



# Alzheimer's disease classification using features extracted from nonsubsampling contourlet subband-based individual networks

Jinwang Feng<sup>a</sup>, Shao-Wu Zhang<sup>a,\*</sup>, Luonan Chen<sup>a,b,c,d,\*</sup>, Jie Xia<sup>b</sup>,  
for the Alzheimer's Disease Neuroimaging Initiative<sup>1</sup>

<sup>a</sup> Key Laboratory of Information Fusion Technology of Ministry of Education, School of Automation, Northwestern Polytechnical University, Xi'an 710072, China

<sup>b</sup> Key Laboratory of Systems Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy Science, Shanghai 200031, China

<sup>c</sup> Key Laboratory of Systems Biology, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Hangzhou 310024, China

<sup>d</sup> School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China

## ARTICLE INFO

### Article history:

Received 9 June 2020

Revised 4 August 2020

Accepted 12 September 2020

Available online 26 September 2020

Communicated by Zidong Wang

### Keywords:

Alzheimer's disease

Magnetic resonance imaging

Individual network

Nonsubsampling contourlet transform

Subband energy feature

## ABSTRACT

Morphological networks constructed with structural magnetic resonance imaging (sMRI) images have been widely investigated by exploring interregional alterations of different brain regions of interest (ROI) in the spatial domain for Alzheimer's disease (AD) classification. However, few attentions are attracted to construct a subband-based individual network with the sMRI image in the frequency domain. In order to verify the feasibility of constructing individual networks with subbands and extract features from the subband-based individual network for AD classification, in this study, we propose a novel method to capture correlations of the abnormal energy distribution patterns related to AD by constructing nonsubsampling contourlet subband-based individual networks (NCSINs) in the frequency domain. Specifically, a 2-dimensional representation of the preprocessed sMRI image is firstly reshaped by down-sampling and reconstruction steps. Then, the nonsubsampling contourlet transform is performed on the 2-dimensional representation to obtain directional subbands, and each directional subband at one scale is described by a column energy feature vector (CV) regarded as a node of the NCSIN. Subsequently, edge between any two nodes is weighted with connection strength (CS). Finally, the concatenation of node and edge features of the NCSINs at different scales is used as a network feature of the sMRI image for AD classification. Meanwhile, the support vector machine (SVM) classifier with a radial basis function (RBF) kernel is applied for categorizing 680 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Experimental results demonstrate that it is feasible to construct the subband-based individual network in the frequency domain and also show that our NCSIN method outperforms five other state-of-the-art approaches.

© 2020 Elsevier B.V. All rights reserved.

## 1. Introduction

It is widely accepted that the human brain is a complicated network and a healthy adult brain has about 100 billion neurons with long and branching extensions [1–3]. Researches associated with the human brain show that a majority of brain diseases are closely

related to the broken connections of neurons distributed different tissue regions [4–11], such as AD and its prodromal stage, mild cognitive impairment (MCI). As a progressive and currently incurable neurodegenerative disease, AD has been a leading cause of dementia [12]. In 2006, AD had affected 26.6 million people worldwide and by 2050, it will affect one in 85 people [13]. Moreover, taking the United States as an example, in 2017, more than 16 million family members and other unpaid caregivers provided an estimated 18.4 billion hours of care to people with AD or other dementias; In 2018, total payments for health care for people age  $\geq 65$  years with dementia are estimated to be \$277 billion [14]. It is obvious that AD can make a great impact on caregivers and society. We have noted that MCI is widely regarded as the transitional stage between healthy control (HC) and AD. Patients with

\* Corresponding authors.

E-mail addresses: [zhangsw@nwpu.edu.cn](mailto:zhangsw@nwpu.edu.cn) (S.-W. Zhang), [lnchen@sibs.ac.cn](mailto:lnchen@sibs.ac.cn) (L. Chen).

<sup>1</sup> Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wpcontent/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

MCI are only characterized by memory impairment, but general cognitive and functional abilities are usually retained. Unfortunately, a comparison between subjects with MCI and HC demonstrates that a conversion rate occurs at an annualized rate of 1% to 2% for HC, whereas for MCI the conversion to full AD occurs at a rate of 10% to 15% per year [15,16]. Therefore, early diagnosis of AD could have important personal and financial benefits. Because a mathematical model has estimated that early and accurate AD diagnosis could save up to \$7.9 trillion in medical and care costs [14].

With the identification of AD biomarkers in recent years, a great deal of studies have been proposed to research structural changes of the human brain, indicating that gray matter atrophy can be obviously identified in MCI and AD patients [17,18]. And these structural changes are reflected in the sMRI image by brain volume atrophy or cortical thickness thinning [19,20]. There is no doubt that the feasibility of discovering clinical biomarkers from the neuroimaging has been demonstrated by these studies [21–23]. Currently, many efforts have been made to analyze atrophy patterns of the brain and extract features from the sMRI image in the spatial domain for AD classification [24–31]. In general, AD classification methods using spatial analysis techniques can be roughly divided into three categories: voxel-based, regions of interest (ROI)-based, and patch-based approaches. In the voxel-based method, features are simply extracted based on statistic or selection for voxels. Such as, Ju et al. [32] used deep learning with brain network and clinical information to make early diagnosis of AD. However, the voxel-based features usually have much higher dimensionality and noise, whose dimensionality need be reduced. In the ROI-based method, the sMRI image is segmented into different ROIs, and then an ROI-based feature vector is used to describe the sMRI image. Such as, Li et al. [33] developed a deep learning method using hippocampal MRI data to predict MCI subjects' progression. However, the dimensionality of the ROI-based features are also relatively high. For the patch-based method, a patch containing multiple ROIs is selected to extract features. Such as, Zhang et al. [34] utilized stationary wavelet entropy to extract texture features of an MRI for AD classification. Jha et al. [35] proposed a novel computer-aided diagnosis cascade model for AD classification. Clearly, multi-resolution techniques, for instance, wavelet and nonsubsampling contourlet transforms which decompose images into low and high frequency components, have been successfully used in medical image analysis [36–38], such as Nazrudeen et al. [39] proposed a framework to fuse CT and MRI from persons affected with Alzheimer, stroke and recurrent tumor.

Nowadays, many efforts have been made to construct a network architecture of the human brain by defining the connection between voxels, vertices, ROIs, or patches in the spatial domain [40–42]. Thus, researches about the brain connectivity have been more and more popular [43], and the purpose of which is to analyze the brain network by means of graph theoretical approaches [44–47]. Generally speaking, methods of constructing the brain network for AD classification can be categorized into: the group-based and the individual-based. In the group-based method, a network for a group of subjects is constructed to investigate the abnormality of network measures. Such as, Liu et al. [48] employed a whole brain hierarchical network for AD classification. In clinic, patients with MCI can be divided into converted (who will convert into AD in their follow-up and are denoted as MCIc in this study) and non-converted (who will not convert into AD in their follow-up and are denoted as MCInc in this study) categories. The MCIc has more AD-like atrophies, while the MCInc has more HC-like atrophies. Moreover, for HC individuals, the extent of brain atrophies varies depending on their age, and for early and later AD patients, the extents of their brain atrophies are also different. Hence, subjects in AD, MCI and HC groups are anisotropic, it is hard

for a network to accurately measure relationships between different atrophy patterns in a group. In the individual-based method, a network for each subject (namely, the individual network) is established to be used as a representation of the sMRI image. Such as, Liu et al. [49] proposed a framework of the combination of multiple kernels to combine edge and node features for AD classification. Even though, all of these studies argue that individual network-based features are more reliable due to the informative network topology, the dimensionality of features extracted from these morphological networks is very high, which makes structural stability of the individual network-based feature poor. Considering that brain is a complex network, we can confirm that the sMRI image cannot be sufficiently described by features extracted from a network constructed in the spatial domain.

In Dong's work [50], they proposed a texture classification and retrieval methods that model adjacent shearlet subband dependences using linear regression in the frequency domain, and in Liu's work [48], they constructed a whole brain hierarchical network using features extracted in the spatial domain. According to the two works, it is feasible to construct subband-based individual networks using features of the sMRI image extracted in the frequency domain for AD classification. Therefore we propose a novel method to construct nonsubsampling contourlet subband-based individual networks (NCSINs) for capturing abnormal energy distribution patterns related to AD. Firstly, after preprocessing to the sMRI image, we can get its 3-dimensional gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) images. The GM image that is mostly related to AD is selected to construct the NCSIN in the following. Considering that the nonsubsampling contourlet is a tool that can efficiently transform a 2-dimensional image, the downsampling and reconstruction steps to the GM image are performed so that we can obtain a 2-dimensional representation of the GM image. We denote the 2-dimensional representation of a GM image as the RGM image. Then, the nonsubsampling contourlet transform is performed on the RGM image to obtain its directional subbands at different scales. For each of the directional subbands at one scale, three energy features of each column of the directional subband are extracted to capture the abnormal energy distribution patterns related to AD, and a column energy feature vector (CV), regarded as a node, is formed to represent the directional subband by concatenating the energy features, following that we use the nodes at one scale to build a NCSIN. The edge between any two nodes in a NCSIN is weighted with connection strength (CS) which is measured by Pearson's correlation coefficient (PCC). Finally, the concatenation of node and edge features of the NCSINs at different scales is used as a network feature of the MRI image. Meanwhile, the support vector machine (SVM) classifier with a radial basis function (RBF) kernel is applied for categorizing 680 subjects (200 AD, 280 MCI, and 200 HC) from the ADNI database. Experimental results demonstrate that it is feasible to construct the subband-based individual network in the frequency domain, and show that our NCSIN method outperforms five other state-of-the-art approaches in terms of accuracy, sensitivity, and specificity. The results also indicate that the NCSIN-based feature can be a promising imaging marker for AD identification.

