



Alzheimer's disease classification from hippocampal atrophy based on PCANet-BLS

Chan Liang¹ · Huan Lao² · Tao Wei¹ · Xuejun Zhang^{1,2,3}

Received: 12 September 2020 / Revised: 18 February 2021 / Accepted: 14 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Alzheimer's Disease (AD) is a neurodegenerative disease of unknown etiology that progresses progressively and is currently incurable. It is more common in the elderly that seriously affects the physical and mental health of patients, thus early detection is very important for the prevention of AD progression. By using PCANet and Broad Learning System (BLS), we propose a novel method to identify Alzheimer's patients according to the clinical symptom of hippocampal atrophy, which is the most important indicator of AD. T1-weighted magnetic resonance images (MRIs) are used in this study, containing 207 patients with AD, 209 patients with mild cognitive impairment (MCI) and 109 cognitively normal (CN) cohorts from ADNI dataset. The left and right hippocampus are segmented from MRI at the first step, then the PACNet is applied to extract features from these images, finally the BLS is used to distinguish the different types of patients. Compared with the traditional machine learning methods, PCANet is able to extract the most informative features inside pictures effectively, while BLS can reach over 95% accuracy rate with lower time consuming. Experimental results have revealed that our method improves the performance of computer-aided diagnosis of Alzheimer's disease in both accuracy and speed of classification task.

Keywords Alzheimer's disease (AD) · Hippocampal atrophy · Magnetic resonance images · PCANet · Broad learning system (BLS) · Computer-aided diagnosis

1 Introduction

Alzheimer's disease (AD) is a degenerative disease of the brain and the most common type of dementia, which manifests in pathology as amyloid β precipitation and nerve fiber

✉ Xuejun Zhang
xjzhang@gxu.edu.cn

¹ School of Computer and Electronic Information, Guangxi University, Nanning, Guangxi, China

² College of Medical, Guangxi University, Nanning, Guangxi, China

³ Guangxi Key Laboratory of Multimedia Communications and Network Technology, Nanning, Guangxi, China

tangles [31]. The patient is characterized by memory impairment, accompanied by aphasia, cognitive impairment, and decreased ability to live. According to the 2020 World Alzheimer's Report, there are currently 5.8 million Alzheimer's patients over 65 years old in the United States, and the mortality rate of AD in 2018 is the sixth leading cause of death in the United States. The cost of patient care reached \$3050 [19]. Worldwide, the number of dementia patients is expected to reach 150 million by 2050. Clinically, the diagnosis of AD is after the appearance of dementia, and at this time most patients are already in the advanced stage of AD, and treatment at this stage is often ineffective. Therefore, improving the clinical treatment effect of AD must start from the early diagnosis of AD.

Mild cognitive impairment (MCI) has slight measurable changes in thinking ability, but it can still carry out daily activities, which is a direct transition period between healthy elderly and AD [23]. About 15% to 20% of people over 65 years of age suffer from MCI. Compared with healthy elderly people, people with MCI have a greater chance of developing AD or other types of dementia. Alex's review [30] shows that about 32% of MCI patients will develop into AD within 5 years, while the conversion rate of the normal elderly is only about 1%. Although MCI has a high risk of developing into AD, it can delay the progress if the patients can be detected and treated early. Therefore, the early diagnosis of MCI has important clinical and social significance for delaying the occurrence of AD.

With the rapid development of neuroimaging technology, it is possible to diagnose AD through neuroimaging. The images commonly used in clinical detection of AD include Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). Deep learning and machine learning are usually used to analyze the image and then build a prediction model. In the study of [4, 20, 26, 28], the researchers used deep neural network (DNN), sparse classifier (SRC), convolutional neural network (CNN) and random forest (RF) methods to analyze and identify AD based on MRI and PET images, and the classification accuracy rates are 94.23%, 92%, 96.86%, and 92.3%, respectively. Literature [25] uses a combination of support vector machine (SVM) and DNN techniques to detect AD, this method only uses MR images for analysis and achieves good results (AD classification accuracy reaches 98.74%). In addition, there are AD diagnosis based on CT images [15] and Electroencephalogram (EEG) [22]. All data used in the above experiments are from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset.

However, in actual diagnosis, it is not realistic to obtain multiple types of medical images at the same time, so how to use one medical image to accurately diagnose AD is still a research difficulty. At present, researchers have analyzed the MR images of AD patients and found that the volume reduction of certain parts of the brain can be used as a biomarker for AD. The work in the literature [5, 14, 21] shows that we can use it as a good biomarker for the diagnosis of AD by observing the reduction in the volume of white matter (WM) in the MR images. Some researchers have focused on the atrophy of other parts of the brain. Karow [18] and other studies found that the medial temporal lobe of AD patients changed significantly, and the hippocampal volume of patients in the MCI stage was reduced by 20%, which means that in the process of MCI to AD, the hippocampal atrophy is the most effective biomarker. Many other studies have shown that compared with healthy elderly people, the hippocampus of AD patients and MCI patients has significant atrophy. Chetelat et al. [9] used voxel-based morphometry to study AD, MCI and cognitive normal state (CN) and verified the research. In addition, their review [8] also pointed out that the most specific and sensitive feature between early diagnosed AD patients and MCI patients and healthy elderly people is hippocampus.

We decided to focus on one of the imaging results like most studies in the field of neuroimaging. Because MR images have the characteristics of high brain tissue resolution,

multi-directional imaging and no interference of craniocerebral artifacts, they can display the hippocampus more clearly, which is the best imaging method to study the morphological changes of the hippocampus. Therefore, this work mainly hopes to use the hippocampus information in the brain MRI image to quickly and accurately distinguish AD, MCI and CN patients to achieve the purpose of assisting clinical diagnosis.

We propose a model called PCANet-BLS, which uses PCANet to extract features from the hippocampus image, and then uses broad learning system (BLS) to perform classification. This is a simple and effective model and has achieved satisfactory diagnostic performance in our work (mainly reflected in the speed and accuracy of diagnosis). After that, we will introduce the methods we used and the results obtained.

