

## Research Article

# Alzheimer's Disease Classification Based on Image Transformation and Features Fusion

Hongfei Jia , Yu Wang , Yifan Duan , and Hongbing Xiao 

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Artificial Intelligence, Beijing Technology and Business University, Beijing 100048, China

Correspondence should be addressed to Yu Wang; wangyu@btbu.edu.cn

Received 16 July 2021; Revised 25 November 2021; Accepted 9 December 2021; Published 28 December 2021

Academic Editor: Huiling Chen

Copyright © 2021 Hongfei Jia et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It has become an inevitable trend for medical personnel to analyze and diagnose Alzheimer's disease (AD) in different stages by combining functional magnetic resonance imaging (fMRI) and artificial intelligence technologies such as deep learning in the future. In this paper, a classification method was proposed for AD based on two different transformation images of fMRI and improved the 3DPCANet model and canonical correlation analysis (CCA). The main ideas include that, firstly, fMRI images were preprocessed, and subsequently, mean regional homogeneity (mReHo) and mean amplitude of low-frequency amplitude (mALFF) transformation were performed for the preprocessed images. Then, mReHo and mALFF images were extracted features using the improved 3DPCANet, and these two kinds of the extracted features were fused by CCA. Finally, the support vector machine (SVM) was used to classify AD patients with different stages. Experimental results showed that the proposed approach was robust and effective. Classification accuracy for significant memory concern (SMC) vs. mild cognitive impairment (MCI), normal control (NC) vs. AD, and NC vs. SMC, respectively, reached 95.00%, 92.00%, and 91.30%, which adequately proved the feasibility and effectiveness of the proposed method.

## 1. Introduction

Alzheimer's disease (AD) [1] is one of the currently incurable brain disorders, characterized by insidious onset and continuous development, which can cause a continuous decline of patient's cognitive and memory abilities, eventually can lead to abnormal life. Studies [2–4] suggested that significant memory concern (SMC) may be an early stage of mild cognitive impairment (MCI) and AD. Clinical symptoms are objectively poor memory and cognitive decline accompanied by changes in brain structure. And it maybe evolves into AD. It is very important to accurately diagnose the disease situation of patient because at present AD cannot be cured completely and can only be slowed or be prevented to further develop by treatment. In addition, the diagnosis of AD requires a mass of medical data which is inhomogeneous, so medical staff bear the heavy burden caused by man-made data analysis.

With the rapid development of deep learning [5–8] and medical imaging technologies, more and more researchers

used medical imaging means such as magnetic resonance imaging (MRI) [9–12], positron emission computed tomography (PET), computed tomography (CT), and deep learning method to assist medical personnel to accurately diagnose and treat AD patients with various stages. Huang et al. [13] improved deep learning network, called VGGNet, which was utilized to classify three-dimensional (3D) images. In the experiments, a classifier was trained using T1-MRI and FDG-PET images, and high precision was achieved. Islam and Zhang [14] proposed a convolutional neural network by combining dense network modules. The experimental results showed that the accuracy of the non-dementia stage was 99%. Zhang et al. [15] designed a convolutional neural network to extract the features of dual modalities including PET and MRI images. The extracted features and information resulted from the minimal status test (MMSE) and the clinical dementia rating (CDR) were fused. The accuracies of AD and normal control (NC), MCI and NC, and AD and MCI were 100%, 96.58%, and 97.43%, respectively. Jain et al. [16] adopted a

mathematical model based on a convolutional neural network (CNN) with transfer learning to diagnose AD. In this model, VGG-16 trained on the ImageNet dataset was used as a feature extractor for classification tasks. The classification accuracies of AD and NC, AD and MCI, and NC and MCI reach separately 99.14%, 99.30%, and 99.22%. Most of the above methods are involved in binary classification in AD, NC, and MCI. However, during the evolution from NC to AD, some other stages such as SMC exist. Therefore, in this paper, some subtype classification research on AD with various stages is made.

PCANet is one of the common convolutional neural networks proposed by Chen et al. [17]. Subsequently, Li et al. [18] improved PCANet network from two-dimensional (2D) CNN to three-dimension (3D) CNN and diagnosed AD patients using structure MRI (sMRI). In this paper based on the 3D PCANet, the max-pooling layer and rectified linear unit (ReLU) layer are added behind each convolution layer to reduce the redundancy of image features. The improved 3D PCANet model is used to extract texture and nonlinear features of brain images. Experimental results demonstrate that the improved method can effectively increase the accuracy of classification.

As a noninvasively imaging technology, fMRI [19] is used to measure spontaneous brain activity which can reflect the status of different brain regions at different times. Many studies suggest that different levels of functional characteristics of fMRI such as amplitude of low-frequency amplitude (ALFF) [20], regional homogeneity (ReHo) [21], and regional functional correlation strength (RFCS) [22] can reflect brain diseases, which can assist medical personnel to diagnose brain diseases. Dai et al. [23] used different types of transforms on fMRI data including ALFF, ReHo, RFCS, and gray matter density (GMD) data and combined with multilevel characterization with multiclassifier (M3) to realize the diagnosis of AD patients. Good results were obtained. However, when multimodal data directly are used, feature redundancy often happens, and classification results will further be influenced. Aiming at the above problem, in this paper, two functional image transforms are selected for fMRI images including ALFF and ReHo and are used to extract features. Then, canonical correlation analysis (CCA) is used to fuse these two features.

Inspired by the above ideas, a method to diagnose AD based on different functional characteristics of fMRI and CCA fusion strategy is proposed in this paper. First, fMRI images are preprocessed and transformed into mReHo (mean ReHo) and mALFF (mean ALFF) images. Then, these two kinds of transformed images are inputted into the improved 3D PCANet model, respectively, for feature extraction. Next, these two features are fused by CCA. Finally, support vector machine (SVM) is utilized to classify. Contributions of this paper are as follows.

- (1) Because fMRI data are four-dimensional (4D) form, and features cannot be directly extracted, 4D fMRI images are converted into 3D form using image transformation such as ALFF and ReHo

- (2) Traditional 3D PCANet network is improved by adding a maximum pooling layer behind each convolution layer which is used to extract image features. So, feature redundancy and human error can be effectively reduced
- (3) CCA is used to fuse two kinds of image features, and the accuracy of the model classification is improved
- (4) AD patients with different stages, especially including SMC, are fully automatically classified, which can assist medical personnel to accurately diagnose and analyze AD

The rest of this paper is organized as follows. In Section 2, we introduced the experimental dataset and proposed method, respectively. In Section 3, we gave the experimental results and analysis of our proposed method and compared methods. A conclusion is drawn in Section 4.

## 2. Methodology

The framework diagram of our proposed method is shown in Figure 1. Specifically, first, fMRI images were preprocessed and transformed. Then, the transformed images were extracted features using the improved 3DPCANet. Then, these two kinds of features were fused by CCA. Finally, SVM was used to classify AD patients with different stages. The detailed steps are explained in the following sections.

*2.1. Data Preprocessing.* fMRI used in this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI). The fMRI dataset includes 34 patients with AD, 26 patients with SMC, 57 patients with EMCI, 35 patients with LMCI, 38 patients with MCI, and 50 NC. Detailed information is shown in Table 1.

fMRI data analysis is used to DPARSF [24] toolkit. Due to the instability of the initial fMRI signal, the first 10 time points of each fMRI data are deleted, and the points are made timing correction, realigned, and normalized. The images are registered to the template proposed by the Montreal Neurological Institute (MNI). The preprocessed images are shown in Figure 2.

