Received: 20 February 2020 | Revised: 6 July 2020 | Accepted: 24 March 2021

IET Image Processing

DOI: 10.1049/ipr2.12218

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

Alzheimer's disease diagnosis based on the visual attention model and equal-distance ring shape context features

Huan Lao[1] | **Xuejun Zhang**[2,3] | **Yanyan Tang**[4] | **Chan Liang**[2]

[1] Medical College, Guangxi University, Nanning, PR China

[2] School of Computer, Electronics and Information, Guangxi University, Nanning, PR China

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

[4] Department of Neurology, Guangxi Medical University First Affiliated Hospital, Nanning, PR China

Correspondence
Xuejun Zhang, School of Computer, Electronics and Information, Guangxi University, Nanning, Guangxi 530004, PR China.
Email: xjzhang@gxu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: Grant No. 81760324; Natural Science Foundation of Guangxi Province, Grant/Award Number: Grant No. 2017JJA170765y; Innovation Project of Guangxi Graduate Education, Grant/Award Number: Grant No.YCBZ2020026

Abstract

Alzheimer's disease (AD) is an irreversible neurodegenerative disease caused by rapid degeneration of brain cells. More and more researchers focus on effective and accurate methods for the diagnosis of AD. In this paper, a method to identify AD by extracting equal-distant ring shape context features from saliency map of structural magnetic resonance imaging (sMRI) is proposed. The experimental results on the thin-layer MR images of the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset showed that our method helped improve the performance of identifying brain diseases. Specifically, the classification accuracy of 94.83% for AD versus CN, 98.31% for AD versus MCI and 85.77% for MCI versus CN, respectively. At the same time, experiments on Open Access Series of Imaging Studies dataset and clinically collected thick-layer MR images verify the classification performance of the method. The results show that this method may have higher application value in clinical application, with classification accuracies of 96.56% and 98.18% for AD versus CN, respectively. Compared with the methods based on gray matter (GM) density, cortical thickness and hippocampal volume, our method achieved higher accuracy of AD (or MCI) and CN classification.

1 | INTRODUCTION

Alzheimer's disease (AD) is a common neurodegenerative disease, mainly manifested as cognitive function impairment, behavioral disorder and mental abnormality, which has seriously affected the daily life of the elderly [1]. At present, there are at least 50 million AD patients or other types of dementia patients in the world. With the aging of the global population, AD patients are expected to double by 2050 [2, 3]. In 2018, the cost of treatment and care for AD patients in the world had reached trillions of dollars, which had brought a heavy economic burden on patient's family and society. Because the damage of central nervous system in patients with AD is irreversible, there is no effective clinical treatment at present, and the only treatment scheme is to delay the progression of the disease as far as possible. Mild cognitive impairment (MCI) is an intermediate stage between normal aging and AD [4–6]. Studies have shown that if MCI is not treated as early as possible, further decline of cognitive function will lead it to develop into AD. Although MCI has a high risk of developing into AD, if it can be detected and treated as soon as possible, the condition of MCI patients does not necessarily develop to AD [7–9]. Therefore, early detection, diagnosis and treatment of MCI can delay the progression of AD, which has important clinical and social significance [10].

Typically, the accurate diagnosis of AD, MCI, and cognitive normal cohorts (CN) by clinicians depend heavily on neuropsychological tests, such as Mini-Mental State Examination (MMSE), Functional Activities Questionnaire (FAQ) and Clinical Dementia Rating (CDR). But neuropsychological tests are more subjective and only suitable for patients who have

some clinical symptoms. With the rapid development of medical imaging modalities, magnetic resonance imaging (MRI), Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) images has played an important in diagnosis of brain diseases [11–20]. However, the diagnosis of AD by clinicians using medical imaging mainly depends on a large number of clinical training and experience, and has a certain degree of subjectivity, which will affect the accuracy of AD diagnosis. Extensive clinical training and experience make it more difficult for new or inexperienced clinicians to diagnose AD. Therefore, with the development of machine learning, the study of computer-aided diagnosis (CAD) methods that accurately classify AD (or MCI) and CN by combining medical imaging and machine learning has become a hot topic.

Although these modalities of new imaging techniques have been used in the auxiliary diagnostic research of AD in recent years, structural MRI (sMRI) is still an effective aid to assist in the early diagnosis of AD. The classification performance of CAD methods based on sMRI and machine learning usually depends on features, and therefore feature extraction becomes a critical step in the classification framework. The most commonly used feature extraction methods for sMRI data include voxel-based methods and region of interest (ROI)-based methods.

Voxel-based methods extract voxel-wise imaging features from the whole brain sMRI to construct classifiers for distinguishing AD (or MCI) from CN. For example, Klöppel et al. [21] used voxel-based morphometry (VBM) to generate a gray matter (GM) density map of the whole brain as the input of support vector machine (SVM) classifier to train a classification model to diagnose AD. Li et al. [22] used a multivariate method to extract six cortical features of each participant, and used a linear SVM model to classify MCI and NC. Chu et al. [23] mainly extracted features from GM segmentation map of T1 MR images, and excluded voxels less than 0.2, which have 299,477 voxels in each subject, served as the input features to the feature selection and SVM classifier. Salvatore et al. [24] proposed a method based on machine learning for AD (or MCI) and CN classification, which used PCA to reduce the density of GM and white matter (WM) images to get the features.

ROI-based methods employ imaging features extracted from brain regions, while these regions are usually pre-determined based on biological prior knowledge or anatomical brain atlases. For example, Zhang et al. [25] and Kim et al. [26] proposed a method to classify AD (or MCI) and CN by SVM. The input features of these methods are calculated as the volume of GM within 93 ROIs on MRI and PET images, as well as the original values of three cerebrospinal fluid (CSF) measurements. Wee et al. [27] constructed a regional cortical thickness similarity map for each subject to describe the relative changes in cortical thickness between ROI pairs, which can significantly improve the classification performance of AD. In the literature, [28] proposed a new method for feature selection of hybrid voxel-wise, which combine *t*-test and genetic algorithm based on Fisher's criterion. The method of Tong et al. first used sparse representation techniques to calculate grading biomarkers for each MCI subject. Then, the grading biomarkers are combined with age and cognitive indicators to provide a more accurate prediction of MCI to AD conversion [29]. Liu et al. [30] first non-linearly register each sMRI separately onto multiple pre-selected atlases, and then extract multiple sets of atlas features for this MR image to construct ensemble classification models for AD/MCI diagnosis. Sørensen et al. [31] extracted the volume and texture features of the hippocampal in sMRI as the input of the SVM classifier to classify AD. Ahmed et al. proposed two AD classification methods based on hippocampal features: in literature [32], they used the circular harmonic functions (CHFs) to extract local features from the hippocampus and posterior cingulate cortex (PCC) to learn SVM classifiers for AD/MCI diagnosis. In literature [33], the classifiers trained independently based on hippocampal and CSF features were combined, followed by another classifier to further refine the diagnostic performance. Zhao et al. [34] also proposed an AD classification method based on hippocampal features. They extracted 56 features of each hippocampus, including intensity, shape, texture and wavelet features.