2 Methodology

In this paper, the method of classifying AD based on PCANet-BLS mainly includes three stages: firstly, the 3D MR original images downloaded from ADNI were processed to obtain 2D images of the left and right hippocampus, and then used these images as the input data of PCANet to extract image features, and finally we classified these features by BLS. Next, we will introduce these steps in detail.

2.1 Pretreatment in MRI

All the 3D MR images in this experiment came from the ADNI dataset (adni.loni.usc.edu) which was founded by Michael W. Weiner in 2003 and concentrated on the study of progress of AD. We selected MR T1 weighted images of 550 subjects from ADNI for this study because the T1 weighted images can better observe the shape of human anatomy than the T2 weighted images. Table 1 shows the details of the subjects used in this experiment.

The original data in NIFTI format downloaded from ADNI were segmented by the open program CAT12, which was developed by Dr. Christian Gaser and Robert Dahnke from the Department of Psychiatry and Neurology at the University Hospital of Jena, Germany. First, registering all the 3D MR images into the MNI space (MNI152 T1 1.5mm brain) by Dartel registration to achieve spatial standardization [11, 16], then the skull of each MR image was removed and we obtained the gray matter (GM) and WM MR image of the voxels of size $121 \times 145 \times 121$. These operations can be implemented via the CAT12 “Segment Data” module. Then we used the AAL Atlas to make the left and right hippocampus masks according to the brain region number [2, 29], and the obtained masks were dotted with the GM maps of all 3D MRIs segmented to obtain the hippocampus of all 3D images. Finally, we used the MRIcro software to save each layer of the left and right hippocampus image as a 2D image in BMP format in the Transverse View section. MRIcro is an interactive software for viewing and analyzing brain images developed by Chris Rorden of the University of

Table 1 Subject information

Group	Males	Females	Age (Mean \pm AD)
n:207AD	97	110	75.34 \pm 8.08
n:209MCI	132	77	75.06 \pm 8.07
n:134CN	50	84	75.84 \pm 3.91

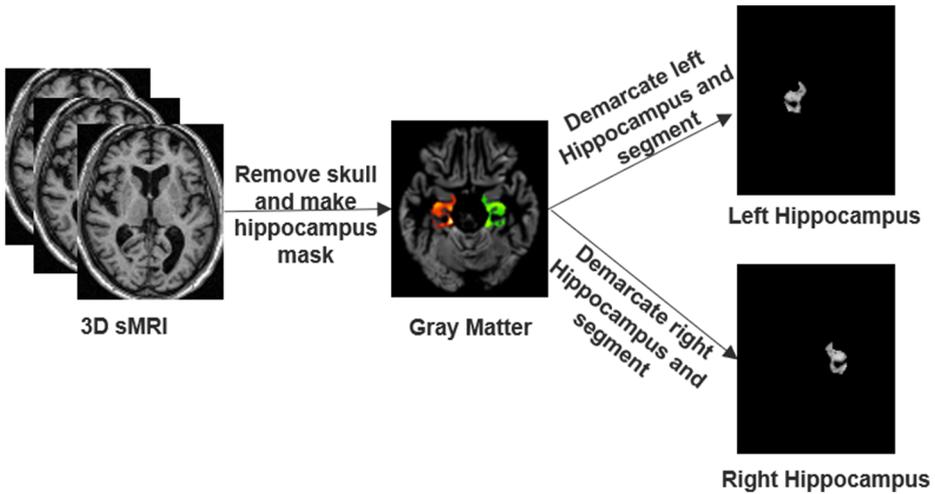


Fig. 1 Pretreatment in MRI

Nottingham. The preprocessing steps and results of the data are shown in the Fig. 1. In order to decrease computation complexity in the subsequent steps, we also cropped the 121×145 size image into a 40×40 size image.

2.2 Feature extraction based on PCANet

PCANet is a simple unsupervised deep learning network proposed by Chan et al. in 2015 [6], which uses principal component analysis of the mapping matrix instead of the convolution kernel in the CNN. Although PCANet has a very simple architecture, compared with CNN, there is no need to adjust a large number of parameters through training, so it can save a lot of time and cost. It has been successfully applied in various aspects such as character recognition [3], face recognition [17], and image classification [27] since it was put forward.

The PCANet used to extract image features in this article is composed of first stage, second stage and output layer. Figure 2 shows the process.

Suppose we have N sets of input images $\{I\}_{i=1}^N$ with $m \times n$ size, in the first stage, we select a block of size $k_1 \times k_2$ to overlay sampling each pixel of the picture and then cascade all the sampling results to obtain the characteristic picture of the picture, where the I_i picture can be expressed as:

$$x_{i,1}, x_{i,2}, \dots, x_{i,\tilde{m}\tilde{n}} \in \mathbf{R}^{k_1 \times k_2} \quad (1)$$

Where $\tilde{m} = m - [k_1/2]$, $\tilde{n} = n - [k_2/2]$.

Next, we need to zero-sample the sampled block, and I_i is expressed as:

$$\overline{\mathbf{X}}_i = [\overline{x_{i,1}}, \overline{x_{i,2}}, \dots, \overline{x_{i,\tilde{m}\tilde{n}}}] \quad (2)$$

After all the pictures are processed in the same way, we obtained the final training sample matrix:

$$\mathbf{X} = [\overline{\mathbf{X}}_1, \overline{\mathbf{X}}_2, \dots, \overline{\mathbf{X}}_N] \in \mathbf{R}^{k_1 \times k_2 \times N\tilde{m}\tilde{n}} \quad (3)$$