Three contributions have been made in the construction of a nonsubsampling contourlet subband-based individual network (NCSIN) in this study, which are listed as follows: First of all, we propose a novel method to construct a subband-based individual network in the frequency domain, at the same time demonstrating the feasibility of constructing the NCSIN for each subject by AD classification experiments. Second, instead of extracting structural features directly from the morphological networks in the spatial domain, the nonsubsampling contourlet is introduced to transform the 2-dimensional representation of a sMRI image for obtaining its directional subbands in the frequency domain, and the NCSINs are

constructed with these subbands for extracting features representing the sMRI image. Third, in order to effectively analyze the energy distributions and precisely capture the abnormal patterns related to AD in the directional subbands, we firstly propose to construct the column energy feature vector by concatenating three energy features of all columns to represent a directional subband.

The rest of this study is organized as follows. Materials and methods are introduced in Section 2. Experimental settings and results are given in Section 3. In the end, a brief conclusion to this study is shown in Section 4.

## 2. Materials and methods

In this section, materials and their preprocessing are introduced firstly, followed by the proposed method of constructing the non-subsampled contourlet subband-based individual network.

### 2.1. Materials

Data used in the preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s disease (AD).

At present, the structural magnetic resonance imaging (sMRI) is one of the widely used modalities as an assistance to diagnose AD and MCI patients from health control (HC) individuals in clinics. At the diagnosis stage, patients with MCI can also be divided into subjects who will convert to AD after 18 months and subjects who will be stable after 18 months, in this study, these two subcategories are separately denoted as MCIC and MCInc. To evaluate the performance of our proposed NCSIN method, 680 sMRI images, including 200 AD, 120 MCIC, 160 MCInc, and 200 HC, are selected from the ADNI database. More detail demographic information about these selected sMRI images are summarized in Table 1. Moreover, MMSE, F, M, and D in Table 1 are abbreviations of Mini Mental State Examination, Female, Male, and Deviation, respectively.

For those selected sMRI images, a four-step preprocessing, including motion correction, registration and skull strap, segmentation, and smoothing, is performed using statistic parametric mapping (SPM8) [51] and voxel-based mapping (VBM8) [52] to remove unrelated tissues and ensure that a certain brain region of different subjects is at the same position. After preprocessing steps, the original sMRI image changes into  $121 \times 145 \times 121$  gray matter (GM), cerebrospinal fluid (CSF), and white matter (WM) tissue images, and the voxel volumes are  $1.5 \times 1.5 \times 1.5 \text{ mm}^3$ . For a visual comparison of the original and the preprocessed sMRI images, an example of four scans of the raw and the GM images is given in Fig. 1. At the experimental stage, the GM image that is

**Table 1**  
The detail demographic information of the 680 subjects selected from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database.

Type	Gender (F/M)	Number	Age (Mean±D)	MMSE (Mean±D)
AD	78/122	200	76.85 ± 7.01	22.15 ± 3.17
MCIC	67/53	120	78.65 ± 9.73	26.38 ± 3.76
MCInc	71/89	160	73.59 ± 7.68	26.21 ± 2.67
MCI	138/142	280	75.76 ± 8.96	26.28 ± 3.17
HC	84/116	200	76.21 ± 4.97	29.09 ± 1.15

mostly related to AD is selected to extract its NCSIN-based features, and four data sets consisted of subjects with AD and HC, subjects with AD and MCI, subjects with MCI and HC, and subjects with MCIC and MCInc are constructed to perform experiments. The four data sets used in this study are described as follows:

1. AD/HC: containing 200 AD subjects and 200 HC subjects;
2. AD/MCI: containing 200 AD subjects and 280 MCI subjects;
3. MCI/HC: containing 280 MCI subjects and 200 HC subjects;
4. MCIC/MCInc: containing 120 MCIC subjects and 160 MCInc subjects.

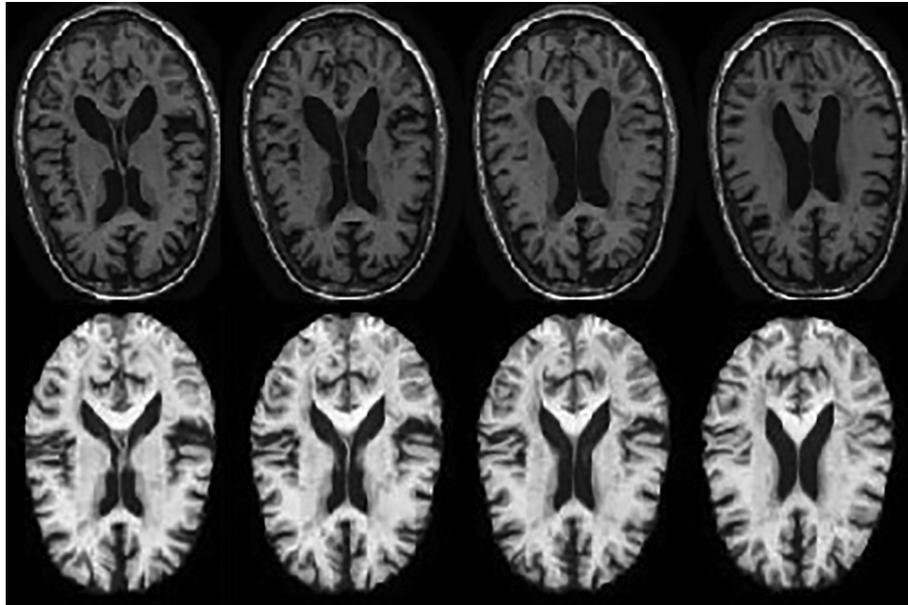
Obviously MCIC and MCInc subjects are patients with MCI, it can be confirmed that differences between their sMRI images are subtle. Hence, MCIC/MCInc is a challenging data set for our proposed NCSIN method in identifying a patient whether converts into AD or not.

### 2.2. Extracting the NCSIN feature

Before doing the nonsubsampling contourlet transform on the GM images, downsampling and reconstruction steps are performed to each of these GM images for getting a 2-dimensional representation. At the step of downsampling, the  $121 \times 145 \times 121$  GM image is downsampled by factor 2 for its each dimensionality, so the GM image convert into a  $60 \times 72 \times 60$  GM image. At the step of reconstruction, non-zero voxels of the  $60 \times 72 \times 60$  GM image are taken out by column, and then each  $M$  selected voxels are regarded as a column of an  $M \times D$  representation of the downsampled GM image, where  $M$  and  $D$  is the number of row and column of the 2D representation, respectively, meanwhile  $M$  and  $D$  satisfy  $M \times D \leq 121 \times 145 \times 121$ . For these reconstructed GM (RGM) images, a RGM image set, denoted as  $\{RGM^i\}, i = 1, 2, \dots, N$ , is constructed to represent them, in this study,  $N = 680$ . In the following, directional subbands of the RGM images contained in  $\{RGM^i\}$  set are obtained by performing the non-subsampled contourlet transform.

Multiscale geometric analysis (MGA) is an effective method for image processing [53–55]. As a kind of implementation of the discrete MGA, the nonsubsampling contourlet transform (NSCT) is widely used in the fields of image denoising, image fusion and image enhancement [56,57]. NSCT is fully shift invariant and efficiently performs the multiscale and multi-direction decomposition of an image, which also resolves the issue of the pseudo Gibbs phenomena effect along the singularities suffered by the contourlet transform [58]. However, its computational efficiency is limited by its high redundancy. That is why the downsampling and reconstruction steps are performed to each of those GM images and get their 2-dimensional representation.

For the usage of directional and multiscale information of the sMRI image, the nonsubsampling contourlet transform is performed on these RGMs. Given the  $i$ -th  $M \times D$  representation (i.e.  $RGM^i$ ), we use the nonsubsampling contourlet with a  $S$ -decomposition-level directional filter bank at each of the  $L$  scales to transform the  $RGM^i$ , that is, the number of directional subbands is  $2^S$  at each scale, and sizes of the directional subbands at different scales are  $M \times D$ . After the nonsubsampling contourlet transform on the  $RGM^i$ , we can obtain  $2^S$  directional subbands at each of the  $L$  scales and a low frequency subband in total. Fig. 2(b) shows an example of 25 subbands of a RGM image that is transformed by the nonsubsampling contourlet with  $S = 3$  levels and  $L = 3$  scales. According to Fig. 2(b), only the contour information of the brain scan is contained in the low frequency subband, which is not related to AD. Therefore, the low frequency subband is abandoned



**Fig. 1.** An example of four scans of the raw and the preprocessed sMRI images. The first row is four scans of the raw sMRI image, and the second row is four corresponding scans of the preprocessed sMRI image.

in constructing the NCSIN. For brevity and simplicity, a set, denoted as  $\{S_{(l,j)}^i\}$ , is used to represent all directional subbands of the  $RGM^i$ , where  $i$  ( $i = 1, 2, \dots, N$ ) is the  $i$ -th subject,  $l$  ( $l = 1, 2, \dots, L$ ) is the  $l$ -th decomposition scale of the  $i$ -th subject, and  $j$  ( $j = 1, 2, \dots, 2^S$ ) is the  $j$ -th directional subband of the  $i$ -th subject at the  $l$ -th scale.

Coefficients (CEs) in a directional subband describe the energy distribution of the spatial structures in the frequency domain. To precisely represent a directional subband, we propose to use three energy features, including Mean (ME), Variance (VA) and Origin moment (OM), to capture the energy distribution patterns of a column in the  $M \times D$  directional subband. For a  $M \times 1$  column in the directional subband, the three energy features are separately represented by

$$ME = \frac{1}{M} \sum_{i=1}^M |CE_i|, \quad (1)$$

$$VA = \frac{1}{M} \sum_{i=1}^M \left( |CE_i| - \frac{1}{M} \sum_{j=1}^M |CE_j| \right)^2, \quad (2)$$

$$OM = \frac{1}{M} \sum_{i=1}^M CE_i^2. \quad (3)$$

Therefore, the  $j$ -th  $M \times D$  directional subband of the  $i$ -th subject at the  $l$ -th scale  $S_{(l,j)}^i$  can be sufficiently described by ME feature vector (MV), VA feature vector (VV) and OM feature vector (OV)

$$MV = [ME_1, ME_2, \dots, ME_D], \quad (4)$$

$$VV = [VA_1, VA_2, \dots, VA_D], \quad (5)$$

$$OV = [OM_1, OM_2, \dots, OM_D]. \quad (6)$$

Then, a column energy feature vector (CV) of the  $j$ -th directional subband is formed by concatenating MV, VV and OV, which is used as a node of the NCSIN and is represented by

$$CV = [MV, VV, OV]. \quad (7)$$

Meanwhile, we use a node set, denoted as  $\{CV_{(l,j)}^i\}$ , to represent all nodes of the  $i$ -th subject at different scales. Obviously, there are  $2^S$  nodes at each of the  $L$  scales.