In general, the features of voxel-based methods usually have high dimensionality, which may cause potential overfitting problems. Therefore, the classification performance of voxel-based methods largely depends on the dimensionality reduction methods. The ROI-based features have low feature dimension and can extract the features of ROIs which are highly related to disease. Therefore, the feature ROI based can improve the classification performance and has been widely used [35]. In this paper, we propose a feature extraction method of equal-distant rings shape context (EDRSC), including saliency map detection, equal-distant rings segmentation and shape context algorithm based on chessboard distance. First, the ROIs of sMRI are extracted, including left and right hippocampus. Second, a visual attention model called PFT (Phase spectrum of Fourier Transform) model is exploited to detect saliency map of ROIs. Then, the shape contour of saliency map is segmented by equal-distant rings, and then the EDRSC features of saliency map are extracted. Finally, Support Vector Classification (SVC) is used to build a disease classification model. The main contributions are that (1) The ring based on equal distance can be uniformly divided the whole shape contour, and the distribution of all points on the image contour can be uniformly reflected. (2) In the rectangular coordinate system, the set of the pixels in each ring are determined by the chessboard distance, and the spatial position information of the whole target shape can be obtained, including the distance information and direction information. (3) Attention selection is an important mechanism of human visual perception. It is a conscious activity in which human chooses and keeps important information from a large amount of information input from the outside world and ignores useless or secondary information. Therefore, in this paper, the visual attention model is used to detect the saliency map of MR images, and the selection of salient regions containing important information to construct EDRSC features can improve the classification performance of our method. (4) The experimental results on the thin-layer MR images of the open datasets (ADNI and Open Access Series of Imaging Studies [OASIS]) and the thick-layer MR images collected in

TABLE 1 Characteristics of the ADNI dataset subjects used in this study

Diagnosed	Number of subjects	Gender(M/F)	Age (Mean ± SD)	MMSE (Mean ± SD)	CDR (Mean ± SD)
AD	95	57/38	75.91 ± 7.38	22.78 ± 2.65	0.88 ± 0.35
MCI	195	132/63	75.50 ± 7.40	27.02 ± 2.39	0.51 ± 0.15
CN	158	85/73	77.06 ± 4.92	29.04 ± 1.27	0.03 ± 0.11

Note. MMSE, Mini-Mental State Examination; SD, standard deviation.

clinic showed that regardless of the thin-layer MR images or the thick-layer MR images, which can get good classification results and have higher application value in clinical application.

2 | MATERIALS AND IMAGE PRE-PROCESSING

All thin-layer 3D images in this study are downloaded from the public dataset Alzheimer's Disease Neuroimaging Initiative (ADNI), which included MR images of subjects with AD, MCI and CN (The URL of the dataset is adni.loni.usc.edu). ADNI dataset was launched in 2003 to connect researchers with research data. The dataset is divided into four stages: ADNI1, ADNI/GO, ADNI2 and ADNI3. It collects a large amount of MRI and PET images, genetic data, blood biochemical indicators and CSF data. The primary goal of ADNI is to verify and determine the relationship between the collected data, determine the progression of AD, and provide the basis for early diagnosis and treatment of AD. ADNI's research protocol was approved by the local institutional review board. The study protocol is specifically as follows: ADNI recruited more than 800 adults including CN subjects, AD subjects and MCI subjects to participate in the study, where age range is 55–90 years (all subjects signed a written informed consent). Specifically, 200 CN subjects were observed for 3 years, and 400 MCI subjects and 200 AD subjects were followed up for 3 years and 2 years, respectively.

2.1 | Subjects

In our method, the images of 448 subjects in the ADNI1/GO stages are mainly selected for the experiments. These MR T1-weighted images are acquired using MPRAGE or equivalent protocols of different resolutions with a slice thickness of 1.2 mm, which have undergone several pre-processed steps of research groups belonged to the ADNI. In detail, first, the geometric distortion of the image caused by the gradient model was corrected,and then the B1 non-uniformity of the image intensity was corrected. Finally, the N3 histogram peak sharpening algorithm was applied to reduce the intensity non-uniformity of the image. The detailed statistical of all research subjects in our method are shown in Table 1, which are divided into three different classes.

CN: They were the normal control group collected by ADNI who did not have depression, MCI, or other dementia. The MMSE score of these subjects is 24 to 30, with the CDR score of 0 [36, 37].

MCI: They have no significant other cognitive impairments and maintain their daily activities. The MMSE score is between 24 and 30, and CDR of 0.5.

AD: They were the ADNI collection of subjects identified as AD who met the NINCDS/ADRDA criteria for possible AD [38]. The MMSE score of AD is between 20 and 26, and CDR of 0.5 to 1.

2.2 | MR image pre-processing

The pre-processing is divided into three steps: tissue segmentation, discriminative ROIs extraction, and image slices generation. In this section, we will describe these three steps in detail.

1) Tissue segmentation: Many studies have shown that the main morphological and structural abnormalities of AD are GM in the brain. Therefore, the accuracy of CAD system is largely dependent on brain tissue or structural segmentation, such as GM or WM tissue section. In this study, all of original 3D MR images downloaded from the ADNI dataset in the NIFTI format are segmented using the CAT12 (dbm.neuro.uni-jena.de/cat/) toolkit running on MATLAB (mathworks.cn) software. CAT12 is a MATLAB toolkit based on SPM12 (fil.ion.ucl.ac.uk/spm/), which developed by Christian Gaser and Robert Dahnke of Departments of Psychiatry and Neurology at Jena University Hospital, Germany. The tissue segmentation procedure can be implemented via the module 'Segment Data'. This is mainly to register all 3D MR images into the MNI space (MNI152 T1 1.5mm brain) by Dartel registration to achieve spatial standardization [39, 40]. Finally, the skull of each MR image is removed, and the GM MR image of size 121 × 145 × 121 voxels are obtained. The results are shown in Figure 1a.

2) Discriminative ROIs extraction: In general, AD and age-matched healthy old people have apparent morphological structural abnormalities compared to their brain structures, including the volume reduction of the hippocampus and the increase in the ventricles. Therefore, in this study, we performed *t*-tests on different categories of subjects to obtain areas of severe atrophic brain atrophy associated with AD. It can be obtained from Figure 1b that the discriminative ROIs are the left and right hippocampus region ($p < 0.0001$). Therefore, the left and right hippocampus are used as a set of distinguish the ROIs in order to better compare AD (or MCI) and CN. The specific ROIs extraction steps are as follows: First, in order to extract the most discriminative ROIs of the hippocampus, we use the AAL Atlas to make the left and right hippocampus mask according to the brain region number [33, 41]. Then, multiply the obtained mask

