis, and reviewing and editing. MT conceptualized the project, developed the technique, designed the study, gathered and analyzed data, and then wrote the first draft. KL gathered and analyzed data as well as to write, evaluate, and edit the final product. SD and XW contributed to the formative stages of analysis and revision. All authors wrote the manuscript and gave their blessing to the final version that was submitted.

FUNDING

This work was supported by Universiti Malaya, and ACU United Kingdom under project number IF 063-2021.

ACKNOWLEDGMENTS

We would like to thank the SRM Institute of Science and Technology, the University of Malaya and Xuzhou Medical University Xuzhou for supporting this research.

REFERENCES

Adams, M., Yang, B., Rahmim, A., and Tang, J. (2018). "Prediction of outcome in Parkinson's disease patients from DAT SPECT images using a convolutional neural network," in *Proceedings of the 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, Sydney, NSW, 1–4. doi: 10.1109/NSSMIC.2018.8824369

Ahlrichs, C., and Lawo, M. (2013). Parkinson's disease motor symptoms in machine learning: a review. *Health Inf.* 2, 1–18. doi: 10.5121/hiij.2013.2401

Bevilacqua, R., Maranesi, E., Di Rosa, M., Luzi, R., Casoni, E., Rinaldi, N., et al. (2020). Rehabilitation of older people with Parkinson's disease: an innovative protocol for RCT study to evaluate the potential of robotic-based technologies. *BMC Neurol.* 20:186. doi: 10.1186/s12883-020-01759-4

Blesa, J., Trigo-Damas, I., Quiroga-Varela, A., and Jackson-Lewis, V. R. (2015). Oxidative stress and Parkinson's disease. *Front. Neuroanat.* 9:91. doi: 10.3389/fnana.2015.00091

Brooks, D. J., Ibanez, V., Sawle, G. V., Playford, E. D., Quinn, N., Mathias, C. J., et al. (1992). Striatal D2 receptor status in patients with Parkinson's disease, striatonigral degeneration, and progressive supranuclear palsy, measured with 11 C-raclopride and positron emission tomography. *Ann. Neurol.* 31, 184–192. doi: 10.1002/ana.410310209

Chakraborty, S., Aich, S., and Hee-Cheol, K. (2020). Detection of Parkinson's disease from 3T T1 weighted MRI scans using 3D convolutional neural network. *Diagnostics* 10:402. doi: 10.3390/diagnostics10060402

Chen, Y., Goorden, M., and Beekman, F. (2021). Convolutional neural network based attenuation correction for 123 I-FP-CIT SPECT with focused striatum imaging. *Phys. Med. Biol.* 66:195007. doi: 10.1088/1361-6560/ac2470

Chien, C. Y., Hsu, S. W., Lee, T. L., Sung, P. S., and Lin, C. C. (2020). Using artificial neural network to discriminate Parkinson's disease from other parkinsonisms by focusing on putamen of dopamine transporter SPECT images. *Biomedicines* 9:12. doi: 10.3390/biomedicines9010012

Dhanalakshmi, S., Sam, M., and Harisudha, K. (2019). Fusion of WPT and MFCC feature extraction in Parkinson's disease diagnosis. *Technol. Health Care* 27, 363–372. doi: 10.3233/THC-181306

Dhanalakshmi, S., and Venkatesh, C. (2016). Classification of carotid artery abnormalities in ultrasound images using an artificial neural classifier. *Int. Arab J. Inf. Technol.* 13, 756–762.

Gao, H., Zhuang, L., Geoff, P., Maaten, L., and Weinberger, K. (2019). Convolutional networks with dense connectivity. *IEEE Trans. Pattern Anal. Mach. Intell.* doi: 10.1109/TPAMI.2019.2918284

Harisudha, K., Dhanalakshmi, S., Ghosh, S., and Gupta, A. S. (2021). Early diagnosis of Parkinson's disease based on non-motor symptoms: a descriptive and factor analysis. *J. Ambient Intell. Humaniz. Comput.* 1–15. doi: 10.1007/s12652-021-02944-0

Jinjin, H., Hongna, T., Jian, C., Minghui, W., Kai, Q., Jingbo, X., et al. (2019). Multi-level features combined end-to-end learning for automated pathological grading of breast cancer on digital mammograms. *Comput. Med. Imaging Graph.* 71, 58–66. doi: 10.1016/j.compmedimag.2018.10.008

Latha, S., Dhanalakshmi, S., and Kumar, R. (2020). Carotid artery ultrasound images analysis: a review of the literature. *Proc. Inst. Mech. Eng. H* 234, 417–443. doi: 10.1177/0954411919900720

Lauretani, F., Ruffini, L., Scaglioni, A., Cidda, C., Guareschi, C., Nardelli, A., et al. (2015). Utilization of the DaT-SCAN SPECT in the diagnosis of Parkinson's disease in older subjects. *Lett. Drug Des. Discov.* 12, 614–621. doi: 10.2174/1570180812999150312124432

Leung, K., Rowe, S., Pomper, M., and Du, Y. (2021). A three-stage, deep learning, ensemble approach for prognosis in patients with Parkinson's disease. *EJNMMI Res.*11:52. doi: 10.1186/s13550-021-00795-6

Magesh, P. R., Myloth, R. D., and Tom, R. J. (2020). An explainable machine learning model for early detection of Parkinson's disease using LIME on DaTSCAN imagery. *Comput. Biol. Med.* 126:104041. doi: 10.1016/j.compbiomed.2020.104041

Marine, A., Geoffroy, P., Roze, E., Degos, B., Garcin, B., and Garcin, B. (2019). Functional motor symptoms in Parkinson's disease and functional parkinsonism: a systematic review. *J. Neuropsychiatry Clin. Neurosci.* 32, 4–13. doi: 10.1176/appi.neuropsych.19030058

Martínez-Murcia, F., Ortiz, A., Gorriz, J., Ramírez, J., Segovia, F., Salas-Gonzalez, D., et al. (2017). "A 3D convolutional neural network approach for the diagnosis of Parkinson's disease," in *Natural and Artificial Computation for Biomedicine and Neuroscience. IWINAC 2017. Lecture Notes in Computer Science*, Vol. 10337, eds J. Ferrández Vicente, J. Álvarez-Sánchez, F. de la Paz López, J. Toledo Moreo, and H. Adeli (Cham: Springer), 324–333. doi: 10.1007/978-3-319-59740-9_32

Matesan, M., Gaddikeri, S., Longfellow, K., Miyaoka, R., Elojeimy, S., Shana, E., et al. (2018). I-123 DaTscan SPECT brain imaging in Parkinsonian syndromes: utility of the putamen-to-caudate ratio. *J. Neuroimaging* 28, 629–634. doi: 10.1111/jon.12530

Minja, B., Vladislava, B., Milica, B., Nikola, S., Milica, Đ., Kostić, V. S., et al. (2019). Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—a review. *Clin. Neurol. Neurosurg.* 184:105442. doi: 10.1016/j.clineuro.2019.105442

Mohammed, F., He, X., and Lin, Y. (2020). An easy-to-use deep-learning model for highly accurate diagnosis of Parkinson's disease using SPECT images. *Comput. Med. Imaging Graph.* 87:101810. doi: 10.1016/j.compmedimag.2020.101810

Nazari, M., Kluge, A., Apostolova, I., Klutmann, S., Kimiaei, S., Schroeder, M., et al. (2022). Explainable AI to improve acceptance of convolutional neural networks for automatic classification of dopamine transporter SPECT in the diagnosis of clinically uncertain parkinsonian syndromes. *Eur. J. Nucl. Med. Mol. Imaging* 49, 1176–1186. doi: 10.1007/s00259-021-05569-9