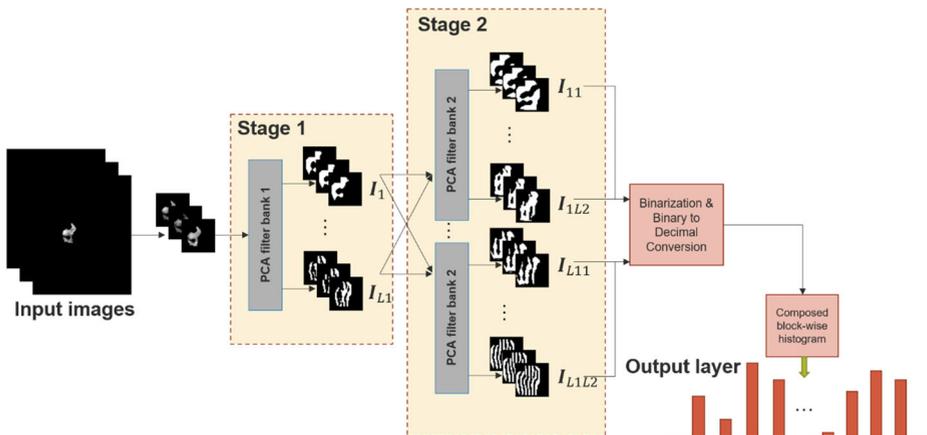


Fig. 2 Illustration of how PCANet extracts features from images through the three simplest processing components: PCA filters, binary hashing, and histograms

Assume that the convolution kernel number of this layer is L_1 , the feature mapping matrix extracted from the L_1 principal eigenvectors of the covariance matrix of \mathbf{X} is:

$$\mathbf{W}_l^1 = \text{mat}_{k_1, k_2} \left(q_l(\mathbf{X}\mathbf{X}^T) \right) \in \mathbf{R}^{k_1 \times k_2}, l = 1, 2, \dots, L_1 \tag{4}$$

\mathbf{W}_l^1 represents the PCA filters obtained by calculation. Finally, the output feature matrices of the first layer are obtained by convolution of the input picture with the PCA filters:

$$\mathbf{I}_i^l = \mathbf{I}_i * \mathbf{W}_l^1, i = 1, 2, \dots, N \tag{5}$$

The mapping process of the second stage is almost the same as the first layer, we used the convolution output of the previous stage as the input of the second stage, and it is also sampled by the same size block, then cascaded the matrix and averaged it:

$$\overline{\mathbf{Y}}_i^l = [\overline{y_{i,l,1}}, \overline{y_{i,l,2}}, \dots, \overline{y_{i,l,m\bar{n}}}] \tag{6}$$

$$\mathbf{Y}^l = [\overline{\mathbf{Y}}_1^l, \overline{\mathbf{Y}}_2^l, \dots, \overline{\mathbf{Y}}_N^l] \in \mathbf{R}^{k_1 \times k_2 \times N\bar{m}\bar{n}} \tag{7}$$

Each input matrix undergoes this process to obtain the block samples of second layer:

$$\mathbf{Y} = [\mathbf{Y}^1, \mathbf{Y}^2, \dots, \mathbf{Y}^{L_1}] \in \mathbf{R}^{k_1 \times k_2 \times N\bar{m}\bar{n}} \tag{8}$$

Similarly, assume that there are L_2 filters in the second layer, we can calculate the PCA filters of the second layer by the following formula:

$$\mathbf{W}_l^2 = \text{mat}_{k_1, k_2} \left(q_l(\mathbf{Y}\mathbf{Y}^T) \right) \in \mathbf{R}^{k_1 \times k_2}, l = 1, 2, \dots, L_2 \tag{9}$$

Because the L_1 output matrices generated by the L_1 filters in the first layer are separately calculated on the L_2 filters in the second layer, the $L_1 \times L_2$ output feature matrices are obtained after each image undergoes two layers of operations:

$$\mathbf{O}_i^l = \left\{ \mathbf{I}_i^l * \mathbf{W}_i^2 \right\}_{l=1}^{L_2} \tag{10}$$

In essence, with the above steps we obtain many large matrices and in order to solve this problem, the output layer performs binarization and hash coding on each output matrix in the front which coding bits is the number of filters in the second stage:

$$\boldsymbol{\tau}_i^l = \sum_{l=1}^{L_2} z^{l-1} \mathbf{H} \left(\mathbf{I}_i^l * \mathbf{W}_i^2 \right) \tag{11}$$

Where $\mathbf{H}(\cdot)$ is a Heaviside step function. Finally, each output matrix from the first stage is divided into B blocks, and the histogram information of the decimal value of each block is calculated, then concatenated all histogram features to obtain a vector as the final feature extraction result.

2.3 Using broad learning system for classification of AD

The BLS was proposed by C. L. Philip Chen et al. in 2018 [7] based on the traditional random vector functional-link neural networks (RVFLNN), compared with the deep neural networks such as CNN, it has unique advantages in time consumption and solving nonlinear, small and medium sample problems. We used this system as a classifier in the experiment.

Assume that we have a K class classification problem and each class contains D-dimensional data which compose the input data $\mathbf{X} \in \mathbf{R}^{K \times D}$, BLS first maps the input data X into n mapping features by the following formula:

$$\mathbf{Z}_i = \Phi \left(\mathbf{XW}_{ei} + \boldsymbol{\beta}_{ei} \right), i = 1, 2, \dots, n \tag{12}$$

Where \mathbf{W}_{ei} and $\boldsymbol{\beta}_{ei}$ are the randomly generated weights, and Φ is the activation function. Using $\mathbf{Z}^n = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n]$ to denote the first n groups of mapping features, m denotes the number of enhancement nodes, then the j-th group of enhancement nodes \mathbf{H}_j can be calculated by the following formula:

$$\mathbf{H}_j = \xi \left(\mathbf{Z}^n \mathbf{W}_{hj} + \boldsymbol{\beta}_{hj} \right), j = 1, 2, \dots, m \tag{13}$$

Where \mathbf{W}_{hj} and $\boldsymbol{\beta}_{hj}$ are also randomly generated weights, and ξ is the activation function. Assume that each feature map and enhancement feature generate p and q nodes respectively, then the result of BLS feature can be expressed as:

$$\mathbf{A} = \left[\mathbf{Z}^n | \mathbf{H}^m \right] \in \mathbf{R}^{N \times (np \times mq)} \tag{14}$$

The target of BLS is to learn to find the best \mathbf{w} so that the sum of squared loss of prediction errors can be minimized, which makes it face the following problems:

$$\min_{\mathbf{W}} J_{BLS} = \frac{1}{2} \|\mathbf{W}\|^2 + \frac{C}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{W}\|^2 \tag{15}$$

The first term is the regularization term to prevent overfitting. Finally, the optimal solution of \mathbf{w} can be calculated by:

$$\mathbf{W}_{BLS}^* = \left(\mathbf{I} + C\mathbf{A}^T\mathbf{A} \right)^{-1} \mathbf{A}^T\mathbf{Y} \tag{16}$$

Where \mathbf{I} is an identity matrix. The basic BLS network is shown in Fig. 3.

