



A CAD system design to diagnose alzheimers disease from MRI brain images using optimal deep neural network

Pemmu Raghavaiah¹ · S. Varadarajan²

Received: 9 October 2020 / Revised: 8 January 2021 / Accepted: 7 April 2021

Published online: 30 April 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Memory related issues in brain are mainly caused by Alzheimer disease (AD) which is the most common form of dementia. This disease must be diagnosed in its prodromal stage known as Mild Cognitive Impairment (MCI) also it needs an accurate detection and classification technique. In this paper, a computer-aided diagnosis (CAD) system is implemented on Magnetic resonance imaging (MRI) data from ADNI database. This disease highly affects the Hippocampus and cerebrum regions which are normally found in the grey matter region of brain. At first, MNI/ICBM atlas space of every three dimensional MRI images are constructed using normalization procedure, then grey matter region of brain is extracted. Subsequently, feature extraction is done by two dimensional Gabor filter in three scales and eight orientations. Then, the proposed optimal Deep Neural Network (DNN) classifier is used to classify the images as Cognitive normal (CN), Alzheimer disease (AD), and Mild Cognitive Impairment (MCI). Here, DNN classifier is optimized by selecting optimal weight parameter using Enhanced Squirrel Search Algorithm. The experimental results prove an efficiency of the proposed method using MR images. The proposed algorithm beats existing techniques in terms of accuracy, sensitivity, and specificity.

Keywords Alzheimer disease · Mild cognitive impairment · Gabor filter · Deep neural network · Enhanced squirrel search algorithm

✉ Pemmu Raghavaiah
raghavaiahpemmu@gmail.com

¹ Department of Electronics and Communication Engineering, University College of Engineering, JNTUK, Kakinada, India

² Department of Electronics and Communication Engineering, SVU College of Engineering, Sri Venkateswara University, Tirupati, India

1 Introduction

Alzheimer Disease (AD) is a permanent brain sickness with dynamic weakness of memory and cognitive functions. In an adult age of humans, AD is the most common case of dementia [14]. Till now, there is no productive treatment is discovered for AD. But Medical community has introduced various treatments in order to detect AD at an early stage. Also, it is very crucial to diagnose the disease in earlier stage. But, early diagnosis of AD and mild cognitive impairment (MCI) still has several challenging issues [20]. In early stage, this disease is called as MCI in which the loss of memory is mild whereas, in the later stage the memory loss is heavy even the patient can't carry on a conversation. Various medical tests are essential to diagnosis the AD which leads to large amount of multivariate heterogeneous data [8]. Because of the heterogeneous nature of the medical tests, it can be definitely exhausted to manually analyze, visualize and compare this data. Neuro-biological and physical tests, clinical dementia rate (CDR), minimal state examination (MMSE) and patient history are the medical assessments which are essential for AD diagnosis [5].

Additionally, the structural and functional changes of the brain are studied by non-invasive techniques like resting-state functional MRI (rs-fMRI) and structural MRI. Because of AD, ventricles in the brain enlarges whereas, the cerebral cortex and the hippocampus shrinks [24]. Judgment, thinking, planning and memory are the tasks which are affected by the changes in the hippocampus and cerebrum regions. The magnitude of change relies on the progression phase of disease in different areas of brain [2]. Using the MRI, it is easy to identify the enlarged ventricles, volume reduction in the cerebral cortex and hippocampus. Developing an approach is essential for the clinicians to differentiate the healthy subjects and brain disorder [13]. Previously encouraging outcomes are accomplished by the medical community and now contributions in the area of detection and classification of AD are developing. In MRI, the AD and normal ageing are differentiated by looking at the visible areas of brain in which it requires a broad experience and knowledge [3, 21].

For the analysis of MCI or AD, the biomedical signal processing approaches are widely employed which includes functional magnetic resonance imaging, positron emission tomography, structural MRI; meanwhile, in several applications some of the machine learning techniques achieved encouraging results [15, 29]. Because of high spatial resolution and clear contrast, the structural MRI analyzes MCI subjects and the patients with AD also, it provides the information regarding the type of brain tissue [22]. Various structural features like hippocampus volumes, cortical thickness and tissue probability maps are considered as features for MCI subjects and AD patients [23]. Detecting the disease based on imaging biomarker is widely investigated by Brain morphometric pattern analysis from structural MRI. Diagnosing AD and its prodromal stage (i.e., MCI) with the help of computer is considered as an interesting and challenging task [19]. When compared with other widely used biomarkers like fluorodeoxyglucose positron emission tomography, and cerebrospinal fluid, the MRI provides a sensitive non-aggressive solution to detect the changes in the abnormal structure of brain [18].

The AD's diagnosis procedure is improved by introducing non-invasive approach known as Computer-Aided Diagnosis (CAD) system [11]. CAD system provides more accurate diagnosis and early information with good performance. In clinical practice, CAD system is very useful for neurologists [25]. Moreover, accurate results from CAD become crucial to provide neuro degeneration care in the early stage of AD in recent years. Previously, the CAD is designed to extract the low level features from neuroimaging data also it is considered as the

classification problem which ignores the latent feature wise relation [26]. According to the recent neuroscience aspects, the multiple brain regions in the human brain are anatomically and functionally interlinked. Therefore the extracted features from different brain regions are extensively related to features of other regions. In this case, the complementary information is useful in multimodal fusion among different neuroimaging modalities [4, 27]. Furthermore, the segmentation and classification processes have several challenges in the field of CAD based AD diagnosis. Hence, there is a need of huge development of accurate approaches for better diagnosis.

Design of a CAD system to estimate the classification performance on grey matter of structural 3D MRI for AD is the main impact of this paper. Here, MR images are initially bias corrected, normalized and segmented using SPM tool. After smoothening, Gabor filter is used to extract texture features from gray matter of brain image. Finally, an optimal DNN is proposed for the classification in order to reach maximum classification accuracy. The optimal DNN is developed by selecting optimal weight using an ESSA in which the parameter is dynamically adjusted using fuzzy logic.

Section 2 briefly describes the literature study and how the present work framed to overcome the research gap in early detection of AD. Section 3 clearly elaborates the proposed methodology. Experimental results and performance evaluation are described in Section 4. Finally, the conclusion of the research work is presented in Section 5.