Oláh, J., Lehotzky, A., Szénási, T., and Ovádi, J. (2021). "A potential innovative therapy for Parkinson's disease: selective destruction of the pathological assemblies of alpha-synuclein," in *Dementia in Parkinson's Disease*, eds L. Zhang and J. M. Olichney (London: IntechOpen). doi: 10.5772/intechopen.97271

Oliveira, F., Faria, D., Costa, D., Castelo-Branco, M., and Tavares, J. (2018). Extraction, selection and comparison of features for an effective automated computer-aided diagnosis of Parkinson's disease based on [123I]FP-CIT SPECT images. *Eur. J. Nucl. Med. Mol. Imaging* 45, 1052–1062. doi: 10.1007/s00259-017-3918-7

Ortiz, A., Munilla, J., Martínez-Ibañez, M., Gorriz, J., Ramírez, J., Salas-Gonzalez, D., et al. (2019). Parkinson's disease detection using isosurfaces-based features and convolutional neural networks. *Front. Neuroinform.* 13:48. doi: 10.3389/fninf.2019.00048

Pahuja, G., Nagabhushan, T. N., and Prasad, B. (2019). Early detection of Parkinson's disease by using SPECT imaging and biomarkers. *J. Intell. Syst.* 29, 1329–1344. doi: 10.1515/jisys-2018-0261

Pianpanit, T., Lolak, S., Sawangjai, P., Sudhawiyangkul, T., and Wilaiprasitporn, T. (2021). Parkinson's disease recognition using SPECT image and interpretable AI: a tutorial. *IEEE Sens. J.* 21, 22304–22316. doi: 10.1109/JSEN.2021.3077949

Porritt, M., Stanic, D., Finkelstein, D., Batchelor, P., Lockhart, S., Hughes, A., et al. (2005). Dopaminergic innervation of the human striatum in Parkinson's disease. *Mov. Disord.* 20, 810–818. doi: 10.1002/mds.20399

Prediger, R. D., Bortolanza, M., de Castro Issy, A. C., dos Santos, B. L., Del Bel, E., and Raisman-Vozari, R. (2014). "Dopaminergic neurons in Parkinson's disease," in *Handbook of Neurotoxicity*, ed. R. Kostrzewa (New York, NY: Springer), 753–788. doi: 10.1007/978-1-4614-5836-4_7

Rui, B., Huang, L., Zhao, K., Donglian, Q., Qiao, W., et al. (2019). "Author recognition of fine-art paintings," in *Proceedings of the 2019 Chinese Control Conference*, Guangzhou. doi: 10.23919/ChiCC.2019.8865492

Rui, M., Ping, Y., and Bowen, X. (2020). Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks. *IEEE Access* 8, 155362–155377. doi: 10.1109/ACCESS.2020.3019327

Rumman, M., Tasneem, A., Pavel, M., Farzana, S., Md, A. A., et al. (2019). "Early detection of Parkinson's disease using image processing and artificial neural network," in *Proceedings of the 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (ICIVPR)*, Kitakyushu. doi: 10.1109/ICIEV.2018.8641081

Shahed, J., and Jankovic, J. (2007). Motor symptoms in Parkinson's disease. *Handb. Clin. Neurol.* 83, 329–342. doi: 10.1016/S0072-9752(07)83013-2

Shiiba, T., Arimura, Y., Nagano, M., Takahashi, T., and Takaki, A. (2020). Improvement of classification performance of Parkinson's disease using shape features for machine learning on dopamine transporter single photon emission computed tomography. *PLoS One* 15:e0228289. doi: 10.1371/journal.pone.0228289

Sivaranjini, S., and Sujatha, C. (2020). Deep learning based diagnosis of Parkinson's disease using convolutional neural network. *Multimed. Tools Appl.* 79, 15467–15479. doi: 10.1007/s11042-019-7469-8

Subhrajit, B., Latha, S., Dhanalakshmi, S., and Muthu, P. (2018). SVM pixel classification on colour image segmentation. *J. Phys. Conf. Ser.* 1000:012110.

Tagare, H., Delorenzo, C., Chelikani, S., Saperstein, L., and Fulbright, R. (2017). Voxel-based logistic analysis of PPMI control and Parkinson's disease DaTscans. *Neuroimage* 152, 299–311. doi: 10.1016/j.neuroimage.2017.02.067

Tufail, A. B., Ma, Y. K., Zhang, Q. N., Khan, A., Zhao, L., Yang, Q., et al. (2021). 3D convolutional neural networks-based multiclass classification of Alzheimer's and Parkinson's diseases using PET and SPECT neuroimaging modalities. *Brain Inf.* 8:23. doi: 10.1186/s40708-021-00144-2

Warren, N. M., Piggott, M. A., Greally, E., Lake, M., Lees, A. J., Burn, D. J. (2007). Basal ganglia cholinergic and dopaminergic function in progressive supranuclear palsy. *Mov. Disord.* 22, 1594–1600. doi: 10.1002/mds.21573

Wenzel, M., Milletari, F., Krüger, J., Lange, C., Schenk, M., Apostolova, I., et al. (2019). Automatic classification of dopamine transporter SPECT: deep convolutional neural networks can be trained to be robust with respect to variable image characteristics. *Eur. J. Nucl. Med. Mol. Imaging* 46, 2800–2811. doi: 10.1007/s00259-019-04502-5

Wolfswinkel, E., Wielaard, J., Lavalaye, J., Hoff, J., Booij, J., Wit, T., et al. (2021). Artificial intelligence-based assistance in clinical 123I-FP-CIT SPECT scan interpretation. *Mapp. Intimacies*. doi: 10.21203/rs.3.rs-721186/v1

Zhou, C., Huang, Y., and Przedborski, S. (2009). Oxidative stress in Parkinson's disease: a mechanism of pathogenic and therapeutic significance. *Ann. N. Y. Acad. Sci.* 1147, 93–104. doi: 10.1196/annals.1427.023

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.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 25 July 2022
doi: 10.3389/fnagi.2022.931729

Alzheimer-Compound Identification Based on Data Fusion and forgeNet_SVM

Bin Yang[1], *Wenzheng Bao*[2*] *and Shichai Hong*[3*]

[1] School of Information Science and Engineering, Zaozhuang University, Zaozhuang, China, [2] School of Information and Electrical Engineering, Xuzhou University of Technology, Xuzhou, China, [3] Department of Vascular Surgery, Zhongshan Hospital, Fudan University, Xiamen, China

Rapid screening and identification of potential candidate compounds are very important to understand the mechanism of drugs for the treatment of Alzheimer's disease (AD) and greatly promote the development of new drugs. In order to greatly improve the success rate of screening and reduce the cost and workload of research and development, this study proposes a novel Alzheimer-related compound identification algorithm namely forgeNet_SVM. First, Alzheimer related and unrelated compounds are collected using the data mining method from the literature databases. Three molecular descriptors (ECFP6, MACCS, and RDKit) are utilized to obtain the feature sets of compounds, which are fused into the all_feature set. The all_feature set is input to forgeNet_SVM, in which forgeNet is utilized to provide the importance of each feature and select the important features for feature extraction. The selected features are input to support vector machines (SVM) algorithm to identify the new compounds in Traditional Chinese Medicine (TCM) prescription. The experiment results show that the selected feature set performs better than the all_feature set and three single feature sets (ECFP6, MACCS, and RDKit). The performances of TPR, FPR, Precision, Specificity, F1, and AUC reveal that forgeNet_SVM could identify more accurately Alzheimer-related compounds than other classical classifiers.