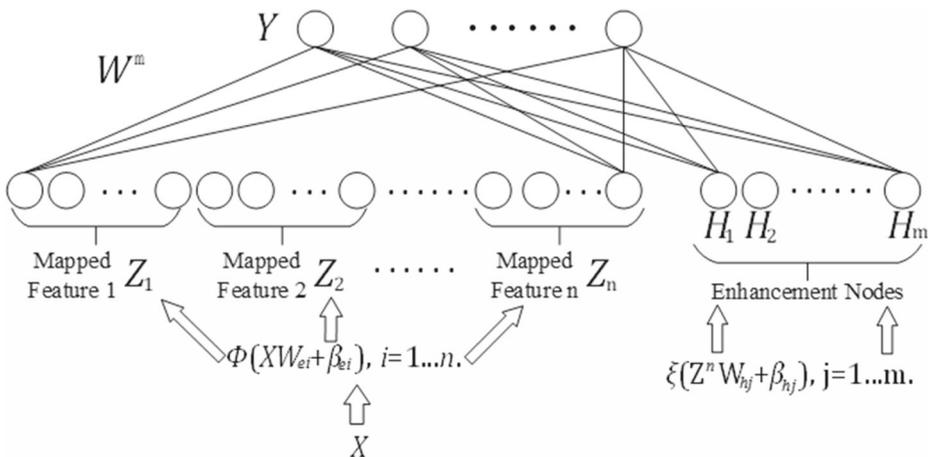


Fig. 3 The structure of BLS

In the experiment, we use comparative experiments to determine the number of layers, the size of the sampling block, the value of B of the PCANet, and the number of feature nodes and enhancement nodes in the BLS to obtain a PCANet-BLS network with appropriate parameters.

3 Experimental results and discussions

3.1 Performance metrics and evaluation method

In this paper, four indicators of accuracy (ACC), sensitivity (SEN), specificity (SPEC), and sample test time were used for verifying the performance of all methods.

1. ACC represents the proportion of the number of subjects who are identified accurately to the total number of subjects;
2. SEN represents the proportion of the number of people who are predicted to be sick and they are actually sick to the number of people who are predicted to be patients;
3. SPEC represents the proportion of the number of people who are predicted to be not sick and they are actually not sick to the number of people who are predicted to be not sick;
4. the sample test time represents the time required to predict the state of a patient using the trained model.

Table 2 is a confusion matrix for the classification results, where 1 represents patients and 0 represents healthy. From the table we can obtain the calculation formulas of ACC, SEN and SPEC:

$$\left\{ \begin{array}{l} ACC = \frac{TP+TN}{TP+TN+FP+FN} \\ SEN = \frac{TP}{TP+FN} \\ SPEC = \frac{TN}{TN+FP} \end{array} \right. \quad (17)$$

Table 2 Classification result confusion matrix

Actual class	Predicted class 1	Actual class 0
1	TP (True Positive)	FN (False Negative)
0	FP (False Positive)	TN (True Negative)

10-fold cross-validation was used as the evaluation method, which represents a method of dividing the data set during training and testing in the experiment. First it randomly divides the data set \mathbf{X} into ten mutually exclusive subsets of similar size of $\mathbf{X} = \mathbf{X}_1 \cup \mathbf{X}_2 \cup \dots \cup \mathbf{X}_{10}$, $\mathbf{X}_i \cap \mathbf{X}_j = \emptyset (i \neq j, i, j \in [1, 10])$, then use the union of the 9 subsets as the training set each time, the remaining subset as the test set, and repeat the previous process 10 times, we take the average of the ten test results as the final result.

3.2 Experimental results and comparison

3.2.1 Experimental results of binary classification

In order to evaluate the recognition performance of the entire network, many traditional classification algorithms were used for comparison in the feature classification experiment. The classification algorithms used in the experiment include SVM, RF, Decision Tree (DT), which were compared with BLS. The following tables show the results of binary classification experiments using different combinations of features (left hippocampus, right hippocampus, left hippocampus + right hippocampus) and different classifiers.

In Table 3, We evaluated the performance of classification experiments between the AD and CN groups. Among them, the SVM model make the most accurate estimation algorithm for the current data set, which achieved the classification accuracy of 97.65%, sensitivity of 98.53%, specificity of 96.17%. The worst results were obtained by the DT. It is worth noting that BLS is not as good as SVM in terms of estimation accuracy, but training time

Table 3 Classification performance of AD-CN

		RH	LH	RH+LH	Execution time(s)	Predict time(s)
SVM	ACC(%)	95.00	97.35	97.65	47.03	0.1156
	SEN(%)	95.42	97.72	98.53		
	SPEC(%)	94.85	96.94	96.17		
RF	ACC(%)	90.87	91.82	93.99	19.97	0.1396
	SEN(%)	94.24	95.00	98.15		
	SPEC(%)	87.50	88.66	89.82		
DT	ACC(%)	76.31	79.96	81.45	1.1795	0.00067
	SEN(%)	82.37	83.44	82.87		
	SPEC(%)	70.25	76.48	80.04		
BLS	ACC(%)	94.13	96.80	97.08	0.7877	0.01974
	SEN(%)	93.61	97.69	98.59		
	SPEC(%)	95.59	96.00	95.09		