For any two nodes  $CV_{(l,1)}$  and  $CV_{(l,2)}$  at the  $l$ -th scale, the connection strength (CS) (edge weight) between them is measured by Pearson's correlation coefficient (PCC)

$$CS = \frac{\sum_{k=1}^{3 \times D} (CV_{(l,1)}(k) - \overline{CV_{(l,1)}})(CV_{(l,2)}(k) - \overline{CV_{(l,2)}})}{\sqrt{\sum_{k=1}^{3 \times D} (CV_{(l,1)}(k) - \overline{CV_{(l,1)}})^2} \sqrt{\sum_{k=1}^{3 \times D} (CV_{(l,2)}(k) - \overline{CV_{(l,2)}})^2}}, \quad (8)$$

where  $CV_{(l,1)}(k)$  and  $CV_{(l,2)}(k)$  are the  $k$ -th features of the  $CV_{(l,1)}$  and  $CV_{(l,2)}$ , and  $\overline{CV_{(l,1)}}$  and  $\overline{CV_{(l,2)}}$  are the mean values of  $CV_{(l,1)}$  and  $CV_{(l,2)}$ . According to nodes  $CV_{(l,j)}, j = 1, 2, \dots, 2^S$  at the  $l$ -th scale and Eq. (8), we can construct an edge weight matrix (namely, a matrix representation of the NCSIN) of nodes at the  $l$ -th scale and obtain  $\frac{2^S \times (2^S - 1)}{2}$  edges. Fig. 2(c) shows a nonsampled contourlet subband-based individual network with  $S = 3$  and  $L = 1$ .

By now, an edge feature vector (EV) can be formed by concatenating the CSs contained in the upper (or lower) triangular part of the edge weight matrix, which is represented by

$$EV = [CS_1, CS_2, \dots, CS_{\frac{2^S \times (2^S - 1)}{2}}]. \quad (9)$$

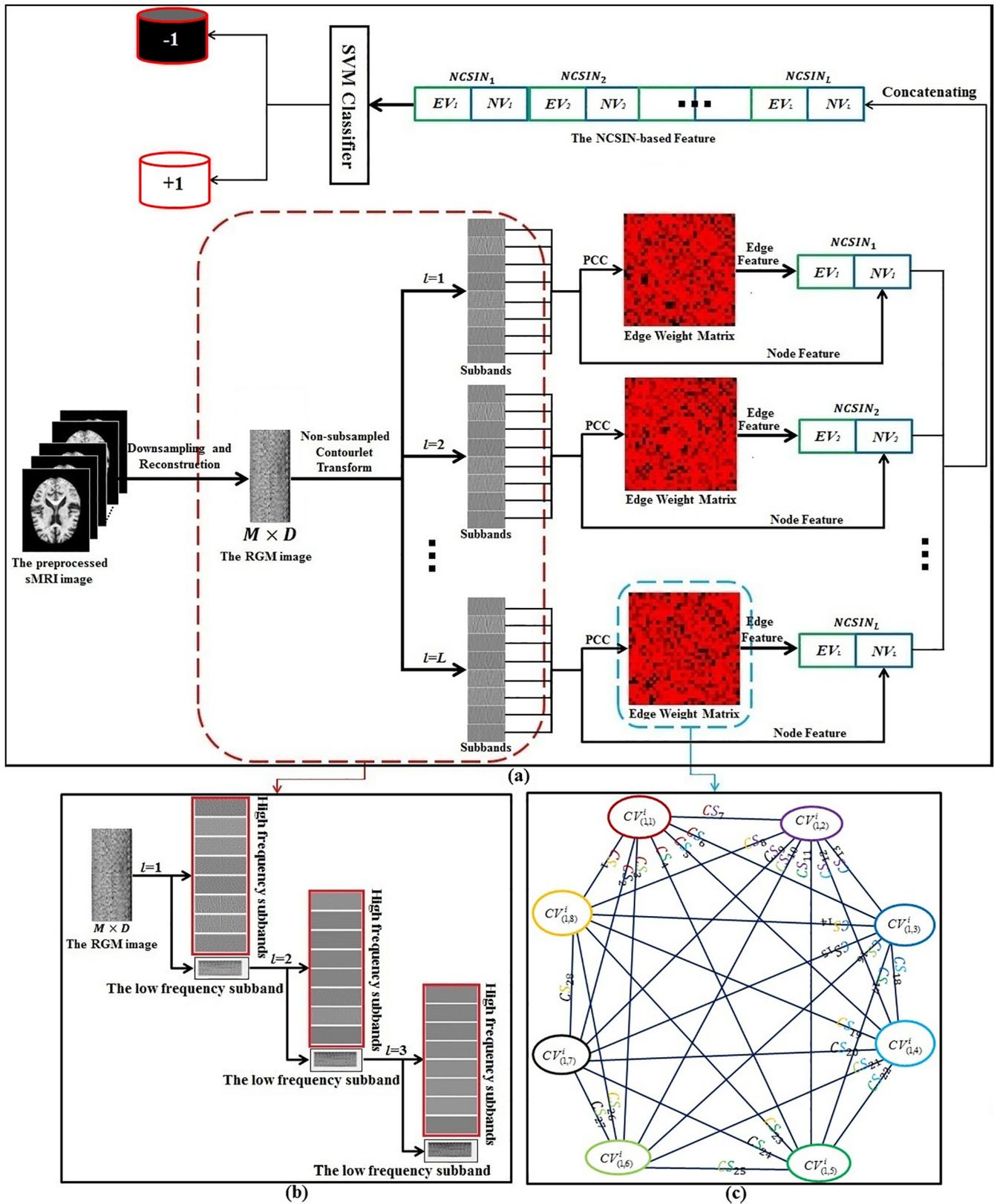
Then, a node feature vector (NV) of the NCSIN is also computed and is represented by

$$NV = [\overline{CV}_1, \overline{CV}_2, \dots, \overline{CV}_{2^S}], \quad (10)$$

where  $\overline{CV} = [\overline{MV}, \overline{VV}, \overline{OV}]$ , here  $\overline{MV} = \frac{1}{D} \sum_{k=1}^D ME_k$ ,  $\overline{VV} = \frac{1}{D} \sum_{k=1}^D VA_k$  and  $\overline{OV} = \frac{1}{D} \sum_{k=1}^D OM_k$ . By concatenating EV and NV, a NCSIN-based feature of the  $l$ -th scale is formed by

$$NCSIN_l = [EV_l, NV_l]. \quad (11)$$

Subsequently, the NCSIN-based feature of the  $i$ -th subject is constructed by jointing  $NCSIN_l, l = 1, 2, \dots, L$ , which is represented by



**Fig. 2.** The flowchart and some detail parts of our proposed method of constructing the nonsubsampled contourlet subband-based individual network (NCSIN). (a) The completed flowchart of constructing the NCSIN. (b) Transformation of the RGM image by the nonsubsampled contourlet with  $S=3$  decomposition levels and  $L=3$  transformation scales. (c) A NCSIN constructed with  $2^3 = 8$  nodes at the  $L$ -th scale.

$$NCSIN^i = [NCSIN_1^i, NCSIN_2^i, \dots, NCSIN_L^i]. \quad (12)$$

For simplicity, a feature set, denoted as  $\{NCSIN^i\}, i = 1, 2, \dots, N$ , is used to represent those selected sMRI images. And finally the NCSIN-based feature is used as an input of the support vector machine (SVM) classifier for AD classification. The SVM classifier is provided by MATLAB software in this study, and we select the radial basis function (RBF) as the kernel of the SVM classifier, which is formulated as

$$k(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right). \quad (13)$$

It is obvious that for the SVM classifier with RBF kernel, two important parameters, the window width of the RBF kernel  $\sigma$  and the penalty coefficient of the SVM classifier  $C$ , need to be estimated based on experiments.  $\sigma$  is used to control the number of support vectors, with the increase of  $\sigma$ , the number of support vectors decreases, and vice versa, the number of support vectors increases. However,  $C$  is used to control the penalty degree for error, with the increase of  $C$ , it is easy to causes over-fitting, and vice versa, it will causes under-fitting. Therefore, parameters  $\sigma$  and  $C$ , need to be estimated based on experiments carefully.

In short, to verify the feasibility of constructing the subband-based individual network in the frequency domain and extract features with relatively low dimension for AD classification, we propose a novel method to construct nonsubsampling contourlet subband-based individual networks (NCSINs) by capturing correlations of abnormal energy distribution patterns related to AD. Specifically, a 2-dimensional representation of the preprocessed sMRI image is firstly obtained by downsampling and reconstruction steps. And then the nonsubsampling contourlet transform is performed on the  $M \times D$  representation of the sMRI image to get its directional subbands at different scales. For directional subbands at one scale, each of these subbands is described by a column energy feature vector that is regarded as a node of the NCSIN. Subsequently, the edge between any two nodes contained in the NCSIN is weighted with connection strength which is measured by Pearson's correlation coefficient. Finally, the concatenation of node and edge features of the NCSINs at different scales is used as a network-based feature of the sMRI image. Meanwhile, the NCSIN-based features of the sMRI images are used as an input of the support vector machine (SVM) classifier with the radial basis function (RBF) kernel for categorizing subjects with AD, MCI and HC. Fig. 2(a) shows the flowchart of our proposed method of constructing the nonsubsampling contourlet subband-based individual network.