*2.2. ALFF Transform Images.* Amplitude of low-frequency fluctuation (ALFF) value is the root mean square of the power spectrum of blood oxygen level dependent (BOLD) signal in the low-frequency band (0.01 Hz-0.08 Hz). The steps are as follows:

- (1) The time series of each voxel after removing the linear drift is passed through a filter with 0.01 Hz-0.08 Hz band
- (2) The filtering result is made fast Fourier transformation to obtain the power spectrum
- (3) The root means square of the power spectrum is calculated
- (4) The average value of step (3) is obtained which is ALFF

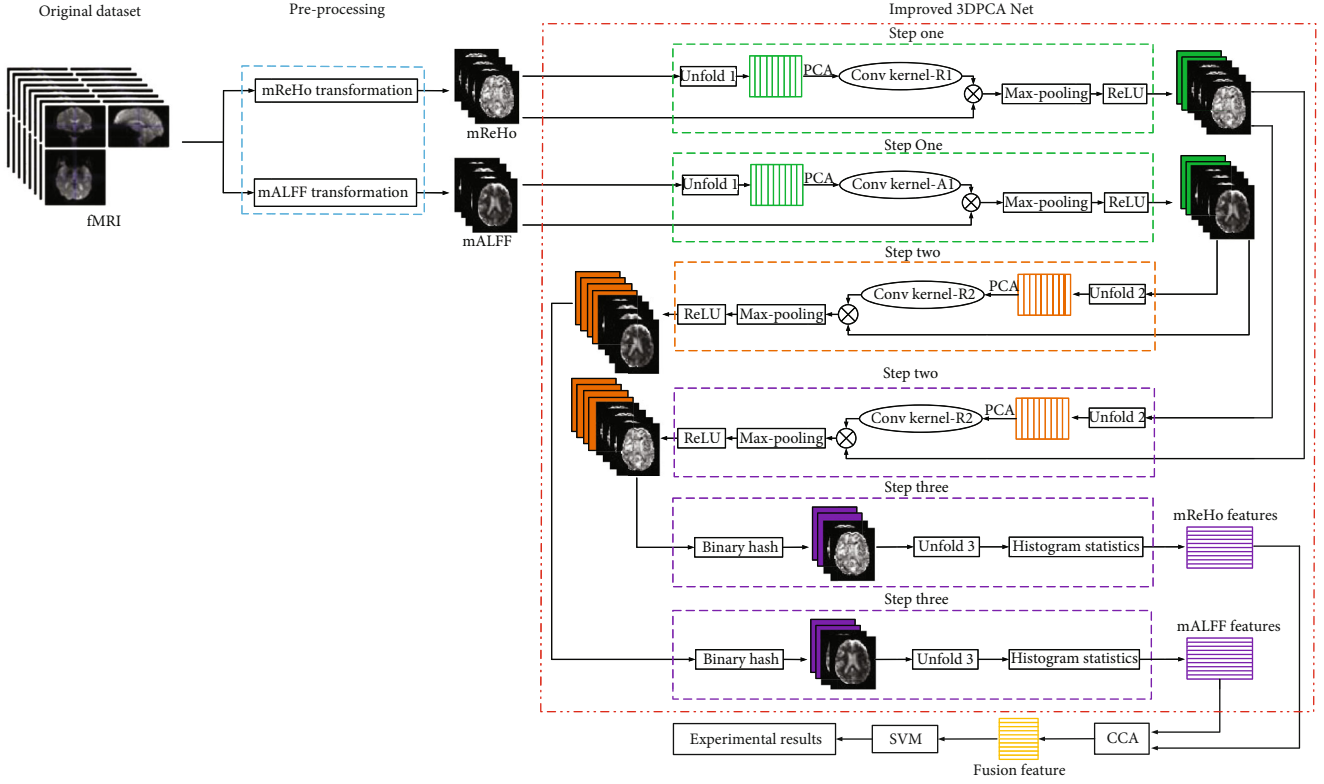


FIGURE 1: Framework diagram of the proposed method.

TABLE 1: Statistical analysis of subject information.

fMRI	Number	Male/female	Age
AD	34	18/16	57~88
SMC	26	14/12	65~83
EMCI	57	34/23	57~90
LMCI	35	14/21	58~88
MCI	38	18/20	57~90
NC	50	28/22	66~91

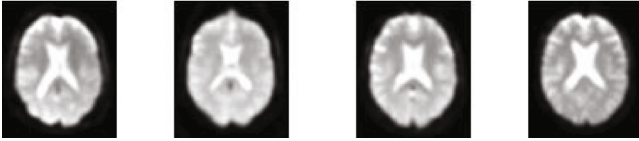


FIGURE 2: Sample images after preprocessing.

Low-frequency signal energy is utilized to represent the activity of neurons in different brain regions. Mean amplitude of low-frequency fluctuations (mALFF) is obtained by dividing the average ALFF of all voxels in the whole brain, because the brain structure of AD patients has changed, and likewise, it is believed that the activity of neurons in each brain area will also change compared with that of the normal control group. The preprocessed fMRI data are calculated mALFF. The after mALFF transformation images are shown in Figure 3.

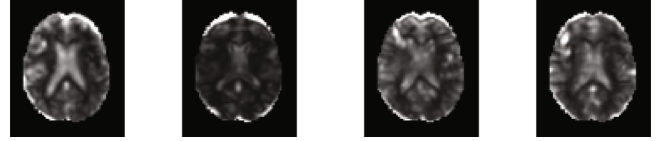


FIGURE 3: Sample images after mALFF image transformation.

**2.3. ReHo Transform Images.** ReHo method was originally proposed by Jiang and Zuo [21], which was used to measure the regional synchronization degree of fMRI time course. ReHo assumes that the selected voxels have temporary similarity with their adjacent voxels which is measured by Kendall's harmony coefficient.

$f(M_1, N_1, O_1, T_1)$  is used to represent fMRI data, where  $M_1$  is the number of rows,  $N_1$  is the number of columns,  $O_1$  is the number of layers, and  $T_1$  is the number of time points for each voxel (the length of time series). The data contains  $M_1 \times N_1 \times O_1$  voxels. Voxel  $j_1$  is denoted by  $V_{j_1}(m_1, n_1, o_1)$  ( $1 \leq m_1 \leq M_1, 1 \leq n_1 \leq N_1, 1 \leq o_1 \leq O_1$ ), and the local consistency of the time series of  $j_1$ th voxel and  $K_1$  nearest neighborhood voxels (usually  $K_1$  is 6, 18, 26) is calculated as follows:

- (1) The time series of voxel  $K_1 + 1$  is expressed as a matrix  $X^1$  with size  $T_1 \times (K_1 + 1)$  pixels, where  $X_1(i, j)$  represents the  $i$ th time point of  $j$ th voxel
- (2) The element value in column vector is replaced by the rank of its column which is the ordinal number

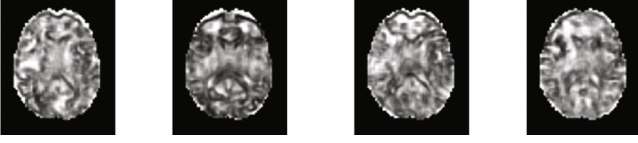


FIGURE 4: Sample images on mReHo.

of the value of the element in  $j_1$ th column, and the matrix  $R_1$  with size  $T_1 \times (K_1 + 1)$  is obtained where  $R_1(i, j)$  represents the rank at the  $i_1$ th time point of the  $j_1$ th voxel

- (3) Kendall harmony coefficient of the time series of  $(K_1 + 1)$ th voxel is calculated, as shown in the following formula

$$W_1 = \frac{\sum_{i=1}^T (S_1 R_1^i)^2 - T \overline{(S_1 R_1^i)}^2}{1/12(K_1 + 1)^2(T_i^3 - T)_i}, \quad (1)$$

where  $S_1 R_1^i = \sum_{j_1=1}^{K_1+1} r_{i,j_1}$  is the rank sum of time point  $i$ , and  $\overline{S_1 R_1^i} = (T_i + 1)(K_1 + 1)/2$  is the average value of  $S_1 R_1^i$ .  $W_1 (0 \leq W_1 \leq 1)$  represents the local consistency of voxel  $V_{j_1}(m_1, n_1, o_1)$ . The closer it is to 1, the greater the similarity of the time series is.