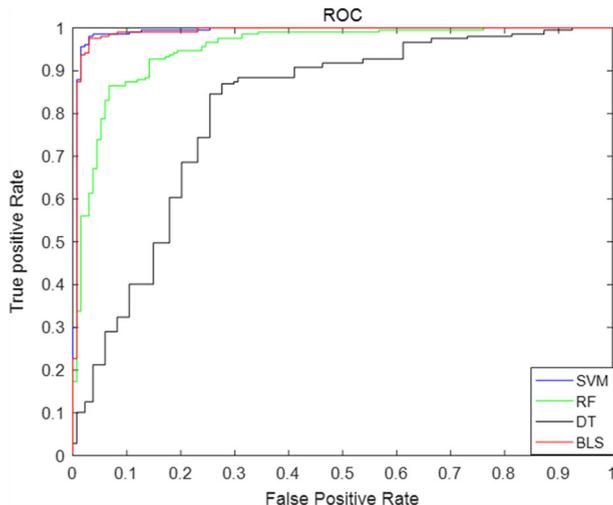


Fig. 4 ROC curves of different methods of AD-CN groups

and test time performance of the method is excellent. In addition, in each algorithm, the left hippocampus provides more effective information than the right hippocampus, but the combination of the left hippocampus and the right hippocampus can get the best algorithm model most of the time, which increase the classification accuracy by 3.15%, 1.49% compared to left hippocampus when we used the RF and DT respectively.

In addition, we also use ROC and AUC values to evaluate various methods. In order to facilitate comparative analysis, we draw the ROC curves of the results of different classification algorithms in the same coordinate graph, as shown in Fig. 4. For the classification

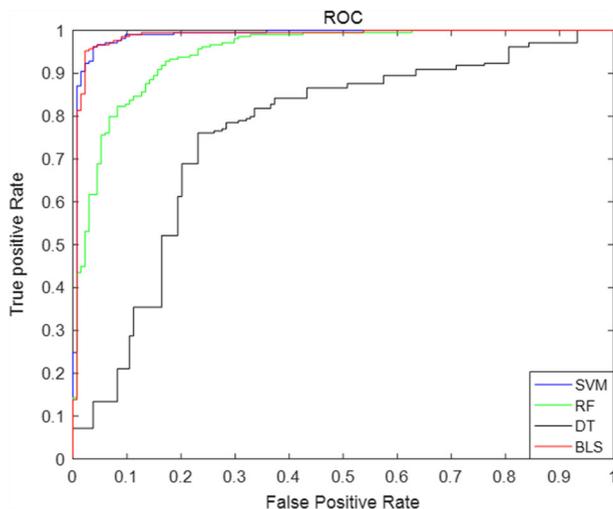


Fig. 5 ROC curves of different methods of MCI-CN groups

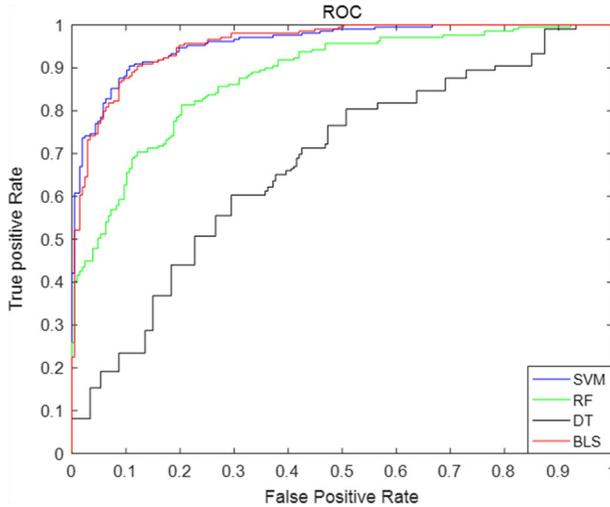


Fig. 6 ROC curves of different methods of AD-MCI groups

results of different groups, we can find that there are some differences in the results of different classification methods. Among them, the ROC curve of SVM and BLS is closer to the upper left corner of the coordinate, that is, the corresponding AUC value is the largest.

Figures 5 and 6 are the ROC curves of the classification results of the MCI-CN and AD-MCI groups, respectively. Judging from the classification results of these three groups, the classification performance of BLS is no worse than that of SVM. Under the premise of ensuring the classification performance, the speed is also greatly improved than SVM. The AUC values of RF and DT are significantly lower than the other two.

Table 4 Classification performance of MCI-CN

		RH	LH	RH+LH	Execution time(s)	Predict time(s)
SVM	ACC(%)	93.88	95.92	96.50	59.44	0.145478
	SEN(%)	96.60	98.00	98.61		
	SPEC(%)	90.74	92.15	93.31		
RF	ACC(%)	84.95	87.36	85.68	18.29	0.1035
	SEN(%)	96.48	93.80	96.23		
	SPEC(%)	73.42	80.91	75.12		
DT	ACC(%)	74.10	77.12	75.53	1.4253	0.00062
	SEN(%)	78.26	81.94	80.61		
	SPEC(%)	69.94	72.31	70.45		
BLS	ACC(%)	94.75	95.06	95.65	0.9329	0.021822
	SEN(%)	97.15	97.40	98.57		
	SPEC(%)	90.91	92.31	91.41		

Table 5 Classification performance of AD-MCI

		RH	LH	RH+LH	Execution time(s)	Predict time(s)
SVM	ACC(%)	87.97	89.41	91.82	107.334	0.2568
	SEN(%)	85.84	87.55	90.49		
	SPEC(%)	90.84	90.82	93.24		
RF	ACC(%)	81.50	82.87	81.10	19.907	0.1191
	SEN(%)	84.72	81.25	80.87		
	SPEC(%)	87.50	88.66	89.82		
DT	ACC(%)	65.63	67.26	64.07	2.153	0.00066
	SEN(%)	66.73	66.34	63.09		
	SPEC(%)	64.53	68.18	65.04		
BLS	ACC(%)	87.96	88.48	91.60	2.854	0.0365
	SEN(%)	85.25	86.28	89.96		
	SPEC(%)	90.51	91.30	93.19		

Table 4 evaluates the algorithm prediction performance of the MCI-CN group. For the performance of different classifiers, we have obtained almost the same results as the AD-CN group. For RF and DT, the feature information of the left hippocampus and the right hippocampus does not enhance the effect of the algorithm, and the left hippocampus is the most effective value. For BLS, the effective information provided by the combination of these two features is basically the same as that of the left hippocampus. In addition, the overall classification accuracy of MCI-CN is lower than that of AD-CN. This is due to the milder symptoms of MCI compared with AD, which makes the diagnosis more difficult. The results we got so far are consistent with literature research.