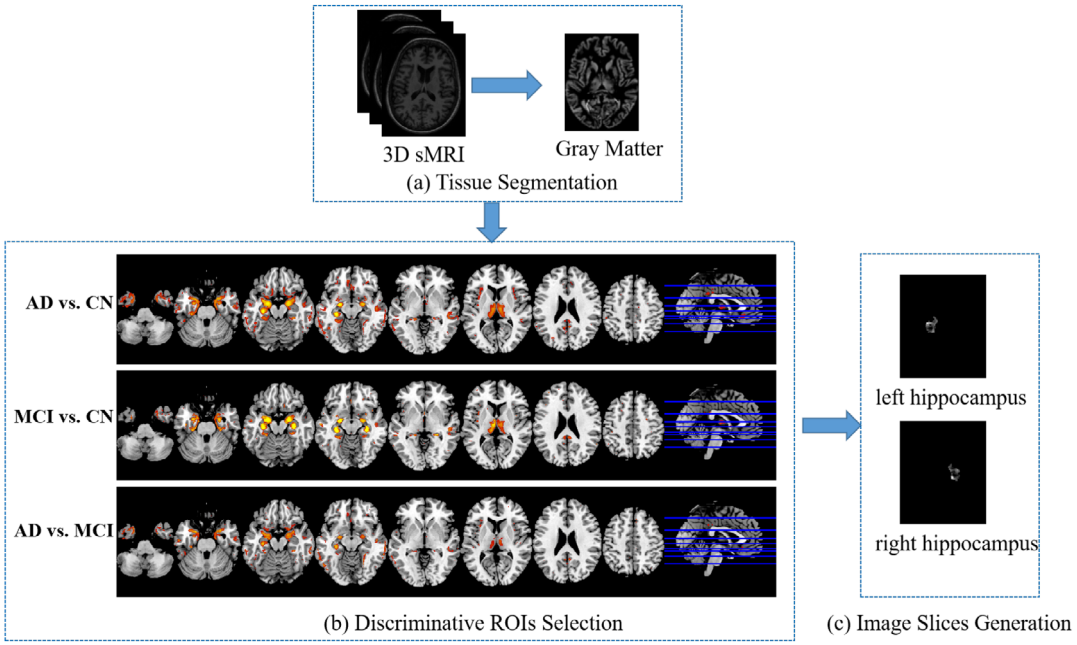


FIGURE 1 MR image pre-processing: (a) Tissue Segmentation (b) Discriminative ROIs Selection (c) Image Slices Generation

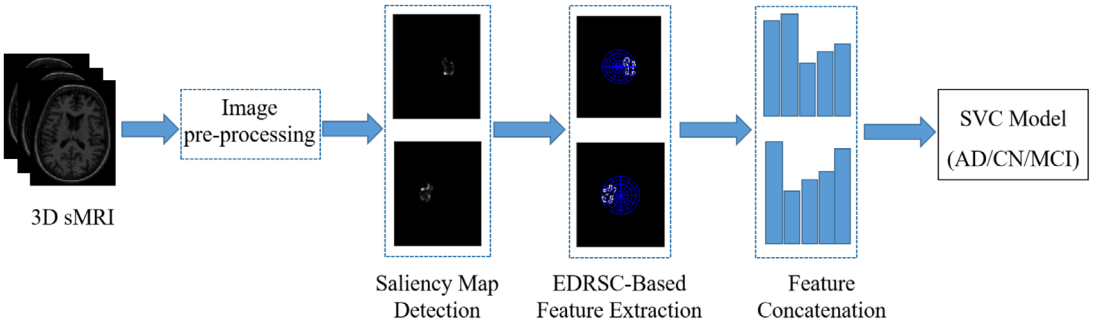


FIGURE 2 Block diagram of our method

with the GM maps to obtain the left and right hippocampus of all 3D images. The results are shown in Figure 1b.

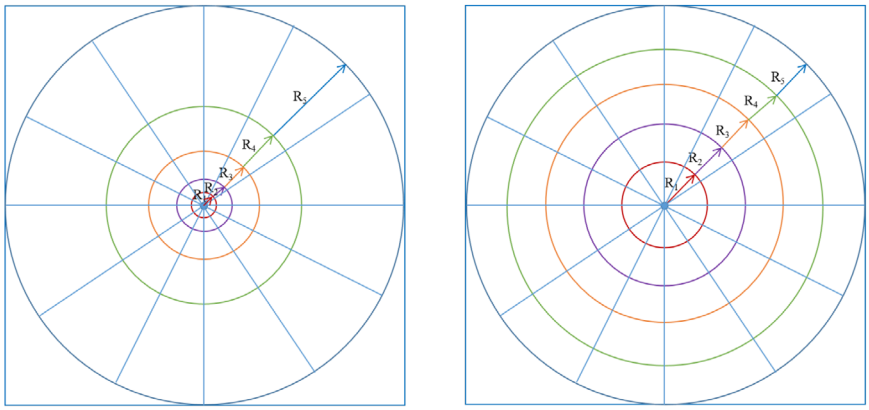
3) Image slices generation: Brain MRI scan is to scan the whole brain one by one along the anatomical axis of the human body. One layer is an image slice, and each slice is a conventional 2D image. All of these 2D images come together to form a 3D MRI. The extraction of 3D brain MRI features are very complex and time consuming. In order to reduce the time of the feature extraction, we use the MRIcro software to save each layer of the Transverse View section of the left and right hippocampus ROIs image as 2D images with BMP format (McCauslandcenter.sc.edu/crnl/mricro/). The results are shown in Figure 1c.

3 | METHOD

This study can be described as the following several main steps. First, all 3D MR image are pre-processed according to the pre-processing steps in Section 2.2 to obtain the BMP image of ROIs. Second, PFT model is exploited to detect saliency map of ROIs BMP image. Then, the shape context of equal-distance ring-based method is used to extract shape features of saliency map. Finally, SVC is used to build the disease classification model. The framework of our method is illustrated in Figure 2, and we will describe the detailed process of feature extraction in later sections.

3.1 | Saliency map detection

Visual attention mechanism is the key to ensure high efficiency of visual cognition process. It can select visual sensory information and only provide the important information to visual perception process, while the other information is rejected, thus making visual cognition process active and selective. With the increasing interest of researchers in visual attention research



(a) Concentric Rings Partition (b) Concentric Rings Partition based on Equal-Distance

FIGURE 3 Schematic diagram of a partition with five rings: (a) Concentric Rings Partition (b) Concentric Rings Partition based on Equal-Distance

and the increasing ability of computer to process information and realize complex computer vision system, many visual attention models are proposed, such as NVT (Neuromorphic Vision Toolkit) model [42], STB (Saliency ToolBox) model [43], SR (Spectral Residual) model [44], and PFT model [45]. In this paper, we use the PFT model for saliency map detection because it runs faster and more effectively than the NVT model, STB model, and SR model. The PFT model is a model that calculates the salient region of the image in the transform domain. It uses the phase spectrum of the Fourier transform to extract the salient region of the image. $I(x,y)$ represents an input image. It is assumed that F and F^{-1} represent Fourier transform and inverse Fourier transform of the image, respectively. Therefore, the saliency map of the detected image using the PFT model can be expressed by the following equation [45]:

$$M(x,y) = g(x,y) * \|F^{-1}[e^{i\cdot p(x,y)}]\|^2. \quad (1)$$

Where $g(x,y)$ represents a two-dimensional Gaussian filter (*sigma*=8), and $p(x,y)$ is determined by the following formula:

$$p(x,y) = P(f(x,y)). \quad (2)$$

Where $P(.)$ represents the phase spectrum obtained by Fourier transform of the input image, and $f(x,y)$ is defined as follows:

$$f(x,y) = F(I(x,y)). \quad (3)$$

Through saliency map detection, an area consistent with the attention focus of human visual system can be obtained. Therefore, we use PFT model to perform saliency map detection on the left and right hippocampus 2D images, respectively. Then, the edge detection of left and right hippocampus are carried out on the saliency map. Because the Canny operator has the advantages of accurate boundary point location, low error rate of edge detection edge and single pixel point, Canny edge detection operator is used to obtain the edge information of saliency map. After obtaining the edges of the left and right hippocampus, image features can be calculated.