2 Related works

A multivariate approach was developed by Yudong et al. [30] to detect the AD. They aimed to diagnose the brain MR images automatically using a new machine learning system. In which, a stationary wavelet entropy was used to extract the textural information of brain image. This method utilized four stage decomposition to provide thirteen stationary wavelet entropy (SWE) features. Then, the neural network with single hidden layer was used to classify the image. The stability of this method was increased by PP-PSO (Predator-Prey Particle Swarm Optimization).

Ruoxuan Cui and Manhua Liu [7] used the benefits of both convolutional and recurrent neural networks to diagnose AD by longitudinal analysis of structural MR images. Initially, they constructed Convolutional Neural Network (CNN) to perform classification by learning MRI's spatial features. Next, a recurrent Neural Network (RNN) was constructed with a cascade structure of three bidirectional gated recurrent units (BGRU) layers to extract longitudinal features for AD classification. The optimal performance was achieved by joining the spatial and longitudinal disease classifier. Additionally, the longitudinal analysis with RNN was modeled at various time points using imaging data.

Shui-Hua et al. [28] detected the Alzheimer's disease by selecting single slice from three dimensional volumetric data with the help of inter-class variance principle. This detection algorithm was worked on the basis of multilayer perceptron (MLP), biogeography-base optimization and wavelet entropy (WE). This WE and MLP were utilized as effectual instruments to analyze the medical images. The best AD detection rate was obtained using WE. The universal approximation method approximated the MLP to satisfy the requirements of AD detection algorithm.

Ronghui Ju et al. [10] described a deep learning network with clinical relevant text information for AD diagnosis at the early stage. Resting-state functional magnetic resonance

imaging (R-fMRI) provided the functional connectivity of brain regions and it was used by the authors to construct a brain network. At the early stage, an auto encoder network was used to distinguish normal aging from MCI. This method effectively exposed the features of discriminative brain network which supports reliable classification of AD detection. Unlike the traditional classifiers, the proposed model was more reliable and stable. The proposed work was useful to predict and prevent AD at an early stage, by excavate the benefits of deep learning network for medical services.

Jialin Peng et al. [22] considered the high-dimensional multi-modality imaging and genetic data for AD diagnosis. A multiple structured kernel learning uses the advantages of both phenotype and genotype information. Particularly, structured feature selection and fusion from heterogeneous modalities were facilitated and feature-wise importance was observed. According to the modalities, kernels were grouped and then each feature was represented with a different kernel. A data-driven approach was used to combine the kernel presentation of multimodal features optimally. The authors evaluated the proposed work using subjects from ANDI database, which clearly discovered the brain regions and SNPs related to AD.

The main contribution of this paper is listed as follows:

- This paper mainly aims to validate the classification accuracy of Gabor texture feature on MRI image to diagnose Alzheimer disease. However, the dimension of the Gabor features are very high. Hence, the dimensionality of the Gabor feature vector is reduced by computing the statistics of mean and standard deviation for all three projections of MRI image. This represents the texture feature of MRI image efficiently.
- To improve the classification accuracy of the Alzheimer disease diagnosis system, an optimal DNN is proposed. Since, the effectiveness of DNN model is very sensitive to its weight parameters because of the direct relationship with the objective function, they need to be selected optimally. In this work, the classification accuracy of the deep stacked sparse auto-encoder is increased by selecting optimal weight parameters using ESSA algorithm.
- To modify the SSA algorithm by introducing a fuzzy logic to find an enhanced solution. In SSA, gliding constant controls the exploration and exploitation of the search space. This parameter need to be adjusted dynamically to get better outputs. In this paper, a fuzzy logic is introduced to adapt the parameters dynamically based on an iteration.

3 Proposed methodology

A computer-aided diagnosis (CAD) is proposed in this paper to detect AD in early stage. This system involves three levels of processing as shown in Fig. 1.

Initially, MRI data preprocessing is done by Statistical Parameter Mapping (SPM) tool. All images are corrected for bias field in similarities also they are normalized and segmented into cerebrospinal fluid (CSF), white matter (WM), and gray matter (GM). Most of the higher order functions of brain are linked by the network within GM. The brain damage caused by Alzheimer are related to higher order parts of the brain. Next to the segmentation, spatial smoothing of GM images is done by Gaussian filtering. This work only uses GM images for further processing. Gabor filter in three scales and eight orientations is used to extract the

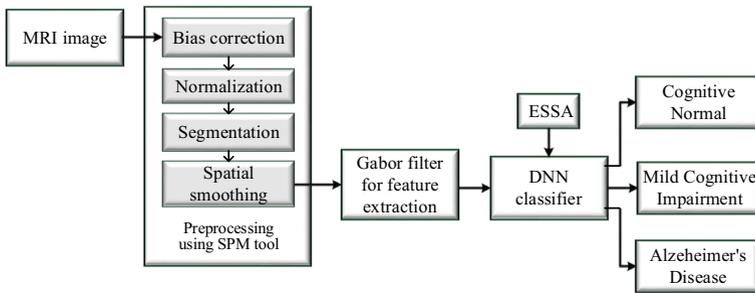


Fig. 1 Proposed methodology

texture features. Mean and standard deviation are obtained from the magnitude of Gabor response in slice by slice manner which represents the text features. Finally the extracted features are given to Deep Neural Network (DNN) for classifying Normal, MCI and AD images. Here, the weight complexities in the classifier will be solved with the Enhanced Squirrel Search Algorithm (ESSA).

3.1 Preprocessing

Coronal, transversal and sagittal views of the subject are presented in the dataset of 3D MRI. It is very difficult to segment the WM, GM and CSF from MRI due to the non-uniformity (bias) of tissue intensities and noise artifacts. A low frequency signal which degrades the MRI image due to the non-homogenous nature of magnetic fields in MRI scanners is known as bias. The intensities of the images may vary while capturing the MRI scan due to the usage of various scanners to scan different subjects or the same subject at a different time. Also, the size of the brain is varied from person to person. Hence, there is a requirement of bias correction and normalization for reducing the non homogeneities present in brain MRI images.

Therefore, the MATLAB software package called Statistical Parameter Mapping (SPM) 12 tool is used to pre-process all MRI images for segmenting the brain tissue accurately. This software is accessible at <http://www.fil.ion.ucl.ac.uk/spm>, also provided by the Wellcome Trust Centre for Neuroimaging, London, UK. This tool implements statistical approaches to analyze the structural neuro images. Here, the MRI image is initially mapped on tissue probability maps for the extraction of mapped portions. Then, the image is corrected for bias and normalized to segment the image into GM, WM and CSF. In which, every tissue classes are allotted with different numbers of Gaussians on the basis of the classical values provided in [6] (i.e., 3 for GM, 2 for WM and 2 for CSF). Total volumes of the tissues are differentiated due to the warping or normalization procedure which maps a collection of images to a template image. A modulation process is applied to update the voxel intensities also it rejects those volume differences.