Keywords: virtual screening, network pharmacology, Alzheimer, data fusion, feature selection, machine learning

OPEN ACCESS

Edited by:
Yizhang Jiang,
Jiangnan University, China

Reviewed by:
Mascot Wang,
University of Waterloo, Canada
Yi Cao,
University of Jinan, China
ZhanHeng Chen,
University of Chinese Academy of Sciences (CAS), China

***Correspondence:**
Wenzheng Bao
baowz55555@126.com
Shichai Hong
xmzshsc@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: *29 April 2022*
Accepted: *24 May 2022*
Published: *25 July 2022*

Citation:
Yang B, Bao W and Hong S (2022) Alzheimer-Compound Identification Based on Data Fusion and forgeNet_SVM. Front. Aging Neurosci. 14:931729. doi: 10.3389/fnagi.2022.931729

INTRODUCTION

Alzheimer's disease (AD) is the most common type of senile dementia, which is a frequently occurring disease of the elderly (Romanelli et al., 1990; Morán et al., 1992; Wang et al., 2014). Its main clinical manifestations are the decline of cognitive function, mental symptoms and behavior disorders, and the decline of daily living ability (Almeida and Crocco, 2000; Daulatzai, 2014; Zhao et al., 2016; Gong et al., 2017). It poses a great threat to the health and quality of life of the elderly and brings a heavy economic burden to society (Rice et al., 1993; Rothstein et al., 1996; Hu, 2006; Wang, 2014). The main reason for the onset of AD is the central nervous system disease in the brain, which causes a series of mental diseases such as learning impairment, memory impairment, and speech impairment (Ogomori et al., 1989; Hao et al., 2013). Family inheritance, physical diseases, and head trauma can cause the onset of this disease (Heyman, 1994; Mehta et al., 1999). However, in the

process of studying the pathogenesis of AD, there are some problems such as unclear pathogenesis, difficult early diagnosis, and no preventable and curable drugs. Therefore, the diagnosis and treatment of AD have been a difficult problem for medical researchers in recent decades.

Alzheimer is a complex disease with multiple factors. At present, the main drugs for the treatment of AD in clinics are acetyl cholinesterase inhibitors, glutamate receptor inhibitors, etc. (Liston et al., 2004; Dong et al., 2005; Sugimoto, 2006). These drugs can alleviate the symptoms caused by the decline of cognitive function, but cannot fundamentally eliminate the pathogeny. Network pharmacology is based on multi-disciplinary knowledge such as system biology, multi pharmacology, bioinformatics, computer technology, and network analysis (Berger and Iyengar, 2009; Chen et al., 2012; Yuan et al., 2019; Li et al., 2020). It systematically studies the drug-target-pathway-disease interaction network and discusses the multi-component, multi-target, and multi-channel pharmacological mechanism of traditional Chinese medicine (TCM) (Li et al., 2014; Xiong et al., 2018; Jiang et al., 2020; Gao et al., 2021). It plays a very important role in exploring treatment approaches and clarifying drug efficacy, especially in finding the effective components of drugs, which is highly consistent with the holistic view emphasized by the theory of traditional Chinese medicine. In recent years, a variety of traditional Chinese medicine prescriptions have been proposed to improve AD by network pharmacology from point of view of multi-component, multi-target, and multi-channel (Sun et al., 2017; An et al., 2020; Wang et al., 2020; Huang et al., 2021). Pang et al. analyzed 25 targets and 13 TCM prescriptions for the treatment of AD and selected 7 representative Chinese medicines (Pang et al., 2016). Naive Bayesian and recursive partitioning was utilized to predict the targets contributing to the chemical components of traditional Chinese medicine in order to construct a compound-target-disease network and explain the synergistic mechanism of multiple effective components of TCM prescriptions. Tao et al. analyzed the compounds of Paeoniae Rubra Radix and Phellodendri Cortex, and the Alzheimer-related targets to reveal the mechanism of these two medicinal materials for intervening AD (Tao et al., 2015). Wang et al. analyzed the main active components of Liuwei Dihuang Decoction and the main action targets of active components and carried out the GO and pathway analyses to give the multi-component, multi-channel and multi-target mechanism of Liuwei Dihuang Decoction in the treatment of AD (Wang et al., 2021). Jiang and Wang utilized network pharmacology to analyze the mechanism of Bajitian for treating AD and obtain that this drug could play an anti-pharmacological role in many aspects, such as neurotransmitter, regulation and regulation of ion channels (Jiang and Wang, 2021).

In network pharmacology, screening the main active compounds of prescriptions is an essential step. In past studies, this step is processed mainly by manually searching public databases. In this study, a novel machine learning method, namely forgeNet_SVM is proposed to identify Alzheimer-related active compounds. The data mining method is utilized to collect Alzheimer-related compounds from the literature. Three molecular descriptors (ECFP6, MACCS, and RDKit) are utilized to obtain the feature sets of compounds respectively, which are fused into an all_feature set. The all_feature set is input to the forgeNet_SVM, in which forgeNet is utilized to give the importance of each feature and select the important features for feature extraction. The selected features are input to support vector machines (SVM) algorithm to identify the new AD-related compounds in TCM prescription.

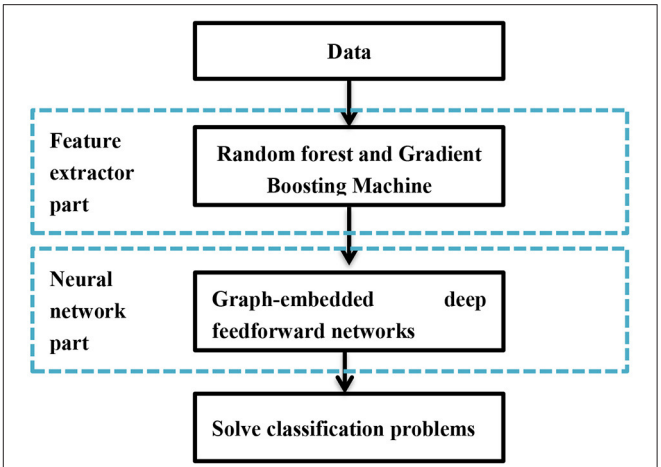


FIGURE 1 | The flowchart of forgeNet algorithm.

METHODS

forgeNet

Forest graph-embedded deep feed forward network (forgeNet) is based on ensemble method and deep learning, which has been utilized for gene regulatory network inference and biology data classification (Kong and Yu, 2020; Yang, 2021). **Figure 1** shows the framework of forgeNet, in which the development of feature graph and classification of deep learning model are contained. Compared to classical deep learning models, forgeNet could solve the dimension imbalance of biomedical data and is more robust (Kong and Yu, 2020).

Development of Feature Graph

With the dimension-imbalance data, the important features of the data are selected for feature extraction. Thus, forgeNet utilizes forest ξ, which includes p decision trees (DTs). According to the training dataset with the classification labels, ξ is fitted and pDTs could be created ($\xi(\theta) = \{T_1(\theta_1),\ T_2(\theta_2),\ \ldots,\ T_p(\theta_p)\}$, $\theta_i$is the coefficient). If a binary tree is considered a special case of a directed graph, the graph set could be obtained as follows.

$$\Phi = \{G_1(V_1,\ E_1),\ \ldots, G_i(V_i,\ E_i),\ \ldots,\ G_N(V_p,\ E_p)\}. \quad (1)$$

Where V_i and $E_i$denote vertex and edge sets of G_i, respectively.