The voxels in the ReHo transform images show similar at the same time series. The greater Kendall's harmony coefficient is, the more similar these time series are. Mean ReHo (mReHo) is obtained by dividing the average ReHo value of the whole brain. After smoothing, the processed images are shown in Figure 4.

**2.4. Improved 3DPCANet.** PCANet [18] is a simple convolutional neural network, which mainly includes the principal component analysis (PCA), convolutional layer, binary hash, and block histogram. The objective of PCA is to achieve the eigenvector of the target matrix, and the eigenvector is taken as the convolution kernel parameter. The role of binary hashing and block histogram is indexing and pooling. In this paper, based on the traditional 3DPCANet, the max-pooling layer and ReLU layer are added behind each convolution layer to reduce the redundancy of image features after convolution and, similarly, to study texture features and nonlinear features of brain images. Subsequently, the classification accuracy improved.  $B$  training images  $\{\Gamma_j\}_{j=1}^B$  with size are considered as the input of 3DPCANet. Features are extracted process using 3DPCANet as follows.

- (a) Input layer

The  $e$  image of all the training samples is sided-cut into a block with the size of  $k \times k \times k$ . The  $L \times W \times H$  patch is produced by the  $e$  image. Subsequently, the patch is vectorized and standardized, i.e.,  $\overline{u_{e,1}}, \overline{u_{e,2}}, \dots, \overline{u_{e,o}}, \dots, \overline{u_{e,L \times H \times W}}$ , where  $\overline{u_{e,o}}$  represents the  $o$ th vector of  $e$ th image (as shown in Equation (2)).

$$\overline{u_{e,o}} = [u_1, u_2, \dots, u_{k \times k \times k}]. \quad (2)$$

All voxel patches are processed by Equation (2) to obtain matrix  $\overline{U}_e$ .

$$\overline{U}_e = [\overline{u_{e,1}}, \overline{u_{e,2}}, \dots, \overline{u_{e,o}}, \dots, \overline{u_{e,L \times H \times W}}], \quad (3)$$

where  $\overline{u_{e,1}}, \overline{u_{e,2}}, \dots, \overline{u_{e,o}}, \dots, \overline{u_{e,L \times H \times W}}$  are expressed different vectors from the same image. Whole training images are processed through Equation (3) to obtain matrix  $U$  (as shown in Equation (4)).

$$U = [\overline{U}_1, \overline{U}_2, \dots, \overline{U}_e, \dots, \overline{U}_B]. \quad (4)$$

Then, the matrix  $U$  is processed dimensionality reduction by PCA, and PCA minimizes the reconstruction error on a group of standard orthogonal filters, which is described as

$$\min_{V \in \mathbb{R}} \|U - VV^T U\|_F^2 \text{ s.t. } V^T V = I_{C_1}, \quad (5)$$

where  $I_{C_1}$  is the identity matrix with size  $C_1 \times C_1$ . The solution  $V$  of this formula is the eigenvector of  $UU^T$ . The expression of the PCA filter is as shown in Equation (6).

$$w_l^1 = \text{mat}(q_l(UU^T)) \in \mathbb{R}, l = 1, 2, \dots, C_1, \quad (6)$$

where  $\text{mat}$  is a function which maps the vector to the matrix  $w \in \mathbb{R}^{k \times k \times k}$ .  $q_l(UU^T)$  represents the  $l$ th feature vector of  $UU^T$ , and  $w_l^1 \in \mathbb{R}^{k \times k \times k}$  is the  $l$ th filter generated in the first step. The PCA filter is convolved with the  $j$ th training image  $\Gamma_e$  in the training image, which is expressed by

$$\Gamma_{e,l}^1 = \Gamma_e * w_l^1, \quad (7)$$

where the symbol “ $*$ ” represents convolution, and the filter  $w_l^1$  is used to convolve all  $B$  training images and to generate  $B * C_1$  images. Then, the max-pooling and ReLU are performed on the image  $\Gamma_{e,l}^1$  generated by formula (7), which is expressed by

$$\Pi_{e,l}^1 = \text{Re LU}(\Gamma_{e,l}^1 \otimes P^1), \quad (8)$$

where the symbol “ $\otimes$ ” represents the max-pooling operation.  $P^1$  denotes the max-pooling layer in the first step, and  $\Pi_{e,l}^1$  represents the image after the maximum pooling layer and ReLU layer are processed.

- (b) Middle layer

In the middle layer of PCA calculation,  $C_1$  images are generated using the  $e$ th image in the  $B$  training images by the first step, among which  $l$ th image is performed similar operations such as Equations (4), and matrix  $Y$  is gotten.

$$Y = [\overline{Y_{1,1}}, \overline{Y_{1,2}}, \dots, \overline{Y_{e,1}}, \overline{Y_{e,2}}, \dots, \overline{Y_{e,l}}, \dots, \overline{Y_{B,1}}, \overline{Y_{B,2}}, \dots, \overline{Y_{B,C_1}}]. \quad (9)$$

The filter  $w_h^2$  is obtained by PCA in the first stage on the matrix  $Y$ . The image  $\Pi_{e,l}^1$  generated by formula (8) in the first step is convolved by the obtained PCA filter, which is described by formula (10).

$$\Pi_{e,l,h}^2 = \Pi_{e,l}^1 * w_h^2. \quad (10)$$

Among  $B * C_1$  images generated in the first step, each image is used to generate  $C_2$  images by formula (8). The max-pooling and ReLU layer processing are performed on the image  $\Pi_{e,l,h}^2$  after convolution, as described by

$$\Omega_{e,l,h}^2 = \text{ReLU}(\Pi_{e,l,h}^2 \otimes P^2), \quad (11)$$

where  $P^2$  denotes the max-pooling layer in the middle layer, and the image  $\Omega_{e,l,h}^2$  is generated by max-pooling and ReLU layer operation.