Here, images are registered to minimize the effect of non-tissue structures like scalp on the segmented image. Then, an 8 mm full-width-half-maximum (FWHM) Gaussian kernel is used to segment the smoothed images. Resultant images are in the Montreal Neurological Institute (MNI) space whose dimensions and resolutions are (121–145–121) and 1.5 mm respectively. The texture features are only extracted from the GM images.

3.2 Feature extraction

Here, 2D Gabor filter in three scales and eight orientations are used to extract the features of the GM image [12]. The standard deviation and mean of Gabor magnitude are calculated to represent the texture features on slice by slice basis. Gabor filter is already known as a linear filter. Sinusoidal signals that are complex in nature are modulated along with Gaussian envelope by this Gabor filter. Frequency resolution and optimal joint space are provided by the Gabor filter. The following expressions define the two dimensional Gabor filter. [31].

$$G(s, t; f_c, \varphi) = \frac{f_c^2}{\prod \xi \chi} \exp\left(-\left(\frac{f_c^2}{\xi^2} \bar{s}^2 + \frac{f_c^2}{\chi^2} \bar{t}^2\right)\right) \times \exp(j2\pi f_c \bar{s}) \quad (1)$$

$$\bar{s} = s \cos \varphi + t \sin \varphi \quad (2)$$

$$\bar{t} = -t \sin \varphi + t \cos \varphi \quad (3)$$

Here, the sharpness along minor axis is mentioned by χ , the sharpness along major axis is denoted by ξ , φ mentions the rotation angle between the x-axis of spatial domain and direction of sinusoidal wave likewise central frequency of filter is denoted by f_c . Gabor filter sets x frequencies also f_c and φ are varied to configure the y orientations which are mentioned in (1)–(3) as:

$$\varphi = \frac{k\pi}{l}, k = \{0, 1, l-1\} \quad (4)$$

$$f_c = m^{-1} f_{\max}, m = \{0, 1, l-1\} \quad (5)$$

At highest frequency, central frequency of filter is mentioned by f_{\max} . Better results are properly obtained by setting various parameters of Gabor filter. Number of orientation, number of scale, and central frequency of the highest frequency f_{\max} are the influenced parameters. This work takes 8 and 3 as number of orientations and number of scales. The following expressions are used to acquire the values of χ and ξ [9].

$$\xi = \frac{1}{\prod} \left(\frac{l+1}{l-1} \right) \sqrt{-\ln q_1} \quad (6)$$

$$\chi = \frac{1}{\prod} \frac{\sqrt{-\ln q_2}}{\prod / 2^v} \quad (7)$$

Here, q_1 , q_2 , and q values are set as 0.5. Although the significant content falls between the filters, the value of q is experimentally confirmed and found that the data loss is moderate. Based on the following expression, the maximum frequency is computed indirectly from ξ .

$$f_{\max} = \frac{\xi}{2 \left(\xi + \frac{\sqrt{\ln 2}}{\Pi} \right)} \tag{8}$$

Here, G_{uv} represents the Gabor filter at scale u and orientation v . Gabor response $GR_{uv}(s, t)$ is provided for an 2d input image $I(s, t)$ by convolving input image with Gabor filter in spatial domain. Imaginary part $I(GR_{uv})$ and real part $R(GR_{uv})$ are two different parts in complex coefficient of GR_{uv} . The following expressions are used to compute the phase and magnitude of Gabor response g_{uv} for every single slice [17].

$$g_{uv}(GR_{uv}) = \sqrt{\text{Re}^2(GR_{uv}) + \text{Im}^2(GR_{uv})} \tag{9}$$

$$\phi(GR_{uv}) = \arctan\left(\frac{\text{Im}(GR_{uv})}{\text{Re}(GR_{uv})}\right) \tag{10}$$

On various frequencies, Gabor filter behaved as scaled version of each other. The better detail is obtained with higher frequency. At different scale, Gabor filter is applied to obtain finite course details of MR images. But, high dimension is obtained by computing Gabor features. Therefore, reduction of dimension is needed in this calculation. These computations use both the standard deviation and mean statistics. Feature vector of the Gabor filter response is obtained by magnitude of mean and standard deviation.

Axial, sagittal and coronal projections are the three feature vectors taken in this study. Here, w_1, w_2 and w_3 are the number of slices of axial, coronal and sagittal projection also, the following expressions explains the corresponding feature vector.

$$F_1 = [m_{00a}(1), s_{00a}(1), m_{01a}(1), s_{01a}(1), \dots, m_{27a}(1), s_{27a}(1), \dots, m_{27a}(w_1), s_{27a}(w_1)] \tag{11}$$

$$F_2 = [m_{00c}(1), s_{00c}(1), m_{01c}(1), s_{01c}(1), \dots, m_{27c}(1), s_{27c}(1), \dots, m_{27c}(w_1), s_{27c}(w_2)] \tag{12}$$

$$F_3 = [m_{00s}(1), s_{00s}(1), m_{01s}(1), s_{01s}(1), \dots, m_{27s}(1), s_{27s}(1), \dots, m_{27s}(w_1), s_{27a}(w_s)] \tag{13}$$

Here, mean of magnitude of Gabor response at scale u , orientation v along axial, coronal and sagittal projections at slice w is represented by $m_{uva}(w), m_{uvc}(w), m_{uvs}(w)$ respectively. Also, $s_{uva}(w), s_{uvc}(w), s_{uvs}(w)$ represents standard deviation of magnitude of Gabor response at scale u , orientation along axial, coronal and sagittal projections respectively at slice w .

3.3 Optimal DNN classifier

The normal image and AD is classified by optimal DNN classifier in this section. The effectual and discriminative features of an image are learned by stacked sparse auto-encoders [31] in deep learning. An auto-encoder is a symmetrical kind of neural network and it is utilized to learn the features of an image from enormous datasets. Auto-encoders minimize the training errors and losses to improve the performances. Input layer, hidden layer and output layers are presented in every stacked auto encoder. Features are coded at every hidden layer and the

coded features are then transferred to next layer. DNN’s output layer is softmax layer which is conjunction with the stacked auto-encoders to classify learned features. The performance of this network can be improved by selecting optimal weight using ESSA. The optimal DNN structure is shown in Fig. 2.