In Table 5, we performed a classification experiment on the AD-MCI group. The results show that our model can accurately classify MCI and AD using information from the left hippocampus and right hippocampus, with an accuracy rate of 91.60% and a sensitivity of 89.96%, the specificity is 93.19%. This result is similar to that obtained by using SVM, but our network execution time and test time are far better than SVM. In addition, DT has the worst performance in this classification, and the performance of using combined features is not as good as using left hippocampus information alone.

3.2.2 Experimental results of multiclass classification

In the research on the early diagnosis of Alzheimer's disease, most of the experiments are binary classification experiments because compared with normal healthy elderly, AD patients and MCI patients are better distinguished. In order to evaluate the performance of the classifier, we also did a multiclass classification experiment of AD, MCI and CN. The experimental results are shown in Table 6.

Table 6 Experiment results of multiclass classification

Method	Features	ACC(%)	SEN(%)	SPEC(%)
PCANet-BLS	RH+LH	87.38	82.24	93.56

3.2.3 Experimental comparison

To evaluate the performance of our approach, we compared the classification accuracy with previous methods using the same dataset. R Jain [24] proposes to use CNN architecture based on transfer learning to detect AD and MCI, researchers use VGG-16 as a feature extractor for the classification task and the accuracy of the 3-way classification using the described method is 95.73% for the validation set. Farooq et al. [12] uses a CNN based pipeline for multi-class classification of AD and late mild cognitive impairment (LMCI) and MCI and CN. Three-dimensional convolutional neural networks (3D-CNNs) were applied with magnetic resonance imaging (MRI) to execute binary and ternary disease classification models [13]. Aderghal et al. [1] and Ebrahim et al. [10] proposed to use an improved CNN network to execute binary classification.

It can be found from the Table 7 that PCANet-BLS has general performance for multi-classification, but shows the best performance for binary classification. Compared with the

Table 7 Comparison with CNN methods

Method	Subjects	Classification type	ACC(%)	SEN(%)	SPEC(%)
R Jain et al. [24]	50AD 50MCI 50CN	AD-MCI-CN	95.73		
Farooq et al. [12]	33AD 22LMCI 49MCI 45CN	AD-LMCI-MCI-CN	98.88	97.90	99.60
Feng et al. [13]	135AD 130MCI 133CN	AD-MCI-CN	93.71	96.82	96.73
Proposed	50AD 50MCI 50CN	AD-MCI-CN	87.38	82.24	93.56
Ebrahim et al. [10]	2519AD 2560CN	AD-CN	97.49	97.46	
Aderghal et al. [1]	48AD 108MCI 58CN	AD-CN MCI-CN AD-MCI	92.50 80.00 85.00	94.70 92.80 93.70	90.40 73.00 79.10
Feng et al. [13]	135AD 130MCI 133CN	AD-CN MCI-CN AD-MCI	99.10 98.90 89.40	99.80 98.90 86.70	98.40 98.80 84.00
Proposed	207AD 209MCI 134CN	AD-CN MCI-CN AD-MCI	97.08 95.65 91.60	98.59 98.57 89.96	95.09 91.41 93.19

Table 8 Experimental results of different numbers of network layers

Layers	Execution time(s)	ACC (%)	SEN (%)	SPEC (%)
1	19.25	90.06	91.79	88.41
2	41.14	97.08	98.59	95.09
3	96.43	97.78	98.96	96.44
4	120.58	97.68	98.65	96.26

current deep learning methods, PCANet-BLS can achieve better classification performance without using operations such as data augmentation and network transfer learning, which means that our method is simple and efficient.

3.3 Parameters discussion

In order to obtain better network performance in this experiment, it is necessary to set appropriate parameters in PCANet and BLS networks. The parameters that PCANet needs to set mainly include the number of network layers, the size of the sampling block k_1 , k_2 , the number of filters per layer L , the block size B of the statistical block histogram information of the last layer, and the block overlap ratio (BOR) size of each block. The parameters that the BLS network needs to set are the number of feature nodes and enhancement nodes.

Because the number of layers of the deep learning network determines the complexity of the network, too many layers will lead to a large feature dimension and increase time complexity. Too few layers will cause the problem of poor data fitting ability, so it is very necessary to choose the appropriate number of network layers. The author pointed out in the original paper that they observed that two-stage PCANet is usually sufficient to achieve good performance, and a deeper architecture does not necessarily lead to further improvement [6]. In this experiment, we study the network layer number based on the AD-CN group, s which set the number of cascaded PCA layers to 1,2,3,4, and the number of filters ($L_1 = L_2 = 8$). The experimental results are shown in the Table 8. After many experiments,

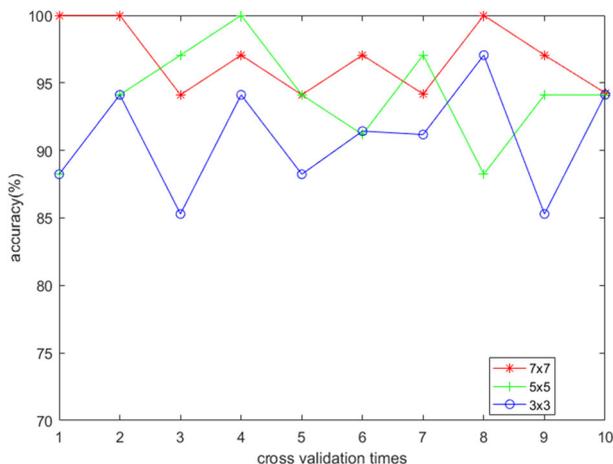
**Fig. 7** Experimental results of different size of sampling block

Table 9 Experimental results of different size of B

BOR = 0, $k_1 \times k_2 = 7 \times 7$			
B	ACC (%)	SEN (%)	SPEC (%)
10×10	95.61	95.94	94.83
14×14	96.18	96.47	96.67
18×18	96.18	96.62	95.26
22×22	95.90	96.85	95.48
26×26	95.88	96.23	96.02
30×30	97.35	97.72	96.94
34×34	94.42	94.79	94.35
38×38	94.72	95.00	95.03

it was found that the accuracy of the three-layer network was 0.7% higher than that of the two-layer network. However, the execution time increased by more than 2 times, so in order to balance the time and accuracy, we choose the two-layer network model finally.