### (c) Output layer

The Heaviside function  $H(\bullet)$  is used to binarize all  $B * C_1 * C_2$  images, and weighted processing is performed to get  $O_{e,l}$ .

$$O_{e,l} = \sum_{h=1}^{C_2} 2^{d-1} \{H(\Omega_{e,l,h}^2)\}. \quad (12)$$

Finally,  $\beta$  blocks with size  $k \times k \times k$  from each image  $O_{e,l}$  are divided in the form of overlapping or nonoverlapping. In the program, we use  $R$  to represent the overlapping rate between blocks. The histogram of each block is made statistics shown by

$$F_e = [\text{Bhist}(O_{e,1}), \text{Bhist}(O_{e,2}), \dots, \text{Bhist}(O_{e,i}), \dots, \text{Bhist}(O_{e,l})], \quad (13)$$

where  $\text{Bhist}(O_{e,i})$  is a function of block division, histogram statistics, and concatenation of image  $O_{e,i}$ .  $F_e$  represents the final eigenvector of the  $e$ th training images  $\Gamma_e$  using 3DPCANet.

The hyperparameters of the improved 3DPCANet include the size  $k \times k \times k$  of the block, filtering parameters  $C_1$  and  $C_2$  in each stage, the number  $\beta$  of blocks, and overlapping rate  $R$  between blocks in the output layer.

In conclusion, convolutional kernel parameters of improved 3DPCANet are studied by PCA, and improved 3DPCANet does not require back propagation. Therefore, improved 3DPCANet need not a host of dataset training. It is suitable for small datasets. Due to the small amount of AD data, improved 3DPCANet is used as the feature extraction model in this paper. We use improved 3DPCANet to extract features of the transformation images mALFF and

mReHo, respectively. Then, these two kinds of features are fused by CCA.

**2.5. Canonical Correlation Analysis.** CCA [25] is one of the algorithms which is used to find correlations between different kinds of data.  $X_2(n_2, m_2)$  and  $Y_2(n_3, m_2)$  is assumed to represent different kinds of datasets, respectively, where  $m_2$  represented the number of samples of the two datasets. Similarly,  $n_2$  and  $n_3$  is represented dimensions of two data features, respectively. CCA is used to reduce the dimension of  $X_2$  and  $Y_2$ . Likewise, the feature vectors  $X_2'(n'_1, m_2)$  and  $Y_2'(n'_1, m_2)$  of  $n'_1$ -dimension are obtained as described by

$$\begin{aligned} X_2' &= a_1^T X_2, \\ Y_2' &= b_1^T Y_2, \end{aligned} \quad (14)$$

where  $a_1$  and  $b_1$  represented the projection vectors of  $X_2$  and  $Y_2$ , respectively. The projection criterion of CCA is that when the number of dimensions of the two sets of data is reduced to  $n'$  dimension, the correlation coefficient of them is the largest. The objective function of CCA is showed by

$$\rho(X_2', Y_2') = \underbrace{\text{argmax}}_{a_1, b_1} \frac{\text{cov}(X_2', Y_2')}{\sqrt{D(X_2')} \sqrt{D(Y_2')}}, \quad (15)$$

where  $a_1$  and  $b_1$  is obtained by maximizing  $\rho(X_2', Y_2')$  to get the corresponding projection vectors  $X_2'$  and  $Y_2'$ .

In this paper, CCA was used to find the correlation features between the transform images including mALFF and mReHo of the same patient. The correlation features were fused. Subsequently, fused features were inputted into SVM classifier to achieve the classification of AD patients with different stages.

**2.6. SVM.** Support vector machine (SVM) [26] is a supervised learning classifier, and the maximal margin hyperplane of learning samples is obtained when making boundary decisions. The decision function of the SVM classifier is expressed by

$$\begin{aligned} T_m &= \min_{\omega, z, \zeta_m} \frac{1}{2} \|\omega\|^2 + C \sum_{m=1}^N \zeta_m, \\ \text{s.t. } T_m(\omega \cdot p_m + z) &\geq 1 - \zeta_m, \zeta_m > 0, m = 1, 2, \dots, N, \end{aligned} \quad (16)$$

where  $T_m$  represents the hyperplane, and  $\omega$  is the normal vector of the hyperplane.  $p_m$  denotes the eigenvector, and  $Z$  represents the bias.  $\zeta_m$  is the relaxation coefficient,  $C$  represents the penalty factor, and  $N$  represents the number of sample. Sequential minimal optimization (SMO) is the most common method to find the global optimal solution of SVM. In addition to SMO algorithm, other methods including

TABLE 2: Experimental results of traditional and improved 3DPCANet.

Methods	Criteria	NC/SMC	NC/MCI	SMC/MCI	SMC/AD	MCI/AD	NC/AD	EMCI/LMCI
mALFF+3DPCANet+SVM	Accuracy	82.61%	85.19%	90.00%	83.33%	77.27%	<b>80.00%</b>	81.48%
	Sensitivity	86.67%	<b>83.33%</b>	<b>100.00%</b>	<b>87.50%</b>	91.67%	86.67%	94.12%
	Specificity	75.00%	86.67%	75.00%	80.00%	60.00%	<b>70.00%</b>	60.00%
	F1	86.67%	<b>83.33%</b>	<b>92.31%</b>	<b>82.35%</b>	81.48%	<b>83.87%</b>	86.49%
	AUC	80.83%	<b>85.00%</b>	<b>87.50%</b>	<b>83.75%</b>	75.83%	<b>78.33%</b>	77.06%
mALFF+improved 3DPCANet+SVM	Accuracy	82.61%	85.19%	90.00%	83.33%	<b>86.36%</b>	76.00%	81.48%
	Sensitivity	86.67%	66.67%	91.67%	75.00%	91.67%	86.67%	94.12%
	Specificity	75.00%	<b>100.00%</b>	<b>87.50%</b>	<b>90.00%</b>	<b>80.00%</b>	60.00%	60.00%
	F1	86.67%	80.00%	91.67%	80.00%	<b>88.00%</b>	81.25%	86.49%
	AUC	76.70%	79.44%	84.38%	82.50%	<b>85.83%</b>	74.70%	77.06%
mReHo (0.01-0.08 Hz)+3DPCANet+SVM	Accuracy	86.96%	<b>77.78%</b>	85.00%	83.33%	77.27%	<b>84.00%</b>	81.48%
	Sensitivity	80.00%	75.00%	83.33%	87.50%	75.00%	86.67%	82.35%
	Specificity	<b>100.00%</b>	<b>80.00%</b>	87.50%	80.00%	80.00%	<b>80.00%</b>	80.00%
	F1	<b>88.89%</b>	75.00%	86.96%	82.35%	78.26%	<b>86.67%</b>	84.85%
	AUC	<b>90.00%</b>	<b>77.50%</b>	85.42%	83.75%	77.50%	<b>83.33%</b>	81.18%
mReHo(0.01-0.04 Hz)+ improved 3DPCANet+SVM	Accuracy	73.91%	68.75%	80.00%	72.22%	81.48%	80.00%	77.78%
	Sensitivity	73.33%	58.82%	82.35%	50.00%	<b>76.47%</b>	86.67%	94.12%
	Specificity	75.00%	80.00%	75.00%	90.00%	<b>90.00%</b>	70.00%	50.00%
	F1	78.57%	66.67%	84.85%	61.54%	<b>83.87%</b>	83.87%	84.21%
	AUC	74.17%	69.41%	78.68%	70.00%	<b>83.24%</b>	78.33%	72.06%
mReHo(0.01-0.08 Hz)+ improved 3DPCANet+SVM	Accuracy	82.61%	74.07%	<b>90.00%</b>	<b>88.89%</b>	77.27%	80.00%	<b>85.19%</b>
	Sensitivity	<b>93.33%</b>	<b>91.67%</b>	<b>91.67%</b>	87.50%	75.00%	86.67%	<b>94.12%</b>
	Specificity	62.50%	60.00%	87.50%	90.00%	80.00%	70.00%	70.00%
	F1	87.50%	<b>75.86%</b>	<b>91.67%</b>	<b>87.50%</b>	78.26%	83.87%	<b>88.89%</b>
	AUC	74.20%	66.67%	<b>92.71%</b>	<b>88.75%</b>	77.50%	76.70%	<b>82.06%</b>

TABLE 3: Experimental results of multimodal data fusion.