Consider the input features vector $F \in R^{M \times 1}$ and the logistic sigmoid function $(1 + \exp(-x))^{-1}$ which is used to encode the vector F. The encoded data is expressed as:

$$s_k = \text{sigm}(w_{k1}F + b_{k1}) \tag{14}$$

Where $w_{k1} \in R^{M \times 1}$ represents the weight matrix which connects the input vector and the code, $b_{k1} \in R^{M \times 1}$ denotes the coding bias. The code of a single layer sparse auto-encoder is mentioned by $s_k \in R^{M \times 1}$. Code s_k is decoded by utilizing another decoder matrix and it is given as:

$$\hat{F} = \text{sigm}(w_{k2}s_k + b_{k2}) \tag{15}$$

Where $w_{k2} \in R^{M \times N}$ denotes the weight matrix which connects the code and the reconstructed data, $b_{k2} \in R^{M \times N}$ represents a deciphering bias. Hence, according to Eqs. (14) and (15) $h_{w_k, b_k}(F_k)$ express \hat{F}_k where the subscript k represents the hidden layer number. Consider

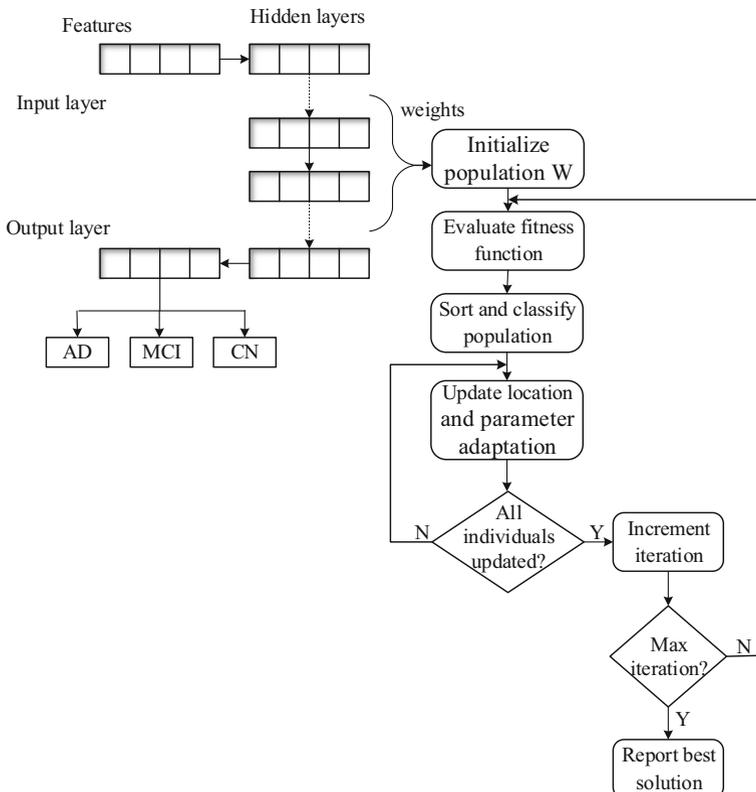


Fig. 2 Optimal DNN classifier

w_k and b_k to represent (w_{k1}, w_{k2}) and (b_{k1}, b_{k2}) respectively. The objective function for diminishing the error between reconstructed data and input data can be written as:

$$J(w_k, b_k, F_k) = \frac{1}{2} \|h_{w_k, b_k}(F_k) - F_k\|^2 \tag{16}$$

Prevention of over complete nonlinear mapping of the input vector and learning trivial mapping are done by the auto encoder which adds two regularization parameters in the weight and a sparse unit. Hence, objective function can be rewritten as:

$$J(w_k, b_k, F_k) = \frac{1}{2} \|h_{w_k, b_k}(F_k) - F_k\|^2 + \alpha (\|w_k\|^2 + \eta \sum_{i=1}^M KL(\rho \|\hat{\rho}_i)) \tag{17}$$

Where α represents the regularization element utilized for balancing the reconstruction objective and weight decline, η denotes the sparse penalty weight, ρ defines the required activation probability, $\hat{\rho}_i$ represents the mean of i^{th} component activation probability in s_k after trained using m data. Kullback–Leibler divergence is expressed as

$$KL(\rho \|\hat{\rho}) = \rho \log \frac{\rho}{\rho_i} + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_i} \tag{18}$$

The performance of this network is affected by the weight parameter (w_k) due to the direct proportionality of this parameter with objective function. Thus, it is important to identify the exact weight parameter (w_k). Hence, the parameter (w_k) is optimized to minimize the objective function $J(w_k, b_k, F_k)$. Here we use ESSA to optimize the weight parameters. The optimization procedure is explained in the following subsection. This optimal DNN classifier’s output layer is a softmax layer which is used to classify data into multiple classes (i.e., normal, MCI and AD classes). The activation function for softmax layer (output layer of DNN) is given as:

$$h_\varphi(F_k) = \begin{bmatrix} p(y_k = 1 | F_k; \varphi) \\ p(y_k = 2 | F_k; \varphi) \\ \vdots \\ p(y_k = n | F_k; \varphi) \end{bmatrix} = \frac{1}{\sum_{i=1}^n \exp^{\varphi_i F_k}} \begin{bmatrix} e^{\varphi_1 F_1} \\ e^{\varphi_2 F_1} \\ \vdots \\ e^{\varphi_n F_1} \end{bmatrix} \tag{19}$$

Where $y_k \in \{1, 2, \dots, n\}$, $\varphi_1, \varphi_2 \dots \varphi_n \in \mathfrak{R}^{m+1}$ and $1 / \sum_{i=1}^n \exp^{\varphi_i F_k}$ is utilized for normalizing the distribution. The one who have maximum probability in all categories is selected as an inputting feature vector.

3.3.1 Enhanced squirrel search algorithm for weight optimization

To find the optimization accuracy, the Squirrel Search Algorithm (SSA) [9] is one of the most recently used optimization algorithm. The gliding constant G_c has an important role in the operation of SSA through the fuzzy system. Also, it improves the exploration and exploitation competency of SSA. The dynamic hunting performance of the southern flying squirrels and their effective way of locomotion named as gliding were imitated by this algorithm. To travel the long distance, small mammals used this efficient gliding mechanism. ESSA begins with

randomly initialized location of flying squirrels. The given below steps of ESSA are used to find the weight optimization.