3.2 | Image feature extraction

Traditional shape context (SC) is mostly used for shape matching and target recognition [46, 47], which is a feature description method based on shape contour proposed by Belongie et al. [48]. The traditional SC is calculated as follows:

Step1: Contour detection. For a given shape I, the edge of the contour is obtained by the edge detection operator, and a set of discrete points $P = \{P_1, P_2, P_3, \ldots, P_N\}$ is obtained by sampling the edge of the contour.

Step2: Shape context calculation. In the log-polar coordinate system, with any point P_i as the reference point, multiple concentric circles are established by Euclidean distance in the local area where P_i is the center of the circle and $R_k (k = 1, 2, 3, 4, 5)$ is the radius. Each concentric circle is divided into multiple divisions in the circumferential direction as shown in Figure 3a. The vector relative position of the point P_i to the other points is reduced to the number of dot distributions within each sector. The statistical distribution histogram of these points, called the shape context of point P_i.

In this study, a new feature extraction method based on shape context of equal-distance ring is introduced. comparing Figure 3a,b, it can be seen that the ring based on equal-distance divides the whole shape contour evenly and can better reflect the distribution of all points on the image contour. In order to be able to quantify, the direction parameter and distance parameter are introduced into the equal-distance rings, where the direction parameter represents the number l of sectors divided into a circle, such as the direction parameter in Figure 3b is 12, and the distance parameter represents the number k of the equal-distance rings, as shown in Figure 3b is 5. Figure 3b is a schematic diagram of a partition with five equal-distance rings.

Suppose an image I, the size of $M \times N$, and all the points on the contour of image are recorded as $P = \{P_1, P_2, P_3, \ldots, P_N\}$, (x,y) is the horizontal and vertical coordinate value of point P_i on the contour, R_k is the radius of the kth ring of equal-distance$(k = 1, 2, 3, 4, 5)$, as shown in Figure 3b. It is obvious that here, the contour image is divided into a set of rings of equal distance, That is, $R_1 = R_2 = R_3 = R_4 = R_5$. Here, the shape context based on the equal-distance rings uses the

chessboard distance instead of the Euclidean distance on the basis of the traditional shape context, and the ring partition can be performed in the following manner:

$$R_k = Max(d_{x,y})/Num(k)(k = 1, 2, 3, 4, 5). \quad (4)$$

Where $d_{x,y}$ represents the checkerboard distance of each point P_i coordinate (x, y) on the contour to the equal-distance rings center coordinates (x_c, y_c), the formula is as follows:

$$d_{x,y} = \sum_{i=1}^{n} Max[|x_i - x_c|, |y_i - y_c|], \quad (5)$$

$$x_c = round(N/2), y_c = round(M/2), \quad (6)$$

In this study, the shape context features based on equal-distance rings are carried out in rectangular coordinate system. First, the chessboard distance between each pixel point and the center point of the image is calculated, and then the chessboard distance is compared with the radius of each ring to determine the set of pixel points in each ring. Finally, the histogram statistics of all the ring sectors can get the spatial position information of all the points of the entire target shape, including distance information and direction information.

Let S_k be the set of pixel values of the kth ring ($k = 1, 2, 3, 4, 5$), and H_l be the number of pixel value sets on the lth sector of each ring ($l = 1, 2, 3, \ldots, 12$), so the feature vector size of each ROI is 1×60. The detailed rules are as follows:

$$S_1 = \{(x, y)|d_{x,y} \le R_1\}, \quad (7)$$

$$S_k = \{(x, y)|R_{k-1} \le d_{x,y} \le R_k\} \ (k = 2, 3, 4, 5), \quad (8)$$

$$H_l = \#\{P_i|(x, y) \in S_k, (y/x) \in [\tan 0°, \tan 360°]\} \quad (9)$$

3.3 | Classification model

In this study, we used LIBSVM to build SVC model for AD classification [49]. LIBSVM is an open source library based on SVM [50–54]. It was developed by Professor Chih-Jen Lin of Taiwan University. It is mainly used for classification (supporting binary classifications and multiple classifications) and regression. LIBSVM is characterized by its simplicity of operation, ease of use, fast and efficient, and relatively few adjustments to the parameters involved in SVM.

4 | EXPERIMENTS

4.1 | Validation methods and evaluation metrics

To estimate the classification performance of this method, 10 times 10-fold cross validation was used to verify the classification experiment results. The basic principle of 10-fold cross validation is to randomly divide the obtained image features into 10 groups, each group is 10% of the total data, of which 9 groups are used each time to build the classification model, and the rest of group is used for testing. Then repeat the above steps for 10 times to calculate the total accuracy. 10 times 10-fold cross verification is to repeat the above steps 10 times. In the experiment, four metrics values such as Accuracy (ACC), Sensitivity (SEN), Specificity (SPE) and area under the ROC curve (AUC) are given to illustrate the classification performance of the method. In generally, the method has a low classification performance in AUC value of 0.5–0.7; at 0.7–0.9, the classification performance can be considered moderate; if it is greater than 0.9, the classification performance is considered to be higher [55]. The evaluation parameters are calculated by the following equation:

$$ACC = (TP + TN)/(TP + TN + FP + FN),$$

$$SEN = TP/(TP + FN),$$

$$SPE = TN/(TN + FP).$$

Where TP is true positive, TN is true negative, FP indicates false positive and FN represents false negative.

4.2 | Classification experimental results

In the experiment, we first test the classification performance of the method. In this study, left and right hippocampus image features are extracted, respectively. Therefore, the input features are divided into three types, including only left hippocampal features (HL for short) , only right hippocampal features (HR for short) and combined with left and right hippocampal features (HC for short), and three classification results were obtained. In order to better evaluate the positive effect of the equal-distant rings segmentation step in this method, we compared the traditional SC method (called SC) with our EDRSC method. At the same time, in order to compare the influence of different distance measurement methods on the method, we use different distance measurement methods to conduct experiments. Among them, EDRSC-based chessboard distance is called EDRSC-CD, EDRSC-based Euclidean distance is called EDRSC-ED, and EDRSC-based city block distance is called EDRSC-BD. Table 2 summaries all the results.

It can be seen from Table 2: (1) The classification performances of EDRSC-ED, EDRSC-BD, and EDRSC-CD are superior to SC, which shows that the features extracted based on the step of equal-distant rings segmentation have better classification results. (2) Among the three classification tasks, the classification results of EDRSC-ED and EDRSC-BD are not much different from EDRSC-CD, but the classification results of EDRSC-CD achieved the highest classification accuracy (The best results are shown in bold in the table). Specifically, in classifying AD from CN, the classification accuracy of HC and HL of all three methods are higher than of HR. Among them, Using HL features, the EDRSC-CD achieved the classification ACC of 97.43%, SEN of 97.82%, SPE of 97.37%, and AUC