Methods	Criteria	NC/SMC	NC/MCI	SMC/MCI	SMC/AD	MCI/AD	NC/AD	EMCI/LMCI
mALFF+mReHo+improved 3DPCANet+tandem+SVM	Accuracy	73.91%	81.48%	85.00%	<b>94.44%</b>	81.82%	80.00%	70.37%
	Sensitivity	80.00%	83.33%	91.67%	<b>100.00%</b>	100.00%	73.33%	64.71%
	Specificity	62.50%	80.00%	75.00%	90.00%	60.00%	<b>90.00%</b>	80.00%
	F1	80.00%	80.00%	88.00%	<b>94.12%</b>	85.71%	81.48%	73.33%
	AUC	61.70%	62.22%	89.58%	<b>95.00%</b>	75.00%	76.70%	84.62%
mALFF+mReHo+improved 3DPCANet+CCA+softmax	Accuracy	73.91%	74.07%	85.00%	77.78%	77.27%	76.00%	77.78%
	Sensitivity	86.67%	83.33%	91.67%	75.00%	91.67%	80.00%	100.00%
	Specificity	50.00%	66.67%	75.00%	80.00%	60.00%	70.00%	40.00%
	F1	81.25%	74.07%	88.00%	75.00%	81.48%	80.00%	85.00%
	AUC	68.33%	75.00%	83.33%	77.50%	75.83%	75.00%	70.00%
mALFF+mReHo+improved 3DPCANet+CCA+SVM	Accuracy	<b>91.30%</b>	<b>88.89%</b>	<b>95.00%</b>	83.33%	<b>86.36%</b>	<b>92.00%</b>	<b>85.19%</b>
	Sensitivity	<b>93.33%</b>	<b>83.33%</b>	<b>91.67%</b>	75.00%	91.67%	<b>93.33%</b>	<b>94.12%</b>
	Specificity	<b>87.50%</b>	<b>93.33%</b>	<b>100.00%</b>	90.00%	<b>80.00%</b>	<b>90.00%</b>	70.00%
	F1	<b>93.33%</b>	<b>86.96%</b>	<b>95.65%</b>	80.00%	<b>88.00%</b>	<b>83.33%</b>	<b>88.89%</b>
	AUC	<b>81.70%</b>	<b>82.22%</b>	<b>92.71%</b>	82.50%	<b>85.83%</b>	<b>91.30%</b>	82.06%

elephant herding optimization [31] and krill herd algorithm [32] also can solve the above same problem. SVM has two design methods including one-to-one and one-to-many. In this paper, we choose one-to-one way for SVM classifier and SMO optimization algorithm.

### 3. Experimental Results and Analysis

In this paper, two kinds of image transformation such as mALFF and mReHo are used. The improved 3DPCANet is used for feature extraction, and SVM is used for

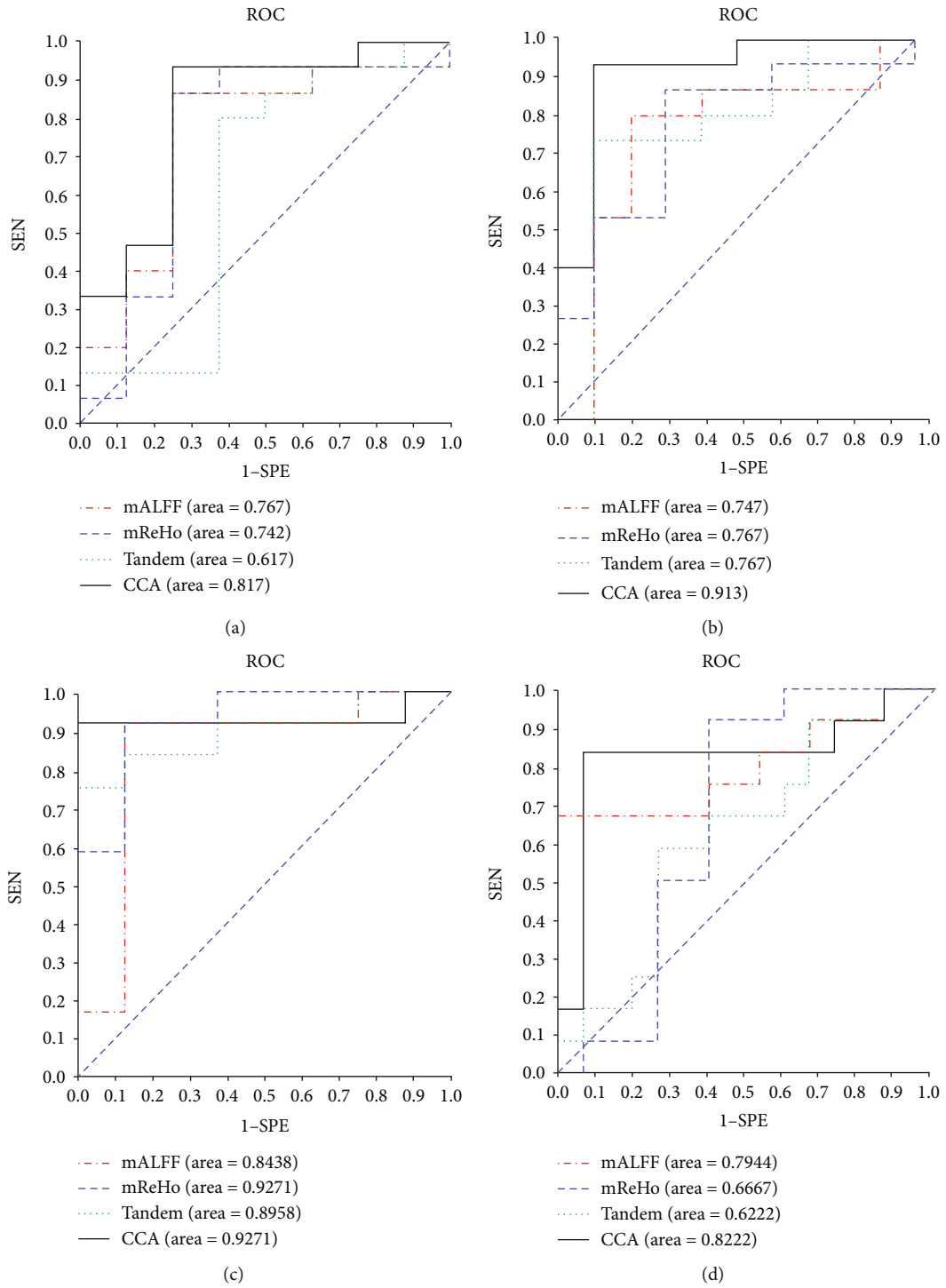


FIGURE 5: The ROC curves. (a) NC vs. SMC, (b) NC vs. AD, (c) SMC vs. MCI, and (d) MCI vs. AD.