(1) Random initialization

In this step, k number of flying squirrels is randomly initialized in the initial population W.

$$W = \begin{bmatrix} W_{1,1} & W_{1,2} & \cdots & W_{1,d} \\ W_{2,1} & W_{2,2} & \cdots & W_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ W_{k,1} & W_{2,2} & \cdots & W_{k,d} \end{bmatrix} \quad (20)$$

Here, i^{th} flying squirrel in the j^{th} dimension (i.e., number of problem variables) is mentioned by $W_{i,j}$. An initial position of every single flying squirrel in the forest is allocated based on a uniform distribution.

$$W_i = W_L + U(0, 1) \times (W_U - W_L) \quad (21)$$

Here, the upper and lower bounds of i^{th} flying squirrel in j^{th} dimension are represented W_U and W_L , also it is uniformly distributed in the range of random number $[0, 1]$ which is denoted by $U(0, 1)$.

(2) Calculation of fitness function and selection

In the total population, expression (17) is used to compute fitness function for every single individual (solution). According to fitness value, these squirrels are allocated with global best, best and other solutions. An optimal selection of weight in DNN classifier is computed based on the fitness function and the following array stores the corresponding values.

$$f = \begin{bmatrix} f(|W_{1,1}, W_{1,2}, \dots, W_{1,d}|) \\ f(|W_{2,1}, W_{2,2}, \dots, W_{2,d}|) \\ \vdots \\ f(|W_{n,1}, W_{n,2}, \dots, W_{n,d}|) \end{bmatrix} \quad (22)$$

Where $f = (J(w_k, b_k, f_k))$. After the storage of fitness value, the array is sorted in ascending order. In the hickory tree $(W_{i,j})_{ht}$ (at the optimal food source), the best one with minimum fitness value is placed. In acorn tree $(W_{i,j})_{at}$ (normal source food), next three best are placed and in normal tree $(W_{i,j})_{nt}$ (no food source), the remaining squirrels are placed. The squirrels that gather their daily food source are then shifted to hickory tree, or else they shift their location to acorn tree to collect the food source.

(3) Update location

Based on the type of evaluated individual, new position of every single flying squirrel is calculated in the population. Presence probability of a predator (P) is used to update the position. Three cases are illustrated to update three types of behavior based on that predator (P).

- Case1: In this case, to obtain optimal food source flying squirrels are moved towards the hickory tree from the acorn tree which is expressed in Eq. (23)

$$(W_{i,j})_{at} = \left\{ \begin{array}{ll} (W_{i,j})_{at} + D_g \times G_c \times ((W_{i,j})_{ht} - (W_{i,j})_{at}) & \text{if } r_1 \geq P \\ \text{random location} & \text{otherwise} \end{array} \right\} \quad (23)$$

- Case 2: In this case, to acquire their food the flying squirrels are move towards the acorn trees from normal trees which is expressed in Eq. (24).

$$(W_{i,j})_{nt} = \left\{ \begin{array}{ll} (W_{i,j})_{nt} + D_g \times G_c \times ((W_{i,j})_{at} - (W_{i,j})_{nt}) & \text{if } r_2 \geq P \\ \text{random location} & \text{otherwise} \end{array} \right\} \quad (24)$$

- Case 3: In this case, some squirrels are found in normal trees for directly move towards the hickory tree to attain optimal food source, where the other squirrels in normal tree will move towards the acorn tree. Also it is expressed in Eq. (25).

$$(W_{i,j})_{nt} = \left\{ \begin{array}{ll} (W_{i,j})_{nt} + D_g \times G_c \times ((W_{i,j})_{ht} - (W_{i,j})_{nt}) & \text{if } r_3 \geq P \\ \text{random location} & \text{otherwise} \end{array} \right\} \quad (25)$$

Here, the random gliding distance D_g is situated in the interval [0.5, 1.11], the random numbers distributed uniformly in the range of [0, 1] which is mentioned by r_1, r_2 and r_3 , G_c is a gliding constant that controls the exploration and exploitation of the search space in the range of [0 2]. During the implementation, fuzzy logic is used in the dynamic adjustment of this parameters which generates good results.

The main difference of using dynamic adjustment of parameters with respect to using fixed parameters is that the selected parameters are used as fuzzy parameters which are modified in the iteration. So, based on iteration, parameter values are dynamically changed to find an enhanced solution. The iterations are the input fuzzy system which produces G_c as output.

Low, medium, and high are taken as input variables in which they are three triangular membership functions used to produce three triangular functions of the same type as low, medium and high with 0–2 range. Expression (26) is used to find out the iteration value.

$$\text{iteration} = \frac{\text{present iteration}}{\text{Total iteration}} \quad (26)$$

In that fuzzy system, the exploration step is achieved in the initial iteration and last iteration achieves the exploitation step. The fuzzy system rules are given as:

- If (Iteration is Low) then (G_c is Low)
- If (Iteration is Medium) then (G_c is Medium)
- If (Iteration is High) then (G_c is High)

(4) Check termination condition

The fitness value of flying squirrels is calculated in every single iteration. Also, the positions are updated until it reaches the maximum number of iterations.

(5) Report the best solution

Final optimal solution is obtained by the location of squirrel on the hickory tree.

The dynamic adaption of parameters in SSA algorithm and the optimal DNN classifier is used to enhance the performance of proposed CAD design.

4 Experimental results and discussion

Experimental setup and their consequences are briefly explained in this section. Three sub-sections named as data set description, evaluation metrics and performance evaluations are comprised in this section. First one explains the details of the experimental dataset which we have used in our work, the prediction measures which are used to determine the performance of proposed method is explained in second sub section and third one explains the quantitative result and comparison with advanced methods.

4.1 Dataset description and evaluation metrics

The performance of our proposed method is evaluated by using ADNI 1 database which is available in <http://adni.loni.usc.edu/>. T1, T2 and PD weighted sequences are used to obtain the MR images in ADNI database using 1.5 T and 3 T MRI systems. AD, normal, and MCI are three different cases of brain which is in the MR image dataset. Totally 280 subjects are selected for our work in which AD comprises 69 subjects, MCI covers 122 subjects and normal class contains 89 subjects. Number of subjects, mini-mental state examination, and age are the selected subject's demographic features which is represented in Table 1.