### 3. Evaluation metrics and experimental results

In this section, we will conduct multiple experiments to evaluate the performance of our NCSIN method for AD classification on AD/HC, AD/MCI, MCI/HC, MCIc/MCInc data sets. Metrics used in this study are firstly introduced, followed by experimental results of our NCSIN method and comparisons with five other state-of-the-art approaches.

#### 3.1. Metrics

In order to obtain an unbiased estimation of the classification performance, ten experiments are conducted on each of the four data sets. In each experiment, the ten-fold cross validation is done to reduce the impact of small sample size and guarantee that each subject can be used as testing. For each ten-fold cross validation, subjects contained in a data set are randomly divided into ten

subsets, nine of the ten subsets are used for training the SVM classifier, and the rest is used as testing. In this study, mean classification accuracy (ACC), sensitivity (Se), and specificity (Sp) of the ten experiments on each of the four data sets are used as the final results, which can be formulated as

$$ACC = \frac{1}{10} \sum_{i=1}^{10} \frac{TP_i + TN_i}{TP_i + FP_i + TN_i + FN_i}, \quad (14)$$

$$Se = \frac{1}{10} \sum_{i=1}^{10} \frac{TP_i}{TP_i + FN_i}, \quad (15)$$

$$Sp = \frac{1}{10} \sum_{i=1}^{10} \frac{TN_i}{FP_i + TN_i}, \quad (16)$$

where  $TP_i$  is the number of correctly classified true positive subjects,  $FP_i$  is the number of incorrectly classified false positive subjects,  $TN_i$  is the number of correctly classified true negative subjects,  $FN_i$  is the number of incorrectly classified false negative subjects, and the subscript  $i$  is the  $i$ -th experiment on a data set.

#### 3.2. Results and discussions

In the step of constructing the  $M \times D$  RGM image, obviously  $D$  determines the number of features contained in a column energy feature vector (CV, namely, a node of the NCSIN), affecting the stability of the NCSIN structure. To get an optimal estimation value, as shown in Fig. 3, we give the ACCs of four experiments on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets under the circumstances of different  $D$  values. According to Fig. 3(a), we can find that ACCs of the four experiments on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets increase gradually when  $D \leq 140$  and slowly decrease when  $D > 140$ , that is to say, the optimal  $D$  value may fall into the interval (130, 150). In order to obtain the optimal  $D$  value, we also show ACCs of the four experiments on AD/HC, AD/MCI, MCI/HC, MCIc/MCInc data sets in Fig. 3(b) when  $130 \leq D \leq 145$ . It can be clearly seen from Fig. 3(b) that ACCs of the four experiments with  $D = 136$  on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets can consistently outperform the ones of the other different  $D$  values. Therefore, we set  $D = 136$  as the optimal number of columns of the RGM image in terms of ACCs shown in Fig. 3 on four data sets.

Decomposition level ( $S$ ) and transformation scale ( $L$ ) are two important parameters for the nonsubsampling contourlet transform. Parameter  $S$  controls the number of directional subbands at each of the  $L$  transformation scales, and parameter  $L$  determines the number of transformation scales. In order to find the optimal combination of  $S$  and  $L$ , we perform estimation experiments on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc four data sets, respectively, and ACCs of each experiment on the four data sets are shown in Fig. 4. It can be seen from Fig. 4 that ACCs of the experiments on the four data sets increase drastically when  $S \leq 3$  under the cases of different decomposition levels, in contrast, ACCs of the experiments on the four data sets begin to decrease gradually when  $S > 3$ . Meanwhile we can consistently get the best ACCs of the experiments on the four data sets when the transformation scale  $L = 1$  with different decomposition levels. In terms of ACCs shown in Fig. 4, we set  $S = 3$  and  $L = 1$  as optimal parameter values of the non-subsampling contourlet transform in this study. In other words, we can totally obtain  $2^S = 8$  directional subbands at  $L = 1$  transformation scale. Therefore, only a NCSIN at the first scale is constructed to represent the sMRI image and the dimensionality of the feature extracted from the proposed NCSIN is  $\left(\frac{2^S \times (2^S - 1)}{2} + 2^S\right) \times L = \left(\frac{8 \times 7}{2} + 8\right) \times 1 = 36$ , which also make the

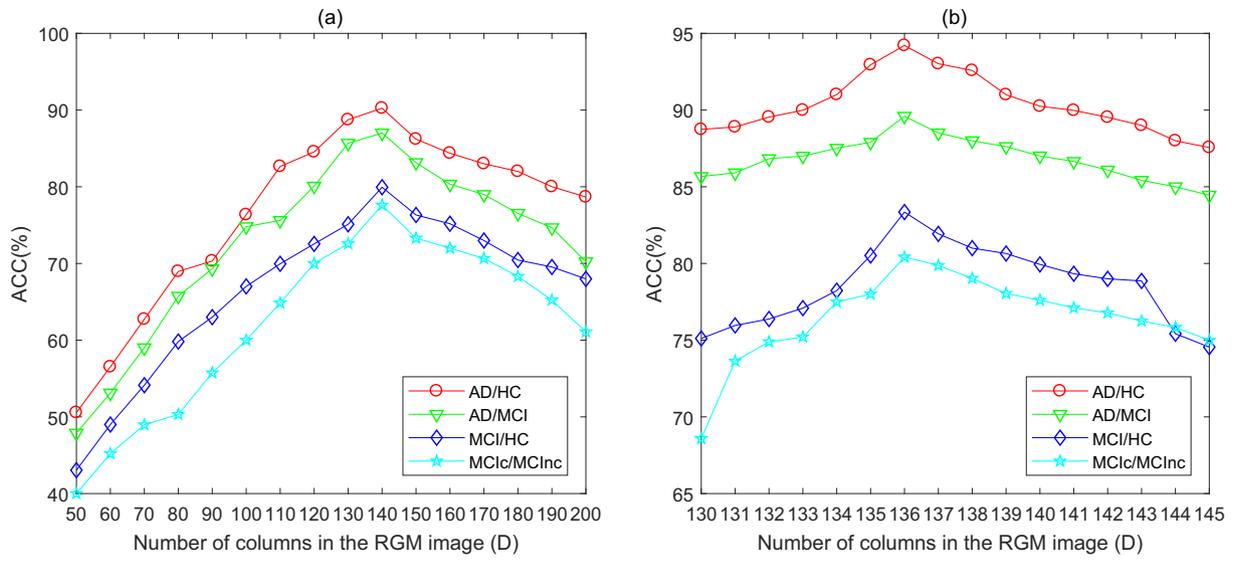


Fig. 3. ACCs of experiments with different D values on AD/Hc, AD/MCI, MCI/Hc, and MCIc/MCInc data sets.

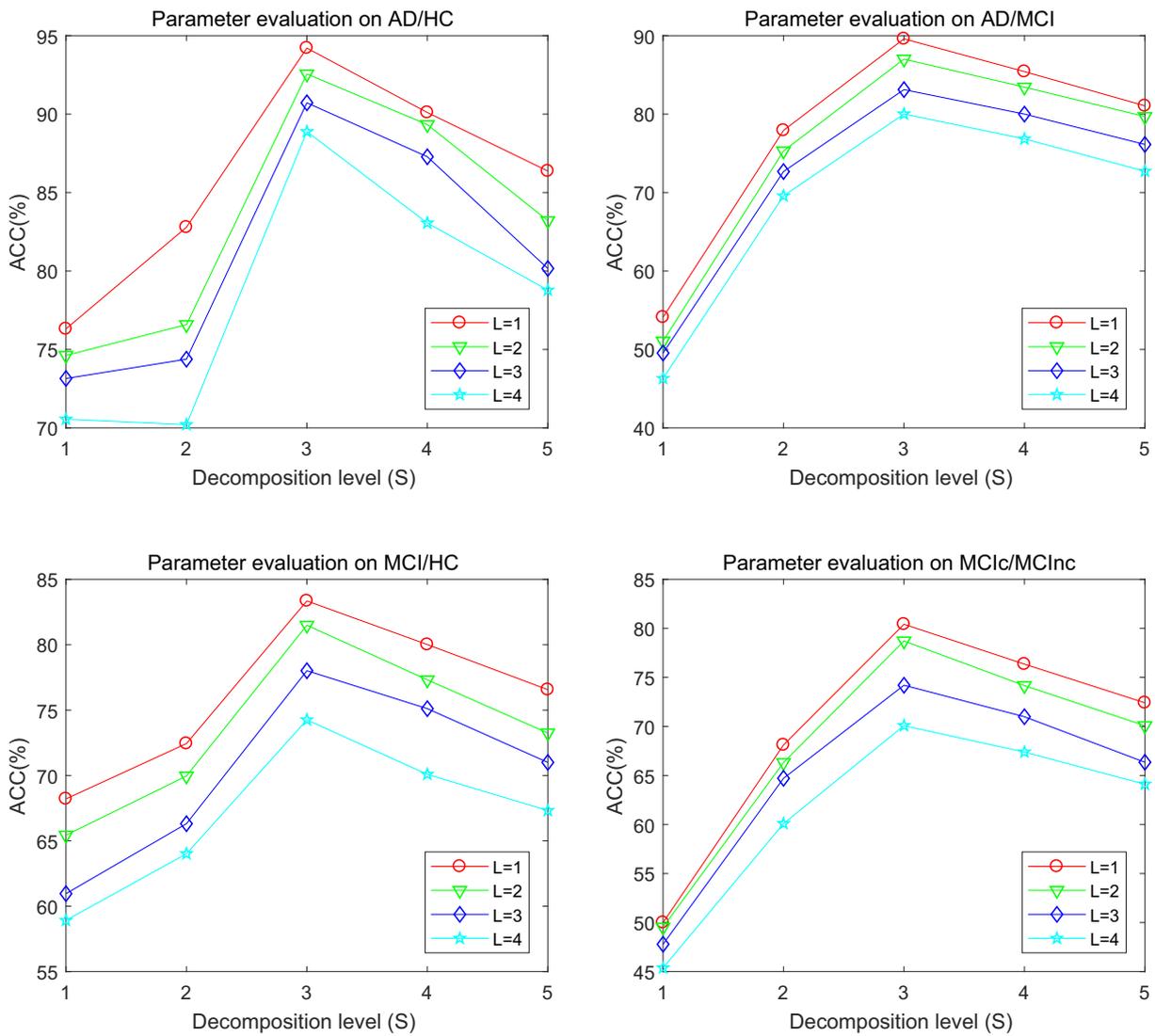


Fig. 4. ACCs of decomposition level (S) and transformation scale (L) evaluation experiments on AD/Hc, AD/MCI, MCI/Hc, and MCIc/MCInc data sets.