In order to combine the directed graph set Φ, we can obtain the final aggregated graph as follows.

$$G = \bigcup_{i=1}^{p} G_i. \tag{2}$$

Classification of Deep Learning Model

According to the feature graph obtained from the previous step, graph-embedded deep feed forward networks (GEDFN) are utilized to train in order to obtain the optimal model, which is utilized to provide the classification results of the unknown data (Yang, 2021). Every layer of GEDFN is given as followed.

$$\begin{aligned} &Z_1 = \sigma(X(W_{in}\Theta G) + b_{in}),\\ &\cdots\\ &Z_{k+1} = \sigma(Z_k W_k + b_k),\\ &\cdots\\ &Z_{out} = \sigma(Z_l W_l + b_l),\\ &y = soft\max(Z_{out} W_{out} + b_{out}). \end{aligned} \tag{3}$$

Where X represents input vector, Z_k denotes the $k-th$ hidden layer, Θ is Hadamard product, W_k and b_k are the weight and bias of the $k-th$ hidden layer, respectively.

forgeNet also gives a feature importance evaluate mechanism, which is based on Graph Connection Weights (GCW) method (Kong and Yu, 2018). The score of $i-th$feature is defined as follows.

$$\begin{aligned} c_i = \sum_{a=1}^{n} |W_{ia}^{(in)} T(A_{ia} = 1)| + \sum_{b=1}^{n} |W_{bi}^{(in)} T(A_{bi} = 1)|\\ + \sum_{c=1}^{b_1} |W_{ic}^{(1)} T(A_{ia} = 1)|. \end{aligned} \tag{4}$$

Where n is the number of features in the dataset, $W^{(in)}$ represent the weights between the input layer and the first hidden layer, and $W^{(1)}$ represent the weights between the first hidden layer and the second hidden layer. After forgeNet is trained, the importance scores for all the features could be computed with the trained weights.

Support Vector Machine

Support vector machine (SVM) is one of the most classical machine learning algorithms, which was proposed in the year 1995 (Cortes and Vapnik, 1995). SVM is suitable for the classification problems of small-medium samples, nonlinear, and high-dimensional pattern recognition. The basic principle of SVM is to find an optimal classification surface (Hyperplane), which can not only separate the samples without errors but also maximize the margin, based on the most classification surface in the case of linear separability (Suykens and Vandewalle, 1999; Saunders et al., 2002). Therefore, the learning process of SVM is an optimization problem.

The training dataset contains N sample points $\{(x_1, y_1), \ldots, (x_N, y_N)\}$, in which x_i is inputting feature vector and y_i is classification label$\{+1, -1\}$. Hyperplane is labeled as $(w \cdot x) + b = 0$. The optimal hyperplane problem is constructed as follows.

$$\begin{aligned} &\min_{\alpha} \tfrac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{N} \alpha_i.\\ &s.t. \quad \sum_{i=1}^{N} \alpha_i y_i = 0,\ \alpha_i \geq 0,\ i = 1, 2, \ldots, N. \end{aligned} \tag{5}$$

By the Lagrange optimization method, the optimal solution $\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_N^*)^T$ is obtained. The optimal classification function can be given as follows.

$$f(x) = \mathrm{sgn}\{\sum_{i=1}^{N} \alpha_i^* y_i (x_i \cdot x) + b^*\}. \tag{6}$$

Where b^* is a classification threshold.

For the linearly separable dataset, linear SVM is suitable. However, for a nonlinear dataset, in order to solve the linear inseparable problem, the kernel function could be utilized to map the characteristics of nonlinear separable data points from a relatively low dimension to a relatively high dimension and calculate the relationship between them. The algorithm process of searching the optimal classification hyperplane in the high-dimensional feature space is similar to linear separable SVM, which utilizes kernel function to replace the point product in the high-dimensional feature space. The common kernel functions contain linear kernel, polynomial kernel, radial basis function (rbf), and Sigmoid kernel function, which are defined as followed.

$$K_{linear}(x_i, x_j) = x_i \cdot x_j. \tag{7}$$

$$K_{polynomial}(x_i, x_j) = ((x_i \cdot x_j) + 1)^d. \tag{8}$$

$$K_{rbf}(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}). \tag{9}$$

$$K_{sigmoid}(x_i, x_j) = \tanh(k(x_i \cdot x_j) + \theta). \tag{10}$$

Where d is an order of polynomial, σ is the radius of radial basis, k is a scalar and θ is a shifting value.

forgeNet_SVM

In order to improve the classification accuracy of SVM, especially for high-dimensional datasets, a new classifier based on forgeNet and SVM (forgeNet_SVM) is proposed in this paper. ForgeNet can not only be utilized for classification but also score the features in the dataset to indicate the importance of the features. Therefore, in forgeNet_SVM algorithm, for high-dimensional datasets, the forgeNet algorithm is used to select important features for feature extraction. In the next step, the important features are input into SVM for learning to solve the classification problem.

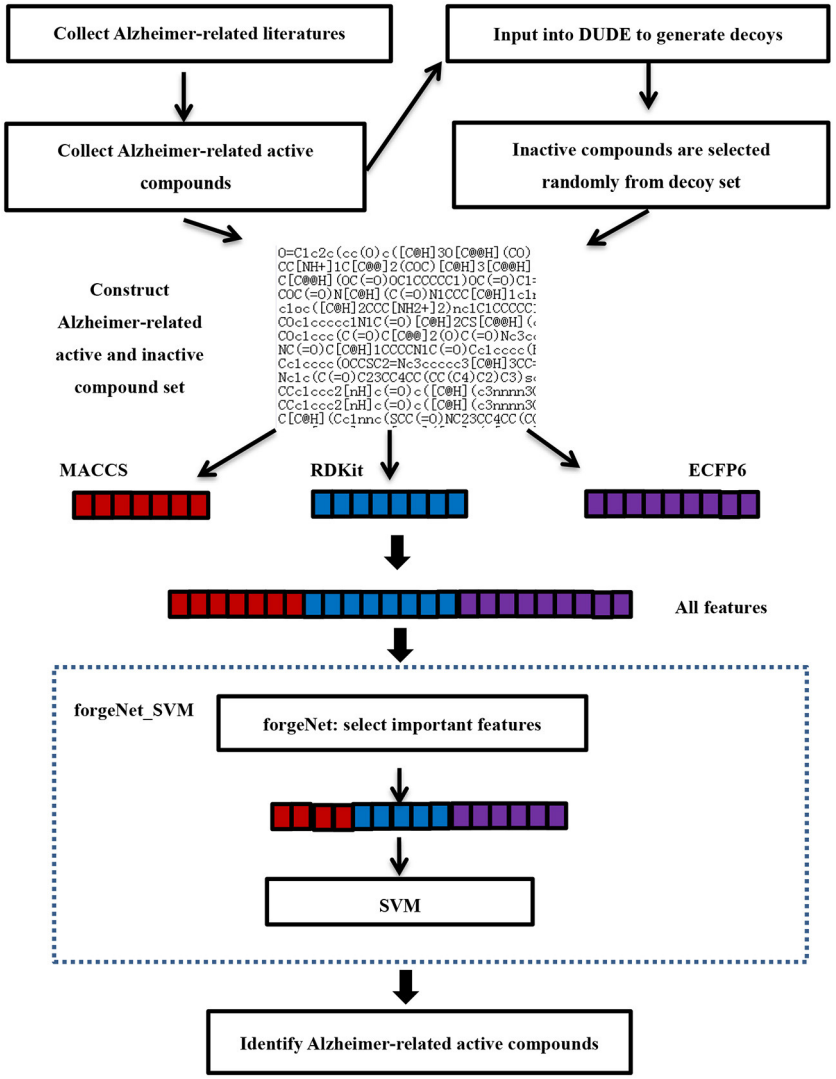


FIGURE 2 | Flowchart of Alzheimer-related active compound identification by forgeNet_SVM.

Alzheimer-Related Active Compound Identification

Figure 2 is the flowchart of Alzheimer-related active compound identification by forgeNet_SVM. The detailed algorithm is given as follows.