AD/CN classification, AD/MCI classification, and MCI/CN classification are three types of AD diagnosis which is considered in the evaluation of proposed method. Accuracy, sensitivity and specificity, False Negative Rate (FNR) and False Positive Rate (FPR) are taken to validate the proposed method.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (27)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (28)$$

$$Specificity = \frac{TN}{TN + FP} \quad (29)$$

$$FPR = \frac{FP}{FP + TN} \quad (30)$$

Table 1 Demographic information of 280 subjects (from ADNI database)

Diagnosis	Number of subjects	Age	MMSE
Normal	89	74.9±5.4	29.5±0.8
MCI	122	71.5±6.7	26.8±1.6
AD	69	76.4±5.9	23.4±2.0

$$FNR = \frac{FN}{FN + TP} \quad (31)$$

Here, FN represents the individuals with disease which are categorized as healthy, FN mentions the healthy individuals which are categorized as diseased, TN denotes the healthy individuals which are correctly classified and TP represents the individuals with disease which are correctly classified. AUC (area under receiver operating characteristic (ROC) curve) is an important index which is used to estimate the overall performance of the classification method. The highest AUC value provides better classification results in proposed classifier.

4.2 Evaluation of the proposed work

Initially, the data set is separated into training set, testing set, also validation set. Training data of optimal DNN is divided with 80: 20 ratio in our work which mentions that 80% data was used for training and testing phase uses 20% data. Effective classification performance is obtained by achieving cross validation of models. All the images are preprocessed and segmented into GM, WM, and CSF tissue in standard MNI space. Except the other models, GM models are analyzed since it is projected to play an insignificant role. Figures 3, 4 and 5 show the preprocessing stages of MRI image taken from ADNI dataset for CN, MCI and AD.

In Figs. 3, 4 and 5, top left block illustrates the axial, sagittal and coronal view of original raw images of CN, MCI and AD and their respective bias corrected images are shown in top right. Bottom left and bottom right blocks represent the normalized and smoothed GM matters respectively. After normalizing, Gabor filter is applied to extract the features from GM matters. Based on these features, proposed optimal DNN classifier is trained to diagnose the AD.

Existing methods such as Individual Hierarchical Network (IHN) [17], Structured Sparsity Regularized Multiple Kernel Learning (SSRMKL) [22], Whole Brain Hierarchical Network (WBHN) [16], and Hippocampal Visual Features (HVF) [1] are used to compare the superiority of our proposed method. Comparative results of AD/CN classification, AD/MCI classification, and MCI/NC classification are provided in Tables 2, 3, and 4.

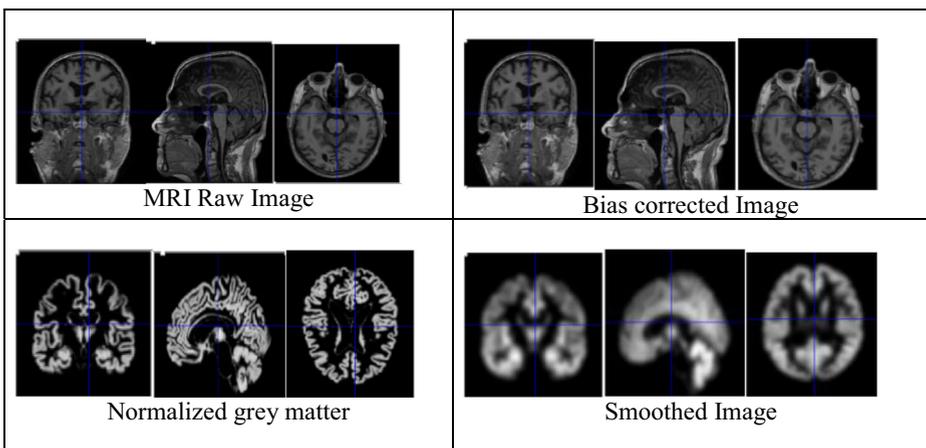


Fig. 3 Preprocessing stages of CN using SPM 12 tool

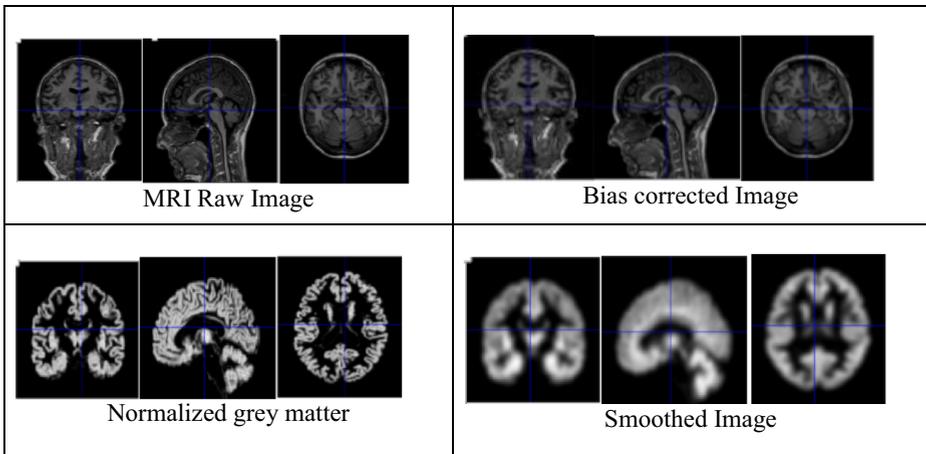


Fig. 4 Preprocessing stages of MCI using SPM 12 tool

The proposed method significantly outperforms other existing methods which are seen in Tables 2, 3, and 4. The accuracy of our method for AD/CN, AD/MCI and MCI/CN are 96.43%, 94.64% and 91.07% respectively. For all classification, our method obtains the best accuracy, specificity (96.43% for AD/CN, 96.43% for AD/MCI and 89.29% for MCI/CN) and sensitivity (96.43% for AD/CN, 92.86% for AD/MCI and 92.86% for MCI/CN) with high value. The overall performance of the diagnostic test is measured by AUC. So, the proposed method is more efficient and robust than existing methods in terms of AUC result.