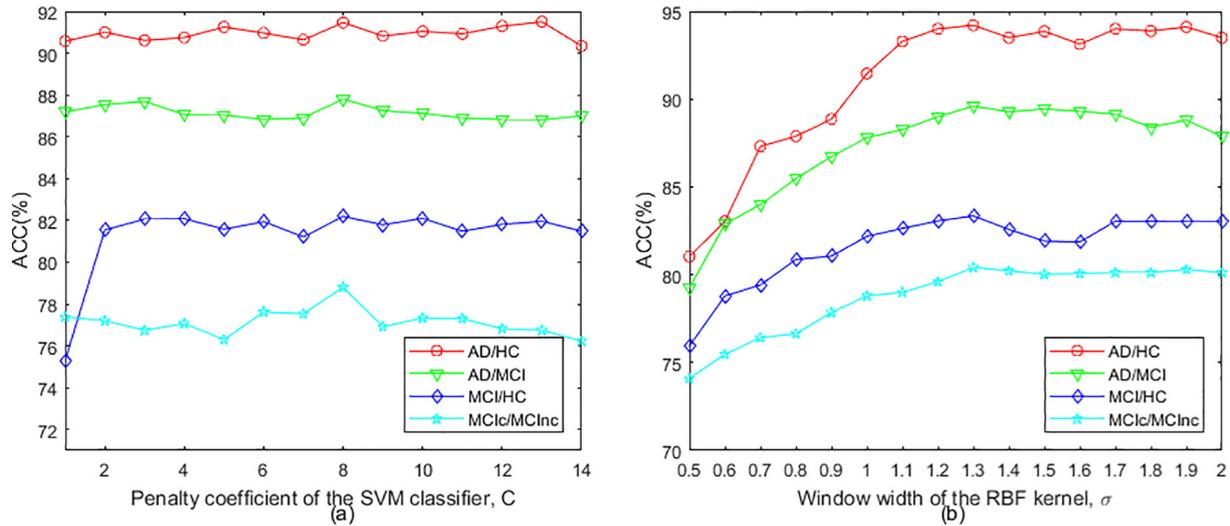


Fig. 5. ACCs of experiments with different C and  $\sigma$  values on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets.

Table 2

Experimental results of our NCSIN method on AD/HC, AD/MCI, MCI/HC and MCIc/MCInc data sets with the number of columns in the RGM image  $D = 136$ , the decomposition level  $S = 3$ , the transformation scale  $L = 1$ , the penalty coefficient of the SVM classifier  $C = 8$ , and the window width of the RBF kernel  $\sigma = 1.3$ .

Data sets	AD/HC	AD/MCI	MCI/HC	MCIc/MCInc
ACC (%)	94.21	90.03	84.64	79.42
Se (%)	96.58	91.00	89.71	82.30
Sp (%)	92.44	89.50	77.45	76.55

Table 3

Experimental results of the five other state-of-the-art comparison approaches and our NCSIN method on the AD/HC data set.

Features	ACC (%)	Se (%)	Sp (%)
Hipp-F (2019) [33]	87.51	87.60	87.42
CDTWS-F (2018) [35]	90.16	90.22	90.15
SWE-F (2018) [34]	92.70	93.67	91.77
RSBN-F (2019) [32]	92.92	94.00	89.85
IHN-F (2018) [49]	93.95	91.07	95.66
Our NCSIN	94.21	96.58	92.44

situation that the nonsubsampling contourlet is a redundant transform alleviated to a large extent.

For the SVM classifier with the RBF kernel, we also estimate its optimal penalty coefficient  $C$  by experiments on the four data sets AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc when the window width of the RBF kernel  $\sigma = 1$  and the optimally estimated parameters of the nonsubsampling contourlet used. ACCs of the experiments with different penalty coefficients  $C$  on the four data sets are shown in Fig. 5(a). It can be clearly observed from Fig. 5(a) that the penalty coefficient of the SVM classifier has few influences on ACCs of the experiments on the four data sets and ACCs of the four experiments fluctuates within a very small range, except the ACC on the MCI/HC data set when  $C \leq 2$ . On the whole of Fig. 5(a), ACCs of the experiments on the four data sets are marginally worse than the ACC when  $C = 8$ . Thus, we set  $C = 8$  as the optimal value of the penalty coefficient of the SVM classifier in this study.

To obtain an optimal estimation of the window width of the RBF kernel  $\sigma$ , we set the penalty coefficient of the SVM classifier  $C = 8$  and use the optimally estimated parameters of the nonsubsampling contourlet for the estimation experiments of the window width of the RBF kernel  $\sigma$ . ACCs of experiments with different window widths of the RBF kernel  $\sigma$  are given in Fig. 5(b). As shown in

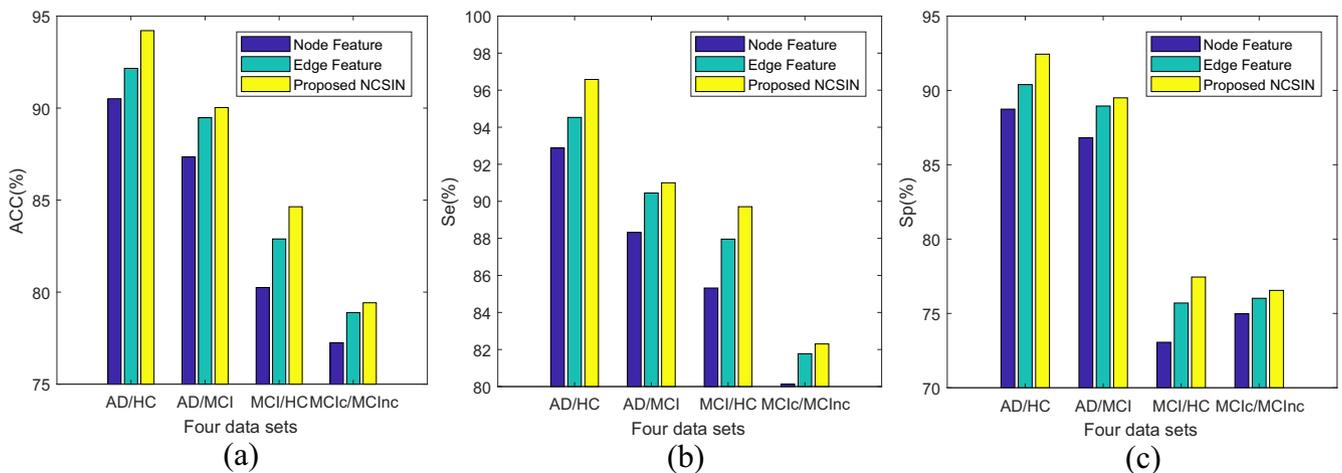


Fig. 6. Experimental results of the node and edge features of the NCSIN and the NCSIN-based feature on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets. (a) ACCs of the node and edge features of the NCSIN and the NCSIN-based feature. (b) Ses of the node and edge features of the NCSIN and the NCSIN-based feature. (c) SpS of the node and edge features of the NCSIN and the NCSIN-based feature.

**Table 4**  
Experimental results of the five other state-of-the-art comparison approaches and our NCSIN method on the AD/MCI data set.

Features	ACC (%)	Se (%)	Sp (%)
Hipp-F (2019) [33]	79.35	79.44	79.26
CDTWS-F (2018) [35]	78.48	75.35	79.98
SWE-F (2018) [34]	81.89	76.26	80.65
RSBN-F (2019) [32]	82.59	84.26	80.11
IHN-F (2018) [49]	89.31	90.83	87.92
Our NCSIN	90.03	91.00	89.53

**Table 5**  
Experimental results of the five other state-of-the-art comparison approaches and our NCSIN method on the MCI/HC data set.

Feature	ACC (%)	Se (%)	Sp (%)
Hipp-F (2019) [33]	77.25	95.79	53.23
CDTWS-F (2018) [35]	81.89	75.79	84.18
SWE-F (2018) [34]	80.67	76.79	86.98
RSBN-F (2019) [32]	83.09	83.46	82.31
IHN-F (2018) [49]	84.43	88.61	81.70
Our NCSIN	84.64	89.71	77.45

**Table 6**  
Experimental results of the five other state-of-the-art comparison approaches and our NCSIN method on the MCIc/MCInc data set.

Features	ACC (%)	Se (%)	Sp (%)
Hipp-F (2019) [33]	69.38	69.47	69.29
CDTWS-F (2018) [35]	69.21	70.74	67.45
SWE-F (2018) [34]	72.86	69.55	75.49
RSBN-F (2019) [32]	72.32	72.21	73.06
IHN-F (2018) [49]	73.96	76.13	72.25
Our NCSIN	79.42	82.30	76.55

Fig. 5(b), ACCs of the four experiment on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets increase gradually when  $\sigma \leq 1.3$ , while, ACCs on the four data sets are going to stabilize when  $\sigma > 1.3$ , but fluctuating within a small range. According to Fig. 5(b), we can find that ACCs of the experiments on the four data sets are slightly inferior to the ACC when  $\sigma = 1.3$ . With the fact of Fig. 5(b) shown, we set  $\sigma = 1.3$  as the optimally estimated value of the window width of the RBF kernel.