classification. All the deep learning models used in this study were built using the pytorch framework, running on a server with a 1.7 GHz Intel Xeon E5-2603 v4 CPU, 16.0 GB RAM, NVIDIA RTX2070 GPU 8 GB, and the Windows 10 (64-bit version) operating system. The evaluation indicators used in this paper include accuracy, sensitivity, and specificity (as shown in Equation (17)).

$$\begin{aligned}
 \text{Accuracy} &= \frac{TP + TN}{TP + FN + FP + TN}, \\
 \text{Sensitivity} &= \frac{TP}{TP + FN}, \\
 \text{Specificity} &= \frac{TN}{FP + TN}.
 \end{aligned} \tag{17}$$

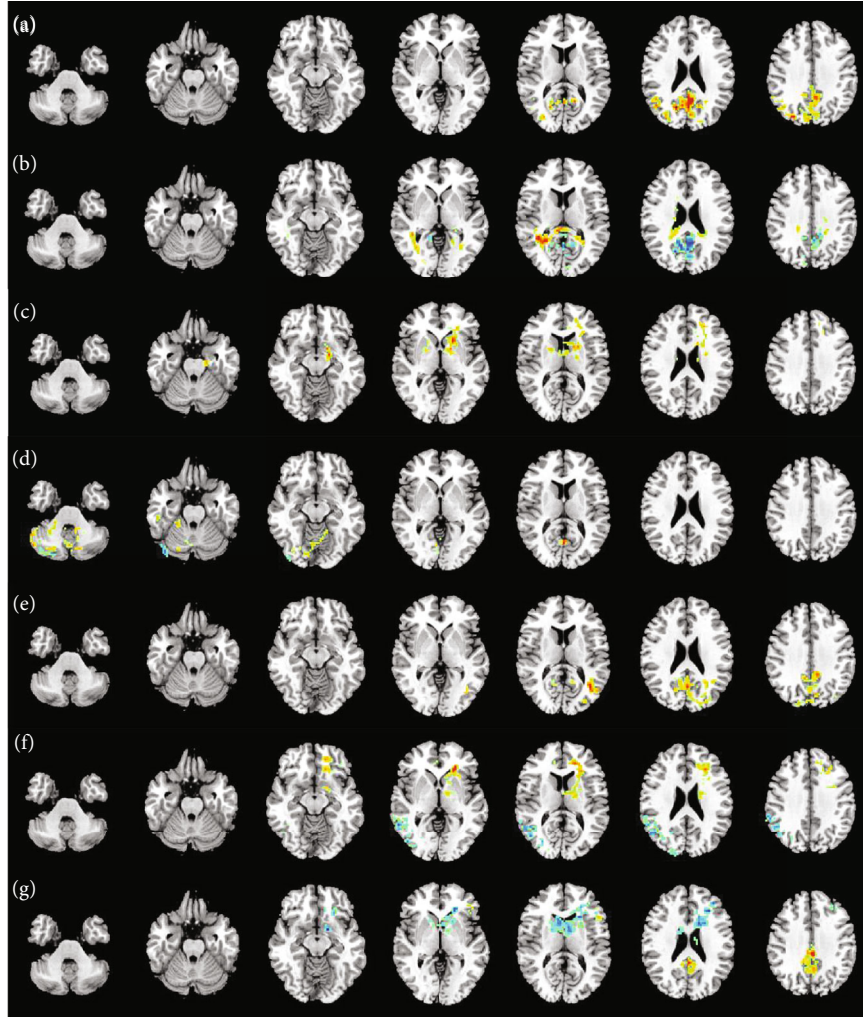


FIGURE 6: mReHo map. (a) MCI vs. AD; (b) NC vs. AD; (c) NC vs. MCI; (d) NC vs. SMC; (e) SMC vs. AD; (f) SMC vs. MCI; (g) EMCI vs. LMCI.

TP and TN, respectively, represent the number of true-positive and true-negative subjects. FP and FN, respectively, represent the number of false-positive and false-negative subjects. The positive class label is 1, and the negative class label is 0.

In the traditional experiment, only the accuracy is used as evaluation criteria of the model, and consequently, the algorithm performance cannot be roundly evaluated in numerous aspects. For original dataset in this paper, the positive class is far greater than the negative class. Imbalance of data is serious. Therefore, the  $F1$  value and the area under curve (AUC) of receiver-operating characteristic (ROC) and the coordinate axis are both used as the evaluation index, among which the  $F1$  is calculated by precision and sensitivity (as shown in Equation (18)).

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FP}, \\
 F1 &= \frac{2 * \text{Precision} * \text{Sensitivity}}{(\text{Precision} + \text{Sensitivity})}.
 \end{aligned} \tag{18}$$

#### (1) The role of improved 3DPCANet

Transformation images mALFF and mReHo were extracted features using improved 3DPCANet, and classifier was used SVM. For demonstrating the performance of improved 3DPCANet, the results, respectively, using traditional and improved 3DPCANet were shown in Table 2. The optimal results are obtained by parameter optimization of grid search method. Similarly, the hyperparameters  $k$ ,  $C_1$ ,  $C_2$ ,  $R$ , and  $\beta$  range of improved 3DPCANet are, respectively (2, 8), (2, 6), (2, 6), (0, 0.6), and (5, 25).

As can be seen from the results in Table 2 that for mALFF transformation, the experimental results were significantly improved with MCI vs. AD, other experiments also have different improvements on the basis of maintaining the original results. For mReHo transformation, results of 3 groups of data are improved significantly on all indicators in SMC vs. MCI, SMC vs. AD, and EMCI vs. LMCI. The experimental results show that using improved 3DPCANet, more discriminative image features can be extracted, and the classification results are effectively promoted because the max-pooling layer and ReLU layer are added behind