Figure 6 displays the accuracy, sensitivity, and specificity result of our multi class proposed classifier with existing classifiers like SVM, Ensemble and KNN [37]. It is evident from Fig. 6 that our classifier based on texture feature representation using Gabor filter outperforms all other classifier which uses GLCM. In order to prove the efficiency of the proposed HRF-DNN, we further plot the ROC curves of different classifiers including SVM, Ensemble and KNN. Figure 7 compares the ROC plots of different classifiers for multiclass classifications problems. The corresponding AUC values of the existing classifier for Normal vs AD vs MCI

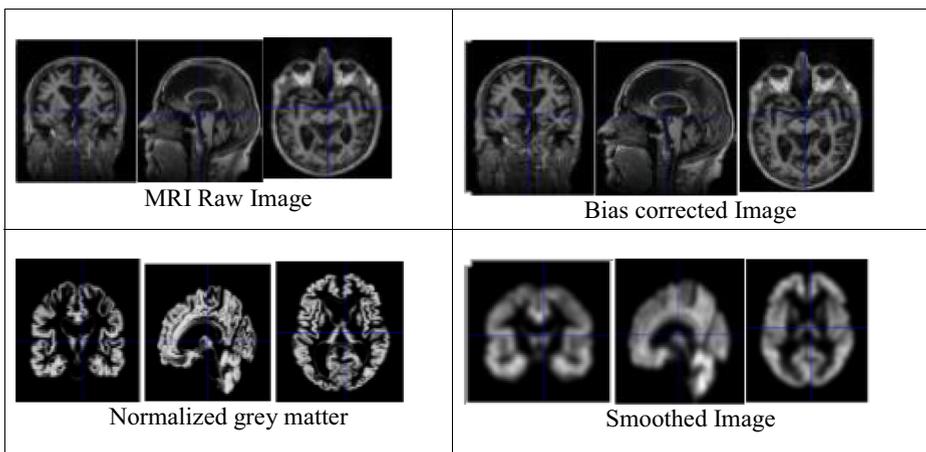


Fig. 5 Preprocessing stages of AD using SPM 12 tool

Table 2 Comparison for AD/CN classification

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC	FPR	FNR
IHN	95.37	92.49	96.08	0.967	0.039	0.075
SSRMKL	96.1	97.3	94.9	0.992	0.051	0.027
WBHN	94.65	95.03	91.76	0.954	0.082	0.049
HVF	80.40	77.61	93.28	0.849	0.067	0.224
Proposed	96.43	96.43	96.43	0.964	0.035	0.035

Table 3 Comparison for AD/MCI classification

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC	FPR	FNR
IHN	90.41	92.83	88.82	0.919	0.111	0.071
SSRMKL	76.9	65.9	82.7	0.808	0.173	0.341
WBHN	88.63	91.55	86.25	0.907	0.137	0.084
HVF	74.51	77.94	71.23	0.756	0.287	0.220
Proposed	94.64	92.86	96.43	0.946	0.035	0.071

Table 4 Comparison for MCI/CN classification

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC	FPR	FNR
IHN	86.56	90.74	84.83	0.854	0.151	0.092
SSRMKL	80.3	85.6	69.8	0.811	0.302	0.144
WBHN	84.79	88.91	80.34	0.826	0.196	0.111
HVF	76.29	72.30	81.53	0.768	0.184	0.277
Proposed	91.07	92.86	89.29	0.910	0.101	0.071

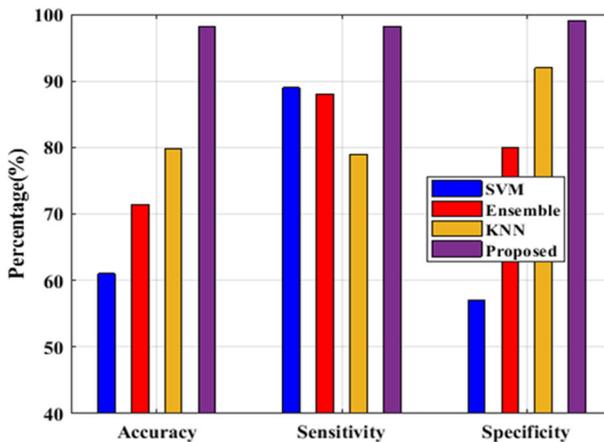


Fig. 6 Classification results of Normal vs AD vs MCI

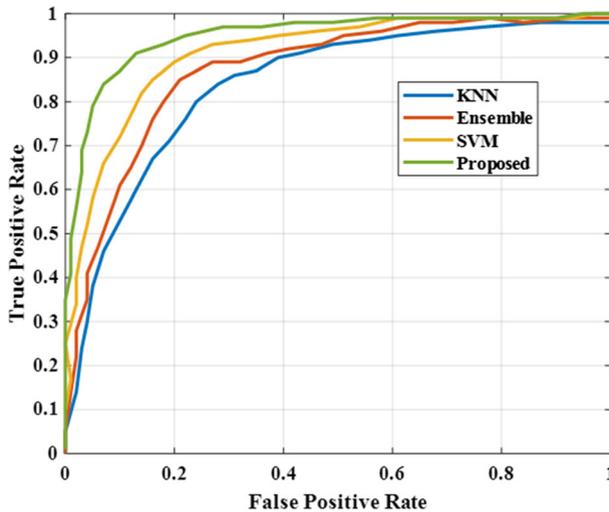


Fig. 7 ROC analysis for Normal vs AD vs MCI classification problem

classification are 0.89, 0.85 and 0.82 respectively. Instead, our proposed classifier obtains AUC of 0.91, showing better classification ability than other methods. In summary, the results show that our proposed method can improve the classification results.

5 Conclusion

The CAD is design with optimal DNN classifier to classify Alzheimer Disease, MCI and cognitive normal in our research work. After preprocessing, the texture features of MRI are extracted using Gabor filter. At different frequency and orientation, details of the image are captured by this filter. These features are used by optimal DNN classifier to diagnose the AD. Similarly, ESSA with dynamic parameter adaptation enhances the performance by choosing optimal weight parameter. The AD of a patient is correctly identified by the proposed method which improves the efficiency of medical treatment.