TABLE 2 Comparison of classification performances (%) of the competing methods

		AD versus CN				AD versus MCI				MCI versus CN			
Method	**Features**	**ACC (%)**	**SEN (%)**	**SPE (%)**	**AUC**	**ACC (%)**	**SEN (%)**	**SPE (%)**	**AUC**	**ACC (%)**	**SEN (%)**	**SPE (%)**	**AUC**
SC	HL	76.69	57.60	88.48	0.810	66.83	0.530	99.22	0.518	66.40	72.81	59.36	0.730
	HR	64.06	11.88	95.51	0.629	66.28	3.520	97.17	0.608	60.00	72.64	44.97	0.657
	HC	66.88	24.17	93.33	0.661	66.72	1.600	98.48	0.554	66.69	73.37	59.20	0.736
EDRSC-ED	HL	91.62	84.94	95.68	0.962	91.66	90.56	92.27	0.975	79.99	84.79	74.07	0.881
	HR	88.80	84.92	91.36	0.945	91.38	86.00	94.32	0.967	83.98	86.49	80.96	0.888
	HC	94.58	91.43	96.48	0.972	96.00	93.67	97.16	0.990	85.94	86.69	85.32	0.933
EDRSC-BD	HL	95.81	94.32	96.49	**0.988**	94.62	96.28	93.66	0.984	76.71	82.48	69.74	0.856
	HR	91.58	89.77	92.53	0.969	91.41	85.27	94.44	0.959	86.03	89.10	82.12	0.860
	HC	93.94	90.82	95.74	0.987	96.79	94.78	97.80	**0.994**	86.54	87.37	**85.70**	0.919
EDRSC-CD	HL	**97.43**	**97.82**	**97.37**	0.987	94.83	92.13	96.25	0.984	80.06	84.61	74.96	0.885
	HR	84.13	75.99	88.66	0.922	82.14	65.40	90.37	0.890	**88.35**	**91.08**	85.42	**0.951**
	HC	94.83	96.10	94.18	0.980	**98.31**	**98.18**	**98.35**	0.993	85.77	88.52	82.80	0.939

TABLE 3 Comparison of classification accuracy with state-of-the-art methods

				ACC (%)	
	Methods	**Modality (AD/MCI/CN)**	**Feature measures**	**AD versus CN**	**MCI versus CN**
Voxel based	Chu et al. [23]	MRI(131/261/188)	Voxel-wise GM	85.00	67.00
	Salvatore et al. [24]	MRI(137/76/162)	WM and GM density maps	76.00	72.00
ROI based	Zhang et al. [25]	MRI+PET+CSF(51/99/52)	Volume of GM within 93 ROI on MRI and PET + CSF	93.20	76.40
	Kim et al. [26]	MRI+PET+CSF(51/99/52)	Volume of GM within 93 ROI on MRI and PET + CSF	93.20	85.40
	Wee et al. [27]	MRI(198/200/200)	Cortical thickness + Hippocampus volumes	92.35	83.75
	Ahmed et al. [32]	MRI(137/162/210)	CHFs hippocampus and posterior cingulate cortex	83.77	69.45
	Ahmed et al. [33]	MRI(35/111/72)	Hippocampal CHF and CSF	87.00	78.22
	Proposed method	MRI(95/195/158)	Hippocampal EDRSC	94.83	85.77

of 0.987; and the ACC, SEN, SPE and AUC of EDRSC-ED and EDRSC-BD are 91.62% versus 95.81%, 84.94% versus 94.32%, 95.68% versus 96.49% and 0.962 versus 0.988, respectively. And using HC features, EDRSC-ED, EDRSC-BD and EDRSC-CD achieved the classification ACC of 94.83% versus 94.58% versus 93.94%, SEN of 96.10% versus 91.43% versus 90.82%, SPE of 94.18% versus 96.48% versus 95.74%, and AUC of 0.980 versus 0.972 versus 0.987, respectively. For the classification of AD and MCI, we can see that the performance of any single type feature of the three methods is lower than the performance of their combination. Among the three types, the HR show the lowest performance. However, when combined with HL, HR can help improve classification. For example, in EDRSC-CD method, the performances using HC demonstrate 3.48% and 16.17% improvements in terms of ACC over the cases of only using HL and HR, respectively. Similar results are obtained for EDRSC-ED and EDRSC-BD methods. In MCI versus CN, EDRSC-CD achieved the ACCs of 80.06%(HL), 88.35%(HR), and 85.77%(HC), while EDRSC-ED and EDRSC-BD achieved the ACCs of 79.99% versus 76.71%(HL), 83.98% versus 86.03%(HR), and 85.94% versus 86.54%(HC), respectively.

4.3 | Comparison with state-of-the-art methods

In Table 3, the classification accuracy of our method is compared with the results of several methods that use sMRI data as the research subject and SVM as the classifier to classify AD versus CN and MCI versus CN, including two voxel-based methods [23, 24] and five ROI-based methods [25–27, 32, 33]. It is

worth noting that the performance evaluation of these methods concerns the feature extraction method rather than the design of the classifier. And in order to ensure the classification performance of each method, the classification results of each method in the table are the best results obtained through experiments on its original dataset [25, 26, 32 56, 57]. Although the sMRI data selected by all methods for experiments are not exactly the same, the sMRI images of all methods come from ADNI and are obtained by MPRAGE or equivalent protocols of different resolutions, which have been uniformly processed by several pre-processing steps of ADNI research groups (See Section 2.1 for details). Therefore, although the results in Table 3 are may not completely comparable, we can roughly comparing our study (i.e. the last row of Table 3) with these state-of-the-art methods to verify the efficacy of our proposed method.

Specifically, of the two voxel-based methods, Chu et al. [23] used 131 AD patients and 188 CN patients for AD diagnosis with an accuracy of 85.00%, and used 261 MCI patients and 188 CN patients for MCI diagnosis with an accuracy of 67.00% . In Salvatore et al. [24], they used 137 AD patients, 76 MCI patients and 162 CN to classify AD and MCI. The accuracy of 76.00% for AD classification, and the accuracy of 72.00% for MCI classification. In the ROI-based method, Zhang et al. [25] and Kim et al. [26] et al. used 51 AD patients, 77 MCI patients and 52 CN patients to classify AD and MCI. By using three modes (MRI + PET + CSF) as features, they reported that the accuracy of 93.20% for AD classification, and the accuracy of 76.40% and 85.40% for MCI classification, respectively. Wee et al. [27] integrated cortical thickness, cortical volume and hippocampal volume as features to classify AD and MCI. Among 198 AD patients, 200 MCI patients, and 200 CN patients, the accuracy of 92.35% for AD classification, and the accuracy of 83.75% for MCI classification. Ahmed et al. also studied the feature extraction of hippocampal ROI. Literature [32] used 137 AD patients, 162 MCI patients and 210 CN patients for AD and MCI classification. By extracting the CHFs coefficients of hippocampus ROI and posterior cingulate cortex ROI as features, they report the accuracy of 83.77% for AD classification, and the accuracy of 69.45% for MCI classification. Literature [33] used 51 AD patients, 111 MCI patients and 72 CN patients to classify AD and MCI. By using the imaging mode and CSF (hippocampal CHF and CSF) as features, they reported that the accuracy of 87.00% for AD classification, and the accuracy of 78.22% for MCI classification. For our method, we use 95 AD patients, 195 MCI patients and 158 CN patients to classify AD and MCI. In AD versus CN, our method achieved the classification accuracy of 94.83%. For classifying MCI from CN, we achieved classification accuracy of 85.77%.

4.4 | Verification on the OASIS dataset

To further verify the classification performance of this study, we conducted training and testing on another public dataset (OASIS dataset).The OASIS is a series of MRI datasets, including OASIS-1, OASIS-2 and OASIS-3, which can be used publicly by researchers. OASIS-3 is a dataset used to classify and diagnose CN and AD, collecting longitudinal neuroimaging, clinical indicators, cognitive scores and biomarkers of subjects. The subjects included 609 CN and 489 AD, with an age range of 42–95 years. The dataset collected more than 2000 MR sessions containing multiple sequences such as T1w, T2w, FLAIR, resting-state BOLD and DTI, and also provided segmentation files generated by MR sessions processed by Freesurfer software. All T1w MR images are available via www.oasis-brains.org. In our method, the T1w MR images of 244 subjects in the OASIS-3 are mainly selected for the experiments, including 114 AD subjects and 130 CN subjects. The detailed statistical of all research subjects are shown in Table 4.