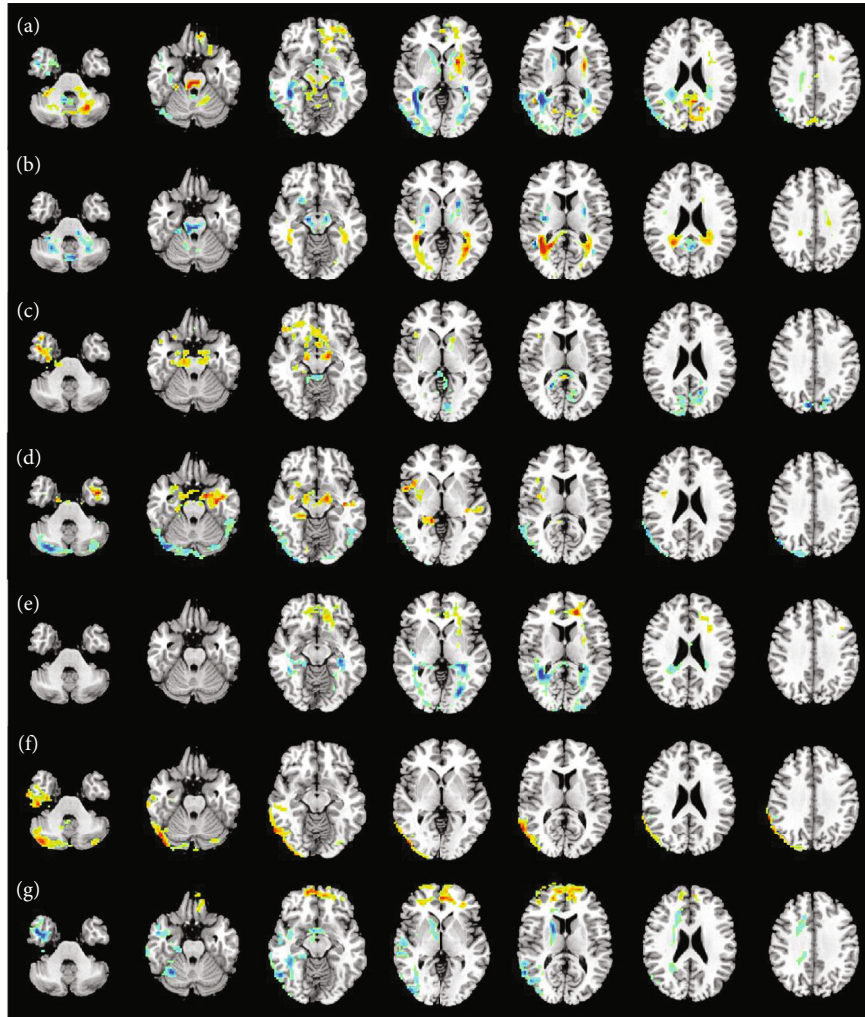


FIGURE 7: mALFF map. (a) MCI vs. AD; (b) NC vs. AD; (c) NC vs. MCI; (d) NC vs. SMC; (e) SMC vs. AD; (f) SMC vs. MCI; (g) EMCI vs. LMCI.

each convolution layer to reduce the redundancy of image features after convolution. So, the improved 3DPCANet is used in subsequent experiments for feature extraction. In addition, because the results on mReHo with 0.01-0.08 Hz are better than those with 0.01-0.04 Hz, in the subsequent experiments, mReHo transformation is used by 0.01-0.08 Hz.

## (2) The role of fusion strategies

To verify the effectiveness of the data fusion, multimodal data fusion and multiple classifiers are used in this paper. The detailed results are shown in Table 3.

As can be seen from the results in Table 3 that compared with single-modal, fusion method can effectively improve the experimental results. However, the results of direct fusion of two data in series are still poor due to feature redundancy. If CCA is used to fuse the features of two transformed images, better results for AD patients at different stages are obtained because CCA can find the most relevant classification features of the two images, and the fusion features enhance the classification discriminative power.

Among these results, the classification accuracy of SMC vs. MCI is 95.00%, and the  $F1$  value and AUC are 95.65% and 92.71%, respectively. Evaluation indicators of NC vs. SMC, NC vs. MCI, NC vs. AD, and MCI vs. AD have been improved compared with those of single modal. Obviously, more effective classification features of AD patients at different stages can be mined by CCA.

SMC and MCI are early stages of AD, so brain structure changes are small and clinical diagnosis can easily result in misdiagnosis. NC vs. MCI and SMC vs. MCI are classified with accuracy of 88.89% and 95.00%.  $F1$  are 86.96% and 95.65%, and AUC are 82.22% and 92.71%, respectively. The experimental results show that the proposed method in this paper can effectively classify the different stages of AD, including the initial stage that is difficult to diagnose. In addition, because the results on SVM are better than those with softmax classifier, in this paper, SVM is used in subsequent experiments.

Figure 5 is the ROC curve of NC vs. SMC, NC vs. AD, SMC vs. MCI, and MCI vs. AD. Each subimage includes 4

TABLE 4: Comparison results of different methods.

Methods	Dataset	Experiment		
		NC vs. AD	MCI vs. AD	NC vs. MCI
Peng et al. [27]	AD (49), MCI (93), NC (47)	<b>96.10%</b>	76.90%	80.30%
Khedher et al. [28]	AD (188), NC (229), MCI (401)	88.49%	85.41%	81.89%
Liu et al. [29]	AD (51), MCI (99), NC (52)	94.37%	—	78.80%
Zhu et al. [30]	AD (51), MCI (99), NC (52)	93.80%	—	79.70%
Li et al. [18]	AD (243), MCI (525), NC (307)	83.95%	82.53%	80.15%
Dai et al. [23]	AD (16), NC (27)	86.84%	—	—
The proposed method	AD (34), MCI (38), NC (50)	92.00%	<b>86.36%</b>	<b>88.89%</b>

groups of experiments, namely, mALFF and mReHo image classification, mALFF and mReHo tandem fusion classification, and mALFF and mReHo CCA fusion classification. Specificity is presented as the abscissa, and sensitivity is represented by the ordinate. It can be seen from the ROC curves in the 4 subimages that the proposed method in this paper has the largest AUC area compared with the single-modal method and direct fusion in series.

In this paper, mReHo map and mALFF map on MCI vs. AD, NC vs. AD, NC vs. MCI, NC vs. SMC, SMC vs. AD, and SMC vs. MCI are visualized by using REST Slice Viewer, respectively, as shown in Figures 6 and 7. By these maps, we can find the differences using two samples  $T$ -test.

It can be seen from Figures 6 and 7 that brain regions are influenced by mReHo transformation including precentral, calcarine, posterior cingulate cortex, cuneus, lingual, medial and paracingulate gyrus, superior occipital gyrus, fusiform, superior parietal gyrus, middle temporal gyrus, and hippocampus. And brain regions are influenced by mALFF transformation including fusiform, inferior temporal gyrus, hippocampus, middle occipital gyrus, calcarine, middle temporal gyrus, precentral, lingual, and cingulate gyrus. So, the brain regions such as fusiform, hippocampus, calcarine, middle temporal gyrus, precentral, and lingual contributes important role for classification, and these regions are also focused on in this paper.

### (3) Comparison results of different methods

In this paper, experimental results were compared with the state-of-the-art methods. As can be seen from the results in Table 4. Fewer experimental datasets were used to obtain better classification results in NC vs. MCI and MCI vs. AD, and the NC vs. AD experimental results are close to the results of other methods. Because CCA is used to perform dimensionality reduction and fusion processing on fMRI transformation images in the proposed method in this paper, it can effectively reduce feature redundancy and image noise. Therefore, the accuracy of image classification is increased, and the best experimental results are obtained.

## 4. Conclusion

In this paper, an AD classification method based on image transformation and features fusion is proposed. The main ideas include that, firstly, fMRI data, respectively, are made image transformation on mALFF and mReHo. Then, an

improved 3DPCANet for feature extraction in two kinds of transformation images is proposed and is, respectively, used to extract features, and these two kinds of features are fused by CCA. Finally, SVM is used to classify AD patients with different stages. SMO was used to find the global optimal solution for SVM. Besides the SMO method, some of the most representatively computational intelligence algorithms can also be used to solve the above problem, like monarch butterfly optimization (MBO), earthworm optimization algorithm (EWA), elephant herding optimization (EHO), moth search (MS) algorithm, slime mould algorithm (SMA), and Harris hawks optimization (HHO). The experimental results show that in the proposed method, improved 3DPCANet reduces feature redundancy and image noise, and texture and nonlinear features of brain images can be extracted, because the maximum pooling layer and ReLU layer are added behind each convolutional layer, which makes the classification features more abundant and robust. Compared with the single model method, if the fusion strategy of two fMRI features like CCA is used, better results can be obtained, which show that fusion strategy can assist medical personnel to accurately diagnose SMC, MCI, EMCI, LMCI, and AD patients.