References

1. Ahmed OB, Benois-Pineau J, Allard M, Amar CB, Catheline G, Alzheimer's Disease Neuroimaging Initiative (2015) Classification of Alzheimer's disease subjects from MRI using hippocampal visual features. *Multimed Tools Appl* 74(4):1249–1266
2. Altaf T, Anwar SM, Gul N, Majeed MN, Majid M (2018) Multi-class Alzheimer's disease classification using image and clinical features. *Biomed Sig Process Control* 43:64–74
3. Babulal GM, Quiroz YT, Albenis BC, Arenaza-Urquijo E, Astell AJ, Babiloni C, Bahar-Fuchs A et al (2019) Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: Update and areas of immediate need. *Alzheimer's Dement* 15(2):292–312
4. Bartos A, Gregus D, Ibrahim I, Tintëra J (2019) Brain volumes and their ratios in Alzheimer's disease on magnetic resonance imaging segmented using Freesurfer 6.0. *Neuroimaging, Psychiatry Research*

5. Bilderbeck AC, Penninx BWJH, Arango C, van der Wee N, Kahn R, Rossum I W-v, Hayden A, Kas MJ, Post A, Dawson GR (2019) Overview of the clinical implementation of a study exploring social withdrawal in patients with schizophrenia and Alzheimer's disease. *Neurosci Biobehav Rev* 97:87–93
6. Çevik A, Weber G-W, Eyüboğlu BM, Oğuz KK, Alzheimer's Disease Neuroimaging Initiative (2017) Voxel-MARS: a method for early detection of Alzheimer's disease by classification of structural brain MRI. *Ann Oper Res* 258(1):31–57
7. Cui R, Liu M, Initiative A's DN (2019) RNN-based longitudinal analysis for diagnosis of Alzheimer's disease. *Comput Med Imaging Graph* 73:1–10
8. Jain R, Jain N, Aggarwal A, Hemanth DJ (2019) Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cogn Syst Res* 57:147–159
9. Jain M, Singh V, Rani A (2019) A novel nature-inspired algorithm for optimization: squirrel search algorithm. *Swarm Evol Comput* 44:148–175
10. Ju R, Hu C, Pan Z, Li Q (2019) Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning. *IEEE/ACM Trans Comput Biol Bioinform (TCBB)* 16(1):244–257
11. Karami V, Nittari G, Amenta F (2019) Neuroimaging computer-aided diagnosis systems for Alzheimer's disease. *Int J Imaging Syst Technol* 29(1):83–94
12. Keserwani P, Pammi V S C, Prakash O, Khare A, Jeon M (2016) Classification of Alzheimer Disease using Gabor Texture Feature of Hippocampus Region. *Int J Image Graph Sig Process* 8, no. 6
13. Li H-C, Chen P-Y, Cheng H-F, Kuo Y-M, Huang C-C (2019) In vivo visualization of brain vasculature in Alzheimer's disease mice by high-frequency micro-Doppler imaging. *IEEE Trans Biomed Eng* 66:3393–3401
14. Li F, Liu M, Initiative A's DN (2018) Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks. *Comput Med Imaging Graph* 70:101–110
15. Lin S-Y, Lin C-P, Hsieh T-J, Lin C-F, Chen S-H, Chao Y-P, Chen Y-S, Hsu C-C, Kuo L-W (2019) Multiparametric graph theoretical analysis reveals altered structural and functional network topology in Alzheimer's disease. *NeuroImage: Clinical* 22:101680
16. Liu J, Li M, Lan W, Wu F-X, Pan Y, Wang J (2016) Classification of alzheimer's disease using whole brain hierarchical network. *IEEE/ACM Trans Comput Biol Bioinforma* 15(2):624–632
17. Liu J, Wang J, Hu B, Wu F-X, Pan Y (2017) Alzheimer's disease classification based on individual hierarchical networks constructed with 3-D texture features. *IEEE Trans Nanobiosci* 16(6):428–437
18. Liu M, Zhang J, Adeli E, Shen D (2018) Joint classification and regression via deep multi-task Multi-Channel learning for Alzheimer's disease diagnosis. *IEEE Trans Biomed Eng* 66(5):1195–1206
19. Liu M, Zhang J, Lian C, Shen D (2019) Weakly supervised deep learning for brain disease prognosis using MRI and incomplete clinical scores. *IEEE Trans Cybern*:1–12
20. Meyer SRA, De Jonghe JFM, Schmand B, Ponds RWHM (2019) Visual associations to retrieve episodic memory across healthy elderly, mild cognitive impairment, and patients with Alzheimer's disease. *Aging Neuropsychol Cognit* 26(3):447–462
21. Pandya MD, Shah PD, Jardosh S (2019) Medical image diagnosis for disease detection: A deep learning approach. In *U-Healthcare Monitoring Systems*, pp. 37–60. Academic Press
22. Peng J, Zhu X, Wang Y, An L, Shen D (2019) Structured sparsity regularized multiple kernel learning for Alzheimer's disease diagnosis. *Pattern Recogn* 88:370–382
23. Platero C, López ME, del Carmen Tobar M, Yus M, Maestu F (2019) Discriminating Alzheimer's disease progression using a new hippocampal marker from T1-weighted MRI: The local surface roughness. *Hum Brain Mapp* 40(5):1666–1676
24. Razavi F, Tarokh MJ, Alborzi M (2019) An intelligent Alzheimer's disease diagnosis method using unsupervised feature learning. *J Big Data* 6(1):32
25. Saravanakumar S, Thangaraj P (2019) A computer aided diagnosis system for identifying Alzheimer's from MRI scan using improved Adaboost. *J Med Syst* 43(3):76
26. Shi Y, Suk H-I, Yang G, Lee S-W, Shen D (2019) Leveraging coupled interaction for multimodal Alzheimer's disease diagnosis. *IEEE Trans Neural Netw Learn Syst*
27. Wang H, Shen Y, Wang S, Xiao T, Deng L, Wang X, Zhao X (2019) Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease. *Neurocomputing* 333:145–156
28. Wang S-H, Zhang Y, Li Y-J, Jia W-J, Liu F-Y, Yang M-M, Zhang Y-D (2018) Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. *Multimed Tools Appl*:1–25
29. Xu L, Yao Z, Li J, Lv C, Zhang H, Bin H (2019) Sparse feature learning with label information for Alzheimer's disease classification based on magnetic resonance imaging. *IEEE Access* 7: 26157–26167

30. Zhang Y, Wang S, Sui Y, Yang M, Liu B, Cheng H, Sun J, Jia W, Phillips P, Gorriz JM (2018) Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization. *J Alzheimers Dis* 65(3):855–869
31. Zhang Y, Zhang E, Chen W (2016) Deep neural network for halftone image classification based on sparse auto-encoder. *Eng Appl Artif Intell* 50:245–255

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.