TABLE 4 Characteristics of the OASIS dataset subjects used in this study

Diagnosed	Gender (M/F)	Age (Mean ± SD)	MMSE (Mean ± SD)
AD	53/61	75.65 ± 8.01	24.07 ± 4.07
CN	49/81	67.01 ± 8.27	29.19 ± 1.06

We compared the results of the AD and CN classification experiments on the OASIS and ADNI datasets, as shown in Figure 4. It can be seen from the Figure 4 that when only the ADNI dataset images are available, the ACC of 94.83%, SEN of 96.10%, SPE of 94.18% and AUC of 0.980. When only the OASIS dataset images are used for training and testing, the ACC of our method was 96.56% (SEN = 96.14%, SPE = 96.85% and AUC = 0.996). When the ADNI and OASIS datasets are used for experiments, the ACC, SEN, SPE and and AUC of 92.65%, 92.36%, 92.86% and 0.973, respectively. From this experiment, we can see that the results of AD and CN classification on OASIS and ADNI datasets are more than 90%, which shows that this method has good classification performance on thin-layer images.

4.5 | Verification on the thick-layer MRI dataset

To validate the classification results of this study in clinical practice, we added thick-layer sMRI collected from clinical practice to ADNI for training and testing. All the thick-layer MR T1-weighted images were collected by the Guangxi Medical University First Affiliated Hospital, and the thickness of the layer is 7 mm (a total of 212 subjects, including 62 AD subjects, 90 MCI subjects, and 60 CN subjects). The regional ethics committee approved the study and obtained written informed consent from all participants. We guarantee that all participants information will be kept confidential and will not be used for commercial purposes.The detailed statistical of all research subjects are shown in Table 5.

From the previous experimental results, it can be seen that the EDRSC method using HC features has better classification performance of AD (or MCI) and CN. As a result, we used the HC features method to test the thick layer of sMRI. The results are shown in Figure 5. Specifically, for classifying AD from CN, when only the ADNI dataset images are available, the ACC of 94.83% (SEN = 96.10%, SPE = 94.18%,

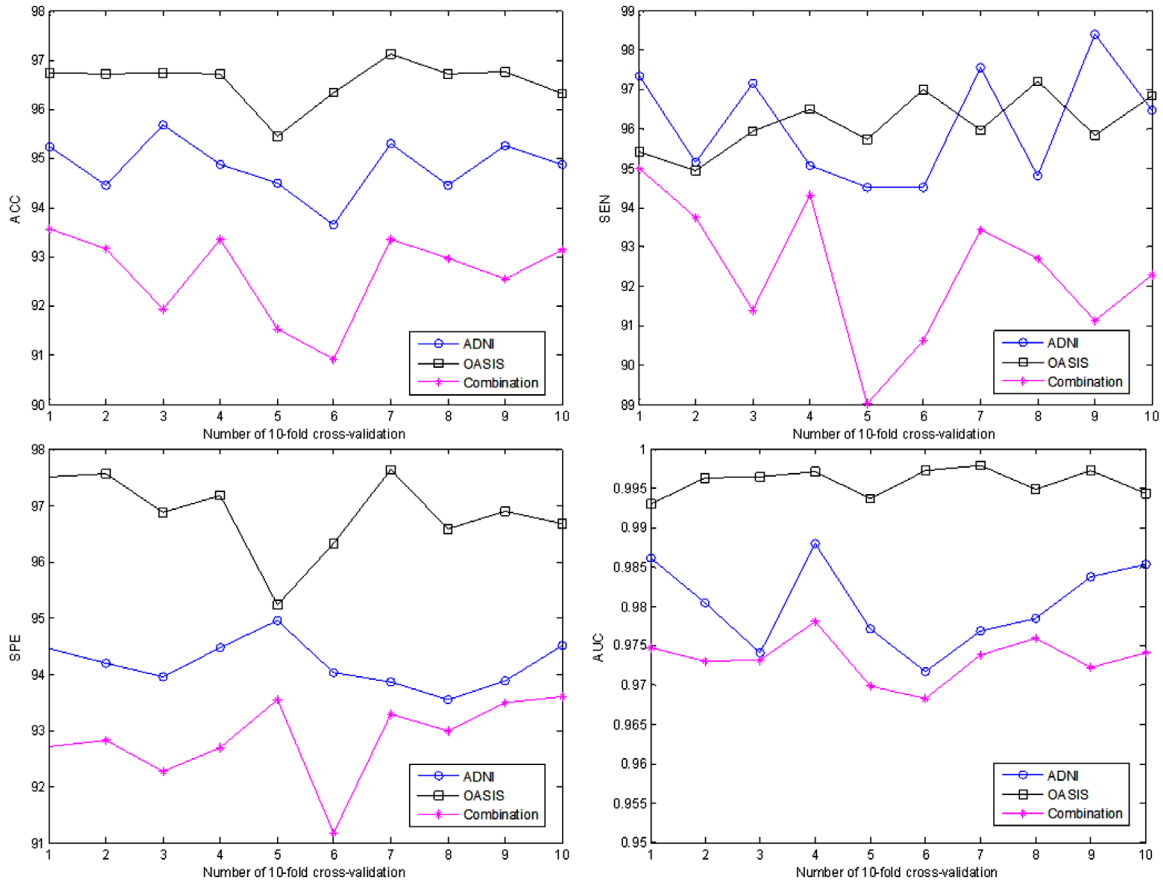


FIGURE 4 The results of verifying the performance of our method on OASIS dataset

TABLE 5 Characteristics of the thick-layer sMRI subjects used in this study

Diagnosed	Number of subjects	Gender (M/F)	Age (Mean ± SD)
AD	62	32/30	76.13 ± 12.23
MCI	90	64/26	71.09 ± 13.26
CN	60	16/44	56.60 ± 4.480

and AUC = 0.980). When only the collected clinical thick-layer sMRI images are used for training and testing, the ACC, SEN, SPE and AUC of 98.18%, 98.13%, 98.49%, 0.999, respectively. When the thick-layer sMRI was collected and added to the ADNI database for experiments, the ACC of 95.12%, SEN of 96.54%, SPE of 94.21% and AUC of 0.986, respectively. For AD and MCI classifications, the results achieved on different datasets are as follows: The ACC of 98.31% (ADNI), 91.55% (thick-layer sMRI) and 93.26% (combination of ADNI and thick-layer MRI), respectively. The SEN of 98.18% (ADNI), 90.24% (thick-layer sMRI) and 90.96% (combination of ADNI and thick-layer MRI), respectively. The SPE of 98.35% (ADNI), 92.49% (thick-layer sMRI) and 94.58% (combination of ADNI and thick-layer MRI), respectively. The AUC of 0.993 (ADNI), 0.977 (thick-layer sMRI) and 0.973 (combination of ADNI and thick-layer MRI), respectively. In MCI versus CN, the proposed method on ADNI dataset achieved the ACC of 85.77%, SEN of 88.52%, SPE of 82.80%, and AUC of 0.939, respectively. On the thick-layer sMRI dataset, the ACC of 98.27%, SEN of 98.36%, SPE of 98.27%, and AUC of 0.999, respectively. When the thick-layer sMRI was collected and added to the ADNI dataset for experiments, the ACC of 89.78%, SEN of 92.07%, SPE of 86.91% and AUC of 0.956. The experimental results showed that regardless of the thin-layer MR images or the thick-layer MR images, the H-EDRSC method can get good classification results, which may have higher application value in clinical application.