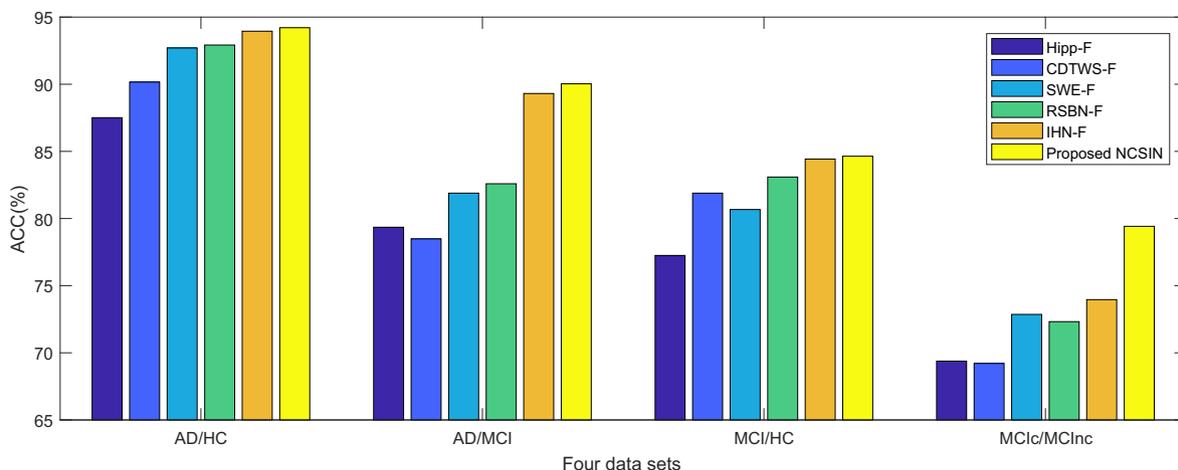
Finally, we list the ACCs, Ses and Sps of our NCSIN method on AD/HC, AD/MCI, MCI/HC and MCIc/MCInc data sets in Table 2 when the number of columns in the RGM image  $D = 136$ , the decomposition level  $S = 3$ , the transformation scale  $L = 1$ , the penalty

coefficient of the SVM classifier  $C = 8$ , and the window width of the RBF kernel  $\sigma = 1.3$ . According to experimental results listed in Table 2, ACCs of our NCSIN method on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets are 94.21%, 90.03%, 84.64% and 79.42%, respectively, and Ses of our NCSIN method consistently outperforms those of Sps, the reason of which is that brain atrophies of the patients with AD and MCI can be captured by energy subband features, and meanwhile, correlations of different subbands are also contained in the NCSIN-based feature. Therefore patients with AD and MCI can be precisely identified.

Meanwhile, we also compare the discriminant power of the NCSIN's node and edge features and the NCSIN-based feature in terms of ACCs, Ses and Sps. Experimental results of the node, edge and NCSIN-based features are shown in Fig. 6, respectively. It can be seen from Fig. 6 that ACCs, Ses and Sps of the proposed NCSIN-based feature consistently outperform those of the node and edge features and ACCs, Ses and Sps of the edge feature consistently outperform those of the node feature. Obviously, the node feature is used to describe information of the directional subbands, while the edge feature is computed to capture correlations of different directional subbands, that is, the proposed NCSIN-based feature is constructed by consisting of the information and correlation of the directional subbands. According to ACCs, Ses and Sps shown in Fig. 6, it can be confirmed that the node feature is an important supplement to the edge feature. And that is why the proposed NCSIN-based feature can achieve a satisfactory result in AD classification. Therefore, we can draw the conclusions that the subband-based individual network constructed in the frequency domain is feasible, correlations and energy distributions of the brain atrophy patterns can be represented by the NCSIN-based feature, and the NCSIN-based feature can be a promising imaging marker for the clinical AD diagnosis via SMRI images.

### 3.3. Comparisons

In this subsection, we will compare our NCSIN method with five other state-of-the-art approaches on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets. For our NCSIN method, parameters are set as the optimal estimated values. For the other state-of-the-art approaches, parameters are selected based on the authors' suggestion in their work. In the following, the hippocampus-based deep learning method proposed by Li et al. [33] is denoted as Hipp-F and uses a constructed deep learning model for classification, the complex dual tree wavelet subband-based method proposed by Jha et al. [35] is denoted as CDTWS-F and uses linear discriminant analysis for classification, the stationary wavelet entropy-based



**Fig. 7.** ACCs of the five other comparison approaches and our NCSIN method on AD/HC, AD/MCI, MCI/HC, and MCIc/MCInc data sets.

method proposed by Zhang et al. [34] is denoted as SWE-F and uses a single-hidden-layer neural network for classification, the resting-state brain network-based method proposed by Ju et al. [32] is denoted as RSBN-F and uses a targeted autoencoder network for classification, and the individual hierarchical network-based method proposed Liu et al. [49] is denoted as IHN-F and uses SVM with default values for classification.

For the AD/HC data set, the experimental results of the five state-of-the-art comparison approaches and our NCSIN method are listed in Table 3. According to ACCs, Ses and Sps shown in Table 3, ACC of our NCSIN method is consistently higher than those of the five state-of-the-art comparison approaches, and the ACC of our NCSIN method has reached to 93.57% and is 0.26% higher than that of the best comparison approach. Additionally, Se of our NCSIN method also consistently outperform those of the five state-of-the-art comparison approaches, the Se of our NCSIN method is 96.58% and is 2.58% higher than that of the best comparison approach. However, the Sp of our NCSIN method is 3.22% lower than that of the best comparison approach, which is 92.44%, the reason of

which is that with aging, older HC individuals also have the same brain atrophy patterns with the early AD patient and our NCSIN method mainly captures brain atrophy patterns by subbands in the frequency domain, so some older HC individuals are identified as patients in classification.

For the AD/MCI data set, the experimental results of the five state-of-the-art comparison approaches and our NCSIN method are listed in Table 4. It can be obviously seen from Table 4 that the ACC, Se and Sp of our NCSIN method consistently outperforms those of the five state-of-the-art comparison approaches. The ACC of our NCSIN method is 90.03%, which is 0.72% higher than that of the best comparison approach. The Se of our NCSIN method is 0.17% higher than that of the five state-of-the-art comparison approaches, which has reached to 91.00%. The Se of our NCSIN method is 89.53% and is 1.16% higher than that of the best comparison approach.

For the MCI/MCIc data set, the experimental results of the five state-of-the-art comparison approaches and our NCSIN method are listed in Table 5. According to ACCs, Ses and Sps shown in

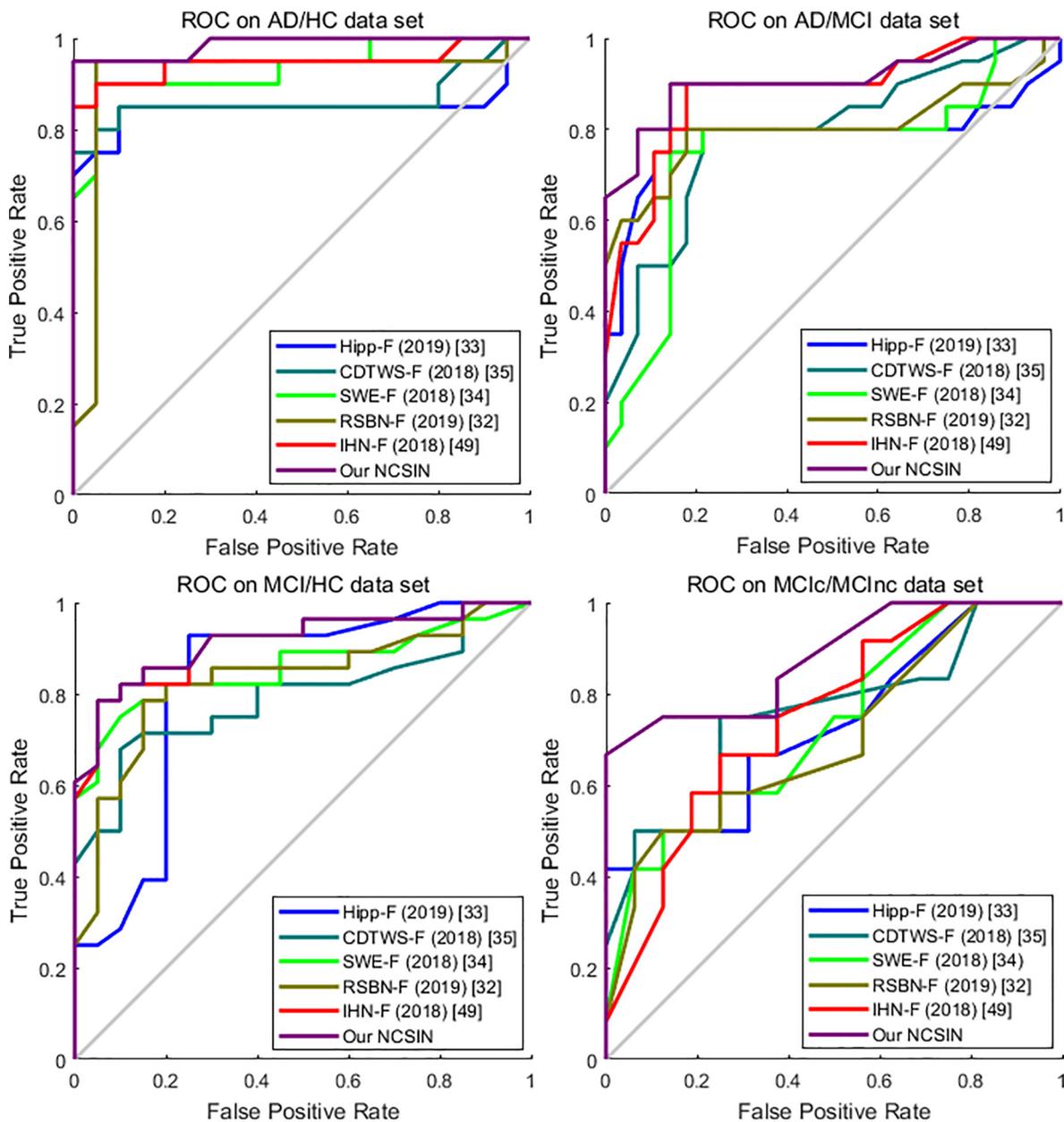


Fig. 8. ROC curves of the five other comparison approaches and our NCSIN method on AD/HC, AD/MCI, MCI/HC, and MCIc/MCIc data sets.