Then we take the classification of AD and CN as an example to set the value of k_1 and k_2 : we use the block sizes of 3×3 , 5×5 , 7×7 respectively to extract features from the left hippocampus pictures of AD and CN in the case of $B=30 \times 30$, BOR=0(blocks do not overlap at all), and then use BLS to perform classification experiments. The results of 10-fold cross-validation are shown in the Fig. 7 and we can observe that the features extracted using the 7×7 block size can obtain better results, so we will use this parameter in the following experiments.

Next consider the influence of the size of block B and the block overlap ratio on the experimental results. We take BOR=0 (blocks do not overlap at all) and $k_1 \times k_2 = 7 \times 7$, B starts from 10 to 38 which interval is 4 to extract the features of the picture, and then use BLS to perform the classification experiment. Table 9 shows the experimental results.

From the Table 9, We can find that the size of B begins from 10. The accuracy, sensitivity, and specificity of the experiments get better as B increases. When B = 30, the result is the best, then the results start to deteriorate with the B increase. So we choose B = 30×30 finally.

And we let B = 30×30 and $k_1 \times k_2 = 7 \times 7$ to discuss the effect of different BOR values on the experimental results. Here, BOR starts from 0 (blocks do not overlap at all), and the interval 0.1 takes a value once, until BOR=0.5 (block overlaps by half). From the experimental results in Table 10, we can find that different BORs do not have a great impact

Table 10 Experimental results of different size of B

B = 30×30 , $k_1 \times k_2 = 7 \times 7$			
BOR	ACC (%)	SEN (%)	SPEC (%)
0	97.35	97.72	96.94
0.1	97.08	97.18	97.20
0.2	97.06	97.60	96.88
0.3	97.07	96.55	97.73
0.4	97.08	97.22	96.67
0.5	96.78	96.71	97.35

on the experimental results. Since the larger the BOR value, the more the number of B, the more the amount of calculation. Considering this problem, we take $BOR = 0$.

As for the parameters setting of BLS, from our own experiments, it was found that the change of the number of feature nodes and enhancement nodes within a certain range does not significantly improve the experimental results, as shown in the Table 11. So we set the number of feature nodes to 100, the number of enhancement nodes is 50. The experiments were all based on the above parameters. Our experiments are implemented by MATLAB R2016b on a notebook computer of 2.5 GHz Intel Core i5-7300HQCPU and 8.0 GB RAM.

3.4 Discussions

The main focus of this work is to use MR imaging for early predictive diagnosis of Alzheimer's disease. The focus of our work lies in how to extract effective information from the images to help the diagnosis. Since the hippocampus is regarded as one of the most affected structures in many types of dementia and changes the most during the disease, we mainly study the impact of hippocampal atrophy on AD.

The process of extracting features used a network called PCANet, which is similar to the deep learning network. And BLS was used as classifier. Judging from the experimental results of the AD-CN, MCI-CN and AD-MCI diagnostic groups, our method got satisfactory classification diagnostic performance. In the experiment, we used different classifiers and different combinations of features to conduct the experiment. In general, SVM achieved the best prediction results in the diagnosis, but the time consumption is also the longest, which is fatal in clinical diagnosis, while the BLS performs well in both accuracy and speed. In addition, we can also observe that the left hippocampus can provide slightly better accuracy than the right hippocampus. In most cases, using the left hippocampus and the right hippocampus together can get the best results. But in the case of our pursuit of speed and efficiency, the parameters of the left hippocampus can be used as effective indicators for predicting and diagnosing AD.

Although our work has achieved good performance, in the future, the following aspects can also be considered: firstly, MCI patients also include stable MCI (sMCI) and progressive MCI (pMCI), of which pMCI will convert to AD. At present, the classification of sMCI and pMCI is still a research problem, we can consider using our method to classify them. Secondly, we can consider applying our method to other aspects. In addition, our work has two main limitations: 1. From the experimental results, our method performs better in the binary classification experiment, and the experimental results in the multi-classification

Table 11 Classification accuracy with different numbers of feature nodes and enhancement nodes

Feature nodes	Enhancement nodes	ACC (%)
100	50	97.08
100	100	97.12
100	150	96.85
100	200	96.67
200	50	97.19
400	50	96.74
1000	50	97.06
2000	50	97.14

task are average. 2. Our method is more suitable for small and medium datasets. For large datasets, the computing time of PCANet will increase significantly.

4 Conclusions

The challenge of AD classification makes us need to use various advanced methods to get better results. This paper proposes a PCANet-BLS method for AD classification. PCANet is similar to deep learning networks. It is a simplified convolutional neural network method that uses the mapping matrix of principal component analysis instead of the convolutional kernel of convolutional neural networks and there is no need to adjust a large number of parameters during the training process, so features can be extracted quickly. BLS uses an incremental learning algorithm, which can improve system performance by adding hidden nodes on the basis of the original broad learning system, without retraining the network and saving a lot of time. In this paper, 550 cases were selected from the ADNI dataset, and the parameters were modified through multiple comparison experiments to identify the most suitable PCANet-BLS. The results of the experiments show that our proposed method has a classification accuracy of 97.08%, 95.65%, and 91.60% for AD-CN, MCI-CN, and AD-MCI in a short time, which is an effective and accurate clinical diagnosis method.

Acknowledgements This work was supported by the Science and Technology Key Projects of Guangxi Province (Grant No.2020AA21077007 and 2021JJYGD170060); and the Innovation Project of Guangxi Graduate Education (Grant No. YCSW2020064).