5 | LIMITATIONS

Our method has the following limitations. First, in this study, we only considered the imaging modality for AD, CN and MCI classifications. However, there are other modalities data in ADNI, for example, genetic data, blood biochemical indicators and CSF data. These modalities data may also contain other supplementary information about diseases, which can further improve the classification performance. Second, studies have shown that MCI includes pMCI and sMCI, where patients with sMCI do not convert to AD, and patients with pMCI convert to AD [58]. We only studied the classification between AD (or CN) and MCI without a more detailed classification, such as further

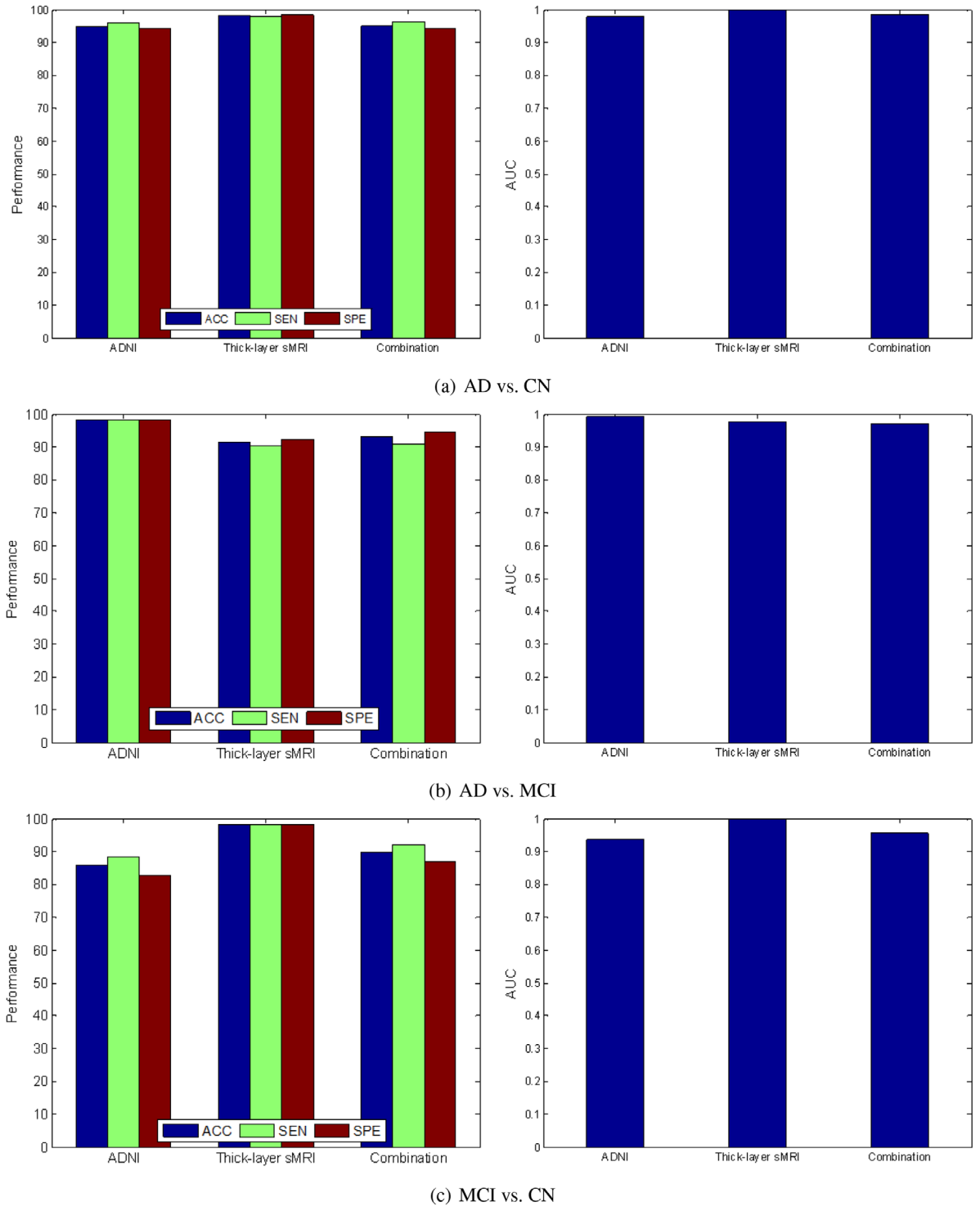


FIGURE 5 The results of verifying the performance of our method on thick-layer sMRI: (a) AD vs. CN (b) AD vs. MCI (c) MCI vs. CN

classification between pMCI and sMCI. Finally, like most studies, we only consider the binary-class classification problem (i.e. AD vs. CN, MCI vs. CN and AD vs. MCI), and do not perform the multi-category classification task. In the future, we will solve the above limitations and further improve the classification performance.

6 | CONCLUSION

sMRI is an effective tool for diagnosing AD. It can be seen from sMRI that the morphological structure of AD patients changes significantly compared with the age-matched CN, such as the decrease of hippocampal volume and the increase of

ventricular volume. Based on the changes of morphology and structure, this paper proposed a method for the diagnosis of AD by extracting EDRSC features based on saliency map of left and right hippocampus from sMRI, respectively. The open datasets (ADNI and OASIS) and the collected clinical thick-layer images were used to carry out the experiments. The experiments showed that this method has higher performance than the existing feature extraction methods (such as GM density, cortical thickness and hippocampal volume or shape).

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 81760324); Natural Science Foundation of Guangxi Province (Grant No. 2017JJA170765y); and Innovation Project of Guangxi Graduate Education (Grant No. YCBZ2020026).