**Table 5**, ACC of our NCSIN method is consistently higher than those of the five state-of-the-art comparison approaches, and the ACC of our NCSIN method has reached to 84.64% and is 0.21% higher than that of the best comparison approach. However, Se and Sp of our NCSIN method are 6.08% and 9.53% lower than those of the best comparison approaches, respectively. The reason is that for AD, hippocampus is the first brain tissue affected by this disease, and thus the Hipp-F extracted from the hippocampus can precisely capture differences between subjects with MCI and HC and obtain the best Se; additionally, MCIc patients have HC-like atrophy patterns, whose identification is mainly based on more detail information, and therefore the SWE-F extracted by multi-level stationary wavelet transform can get the best Sp. Based on ACCs of our NCSIN method shown in **Tables 4 and 5**, we can find that subjects with MCI have more HC-like and in contrast less AD-like energy distribution patterns in their sMRI images, that is, brain tissues of patients with MCI are only partly destroyed to a small extent and this is the best time to benefit from the clinical treatment. Therefore, MCI is a key stage for the progression of a patient and the identification of patients with MCI can make significant sense in clinics.

For the challenging MCIc/MCIinc data set, the experimental of the five state-of-the-art comparison approaches and our NCSIN method are listed in **Table 6**. It can be clearly observed from **Table 6** that the ACC, Se and Sp of our proposed NCSIN method consistently outperform those of the five state-of-the-art comparison approaches. The ACC of our NCSIN method is 79.42%, which is obviously higher than 73.96% of the best comparison approach. The Se of our NCSIN method is 6.17% higher than that of the five state-of-the-art comparison approaches, which has reached to 82.30%. The Sp of our NCSIN method is 76.55% and is 1.06% higher than that of the best comparison approach.

According to **Tables 3–6**, we can conclude that identification performance of the NCSIN method outperforms these of the five comparison approaches and the individual network constructed with subbands is feasible, meaning that features extracted from the NCSIN can capture correlations and energy distributions of the brain atrophy patterns distributed in the subbands.

For the visual comparison, ACCs of five other comparison approaches and our proposed NCSIN method on AD/HC, AD/MCI, MCI/HC, and MCIc/MCIinc data sets are shown in **Fig. 7**. Obviously, the ACCs of our NCSIN consistently outperform the five other comparison approaches on the four data sets. For the AD/HC, AD/MCI and MCI/HC data sets, the ACCs of our NCSIN marginally higher than those of the five other comparison approaches. While, the ACC of our NCSIN is significantly higher than those of the five other comparison approaches on the challenging MCIc/MCIinc data sets, which means that the subtle differences of brain atrophy patterns between subjects with MCIc and MCIinc can be easily captured by subbands in the frequency domain. In order to better evaluate our NCSIN and make the evaluation indexes more convincing, we show the receiver operating characteristic (ROC) curves of the five other comparison approaches and our NCSIN method on AD/HC, AD/MCI, MCI/HC, and MCIc/MCIinc data sets in **Fig. 8**. It can be clearly observed from **Fig. 8** that areas under the ROC curves of our NCSIN method are consistently larger than those of the five state-of-the-art approaches on the four data sets, indicating that the proposed NCSIN feature has relatively high diagnostic accuracy in terms of AD classification.

In summary, experimental results on four data sets have verified the feasibility of constructing the individual network with subbands in the frequency domain, and also demonstrate that our NCSIN method outperforms the five other state-of-the-art approaches in terms of the accuracies, sensitivities and specificities, indicating that features extracted from our constructed NCSIN

can be a promising imaging marker for the clinical AD diagnosis via MRI images.

#### 4. Conclusions

In this study, we propose a novel method to construct nonsub-sampled contourlet subband-based individual networks (NCSINs) by capturing correlations of abnormal energy distribution patterns related to AD. Firstly, the 2D representation of the preprocessed sMRI image is obtained by downsampling and reconstruction steps. Then, the nonsubsampled contourlet transform is performed on the 2D representation to get its directional subbands at different scales. For directional subbands at one scale, each of them is described by a column energy feature vector that is regarded as a node of the NCSIN. Subsequently, the NCSIN is constructed using edges computed by PCC. Finally, the concatenation of node and edge features of the NCSINs at different scales is used as a network-based feature of the sMRI image for AD classification. Experimental results have verified the feasibility of constructing the subband-based individual network in the frequency domain for AD classification, which provides a new clue to analyze the sMRI image in frequency domain. Moreover, features extracted from the subband-based individual network have the relatively low dimension, which can be used as an imaging marker for AD diagnosis via MRI images. However, extracted features with our NCSIN cannot give a clear biological meaning and have no directly correlation with brain atrophy regions related to AD. Therefore, in the future work, we will construct a subband-based individual network that can extract features from brain atrophy regions for AD classification.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

This paper was supported by the National Natural Science Foundation of China (Grant No. 61873202, 61473232, 91430111, 31771476, 81471047, 31930022, and 11871456), National Key R&D Program (Grant No. 2017YFA0505500 and 2016YFC0903400), Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB38040400), and Shanghai Municipal Science and Technology Major Project (Grant No. 2017SHZDZX01).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimers Association; Alzheimers Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics.

The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimers Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.neucom.2020.09.012>. Code and sample IDs of this paper can be downloaded from <https://github.com/NWPU-903PR/NCSIN.git>.