References

1. Aderghal K, Khvostikov A, Krylov A, Benois-Pineau J, Afdel K (2018) Catheline, G.: Classification of alzheimer disease on imaging modalities with deep cnns using cross-modal transfer learning. pp 345–350
2. Ahmed OB, Benois-Pineau J, Allard M, Amar CB, Catheline G (2015) Classification of alzheimer's disease subjects from mri using hippocampal visual features. *Multimed Tools Appl* 74(4):1249–1266
3. Aly S, Mohamed A (2019) Unknown-length handwritten numeral string recognition using cascade of pca-svmnet classifiers. *IEEE Access* 7(4):52024–52034
4. Andres O, Lozano F, Gorriz JM, Javier R, Martinez MFJ (2017) Discriminative sparse features for alzheimer's disease diagnosis using multimodal image data. *Curr Alzhmer Res* 14(1)
5. Balthazar MLF, Yasuda CL, Pereira FR, Pedro T, Cendes F (2010) Differences in grey and white matter atrophy in amnesic mild cognitive impairment and mild alzheimer's disease. *Eur J Neurol* 16(4):468–474
6. Chan TH, Jia K, Gao S, Lu J, Zeng Z, Ma Y (2015) Pcanet: A simple deep learning baseline for image classification? *IEEE Trans Image Process* 24(12):5017–5032
7. Chen CLP, Liu Z (2018) Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE Trans Neural Netw Learn Syst* 29(99):10–24
8. Chetelat GAL, Baron JC (2003) Early diagnosis of alzheimer's disease: contribution of structural neuroimaging. *Neuroimage* 18(2):525–541
9. Chetelat GAL, Desgranges B, Vincent DLS, Viader F, Eustache F, Baron JC (2002) Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment. *Neuroreport* 13(15):1939–1943
10. Ebrahim D, Ali-Eldin AMT, Moustafa HE, Arafat H (2020) Alzheimer disease early detection using convolutional neural networks. In: 2020 15th international conference on computer engineering and systems (ICCES)
11. Erik O, Newell F, Mitchell K (2013) Combined structural and functional imaging reveals cortical deactivations in grapheme-color synaesthesia. *Front Psychol* 4(2):755
12. Farooq A, Anwar SM, Awais M, Rehman S (2017) A deep cnn based multi-class classification of alzheimer's disease using mri. In: 2017 IEEE international conference on imaging systems and techniques (IST)

13. Feng W, Halm-Lutterodt NV, Tang H, Mecum A, Guo X (2020) Automated mri-based deep learning model for detection of alzheimer's disease process. *Int J Neural Syst* 30(6):2050032
14. Frings L, Yew B, Flanagan E, Lam B, Hüll M, Huppertz H, Hodges J, Hornberger M (2014) Longitudinal grey and white matter changes in frontotemporal dementia and alzheimer's disease. *PLoS ONE* 9(4):e90814
15. Gao XW, Hui R (2016) A deep learning based approach to classification of ct brain images. In: *Sai computing conference*
16. Grabner G (2006) Symmetric atlas and model based segmentation: an application to the hippocampus in older adults
17. Huang J, Yuan C (2015) Weighted-pcanet for face recognition. In: *International conference on neural information processing*
18. Karow DS, Mcevoy LK, Fennema-Notestine C, Hagler DJ, Jennings RG, Brewer JB, Hoh CK, Dale AM (2010) Relative capability of mr imaging and fdg pet to depict changes associated with prodromal and early alzheimer disease. *Radiology* 256(3):932
19. Longhe Z (2020) 2020 alzheimer's disease facts and figures. *Alzheimer's Dement* 16(3)
20. Lu D, Popuri K, Ding GW, Balachandar R, Beg MF, Initiative ADN (2018) Multimodal and multiscale deep neural networks for the early diagnosis of alzheimer's disease using structural mr and fdg-pet images. *Entropy* 20(1):5697
21. Migliaccio R, Agosta F, Possin KL, Rabinovici GD, Miller BL, Gorno-Tempini ML (2012) White matter atrophy in alzheimer's disease variants. *Alzheimer's Dement* 8(5):S78–S87.e2
22. Morabito FC, Campolo M, Ieracitano C, Ebadi JM, Bramanti P (2016) Deep convolutional neural networks for classification of mild cognitive impaired and alzheimer's disease patients from scalp eeg recordings. In: *IEEE international forum on research and technologies for society and industry leveraging a better tomorrow*
23. Morris JC, Storandt M, Miller JP, Mckeel DW, Berg L (2001) Mild cognitive impairment represents early-stage alzheimer disease. *JAMA Neurol* 58(3):397–405
24. Jain NJ, Aggarwal A, Hemanth D (2019) Convolutional neural network based alzheimer's disease classification from magnetic resonance brain images. *Cogn Syst Res* 57:147–159
25. Raza M, Awais M, Ellahi W, Aslam N, Le-Minh H (2019) Diagnosis and monitoring of alzheimer's patients using classical and deep learning techniques. *Expert Syst Appl*
26. Sarraf S, Tofighi G (2016) Classification of alzheimer's disease using fmri data and deep learning convolutional neural networks
27. Shi J, Wu J, Li Y, Zhang Q, Ying S (2017) Histopathological image classification with color pattern random binary hashing based pcnet and matrix-form classifier. *IEEE J Biomed Health Inform* 1–1
28. Tripoliti EE, Fotiadis DI (2008) Argyropoulou M.: A supervised method to assist the diagnosis and classification of the status of alzheimer's disease using data from an fmri experiment
29. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage* 15(4):273–289
30. Ward A, Tardiff S, Dye C, Arrighi HM (2013) Rate of conversion from prodromal alzheimer's disease to alzheimer's dementia: A systematic review of the literature. *Dement Geriatr Cogn Disorders Extra* 3(1):320–332
31. Wilson RS, Segawa E, Boyle PA, Anagnos SE, Hizez LP, Bennett DA (2012) The natural history of cognitive decline in alzheimer's disease. *Psychol Aging* 27(4):1008–1017