REFERENCES

1. Martin, C.R., Preedy, V.R., Hunter, R.J.: Nanomedicine and the Nervous System. CRC Press, Boca Raton (2012)
2. Alzheimer's Association,: 2018 Alzheimer's disease facts and figures. Alzheimer's & Dementia 14(3), 367–429 (2018)
3. Khan, A., Corbett, A., Ballard, C.: Emerging treatments for Alzheimer's disease for non-amyloid and non-tau targets. Expert Rev. Neurother. 17(7), 683–695 (2017)
4. Sperling, R.A., et al.: Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia 7(3), 280–292 (2011)
5. McKhann, G., et al.: Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34(7), 939–939 (1984)
6. Dubois, B., et al.: Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria. Lancet Neurol. 6(8), 734–746 (2007)
7. Reitz, C., et al.: A summary risk score for the prediction of Alzheimer disease in elderly persons. Arch. Neurol. 67(7), 835–841 (2010)
8. Barnes, D.E., Yaffe, K.: The projected effect of risk factor reduction on Alzheimer's disease prevalence. Lancet Neurol. 10(9), 819–828 (2011)
9. Reitz, C., Mayeux, R.: Alzheimer disease: Epidemiology, diagnostic criteria, risk factors and biomarkers. Biochem. Pharmacol. 88(4), 640–651 (2014)
10. Sevigny, J., et al.: The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. Nature 537(7618), 50–56 (2016)
11. Leon, M.J., et al.: Longitudinal CSF isoprostane and MRI atrophy in the progression to AD. J. Neurol. 254(12), 1666–1675 (2007)
12. Du, A.T., et al.: Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. Brain 130(4), 1159–1166 (2006)
13. McEvoy, L.K., et al.: Alzheimer disease: Quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. Radiology 251(1), 195–205 (2009)
14. Fjell, A.M., et al.: CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. J. Neurosci. 30(6), 2088–2101 (2010)
15. Chételat, G., et al.: Mild cognitive impairment: Can FDG-PET predict who is to rapidly convert to Alzheimer's disease? Neurology 60(8), 1374–1377 (2003)
16. Higdon, R., et al.: A comparison of classification methods for differentiating fronto-temporal dementia from Alzheimer's disease using FDG-PET imaging. Stat. Med. 23(2), 315–326 (2004)
17. Foster, N.L., et al.: FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. Brain 130(10), 2616–2635 (2007)
18. Bouwman, F.H., et al.: CSF biomarkers and medial temporal lobe atrophy predict dementia in mild cognitive impairment. Neurobiol. Aging 28(7), 1070–1074 (2007)
19. Mattsson, N.: CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. JAMA 302(4), 385–393 (2009)
20. Shaw, L.M., et al.: Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. Ann. Neurol. 65(4), 403–413 (2009)
21. Klöppel, S., et al.: Automatic classification of MR scans in Alzheimer's disease. Brain 131(3), 681–689 (2008)
22. Li, S., et al.: Abnormal changes of multidimensional surface features using multivariate pattern classification in amnestic mild cognitive impairment patients. J. Neurosci. 34(32), 10541–10553 (2014)
23. Chu, C., et al.: Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. NeuroImage 60(1), 59–70 (2012)
24. Salvatore, C., et al.: Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach. Front. Neurosci. 9, 307–319 (2015)
25. Zhang, D., et al.: Multimodal classification of Alzheimer's disease and mild cognitive impairment. NeuroImage 55(3), 856–867 (2011)
26. Kim, J., Lee, B.: Identification of Alzheimer's disease and mild cognitive impairment using multimodal sparse hierarchical extreme learning machine. Hum. Brain Mapp. 39(9), 3728–3741 (2018)
27. Wee, C.Y., Yap, P.T., Shen, D.: Alzheimer's disease neuroimaging initiative. Prediction of Alzheimer's dis ease and mild cognitive impairment using cortical morphological patterns. Hum. Brain Mapp. 34(12), 3411–3425 (2013)
28. Beheshti, I., Demirel, H., Matsuda, H.: Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. Comput. Biol. Med. 83, 109–119 (2017)
29. Tong, T., et al.: A novel grading biomarker for the prediction of conversion from mild cognitive impairment to Alzheimer's disease. IEEE Trans. Biomed. Eng. 64(1), 155–165 (2017)
30. Liu, M., Zhang, D., Shen, D.: Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment. IEEE Trans. Med. Imaging 35(6), 1463–1474 (2016)
31. Sørensen, L., et al.: Early detection of Alzheimer's disease using MRI hippocampal texture. Hum. Brain Mapp. 37(3), 1148–1161 (2016)
32. Ahmed, O.B., et al.: Alzheimer's disease diagnosis on structural MR images using circular harmonic functions descriptors on hippocampus and posterior cingulate cortex. Comput. Med. Imaging Graph. 44, 13–25 (2015)
33. Ahmed, O.B., et al.: Classification of Alzheimer's disease subjects from MRI using hippocampal visual features. Multimedia Tools Appl. 74(4), 1249–1266 (2015)
34. Zhao, K., et al.: Early classification of Alzheimer's disease using hippocampal texture from structural mri. Proc. SPIE 137, 101372E (2017)
35. Shi, J., et al.: Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's Disease. IEEE J. Biomed. Health Inform. 22, 173–183 (2018)
36. Folstein, M.F., Folstein, S.E., McHugh, P.R.: Mini-mental state. J. Psychiatr. Res. 12(3), 189–198 (1975)
37. Morris, J.C.: The clinical dementia rating (CDR): Current version and scoring rules. Neurology 43(11), 2412–2412 (1993)
38. McKhann, G.M., et al.: 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. Alzheimer's & Dementia 7(3), 263–269 (2011)
39. Grabner, G., et al.: Symmetric Atlasing and Model Based Segmentation: An Application to the Hippocampus in Older Adults, Lecture Notes in Computer Science, pp. 58–66, Springer Science+Business Media, Berlin (2006)
40. O'Hanlon, E., Newell, F.N., Mitchell, K.J.: Combined structural and functional imaging reveals cortical deactivations in grapheme-color synaesthesia. Front. Psychol. 4, 755 (2013)

41. Tzourio-Mazoyer, N., et al.: Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15(1), 273–289 (2002)
42. Itti, L., Koch C., Niebur, E.: A model of saliency based visual attention for rapid scene analysis. IEEE Trans. Patt. Anal. Mac. Intell. 20(11), 1254–1259 (1998)
43. Walther D., Koch, C.: Modeling attention to salient proto-objects. Neural Netw. 19, 1395–1407 (2006)
44. Hou, X., Zhang, L.: Saliency detection: A spectral residual approach. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8 (2007)
45. Guo, C., Ma, Q., Zhang, L.: Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8 (2008)
46. Xiao, D., et al.: An improved 3D shape context based non-rigid registration method and its application to small animal skeletons registration. Comput. Med. Imaging Graph. 34(4), 321–332 (2010)
47. Roman-Rangel, E., et al.: Retrieving ancient Maya glyphs with Shape Context. In: 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops (2009)
48. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. 24(4), 509–522 (2002)
49. Ulaby, F.T., et al.: Learning with Kernels: Support vector machines, regularization, optimization, and beyond. IEEE Trans. Neural Networks 16(3), 781–781 (2005)
50. Fan, Y., Shen, D., Davatzikos, C.: Classification of Structural Images via High-Dimensional Image Warping, Robust Feature Extraction, and SVM, Medical Image Computing and Computer-Assisted Intervention-MICCAI 2005, pp. 1–8 (2005)
51. Mourão-Miranda, J., et al.: Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data. NeuroImage 28(4), 980–995 (2005)
52. Craddock, R.C., et al.: Disease state prediction from resting state functional connectivity. Magn. Reson. Med. 62(6), 1619–1628 (2009)
53. Kloppel, S., et al.: Automatic classification of MR scans in Alzheimer's disease. Brain 131(3), 681–689 (2008)
54. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2(3), 1–27 (2011)
55. Fawcett,T.,: An introduction to ROC analysis. Pattern Recognit. Lett. 27(8), 861–874 (2006)
56. Lian, C., et al.: Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. IEEE Trans. Pattern Anal. Mach. Intell. 42, 880–893 (2018)
57. Liu, F., et al.: Inter-modality relationship constrained multi-modality multi-task feature selection for alzheimer's disease and mild cognitive impairment identification. Neuroimage 84, 466–475 (2014)
58. Cheng, B., et al.: Domain transfer learning for MCI conversion prediction. IEEE Trans. Biomed. Eng. 62(7), 1805–1817 (2015)

How to cite this article: Lao, H., et al.: Alzheimer's disease diagnosis based on the visual attention model and equal-distance ring shape context features. *IET Image Process.* 1–12 (2021).
https://doi.org/10.1049/ipr2.12218