1. Studies on TCM in the treatment of AD have to be searched in the literature databases. The queried works of literature need to be analyzed and then collected and mined for important drugs and prescriptions for the treatment of AD, which contains Epimedii Folium, *Anemarrhena asphodeloides*, Radix Ginseng-Poria drug pair, Bajitian, and Polygni Multiflori Caulis. Next, *m*Alzheimer-related closely active compounds, such as naringin, quercetin, Kaempferol, β-Sitosterol, Isorhamnetin, Stigmasterol, and Icariin have to be retrieved. These important compounds have been verified by biological experiments or the molecular docking method. *m* active compounds are utilized as positive samples for further data analysis. In order to determine the negative sample, *m* active compounds are input to the UDU-E website to generate the corresponding decoys (Mysinger et al., 2012). In order to set up the inactive compound set (negative samples), the random decoy selection is performed 3 *m* times from the obtained decoy sets without putting it back. Thus, the inactive compound set contains 3 *m* compounds. The sets of active and inactive compounds constitute the compound sample dataset.
2. The molecular structures of compounds in the dataset collected are SMILES (simplified molecular input line entry system). According to the SMILES structures, three molecular descriptors (ECFP6, MACCS, and RDKit) are utilized to obtain the feature sets of compounds respectively. ECFP6 $(e_1, e_2, \ldots, e_{n_e})$, MACCS $(m_1, m_2, \ldots, m_{n_m})$ and RDKit $(r_1, r_2, \ldots, r_{n_r})$ feature sets of each compound are fused into an

all_feature set $(e_1, e_2, \ldots, e_{n_e}, m_1, m_2, \ldots, m_{n_m}, r_1, r_2, \ldots, r_{n_r})$, where n_e, n_m, and $n_r$are the numbers of ECFP6, MACCS, and RDKit feature sets, respectively. The forgeNet_SVM is utilized to identify Alzheimer-related compounds according to the dataset collected. In order to improve the classification performance of the classifier, all features are input to the forgeNet, which could be utilized to provide the importance of each feature. According to the score of each feature, the important features for classification are selected in order to achieve the purpose of feature extraction. The selected feature set is give as $[d_1, d_2, \ldots, d_n]$.Next, the selected features are input to SVM algorithm for learning. The features of new compounds in TCM prescription are extracted with the same method, which are input to SVM in order to be identified.

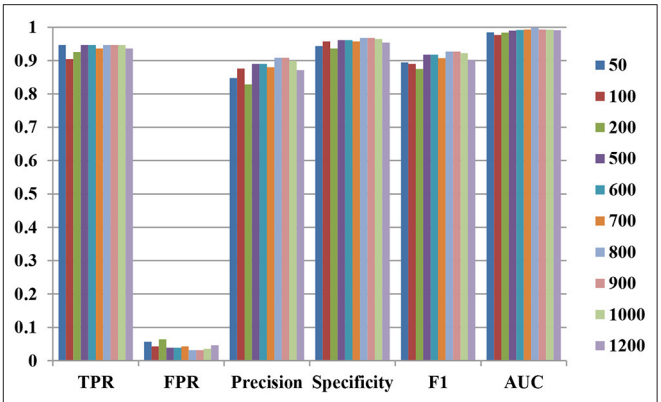


FIGURE 3 | Performances of forgeNet_SVM with the different numbers of features.

EXPERIMENTS AND DISCUSSIONS

In order to test the effectiveness of the proposed method in this paper, the prescriptions and drugs for treating AD are searched. In total 94 Alzheimer-related active compounds are collected and 282 unrelated compounds are also obtained. Each compound is extracted by ECFP6, MACCS, and RDKit to obtain three feature sets (ECFP6, MACCS, and RDKit), respectively. These three feature sets are combined, and a total of 2,423 features are obtained for each compound as the all_feature set. In order to evaluate the performance of the method, TPR, FPR, Precision, Specificity, F1, ROC, and AUC are applied. Seven classical classifiers containing AdaBoost (Cao et al., 2013), Gradient Boosting Decision Tree (GBDT) (Hu and Min, 2018), K-Nearest Neighbor (KNN) (Denoeux, 1995), logistic regression (LR) (Maalouf, 2011), naive Bayes (NB) (Rish, 2001), random forest (RF) (Breiman, 2001), and decision tree (DT) (Breiman et al., 1984)) are also utilized to identify the compounds about Alzheimer. In forgeNet_SVM, the number of trees is set to 1,000, random forest is utilized, three hidden layers are contained, the learning rate is set as 0.0001, the number of training epochs is set to 50, and the linear kernel is selected as the kernel function. In GBDT, the maximum number of weak learners is set to 200. In LR, *L*2 norm is utilized to constrain the arguments. In RF, the number of decision trees is set to 100, the bootstrap method is utilized and the number of features is set to $\sqrt{n\_features}$(*n_features* is the number of features) when searching for the best segmentation.

For forgeNet_SVM, forgeNet can select the important features from a large number of feature sets. First, the different numbers of features are tested for affecting the performance of our method. The numbers of important features selected by forgeNet are 50, 100, 200, 500, 600, 700, 800, 900, 1,000, and 1,200. With the different numbers of feature sets, by 10-cross validation method, the performances of TPR, FPR, Precision, Specificity, *F1*, ROC, and AUC obtained are shown in **Figure 3**. The 10-cross validation method is utilized to divide the training and testing datasets in order to evaluate the model. From **Figure 3**, we can see that our method performs best in terms of TPR when selecting 50, 500, 600, 800, 900, and 1,000 features. In terms of FPR, Precision, Specificity, and F1, our method performs best when selecting 800 and 900 features. Through the results, we can see that our method performs best when 800 and 900 features are selected. In the following experiment, we select the first 900 important features as feature set by forgeNet.

We compare the effects of different feature sets on the performance of the algorithm. The feature sets include ECFP6, MACCS, and RDKit, and all features and selected features are obtained by forgeNet. Two datasets are utilized. The first dataset contains all the compounds (**Dat1**), and another one is obtained by random division (**Dat2**) in which 70% of compounds are used as the training set and the remaining compounds are as the testing set. With Dat1, using the 10-cross validation method, the performances of our method with different feature sets for Alzheimer-related compound identification are shown in **Figure 4** and **Table 1**. From **Figure 4**, it could be seen that the selected feature set has better ROC curves than three single feature sets (ECFP6, MACCS, and RDKit) and all features. Furthermore, in terms of AUC, the selected feature set is 4% higher than ECFP6, 6% higher than MACCS, 4.1% higher than RDKit, and 0.4% higher than the all_feature set. From **Table 1**, it could be seen that in terms of TPR, FPR, Precision, Specificity, and F1, the selected feature set performs better than ECFP6, MACCS, RDKit, and the all_feature sets. With Dat2 and the different feature sets, the identification results of active compounds are shown in **Figure 5** and **Table 2**. From **Figure 5**, the selected features are utilized to obtain a better ROC curve than the other four feature sets. In terms of AUC, the selected feature set is 4, 6, 4.1, and 0.37% higher than ECFP6, MACCS, RDKit, and the all_feature sets, respectively. **Table 2** shows that our selected features could make SVM obtain the best performances of TPR, FPR, Precision, Specificity, and F1. From all the results, it could be seen that the merged feature set (all features) performs better than the three single feature sets (ECFP6, MACCS, and RDKit). Using the forgeNet, the important features could be selected, so the selected feature set could obtain better performances than the merged feature set in terms of TPR, FPR, Precision, Specificity, and F1. Thus the