## References

- [1] T. Tong, Q. Gao, R. Guerrero, 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) (2017) 155–165.
- [2] S. Luo, X. Li, J. Li, Automatic Alzheimer's disease recognition from MRI data using deep learning method, *J. Appl. Math. Phys.* 5 (9) (2017) 1892.
- [3] C.A. Budgeon, K. Murray, B.A. Turlach, et al., Constructing longitudinal disease progression curves using sparse, short-term individual data with an application to Alzheimer's disease, *Stat. Med.* 36 (17) (2017) 2720–2734.
- [4] M. Moetesum, I. Siddiqi, N. Vincent, F. Cloppet, Assessing visual attributes of handwriting for prediction of neurological disorders-A case study on Parkinson's disease, *Pattern Recognit. Lett.* 121 (2019) 19–27.
- [5] M.A. El-Yacoubi, S. Garcia-Salicetti, C. Kahindo, A.S. Rigaud, V. Cristancho-Lacroix, From aging to early-stage Alzheimer's: Uncovering handwriting multimodal behaviors by semi-supervised learning and sequential representation learning, *Pattern Recognit.* 86 (2019) 112–133.
- [6] A.F. Mendelson, M.A. Zuluaga, M. Lorenzi, et al., Selection bias in the reported performances of AD classification pipelines, *NeuroImage* 14 (2017) 400–416.
- [7] J.W. Vogel, E. Vachon-Preseau, A.P. Binette, et al., Brain properties predict proximity to symptom onset in sporadic Alzheimer's disease, *Brain* 141 (6) (2018) 1871–1883.
- [8] B. Cheng, M. Liu, D. Zhang, et al., Robust multi-label transfer feature learning for early diagnosis of Alzheimer's disease, *Brain Imag. Behav.* 13 (1) (2019) 138–153.
- [9] S.E. Viswanath, P. Tiwari, G. Lee, et al., Dimensionality reduction-based fusion approaches for imaging and non-imaging biomedical data: Concepts, workflow, and use-cases, *BMC Med. Imag.* 17 (1) (2017) 2.
- [10] G. Giulietti, M. Torso, L. Serra, et al., Whole brain white matter histogram analysis of diffusion tensor imaging data detects microstructural damage in mild cognitive impairment and Alzheimer's disease patients, *J. Magnet. Res. Imag.* 48 (3) (2018) 767–779.
- [11] O. Geman, O.A. Postolache, I. Chiuchisan, et al., An intelligent assistive tool using exergaming and response surface methodology for patients with brain disorders, *IEEE Access* 7 (2019) 21502–21513.
- [12] J. Jia, C. Wei, S. Chen, et al., The cost of Alzheimer's disease in China and re-estimation of costs worldwide, *Alzheimer's & Dementia* 14 (4) (2018) 483–491.
- [13] M.C. Carrillo, L.J. Bain, G.B. Frisoni, et al., Worldwide Alzheimer's disease neuroimaging initiative, *Alzheimer's & Dementia* 8 (4) (2012) 337–342.
- [14] A. Association, Alzheimer's association report: 2018 Alzheimer's disease facts and figures, *Alzheimer's & Dementia* 14 (2018) 367–429.
- [15] J. Dukart, F. Kherif, K. Mueller, et al., Generative FDG-PET and MRI model of aging and disease progression in Alzheimer's disease, *PLoS Comput. Biol.* 9 (4) (2013) e1002987–e1002997.
- [16] R. Cuingnet, E. Gerardin, J. Tessieras, et al., Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database, *NeuroImage* 56 (2) (2011) 766–781.
- [17] I. Álvarez Illán, J.M. Góriz, J. Ramírez, et al., Projecting independent components of SPECT images for computer aided diagnosis of Alzheimer's disease, *Pattern Recogn. Lett.* 31 (11) (2010) 1342–1347.
- [18] F.E.A. El-Gamal, M.M. Elmogy, M. Ghazal, et al., A novel early diagnosis system for mild cognitive impairment based on local region analysis: A pilot study, *Front. Human Neurosci.* 11 (2018) 643.
- [19] C.R. Jack, M.A. Bernstein, N.C. Fox, The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods, *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for, Magnet. Res. Med.* 27 (4) (2008) 685–691.
- [20] S. Guo, C. Lai, C. Wu, et al., Conversion discriminative analysis on mild cognitive impairment using multiple cortical features from MR images, *Front. Aging Neurosci.* 9 (2017) 146–159.
- [21] M. Liu, J. Zhang, D. Nie, et al., Anatomical landmark based deep feature representation for MR images in brain disease diagnosis, *IEEE J. Biomed. Health Inform.* 22 (5) (2018) 1476–1485.
- [22] I. Beheshti, H. Demirel, H. Matsuda, et al., 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 (2017) 109–119.
- [23] T. Glzman, J. Solomon, F. Pestilli, et al., Shape-attributes of brain structures as biomarkers for Alzheimer's disease, *J. Alzheimer's Disease* 56 (1) (2017) 287–295.
- [24] J. Liu, J. Wang, Z. Tang, et al., Improving Alzheimer's disease classification by combining multiple measures, *IEEE/ACM Trans. Comput. Biol. Bioinf.* 15 (5) (2018) 1649–1659.
- [25] A. Cevik, G.W. Weber, B.M. Evuboglu, et al., Voxel-MARS: a method for early detection of Alzheimer's disease by classification of structural brain MRI, *Ann. Oper. Res.* 258 (1) (2017) 31–57.
- [26] S. Lu, Y. Xia, W. Cai, et al., Early identification of mild cognitive impairment using incomplete random forest-robust support vector machine and FDG-PET imaging, *Comput. Med. Imag. Graphic.* 60 (2017) 35–41.
- [27] J. Peng, X. Zhu, Y. Wang, L. An, D. Shen, Structured sparsity regularized multiple kernel learning for Alzheimer's disease diagnosis, *Pattern Recognit.* 88 (2019) 370–382.
- [28] D. Cárdenas-Pena, D. Collazos-Huertas, G. Castellanos-Dominguez, Enhanced data representation by kernel metric learning for dementia diagnosis, *Front. Neurosci.* 11 (2017) 413.
- [29] D.P. Varikuti, S. Genon, A. Sotiras, et al., Evaluation of non-negative matrix factorization of grey matter in age prediction, *NeuroImage* 173 (2018) 394–410.
- [30] C.D. Stefano, F. Fontanella, D. Impedovo, G. Pirlo, A.S.D. Freca, Handwriting analysis to support neurodegenerative disease diagnosis: A review, *Pattern Recognit.* 121 (2019) 37–45.
- [31] Z. Wang, X. Zhu, E. Adeli, et al., Multi-modal classification of neurodegenerative disease by progressive graph-based transductive learning, *Medic. Imag. Anal.* 39 (2017) 218–230.
- [32] R. Ju, C. Li, Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning, *IEEE/ACM Trans. Comput. Biol. Bioinf.* 16 (1) (2019) 244–257.
- [33] H. Li, M. Habes, D.A. Wolk, et al., A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data, *Alzheimer's & Dementia* 15 (8) (2019) 139–150.
- [34] Y. Zhang, S. Wang, Y. Sui, et al., Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization, *J. Alzheimer's Disease* 65 (3) (2018) 855–869.
- [35] D. Jha, S. Alam, J.Y. Pyun, et al., Alzheimer's disease detection using extrame learning machine, complex dual tree wavelet principal coefficients and linear discriminant analysis, *J. Medic. Imag. Health Inf.* 8 (5) (2018) 881–890.
- [36] P. Ganasala, V. Kumar, CT and MR image fusion scheme in nonsubsampling contourlet transform domain, *J. Digit. Imag.* 27 (3) (2014) 407–418.
- [37] Y. Shen, E. Ren, J.W. Wang, et al., A nonsubsampling contourlet transform based medical image fusion method, *Inf. Tech. J.* 12 (4) (2013) 749–755.
- [38] G. Yang, M. Li, L. Chen, et al., The nonsubsampling contourlet transform based statistical medical image fusion using generalized gaussian density, *Comput. Math. Method Med.* (2015) 262819.
- [39] M. Nazrudeen, M. Rajalakshmi, S. Sureshkumar, Medical image fusion using non-subsampling contourlet transform, *Int. J. Eng. Res. Tech.* 3 (3) (2014) 1248–1252.
- [40] Y. Li, Y. Wang, G. Wu, et al., Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features, *Neurobiol. Aging* 33 (2) (2012) 427.e15–427.e30.
- [41] B. Jie, M. Liu, D. Shen, Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease, *Med. Imag. Anal.* 47 (2018) 81–94.
- [42] D. Lu, K. Popuri, G.W. Ding, R. Balachandar, et al., Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FEG-PET images, *Sci. Rep.* 8 (2018) 5697–5709.
- [43] J. Zhao, Y. Du, X. Ding, et al., Alteration of functional connectivity in patients with Alzheimer's disease revealed by resting-state functional magnetic imaging, *Neur. Regener. Res.* 15 (2) (2020) 285–292.
- [44] J. Shi, X. Zheng, Y. Li, et al., Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease, *IEEE J. Biomed. Health Inf.* 22 (1) (2018) 173–183.
- [45] C.Y. Wee, P.T. Yap, K. Denny, et al., Resting-state multi-spectrum functional connectivity networks for identification of MCI patients, *PLoS One* 7 (5) (2012) e37828–e37838.
- [46] B. Duraisamy, J.V. Shanmugam, J. Annamalai, Alzheimer disease detection from structural MR images using FCM based weighted probabilistic neural network, *Brain Imag. Behav.* 13 (1) (2019) 87–110.
- [47] C.C. Huang, S.H. Lin, C.P. Lin, The effect of network template from normal subjects in the detection of network impairment, *Acta. Neurobiol. Exp.* 76 (2016) 294–303.
- [48] J. Liu, M. Li, W. Lan, et al., Classification of Alzheimer's disease using whole brain hierarchical network, *IEEE/ACM Trans. Comput. Biol. Bioinf.* 55 (2) (2018) 624–632.
- [49] J. Liu, J. Wang, B. Hu, F.X. Wu, Y. Pan, Alzheimer's disease classification based on individual hierarchical networks constructed with 3-D texture features, *IEEE Trans. Nanobiosci.* 16 (6) (2018) 428–437.

- [50] Y. Dong, D. Tao, X. Li, et al., Texture classification and retrieval using shearlets and linear regression, *IEEE Trans. Cybern.* 45 (3) (2015) 358–369.
- [51] J. Ashburner, T. Group, Statistical parametric mapping (SPM8), Functional Imaging Laboratory, Institute of Neurology 12, Queen Square, London WC1N 3BG, UK.
- [52] J. Ashburner, K.J. Friston, Voxel-based morphometry—the methods, *Neuroimage* 11 (6) (2000) 805–821.
- [53] J. Feng, X. Liu, Y. Dong, et al., Structural difference histogram representation for texture image classification, *IET Image Process.* 11 (2) (2016) 118–125.
- [54] Y. Dong, J. Feng, L. Liang, et al., Multiscale sampling based texture image classification, *IEEE Signal Process. Lett.* 24 (5) (2017) 614–618.
- [55] M.N. Do, M. Vetterli, The contourlet transform: An efficient directional multiresolution image representation, *IEEE Trans. Image Process.* 14 (12) (Dec. 2005) 2091–2106.
- [56] N. An, H. Ding, J. Yang, et al., Deep ensemble learning for Alzheimers disease classification, *J. Biomed. Infor.* 105 (2020) 103411.
- [57] R. Rodriguezsanchez, J.A. Garcia, J. Fdezvaldivia, et al., Image inpainting with nonsubsampling contourlet transform, *Pattern Recog. Lett.* 34 (13) (2013) 1508–1518.
- [58] A.L. Cunha, J. Zhou, M.N. Do, The nonsubsampling contourlet transform: Theory, design, and applications, *IEEE Trans. Image Process.* 15 (10) (2006) 3089–3100.

pattern recognition theory and machine learning to bioinformatics and biological systems.



**Luonan Chen:** He received the BS degree from Huazhong University of Science and Technology, Wuhan, China, in 1984, and the MS and PhD degrees from Tohoku University, Sendai, Japan, in 1988 and 1991, respectively. Since 2010, he has been a professor and executive director at Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. His fields of interest are computational systems biology, bioinformatics, and nonlinear dynamics. In recent years, he published over 350 journal papers and two monographs in the area of computational systems biology.



**Jie Xia:** She received the BS degree from Nanjing Normal University, Nanjing, China, in 2007, and the MS degrees from Yangzhou University, Yangzhou, China, in 2011. Since 2011, she has been an assistant research fellow at Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. Her fields of interest are computational systems biology, bioinformatics.



**Jinwang Feng:** He received his BS degree and MS degree from Luoyang Institute of Science and Technology and Henan University of Science and Technology in 2010 and 2017, respectively. He was a teacher of Zhengzhou Technical College between 2017 and 2018. From 2018 on, he started pursuing his PhD degree of Northwestern Polytechnical University. Now he is studying in Key Laboratory of Systems Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy Science. His current research interests include medical image processing, pattern recognition, machine learning, and computer vision.



**Shao-Wu Zhang:** He received his PhD Degree in Control Theory and Control Engineering from Northwestern Polytechnical University, China. He is currently a Professor in the School of Automation, Northwestern Polytechnical University, China. He worked as a visiting scholar in the Department of Electrical and Computer Engineering, University of Texas San Antonio, USA in May, 2016; as a visiting scholar and staff in the Computational Biology and Bioinformatics Group, University of Southern California, USA from 2008 to 2009; as a staff in the Department of Computing, Hong Kong Polytechnic University in Aug., 2008. His recent interests include

RNA methylation pattern mining, complex biological networks, Brain imaging genomics, lncRNA prediction, drug-target prediction, metagenomics, applications of