## Data Availability

fMRI dataset used in this study comes from the Alzheimer’s Disease Neuroimaging Initiative (ADNI). The fMRI dataset includes 34 AD patients, 57 EMCI patients, 35 LMCI patients, 26 SMD patients, and 50 NC. Experimental data is obtained by sending an email to the ADNI and signing the related agreement. Since, in this laboratory, the classification of Alzheimer’s disease is studied by the fusion of fMRI and sMRI image information, the subjects possessing fMRI images and sMRI images are selected in the ADNI dataset. The link on ADNI dataset is <http://adni.loni.usc.edu/>.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

## Acknowledgments

This work is supported by the Joint Project of Beijing Natural Science Foundation and Beijing Municipal Education Commission (no. KZ202110011015).

## References

- [1] P. Scheltens, K. Blennow, M. M. B. Breteler et al., "Alzheimer's disease," *The Lancet*, vol. 388, no. 10043, pp. 505–517, 2006.
- [2] B. Reisberg, M. B. Shulman, C. Torossian, L. Leng, and W. Zhu, "Outcome over seven years of healthy adults with and without subjective cognitive impairment," *Alzheimers & Dementia*, vol. 6, no. 1, pp. 11–24, 2010.
- [3] A. J. Mitchell, H. Beaumont, D. Ferguson, M. Yadegarfar, and B. Stubbs, "Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: meta-analysis," *Acta Psychiatrica Scandinavica*, vol. 130, no. 6, pp. 439–451, 2014.
- [4] M. Rönnlund, A. Sundström, R. Adolfsson, and L.-G. Nilsson, "Self-reported memory failures: associations with future dementia in a population-based study with long-term follow-up," *Journal of the American Geriatrics Society*, vol. 63, no. 9, pp. 1766–1773, 2015.
- [5] E. Hosseini-Asl, G. Gimel'Farb, and A. El-Baz, "Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network," *Frontiers in Bioscience*, vol. 23, no. 3, pp. 584–596, 2016.
- [6] C. Lian, M. Liu, J. Zhang, and D. Shen, "Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 4, pp. 880–893, 2020.
- [7] H. Lao, X. Zhang, Y. Tang, and C. Liang, "Alzheimer's disease diagnosis based on the visual attention model and equal-distance ring shape context features," *IET Image Processing*, vol. 15, no. 4, pp. 1–12, 2021.
- [8] H. T. Gorji and N. Kaabouch, "A deep learning approach for diagnosis of mild cognitive impairment based on MRI images," *Brain Sciences*, vol. 9, no. 9, p. 217, 2019.
- [9] S. Moccia, R. Banali, C. Martini et al., "Development and testing of a deep learning-based strategy for scar segmentation on CMR-LGE images," *Magma Magnetic Resonance Materials in Physics Biology & Medicine*, vol. 32, no. 2, pp. 187–195, 2019.
- [10] A. Khazaei, A. Ebrahimzadeh, and A. Babajani-Feremi, "Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory," *Clinical Neurophysiology*, vol. 126, no. 11, pp. 2132–2141, 2015.
- [11] Y. Shi, W. Zeng, J. Deng, W. Nie, and Y. Zhang, "The identification of Alzheimer's disease using functional connectivity between activity voxels in resting-state fMRI data," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 8, pp. 1–11, 2020.
- [12] R. Gui, T. Chen, and H. Nie, "Classification of task-state fMRI data based on circle-EMD and machine learning," *Computational Intelligence and Neuroence*, vol. 2020, no. 3, pp. 1–10, 2020.
- [13] Y. Huang, J. Xu, Y. Zhou, T. Tong, X. Zhuang, and the Alzheimer's Disease Neuroimaging Initiative (ADNI), "Diagnosis of Alzheimer's disease via multi-modality 3D convolutional neural network," *Frontiers in Neuroence*, vol. 13, pp. 509–521, 2019.
- [14] J. Islam and Y. Zhang, "A novel deep learning based multi-class classification method for Alzheimer's disease detection using brain MRI data," *Springer*, vol. 10654, pp. 213–222, 2017.
- [15] F. Zhang, Z. Li, B. Zhang, H. du, B. Wang, and X. Zhang, "Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 361, pp. 185–195, 2019.
- [16] R. Jain, N. Jain, A. Aggarwal, and D. J. Hemanth, "Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images," *Cognitive Systems Research*, vol. 57, pp. 147–159, 2019.
- [17] T. H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: a simple deep learning baseline for image classification?," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5017–5032, 2015.
- [18] L. Shutong, X. Bin, and L. Weisheng, "Alzheimer's disease assisted diagnosis based on 3D-PCANet," *Computer Science*, vol. 45, no. 6, pp. 153–155, 2018.
- [19] E. A. DeYoe, P. Bandettini, J. Neitz, D. Miller, and P. Winans, "Functional magnetic resonance imaging (fMRI) of the human brain," *Journal of Neuroscience Methods*, vol. 54, no. 2, pp. 171–187, 1994.
- [20] Y. Han, J. Wang, Z. Zhao et al., "Frequency-dependent changes in the amplitude of low-frequency fluctuations in amnesic mild cognitive impairment: a resting-state fMRI study," *NeuroImage*, vol. 55, no. 1, pp. 287–295, 2011.
- [21] L. Jiang and X. N. Zuo, "Regional homogeneity," *The Neuroscientist*, vol. 22, no. 5, pp. 486–505, 2016.
- [22] T. Jiang, Y. He, Y. Zang, and X. Weng, "Modulation of functional connectivity during the resting state and the motor task," *Human Brain Mapping*, vol. 22, no. 1, pp. 63–71, 2004.
- [23] Z. Dai, C. Yan, Z. Wang et al., "Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3)," *NeuroImage*, vol. 59, no. 3, pp. 2187–2195, 2012.
- [24] C. G. Yan, X. D. Wang, X. N. Zuo, and Y. F. Zang, "DPABI: data processing & analysis for (resting-state) brain imaging," *Neuroinformatics*, vol. 14, no. 3, pp. 339–351, 2016.
- [25] R. M. Thorndike, "Canonical correlation analysis," in *Handbook of Applied Multivariate Statistics and Mathematical Modeling*, pp. 237–263, 2000.
- [26] C. Saunders, M. O. Stitson, and J. Weston, "Support vector machine," *Computer Science*, vol. 1, no. 4, pp. 1–28, 2002.
- [27] J. Peng, X. Zhu, Y. Wang, and D. S. Le An, "Structured sparsity regularized multiple kernel learning for Alzheimer's disease diagnosis," *Pattern Recognition*, vol. 88, pp. 370–382, 2019.
- [28] L. Khedher, J. Ramírez, J. M. Górriz, A. Brahim, and F. Segovia, "Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images," *Neurocomputing*, vol. 151, no. 1, pp. 139–150, 2015.
- [29] F. Liu, C. Y. Wee, H. Chen, and D. Shen, "Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's disease and mild cognitive impairment identification," *NeuroImage*, vol. 84, pp. 466–475, 2014.
- [30] X. Zhu, H. I. Suk, and D. Shen, "A novel matrix-similarity based loss function for joint regression and classification in AD diagnosis," *NeuroImage*, vol. 100, pp. 91–105, 2014.
- [31] J. Li, H. Lei, A. H. Alavi, and G. G. Wang, "Elephant herding optimization: variants, hybrids, and applications," *Mathematics*, vol. 8, no. 9, p. 1415, 2020.
- [32] G.-G. Wang, A. H. Gandomi, A. H. Alavi, and D. Gong, "A comprehensive review of krill herd algorithm: variants, hybrids and applications," *Artificial Intelligence Review*, vol. 51, no. 1, pp. 119–148, 2019.