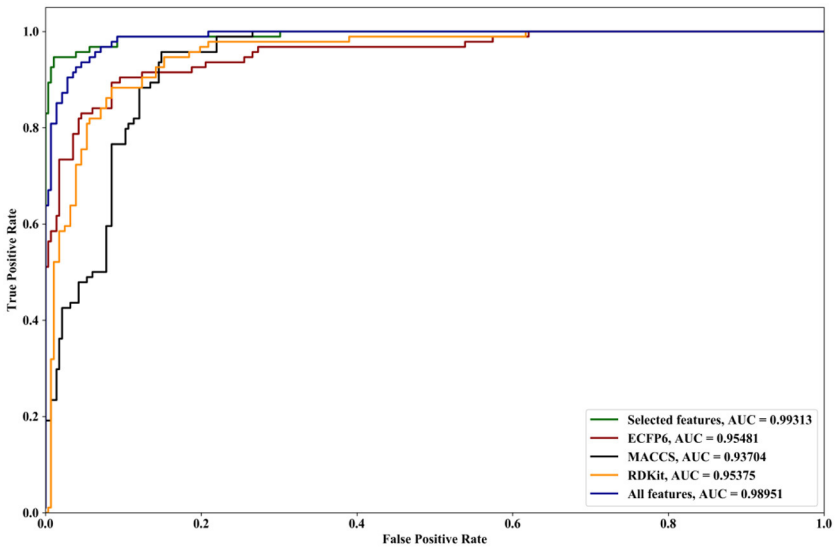


FIGURE 4 | ROC curves and AUC performances of our method with different feature sets for Alzheimer-related compound identification with Dat1.

TABLE 1 | Performances of our method with different feature sets for Alzheimer-related compound identification with Dat1.

Feature sets	TPR	FPR	Precision	Specificity	F1
Selected features	**0.946809**	**0.031915**	**0.908163**	**0.968085**	**0.927083**
ECFP6	0.829787	0.060284	0.821053	0.939716	0.825397
MACCS	0.882979	0.124113	0.70339	0.875887	0.783019
RDKit	0.882979	0.106383	0.734513	0.893617	0.801932
All features	0.93617	0.056738	0.846154	0.943262	0.888889

The bold values denote the best performances.

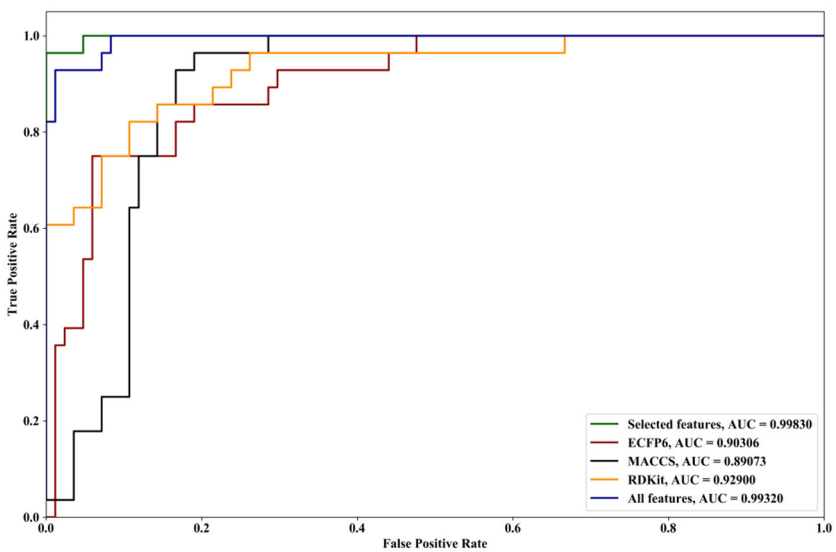


FIGURE 5 | ROC curves and AUC performances of our method with different feature sets for Alzheimer-related compound identification with Dat2.

TABLE 2 | Performances of our method with different feature sets for Alzheimer-related compound identification with Dat2.

Feature sets	TPR	FPR	Precision	Specificity	F1
Selected features	**0.964286**	**0**	**1**	**1**	**0.981818**
ECFP6	0.678571	0.059524	0.791667	0.940476	0.730769
MACCS	0.821429	0.142857	0.657143	0.857143	0.730159
RDKit	0.857143	0.214286	0.571429	0.785714	0.685714
All features	0.678571	0.059524	0.791667	0.940476	0.730769

The bold values denote the best performances.

TABLE 3 | Performances of 15 methods for Alzheimer-related compound identification with Dat1.

Methods	TPR	FPR	Precision	Specificity	F1	AUC
forgeNet_SVM	0.946809	0.031915	0.908163	0.968085	**0.927083**	**0.99313**
AdaBoost	0.914894	0.035461	0.895833	0.964539	0.905263	0.974083
forgeNet_AdaBoost	0.914894	0.035461	0.895833	0.964539	0.905263	0.974083
GBDT	0.904255	0.039007	0.885417	0.960993	0.894737	0.981326
forgeNet_GBDT	0.914894	0.028369	0.914894	0.971631	0.914894	0.982383
KNN	**0.989362**	0.77305	0.299035	0.22695	0.459259	0.798759
forgeNet_KNN	0.893617	0.028369	0.913043	0.971631	0.903226	0.978101
LR	**0.989362**	0.56383	0.369048	0.43617	0.537572	0.942282
forgeNet_LR	0.93617	0.042553	0.88	0.957447	0.907216	0.942282
NB	0.287234	**0**	**1**	**1**	0.446281	0.643617
forgeNet_NB	0.946809	0.039007	0.89	0.960993	0.917526	0.962464
RF	0.882979	0.031915	0.902174	0.968085	0.892473	0.98823
forgeNet_RF	0.904255	0.031915	0.904255	0.968085	0.904255	0.986457
DT	0.87234	0.109929	0.725664	0.890071	0.792271	0.881206
forgeNet_DT	0.946809	0.060284	0.839623	0.939716	0.89	0.943262

The bold values denote the best performances.

TABLE 4 | Performances of 15 methods for Alzheimer-related compound identification with Dat2.

Methods	TPR	FPR	Precision	Specificity	F1	AUC
forgeNet_SVM	0.964286	**0**	**1**	**1**	**0.981818**	**0.998299**
AdaBoost	0.357143	0.309524	0.277778	0.690476	0.3125	0.991071
forgeNet_AdaBoost	0.892857	**0**	**1**	**1**	0.943396	0.995748
GBDT	0.821429	0.607143	0.310811	0.392857	0.45098	0.997449
forgeNet_GBDT	0.928571	**0**	**1**	**1**	0.962963	0.993197
KNN	**1**	**1**	0.25	0	0.4	0.742347
forgeNet_KNN	0.892857	0.035714	0.892857	0.964286	0.892857	0.94494
LR	**1**	0.678571	0.329412	0.321429	0.495575	0.964711
forgeNet_LR	0.928571	0.071429	0.8125	0.928571	0.866667	0.985544
NB	0	**0**		**1**		0.5
forgeNet_NB	0.964286	0.059524	0.84375	0.940476	0.9	0.951743
RF	0.535714	0.130952	0.576923	0.869048	0.555556	0.987724
forgeNet_RF	0.928571	**0**	**1**	**1**	0.962963	0.996173
DT	0.857143	0.630952	0.311688	0.369048	0.457143	0.839286
forgeNet_DT	0.964286	0.011905	0.964286	0.988095	0.964286	0.97619

The bold values denote the best performances.

feature extraction method can improve the accuracy of active compound recognition.

AdaBoost, GBDT, KNN, LR, NB, RF, and DT are also directly utilized to predict Alzheimer-related compounds with Dat1 and Dat2. In forgeNet_SVM, SVM is also replaced with these seven classifiers in order to constitute forgeNet_AdaBoost, forgeNet_GBDT, forgeNet_KNN, forgeNet_LR, forgeNet_NB, forgeNet_RF, and forgeNet_DT, which are utilized to identify

compounds. With Dat1 and Dat2, the performances of 15 methods for Alzheimer-related compound identification are listed in **Tables 3**, **4**, respectively. From **Table 3**, KNN and LR could obtain the best TPR performance, which shows that KNN and LR could identify the most active compounds. But these two methods shave the worst FPR performances, which are 0.77305 and 0.56383, respectively. The results reveal that LR identifies most of the compounds as active compounds. In terms of FPR, Precision, and Specificity, NB performs best. But NB has the worst TPR performance, which shows that NB identifies most of the compounds as inactive compounds. In terms of F1 and AUC, forgeNet_SVM could obtain the best performances among the 15 methods. From **Table 4**, KNN and LR could gain the best TPR performance, which reveals that these two methods could identify all true active compounds. forgeNet_SVM, forgeNet_NB, and forgeNet_DT could obtain the second better TPR performance. ForgeNet_SVM could gain the best FPR performance, which shows that our proposed method can identify all true inactive compounds. In terms of Precision, Specificity, F1, and AUC, forgeNet_SVM also performs best. On the whole, our proposed method could infer more true active and inactive compounds than other methods.

CONCLUSION

In this study, a novel Alzheimer-related compound identification algorithm based on data fusion and forgeNet_SVM is proposed. Three feature description methods (ECFP6, MACCS, and RDKit) are utilized to obtain the feature sets of Alzheimer related and unrelated compounds, which are fused into the all_feature set. In forgeNet_SVM, all_feature set is input to forgeNet, which could evaluate the importance of each feature and extract the important features according to the given scores. The selected features are input to SVM algorithm to identify the new compounds in a TCM prescription. The Alzheimer-related dataset collected is utilized, and the experiment results show that forgeNet_SVM could identify more true-positive compounds and fewer false-positive compounds than other classical classifiers, such as AdaBoost, GBDT, KNN, LR, NB, RF, and DT. We make the comparison experiments that give the optimal number of the selected features for forgeNet_SVM. In terms of TPR, FPR, Precision, Specificity, F1, and AUC, the selected feature set performs better than the all_feature set and three single feature sets (ECFP6, MACCS, and RDKit).

In the future, we will apply forgeNet_SVM to identify other diseases related compounds, such as cancer, COVID-19, and cardiovascular diseases.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

WB conceived the method. BY designed the method and conducted the experiments. WB and SH wrote the main manuscript text. All authors reviewed the manuscript.

FUNDING

This work was supported by the Talent Project of Qingtan Scholar of Zaozhuang University, the Natural Science Foundation of China (No. 61902337), the Fundamental Research Funds for the Central Universities (2020QN89), Xuzhou Science and Technology Plan Project (KC19142 and KC21047), Shandong Provincial Natural Science Foundation, China (No. ZR2015PF007), Jiangsu Provincial Natural Science Foundation (No. SBK2019040953), Natural Science Fund for Colleges and Universities in Jiangsu Province (No. 19KJB520016), and Young Talents of Science and Technology in Jiangsu.

REFERENCES

Almeida, O. P., and Crocco, E. I. (2000). Perception of cognitive deficits and behavior disorders in patients with Alzheimer's disease. *Arq. Neuropsiquiatr.* 58, 292–299. doi: 10.1590/S0004-282X2000000200015

An, H. M., Huang, D. R., Yang, H., Liu, X. G., Du, J., Li, Y., et al. (2020). Comprehensive chemical profiling of Jia-Wei-Qi-Fu-Yin and its network pharmacology-based analysis on Alzheimer's disease. *J. Pharm. Biomed. Anal.* 189, 113467. doi: 10.1016/j.jpba.2020.113467

Berger, S. I., and Iyengar, R. (2009). Network analyses in systems pharmacology. *Bioinformatics.* 25, 2466–2472. doi: 10.1093/bioinformatics/btp465

Breiman, L. (2001). Random forest. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J. (1984). Classification and Regression Trees (CART). *Biometrics.* 40, 358. doi: 10.2307/2530946

Cao, Y., Miao, Q. G., Liu, J. C., Gao, L. (2013). Advance and prospects of AdaBoost algorithm. *Zidonghua Xuebao/Acta Automatica Sinica.* 39, 745–758. doi: 10.1016/S1874-1029(13)60052-X

Chen, Y., Liu, Z. L., and Xie, Y. B. A. (2012). knowledge-based framework for creative conceptual design of multi-disciplinary systems. *Comput. Aided Des.* 44, 146–153. doi: 10.1016/j.cad.2011.02.016

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018

Daulatzai, M. A. (2014). Role of stress, depression, and aging in cognitive decline and Alzheimer's disease. *Curr. Top. Behav. Neurosci.* 18, 265–296. doi: 10.1007/7854_2014_350

Denoeux, T. (1995). A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE Trans. Syst. Man Cybern.* 25, 804–813. doi: 10.1109/21.376493

Dong, H., Csernansky, C. A., Martin, M. V., Bertchume, A., Vallera, D., and Csernansky, J. G. (2005). Acetylcholinesterase inhibitors ameliorate behavioral deficits in the Tg2576 mouse model of Alzheimer's disease. *Psychopharmacology.* 181, 145–152. doi: 10.1007/s00213-005-2230-6

Gao, Q., Han, Z. Y., Tian, D. F., Liu, G. L., Wang, Z. Y., Lin, J. F., et al. (2021). Xinglou Chengqi Decoction improves neurological function in experimental stroke mice as evidenced by gut microbiota analysis and network pharmacology. *Chin. J. Nat. Med.* 12, 881–899. doi: 10.1016/S1875-5364(21)60079-1

Gong, X. Q., Luo, L. J., and Neurology, D. O. (2017). Comparative analysis on cognitive function and behavioral and psychological symptoms between vascular dementia and Alzheimer disease. *Neural Repair.* 12, 122–123. doi: 10.16780/j.cnki.sjssgncj.2017.02.008

Hao, C., Freeman, C., Jacobson, G. A., and Small, D. H. (2013). Proteoglycans in the central nervous system: role in development, neural repair, and Alzheimer's disease. *IUBMB*. 65, 108–120. doi: 10.1002/iub.1118

Heyman, A. (1994). Head trauma as a risk factor for Alzheimer's disease. *J. Neurol. Sci.* 127, 6–6. doi: 10.1016/0022-510X(94)90119-8

Hu, J., and Min, J. (2018). Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model. *Cogn. Neurodyn.* 12, 431–440. doi: 10.1007/s11571-018-9485-1

Hu, R. D. (2006). The effects of health education to the family members of elderly patients with alzheimer's disease on the quality of life. *Med. J. Chin. People's Lib.* 23, 7–9. doi: 10.3969/j.issn.1008-9993.2006.06.003

Huang, X. Y., L,i T. T., Zhou, L., Liu, T., Xiong, L. L., and Yu, C. Y. (2021). Analysis of the potential and mechanism of Ginkgo biloba in the treatment of Alzheimer's disease based on network pharmacology. *Ibrain* 7, 21–28. doi: 10.1002/j.2769-2795.2021.tb00060.x

Jiang, R., Zhang, X., Li, Y., Zhou, H., Wang, H., Wang, F., et al. (2020). Identification of the molecular mechanisms of Salvia miltiorrhiza relevant to the treatment of osteoarthritis based on network pharmacology. *Discov. Med.* 30, 83–95.

Jiang, Z., and Wang, Z. (2021). Material basis and mechanism of bajitian (morindae officinalis radix) treating Alzheimer's disease. *J. Tradit. Chin. Med.* 39, 255–258W. doi: 10.13193/j.issn.1673-7717.2021.03.061

Kong, Y., and Yu, T. (2018). A graph-embedded deep feedforward network fordisease outcomeclassification and feature selection using gene expressiondata. *Bioinformatics.* 34, 3727–3737. doi: 10.1093/bioinformatics/bty429

Kong, Y., and Yu, T. (2020). forgeNet: a graph deep neural network model using tree-based ensemble classifiers for feature graph construction. *Bioinformatics.* 36, 3507–3515. doi: 10.1093/bioinformatics/btaa164

Li, R., Li, Y., Liang, X., Yang, L., Su, M., and Lai, K. P. (2020). X, et al. Network Pharmacology and bioinformatics analyses identify intersection genes of niacin and COVID-19 as potential therapeutic targets. *Brief. Bioinformatics.* 22, 1279–1290. doi: 10.1093/bib/bbaa300

Li, X., Wu, L. H., Liu, W., Jin, Y. C., Chen, Q., Wang, L. L., et al. (2014). A Network Pharmacology Study of Chinese Medicine QiShenYiQi to Reveal Its Underlying Multi-Compound, Multi-Target, Multi-Pathway Mode of Action. *PLoS ONE* 9, e95004. doi: 10.1371/journal.pone.0095004

Liston, D. R., Nielsen, J. A., Villalobos, A., Chapin, D., Jones, S. B., Hubbard, S. T., et al. (2004). Pharmacology of selective acetylcholinesterase inhibitors: implications for use in Alzheimer's disease. *Eur. J. Pharmacol.* 486, 9–17. doi: 10.1016/j.ejphar.2003.11.080

Maalouf, M. (2011). Logistic regression in data analysis: an overview. *Int. J. Data Anal. Tech. Strateg.* 3, 281–299. doi: 10.1504/IJDATS.2011.041335

Mehta, K. M., Ott, A., Kalmijn, S., Slooter, A. J., Duijn, C. M. V., Hofman, A., and Breteler, M. M. (1999). Head trauma and risk of dementia and Alzheimer'sdisease: the Rotterdam study. *Neurology.* 53, 1959–1962. doi: 10.1212/WNL.53.9.1959

Morán, M. A., Cebrián, J. L., Gómez-Ramos, P., Cabello, A., Madero, S., and Mufson, E. J. (1992). Diagnosis of Alzheimer's disease. Evaluation of senile plaques of the diffuse type. *Medicina Clínica.* 98, 19–23.

Mysinger, M. M., Carchia, M., Irwin, J. J., Shoichet, B. K. (2012). Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* 55, 6582. doi: 10.1021/jm300687e

Ogomori, K., Kitamoto, T., and Tateishi, J. (1989). Beta-protein amyloid is widely distributed in the central nervous system of patients with Alzheimer's disease. *Am. J. of Pathol.* 134, 243–251.

Pang, X. C., Wang, Z., Fang, J. S., Lian, W. W., Zhao, Y., Kang, D., et al. (2016). Network pharmacology study of effective constituents of traditional Chinese medicine for Alzheimer's disease treatment. *Acta Pharmaceutica Sinica.* 51, 725.—731. doi: 10.16438/j.0513-4870.2015-0950

Rice, D. P., Fox, P. J., Max, W., Webber, P. A., Hauck, W. W., Lindeman, D. A., et al. (1993). The economic burden of caring for people with Alzheimer's disease. *Health Aff.* 12, 164–176. doi: 10.1377/hlthaff.12.2.164

Rish, I. (2001). An empirical study of the naive Bayes classifier. *J. Universal Comp. Sci.* 3, 41–46. doi: 10.1002/9781118721957.ch4

Romanelli, M. F., Ashkin, K., Morris, J. C., and Coben, L. A. (1990). Advanced Alzheimer's disease is a risk factor for late-onset seizures. *Arch. Neurol.* 47, 847–850. doi: 10.1001/archneur.1990.00530080029006

Rothstein, Z., Prohovnik, I., Davidson, M., Beeri, M. S., and Noy, S. (1996). The economic burden of Alzheimer's disease in Israel. *Isr. J. Med. Sci.* 32, 1120–1123.

Saunders, C. Stitson, M. O., Weston, J. Holloway, R. Bottou, L. Scholkopf, B. et al. (2002). Support vector machine. *Comp. Sci.* 1, 1–28. doi: 10.1007/978-3-642-27733-7_299-3

Sugimoto, M. (2006). Acetylcholinesterase inhibitors used in treatment of Alzheimer's disease prevent glutamate neurotoxicity via nicotinic acetylcholine receptors and phosphatidylinositol 3-kinase cascade. *Neuropharmacology.* 51, 474–486. doi: 10.1016/j.neuropharm.2006.04.007

Sun, L. M., Liu, L. F., Zhu, H. X., Zhu B J, Zhang Q C. (2017). Network pharmacology-based study on intervention mechanism of Huanglian Jiedu decoction in the treatment of Alzheimer's disease. *Acta Pharmaceutica Sinica.* 8, 1268–1275. doi: 10.16438/j.0513-4870.2017-0144

Suykens, J. A. K., and Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Process. Lett.* 9, 293–300. doi: 10.1023/A:1018628609742

Tao, X. Q., Zhang, X. Z,, Li, N., Cao, L., Ding, G., Wang, Z. Z., et al. (2015). Study on molecular mechanism of Paeoniae Rubra Radix and Phellodendri Cortex intervening Alzheimer's disease using network pharmacology methods. *Chin. Tradit. Herb. Drugs.* 46, 1634–1639. doi: 10.7501/j.issn.0253-2670.2015.11.013

Wang, M., Wang, S., Li, Y., Cai, G. M., Cao, M., Li, L. F. (2020). Integrated analysis and network pharmacology approaches to explore key genes of Xingnaojing for treatment of Alzheimer's disease. *Brain Behav.* 10, e01610. doi: 10.1002/brb3.1610

Wang, R., Jia, Y., Song, J., Liu, L. J., Zhan, X. H., Hou, J. L., et al. (2021). Mechanism of Liuwei Dihuang decoction in treatment of dementia based on network pharmacology. *J. Henan University (Medical Science).* 40, 84–92.

Wang, X., Kim, J. R., Lee, S. B., Kim, Y. J., Joung, M., Kwon, H. W., et al. (2014). Effects of curcuminoids identified in rhizomes of Curcuma longa on BACE-1 inhibitory and behavioral activity and lifespan of Alzheimer's disease Drosophila models. *BMC Complement. Med. Ther.* 14, 88. doi: 10.1186/1472-6882-14-88

Wang, Z. X. (2014). Effects of extended care on the quality of life of the elderly patients with Alzheimer's disease. *Practical Geriatr.* 28, 254–259. doi: 10.3969/j.issn.1003-9198.2014.03.024

Xiong, D. D., Qin, Y., Xu, W. Q., He, R. Q., Wu, H. Y., Wei, D. M., et al. (2018). A network pharmacology-based analysis of multi-target, multi-pathway, multi-compound treatment for ovarian serous cystadenocarcinoma. *Clin. Drug Investig.* 38, 909–925. doi: 10.1007/s40261-018-0683-8

Yang, B. (2021). *Gene Regulatory Network Identification based on Forest Graph-embedded Deep Feedforward Network. 6th International Conference on Cloud Computing and Internet of Things.* Okinawa, p. 68–72. doi: 10.1145/3493287.3493297

Yuan, C. Y., Liu, B. T., Huang, J. Y., Yan Z. S., Chen, R., and Huo, L. N. (2019). Application of network pharmacology on screening and mechanism of pharmacodynamic substances of traditional Chinese medicine. *Guangzhou Chem. Indust.* 47, 20–22.

Zhao, W. N., Bi, P. X., Li, S. O., Yin, C. H., Yang, Y. D., and Sun, L. (2016). Comparative study of damage to cognitive function and mental behavior in patients with general paresis of the insane, Alzheimer's disease, and frontotemporal dementia. *Int. J. Clin. Exp. Med.* 9, 7374–7380.

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership