IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 8, 2020, accepted June 13, 2020, date of publication June 18, 2020, date of current version July 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3003424

Resting State fMRI and Improved Deep Learning Algorithm for Earlier Detection of Alzheimer's Disease

HAIBING GUO[1] **AND YONGJIN ZHANG**[2]
[1]School of Science, Jiangsu Ocean University, Lianyungang 222005, China
[2]Neurology Department, The First People's Hospital of Lianyungang, Lianyungang 222005, China

Corresponding author: Haibing Guo (rdguoguo123@163.com)

ABSTRACT The development of computerized healthcare has been powered by diagnostic imaging and machine learning techniques. In particular, recent advances in deep learning have opened a new era in support of multimedia healthcare distribution. For earlier detection of Alzheimer's disease, the study suggested the Improved Deep Learning Algorithm (IDLA) and statistically significant text information. The specific information in clinical text includes the age, sex and genes of the person and apolipoprotein E; the brain function is established using resting-state functional data (MRI) for the measurement of connectivity in the brain regions. A specialized network of autoencoders is used in earlier diagnosis to distinguish between natural aging and disorder progression. The suggested approach incorporates effectively biased neural network functionality and allows a reliable Alzheimer's disease recognition. In comparison with conventional classifiers depends on time series R-fMRI results, the proposed deep learning algorithm has improved significantly and, in the best cases, the standard deviation reduced by 45%, indicating the forecast model is more reliable and efficient in relation to conventional methodologies. The work examines the benefits of improved deep learning algorithms from recognizing high-dimensional information in healthcare and can lead to the early diagnosis and prevention of Alzheimer's disease.

INDEX TERMS Alzheimer's disease, autoencoder network, improved deep learning algorithm (IDLA), R-fMRI data.

I. INTRODUCTION

The dementia type Alzheimer's disease (AD) is described by traditional middle and old-aged thought and learning disabilities. Neuritic statues are found in the brain and degeneration of specific brain cells are the pathological features of this process [1]. The problems usually grow slowly and become severe enough to interact with everyday life. Although elderliness is the major risk factor, AD is an old-age illness [2]. In its early stages, memory loss is minimal, while the communication and capacity of the patient to respond significantly degrades during initial stages [3], [4], as shown in Figure.1. Current approaches cannot delay the progression of Alzheimer's disease (AD), but early diagnosis can help prevent the disease's occurrence and help patients improve their quality of life. The number of individuals with AD has been estimated to be increased in the next 20 years, with one of 85 individuals in 2050 [5]. The correct diagnosis is therefore very important especially in the early diagnosis of AD.

Deep learning is used for data interpretation and analysis. In addition, variations and data models can be classified. This allows for decisions that cannot usually be taken using standard processes of time saving and efforts [6].

FMRI is an effective imaging mode for assessing the interconnection of structurally separated and functionally different brain networks, especially in resting-state fMRI. The neural network which is the basis of the interactive pathogenesis can thus be established during the neural degenerative stages. AD can be diagnosed with fMRI data [7], [8]. By measuring the connectivity of the brain network in some brain regions, a network focused on each resting-state operations may be built and the loss of brain function between AD and healthy patients may be assessed [9], [10].

The MCI detection results are based on the coefficient of correlation. In addition, the clinical test data allow us to analyze and diagnosis MCI from a different perspective,

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

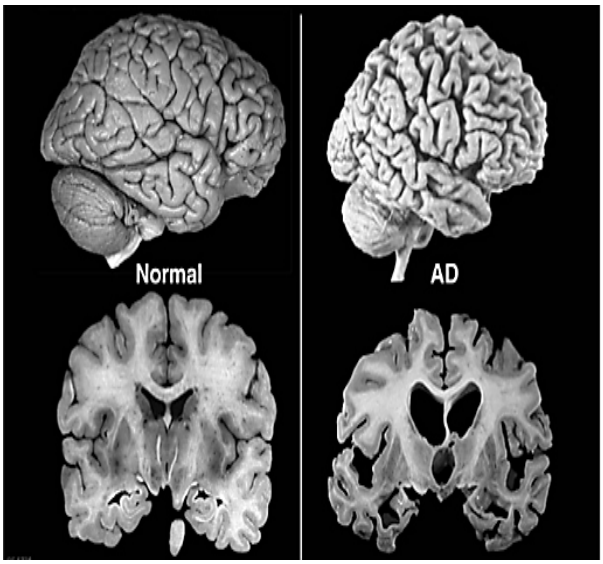


FIGURE 1. Alzheimer's disease prediction.

the relationship between MCI and other physiological factors. A deep network encoder model is then configured for these associations to be categorized [11]. A deep learning model based on stacked auto-encoders is built to remove high-dynamic data in hierarchy. The deep network of autoencoders thus effectively eliminates the biased features of brain networks and provides high precision [12], [13]. The technique significantly enhances the ability to differentiate MCIs and NCs from the traditional approach and offers support for clinical decisions on neurodegenerative diseases, in particular AD [14], [15].

The major objectives of this paper are discussed as follows

- The brain connectivity derived from image data on the brain and many other physical markers that are obtained from healthcare systems.
- The correlation coefficient data aspect is even less disease sensitive as matched with the R-fMRI Data.
- Next architecture consists of a custom autoencoder to classify MCIs of NCs.
- Distinct brain activity and physical condition can be removed on different scales for higher performance accuracy.

II. LITERATURE SURVEY

Lorenzi *et al.* [16] introduced the Disease Progression Modeling (DPM) of Alzheimer's to show short-term clinical data on the long term pathologic trajectories. DPM has the ability to show an important clinical device for automatic diagnosis, together with the capacity to provide a data-driven description of the natural development of the disease through an explicit description of biomarker transitions from normal to pathological phases across the disease axis. Throughout this analysis, they had reformulated DPM in a probability system to quantification the diagnostic complexity of the seriousness of individual diseases with regards to missing metrics, biomarkers, and follow-up data in a hypothetical clinical scenario. They indicate that the step performed in 582 amyloid positive test entities is of high diagnostic reliability. This measurement significantly minimizes the uncertainty of the forecast. The change from normal to pathologic stages is mainly related to increasing hypo-metabolism of the brain, temporal atrophy, and decreasing clinical levels.

Padilla *et al.* [17] introduced the early detection of Alzheimer's disease (AD) computer-aided diagnosis (CAD) technique on the basis of nonnegative matrix factor (NMF) and trustworthy vector machines (SVM). The CAD technique was designed to analyze and classify functional brain images. For this reason, two separate brain image databases are chosen: a single-photon emission computed tomography (SPECT) and an image of Positrons emitting tomography (PET), each with reference data of Alzheimer's disease (AD) patients and health checkups. The Fisher Discriminant ratios (FDRs) and Non-negative Matrix Factor (NMF) for the collection of the most important features are analyzed in those databases. Through the SVM-based classifier the resulting NMF-transformed data contexts containing a reduced number of functions with confidence limits for decision making.

The study of functional neuroimaging information was very critical to understand neurodegenerative diseases such as Alzheimer's disease (AD) in human brain research. The most common approach in AD neuroimaging research has been the hierarchical design, where different areas of the brain/voice are automatically evaluated. The author [18] suggested a machine-learning technique called the Sparse Inverse Covariance Analysis (SICA) in terms of developing interactions with the brain region at minimal level computation costs and an adequate level of sparsity. Each dimension of an inverse matrix covariance is a function of the component pair of variables, with all other variables, for Gaussian assumptions. Implementing a sparsity limit eliminates excessive/noisy functional dependencies through the zero setting of the component and thus ensuring dependent independence for the variable pair.

Cui and Liu [19] recently suggested an approach to the integration of global and local characteristics with the use of three-dimensional networking and structural analysis to diagnose AD from Hippocampus Analysis (HA). The proposed method can enhance classification with local visual and global formal features. Lulu Yue *et al.* [20] suggested a VHFE (Voxel-based Hierarchical Function Extraction) process for early diagnosis of AD. Next, the entire brain is wrapped in 90 regions of interest (ROIs) based on the anatomical automatic labeling (AAL) design. In order to divide uninformative results, they select the informative voxels of each ROI and position them on a vector-based on their values. The first stage characteristics are chosen according to the correlation of voxels between the various groups. First, each subject consisting of the fetched voxel was fed into its brain feature maps into a convolutional neural network (CNN) to learn the hidden characteristics.

III. MATERIALS

A. DATA SOURCE DESCRIPTION

In this paper, the database of the Alzheimer's neuroimaging disease (ADNI) association is used for choosing all the subjects. ADNI's aim is to examine the possible integration into MRI, fMRI, PET, genetic data and clinical tests of human brain structures and functionalities, as well as to minimize or cure AD.

Select all ADNI-2 fMRI data in the ADNI database (http://adni.loni.usc.edu/) because the most fMRI data is provided by ADNI-2. According to ADNI-2, the data is the most commonly accessed. The first visit by each subject is then data from the fMRI. Therefore, data is filtered according to image registration quality. This results in a total of 91 MDIs and a sample of 79 NCDs. The ages of MCI respondents aged between 66.5 and 86.3 years. The distribution of gender and age between the two groups is not noticeable. The sample images are taken from ADNI datasets which is shown in figure 2.

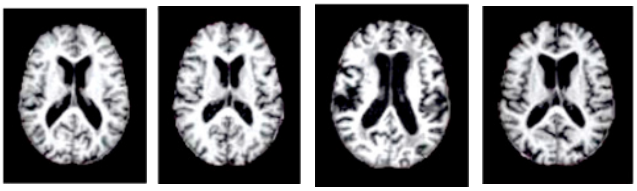


FIGURE 2. Sample images taken from ADNI datasets.

B. DATA PREPROCESSING AND ACQUISITION

Here it analyzed both the ADNI Time Series R-fMRI and the pre-processing data for the correlation coefficient. ADNI is a text of the clinical test including age, sex. The MCI estimates and analyses are based on these data. The ApoE E4 gene expression is the genetic information that is focused in this paper. The ApoE gene encodes apolipoprotein E protein, which is a brain as well as other cholesterol carrier. In-text format age, sex, and gene information is stored and ADNI downloads the DICOM data.

In order to manage original R-fMRI (http:/rfmri.org / dpabi) data, a DPABI (DataProcessing Assistant for Resting-Status fMRI) toolbox has been developed for Brain Image Data Processing & Analysis. DPABI is a toolkit based on fMRI data processing, the program SPM8 and the tools fMRI Data Analysis (REST) Resting-State toolkit. First, it has regressed a large number of confounding factors before folding changes (FC) analysis to which the impact of physiological artifacts.

IV. METHODOLOGY

A. CLASSIFICATION OF R-fMRI MODEL

Initially, it processes raw R-fMRI data in the early diagnostic method of AD as described in Section 3. The time-series (90×130) matrix are collected, which shows that blood levels have changed over the years in each brain region. The next step is to build the network of the brain. This matrix shows how different brain areas are connected with a strong brain connectivity network, which accurately and effectively represents the health situation of the brain.

Finally, together with the clinical examination information, it has the time series data and matrices in different models of extraction and comparing the classification results. Three-layer architecture is used to represent the cognitive development of the nervous system and to isolate the features of the brain network precisely. It is carried out to ensure that there is no over-fitting for limited data sample numbers. Figure 3 shows the classification processes for the detection of MCI for R-fMRI data.

B. BRAIN NETWORKS DEVELOPMENT

Different areas of the human brain are not different. The neurons are connected by a difficult network to perform advanced functions with chemistry and electrical gap junctions as the basic types of connections. However, neural connections in the brain are difficult to observe directly. Throughout this research, it is proposed to evaluate the connections between R-fMRI data in some brain regions in a simple and effective way. The correlation coefficient of the Pearson is generally used to measure the functional connection strength. This paper measures the connectivity of different regions of the brain. Check the R-fMRI results measurement before the correlation is determined. The most common and most easy to monitor is the QQ plot (figure 4) and the Shapiro-Wilk to verify the data is normally distributed.

The Pearson correlation coefficient is determined using the gaussianity of information of the time series. It can show that different brain regions are functionally related. The correlation matrix could be provided with a false discovery rate (FDR). It helps us to understand the functional connection between different brain areas with a significant impact on the study of mechanisms of brain activity, assessment of interoperability and brain health. At the end of the day, a matrix 90×90 is obtained from 90 ROIs. The classification is only made from the upper triangular matrix since the matrix of the correlation coefficient is symmetrical. Figure 5 is a brain network system which shows the ROI and connections of brain nodes and edges.

The network properties enter into the connection matrices to the auto-encoder after the above steps are performed. Because of its higher sensitivity to brain conditions, this network function can improve the deep learning algorithm's performance.

C. AUTOENCODER PRE-TRAINING

The design of an autoencoder comprises of the *ms* layers, which do not independently pre-train each layer. The following three steps have therefore been taken to attain the optimal parameters of each layer of the auto-encoder. First, the weight parameters (V, d) are initialized to close to zero values. Furthermore, with a cost minimization function,

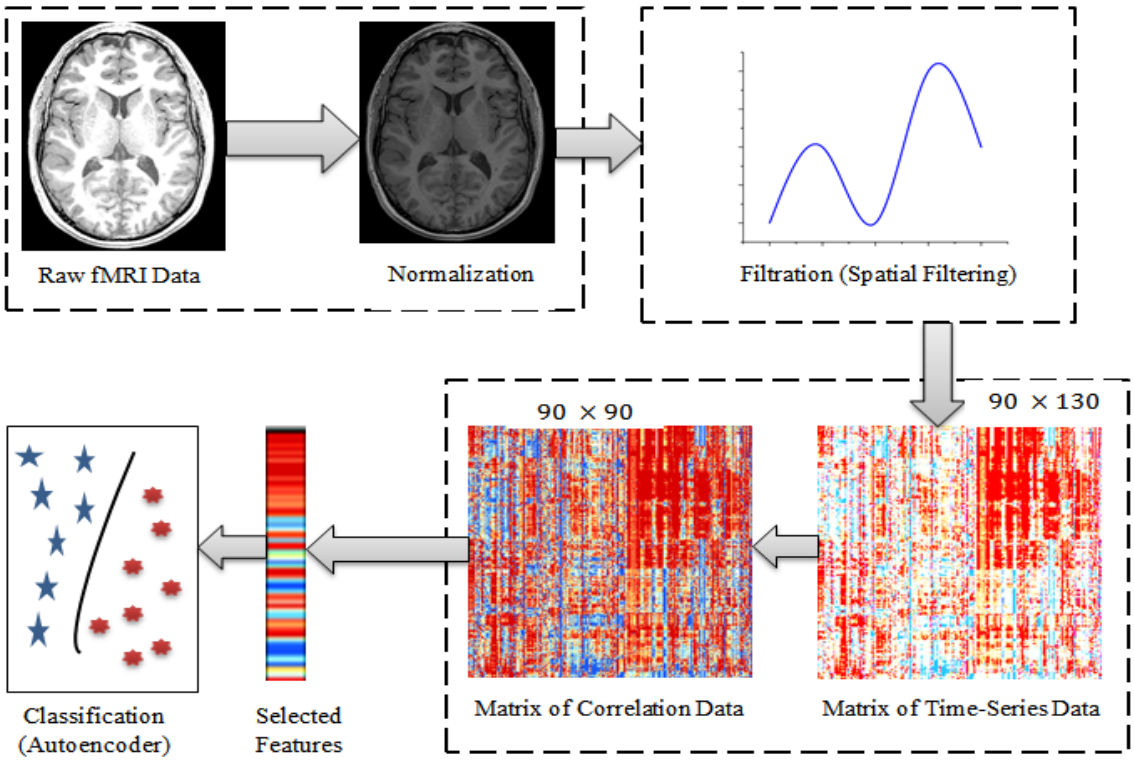


FIGURE 3. Classification processes for the detection of MCI for R-fMRI data.

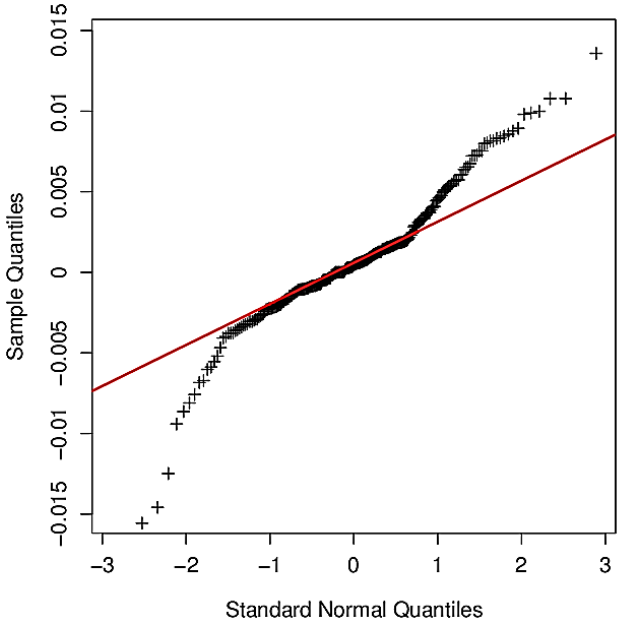


FIGURE 4. QQ plot for gaussianness data testing.

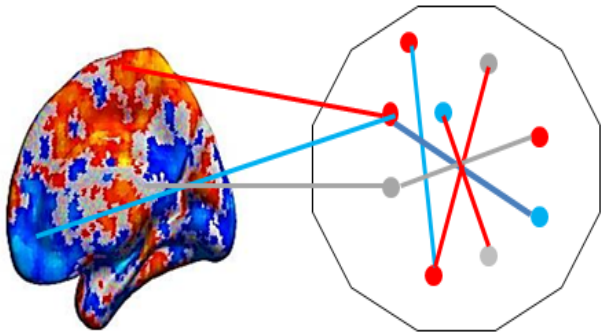


FIGURE 5. Brain network schematic diagram.

it optimizes parameters. Thirdly, the output activation vector of the layer is determined and is provided by the layer $q+1$. If q is equal to mq which is the last layer of the model, then the softmax function input is saved. The following are detailed processes:

1) PARAMETER INITIALIZATION

Pre-training intends at decreasing the cost function $J_s(V, d)$ and improving the weight of the process. Different forecasts of the total maximum of a space variable are better weight parameters. Random numbers near zero in each layer should be initiated with parameters.

Where

$$J_{sparse} = J_s \tag{1}$$

The Broyden-Fletcher-Goldfarb-Shanno Method (BFGSM) uses only limited memory to minimize cost functions. BFGSM is a technique almost new. Considering the importance of Newton's technology of partial derivatives and

inverse hessian derivatives, a lot of computational resources and storage are needed for BFGSs. The newton approach can not generally be used for practical scalability. The quasi-Newton method, on the other hand, overcomes the situation by computing the reverse Hessian matrix.

L-BFGSM needs less space for development than BFGSM. Instead of entirely measuring and analyzing the inverse Hessian matrix. The autoencoding processes will be introduced below.

The $J_s(V, d)$ cost function comprises of 3 aspects: square errors, weight losses and the consequence of availability. The following will be a detailed description.

A mean square error, which explains the differences between developments and expectations, is the first dimension of $J_s(V, d)$. A data set of n samples of training, $\left(u^1, v^1\right), \ldots\ldots (u^n, v^n)$ in which j indicates the jth input and v^j demonstrates the label. The training sample cost function $\left(u^j, v^j\right)$ can be defined as follows in eq(2):

$$J\left(V, d; u, v\right) = \frac{1}{2}\left\|h_{V,d}\left(u^j\right) - v^j\right\|^2 \tag{2}$$

Consequently, all n samples of training could be expressed as a mean error:

$$J_s\left(V, d\right) = \frac{1}{n}\sum\nolimits_{j=1}^{n}\frac{1}{2}\left\|h_{V,d}\left(u^j\right) - v^j\right\|^2 \tag{3}$$

The second aspect of costs is the decay or canonical element of weight. This item is aimed at decreasing the weight range in order to reduce the risk of over-fitting. A small ρ can overfit, whereas a large ρ can lead to underfitting. In each of these cases, the prediction could be unsatisfactory.

In this experiment, it defined an estimated value for ρ first, then it fine-tuned to make sure the network of the auto-encoder fits. To optimize this function, grid search is applied. The weight decay cost function is described below:

$$\begin{aligned} J_s\left(V, d\right) &= \frac{1}{n}\sum\nolimits_{j=1}^{n}\frac{1}{2}\left\|h_{V,d}\left(u^j\right) - v^j\right\|^2 \\ &\quad + \text{Weight Decay} \end{aligned} \tag{4}$$

$$\text{Weight Decay} = \frac{\rho}{2}\sum\nolimits_{q=1}^{mq-1}\sum\nolimits_{j=1}^{rq}\sum\nolimits_{k=1}^{rq+1}\left(V_{jk}^{q}\right)^2 \tag{5}$$

$$\begin{aligned} J_s\left(V, d\right) &= \frac{1}{n}\sum\nolimits_{j=1}^{n}\frac{1}{2}\left\|h_{V,d}\left(u^j\right) - v^j\right\|^2 \\ &\quad + \frac{\rho}{2}\sum\nolimits_{q=1}^{mq-1}\sum\nolimits_{j=1}^{rq}\sum\nolimits_{k=1}^{rq+1}\left(V_{jk}^{q}\right)^2 \end{aligned} \tag{6}$$

Sparsity punishment is the third term for cost. Many recent studies have proven the sparseness of brain activity, meaning that only a few regions participate at a time in brain activity.

Enable $c_k^2(u^j)$ is the first unit output of kth hidden layer at the feed input

$$\check{\sigma}k = \frac{1}{n}\sum\nolimits_{j=1}^{n} c_k^2(u^j) \tag{7}$$

It describes the kth hidden unit's average active value in all m samples of training. A value close to zero commonly initializes the small network parameter β to ensure the number of neurons enabled is sparse throughout the network. It have set $\check{\sigma}k$ alternatively as the target value. For example, if $\sigma = 0.1$, most hidden devices will maintain their network average active value close to 0.1 in order to be able to function. The term ' sparsity punishment' based on the difference between the Kullback and Leibler has been added to make it near to σ:

$$\begin{aligned} &\sum\nolimits_{k=1}^{rq} KL(\sigma \,\|\check{\sigma}k) \\ &\quad = \sum\nolimits_{k=1}^{rq}\left[\sigma\log\frac{\sigma}{\check{\sigma}k} + (1-\sigma)\log(\frac{1-\sigma}{1-\check{\sigma}k})\right] \end{aligned} \tag{8}$$

The cost function is presented as follows while introducing the sparse constraint:

$$\begin{aligned} J_s\left(V, d\right) &= \frac{1}{n}\sum\nolimits_{j=1}^{n}\frac{1}{2}\left\|h_{V,d}\left(u^j\right) - v^j\right\|^2 \\ &\quad + \frac{\rho}{2}\sum\nolimits_{q=1}^{mq-1}\sum\nolimits_{j=1}^{rq}\sum\nolimits_{k=1}^{rq+1}\left(V_{jk}^{q}\right)^2 \\ &\quad + \alpha\sum\nolimits_{k=1}^{rq} KL(\sigma \,\|\check{\sigma}k) \end{aligned} \tag{9}$$

The parameter α is the same as ρ and form (7) represents the total cost function of the tests

The gradient is a key component of an autoencoder in addition to the cost function. The cost function gradient matrix is measured by back propagation. Back propagation is a supervised learning method when the results are compared to expectations. The gradient cost function is represented as follows

$$\begin{cases} \dfrac{\delta}{\delta(V_{jk}^{q})}J_s\left(V, d\right) = \dfrac{1}{n}\sum\nolimits_{j=1}^{n}\dfrac{\delta}{\delta(V_{jk}^{q})}J\left(V, d, u^j, v^k\right) + \rho(V_{jk}^{q}) \\ \dfrac{\delta}{\delta(d_{jk}^{q})}J_s\left(V, d\right) = \dfrac{1}{n}\sum\nolimits_{j=1}^{n}\dfrac{\delta}{\delta(d_{jk}^{q})}J\left(V, d, u^j, v^k\right) \end{cases} \tag{10}$$

The last layer error is calculated and either propagate to the last layer before. The last layer error can be described as follows:

$$\delta_j^{m_q} = \frac{\delta}{\delta(W_j^{m_q})}\frac{1}{2}\left\|v - h_{V,d}\left(u\right)\right\|^2 \tag{11}$$

$$\delta_j^{m_q} = -\left(v_i - c_j^{m_q}\right).f'(W_j^{m_q}) \tag{12}$$

where W_j^{mq} is a weighted sum of all jth unit inputs in the m_qth layer. An error can be determined on the basis of the (q + 1)th layer error with other network layers, including the qth layer. Given the constraint of sparsity and it can be represented as follows:

$$\delta_j^{q} = \left[\sum\nolimits_{k=1}^{m_q+1} W_j^{m_q}\delta^{q+1} + \alpha\left(-\frac{\sigma}{\check{\sigma}k} + \frac{1-\sigma}{1-\check{\sigma}k}\right)\right].f'(W^q) \tag{13}$$

Finally, with the L-BFGS algorithm, it can use the cost-function gradient vectors and its partial derivatives to update the entire network.

2) ACTIVATED VECTOR CALCULATION

The algorithm returns the optimal parameters that can be forwarded to the next layer once the cost function of the existing layer is lowered. The last output layer is the next input layer. The process of forward propagation will continue until all m_q layers in the network are pre-trained. Finally, the value of the network's last layer c^{m_q} will be acquired.

D. SOFTMAX CLASSIFICATION

The Softmax regression is an extension of the logical regression in the Multi-Class Classification and is a deep-Neural network output layer.

Where $Softmax = Smax$

$$J_{Smax(\theta)} = -\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{p} 1\left\{v^j = k\right\} \log g(v^j = k|u^j;\theta) + \frac{\rho}{2}\sum_{j=1}^{p}\sum_{p=0}^{m}\theta_{jk}^2 \quad (14)$$

Since p is the amount of labeling, $\frac{\rho}{2}\sum_{j=1}^{p}\sum_{p=0}^{m}\theta_{jk}^2$ is the term weight decay used for numerical problems with a softmax regression representation, and $g(v^j = k \left| u^j;\theta\right)$ is the probability.

Where,

$$g(v^j = k \left| u^j;\theta\right) = \frac{e^{\theta_K^T u^j}}{\sum_{s=1}^{g} e^{\theta_K^T u^j}} \quad (15)$$

The gradient of the eq(15) is determined accordingly in order to minimize $J_{Smax(\theta)}$:

$$\nabla_{\theta_k} J\,(\theta) = -\frac{1}{n}\sum_{j=1}^{n}\left[u^j(1\left(v^j = k\right) - g(v^j = k \left| u^j;\theta\right))\right] + \rho\theta_k \quad (16)$$

Again, it uses the cost-optimization algorithm for L-BFGS. Because of its cost and its gradient, a softmax classifier is being used as the input in the last auto encoder vector hidden layer (figure 6).

E. AUTOENCODER FINE TUNING

The complete auto-encoder network performance of Figure 5 is enhanced through fine-tuning. The pre-training approach is used to define input data properties carefully, while the fine-tuning step modifies the functions to change limits in different classes. The advanced, cost-reduction approach and L-BFGS methods are used to reduce the difference in expectation with the real performance of the Softmax model.

F. K-FOLD CROSS-VALIDATION

Autoencoder model may overcome the problem with the small sample data. The subject sample is divided into ten equivalent samples randomly, with a sample retained for autoencoding, whereas the other 9 sub-samples are used to check. The next step is to repeat the cross-validation process 10 times and the validation of each sub-sample is used once. After that, the over-fitting problem is addressed in the first place.

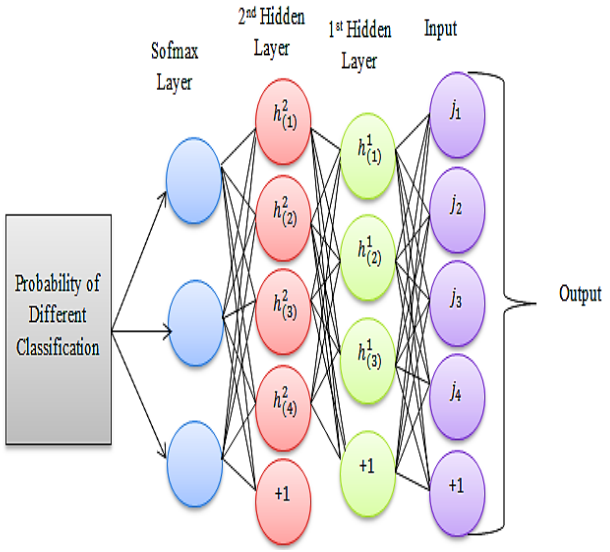


FIGURE 6. Autoencoder network.

G. IMPROVED DEEP LEARNING ALGORITHM

Improved deep learning algorithms conserve transitions between sparsely distributed descriptions in clear areas. In some cases, the transitions might seem like a linear series with the notes in a melody, but many workable future inputs can be expected in the known case. An IDLA location makes specific predictions based largely on a background that would stretch far in time again. Most of the reminiscent people in IDLA reminiscence or transitions between spatial patterns are concerned with sequence reminiscence. The algorithm shows a pseudo-code of the proposed algorithm.

V. RESULTS AND DISCUSSION

Here, the best auto-encoder network structure with the structure data is identified in 10-fold cross-validation. The network of autoencoders with several hidden layers has also been created. Perform an experiment of classification ten times between each network structure [21] and evaluate the accuracy of the test rates, as in fig. 7(a). This figure indicates that in six experiments out of the ten experiments, the autoencoder network with two data structures offers maximum accuracy [22].

The number of hidden nodes usually depends on experiments. If the network is too small, complicated information can't fit, but the training time increases and overfit. The Autoencoder network with two hidden layers, i.e. 200 nodes, is used. It relates the convergence of loss function to show the benefit of using data with coefficients of correlation. Figure 7(b) displays the loss function for the number of tuning points iterations. This figure shows that the loss functions for both types of input decrease at the beginning of the training stage quickly. The loss function convergence slows down and decreases after about 50 iterations.

The classification's accuracy varies if the data can be produced randomly, so the experiment is repeated ten

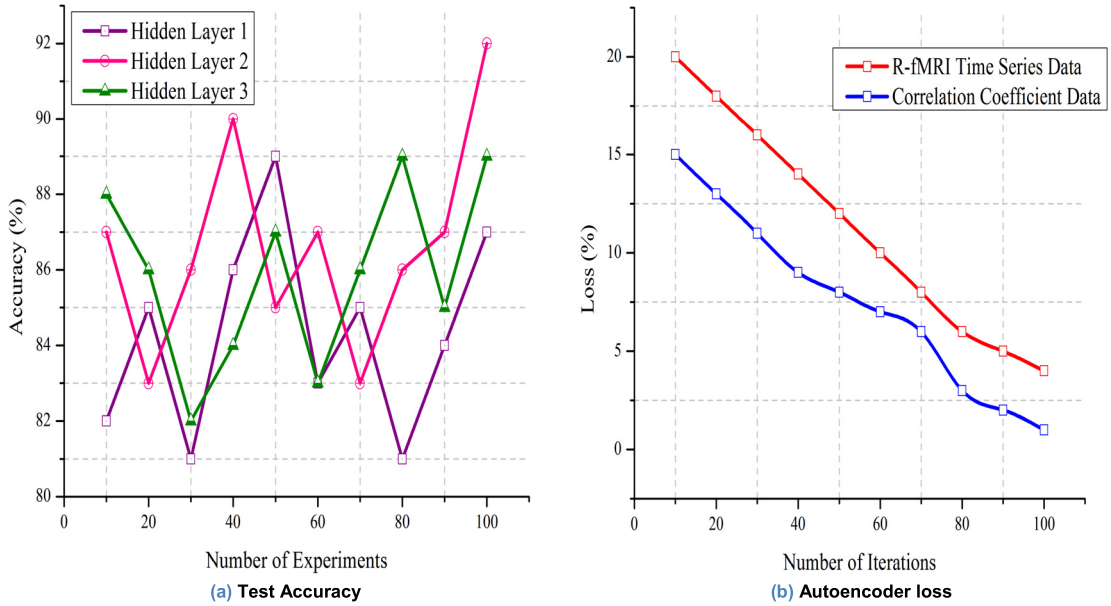


FIGURE 7. (a) Test accuracy. (b) Autoencoder loss.

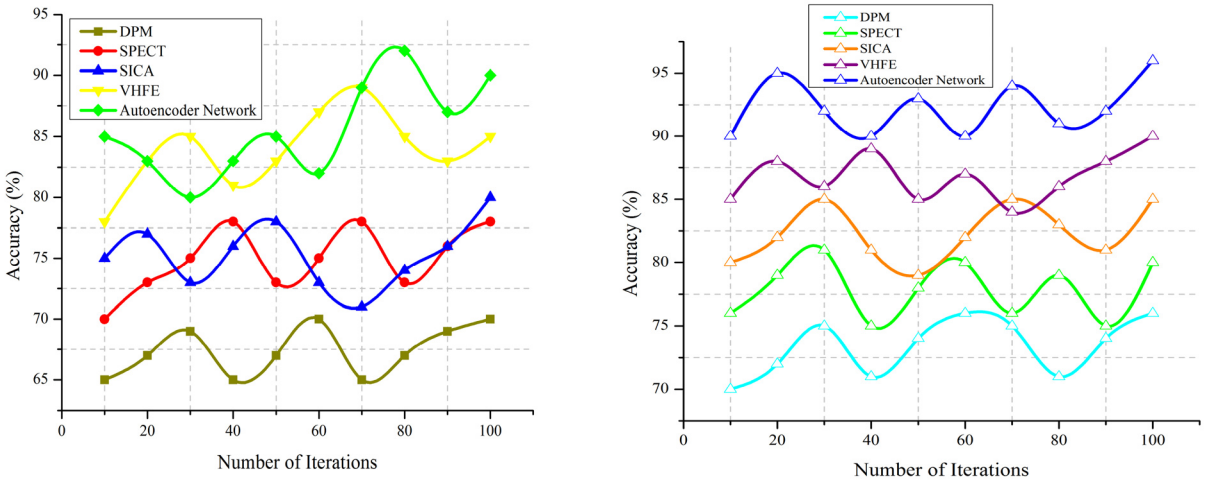
Algorithm 1 Improved Deep Learning Algorithm

```
Use Deep Learning Algorithm
A and B for training data set as an input
a_test of test dataset as an input
for r in active_ Columns(l)
Predicted = false
for u = 0 to cells per column – 1
If predictive state(r, u, l-1) == true then
p = get active segment(r, u, l-1, active_State)
if p.sequence segment == true then
model = Deep learning algorithm()
model.fit(A, B) active state(r, u, l) = 1
if predicted == false then
for u = 0 to cells per column - 1 active state(r, u, l) = 1
for r, u in cells for p in segments(r, u)
if segment active(r, u, p, l) then
predicted = model. predict(a_test)
function Predict(a) l = l0
while t is not a terminal node do
l = the child node l0 of l
end while
return l
end function
```

times. It perform a cross-validation ten times in each experiment. Figure 8(a) showed an algorithm's accuracy difference when R-fMRI time series data are used in different experiments [23], [24]. The graph shows that the autoencoder check accuracy is relatively high. The test is used for determining the significance of the disparity in classification between the autoencoder and the traditional SPECT, DPM, SICA and VFHE models. The testing accuracy is shown for each algorithm in Figure 8(b) in all experimental studies. The correlation data shall be when the input data is. It shows that the testing accuracy of each model is improved considerably by the changes in the input to the correlation data from R-fMRI data [25].

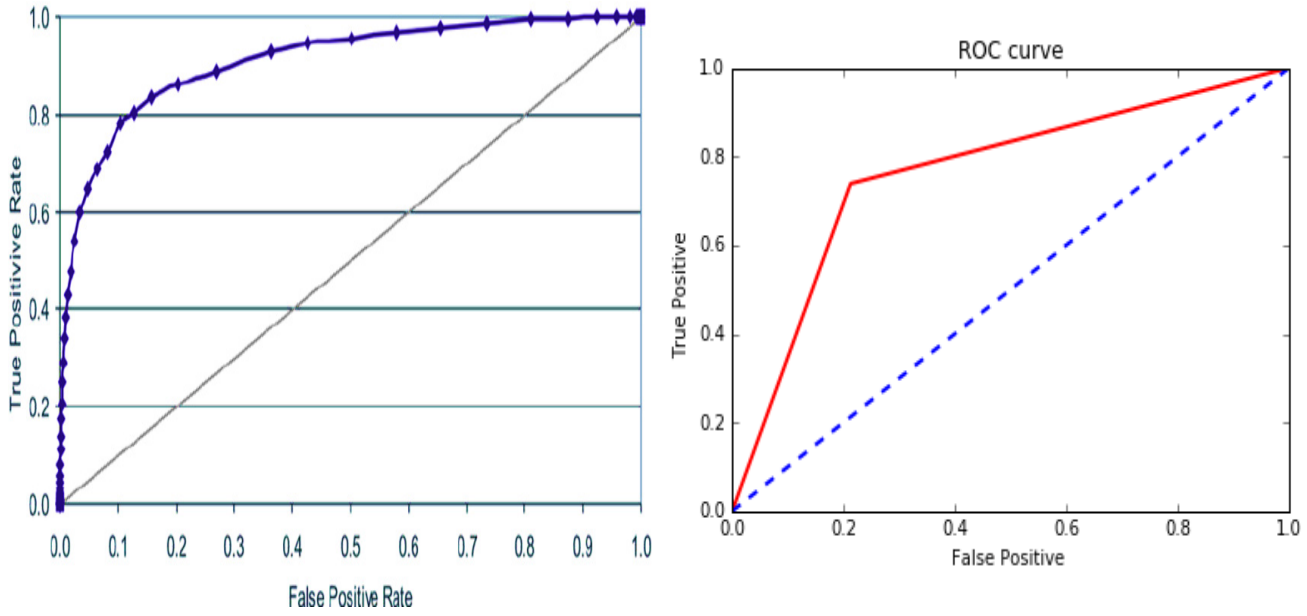
However, the ROC curves should be used to determine how effectively MCIs of different models with different data sets can be identified. If the AUC is close, the effectiveness of MCI detection is higher. Figure 9 shows the ROC curves for different MCI models when data from the R-fMRI correlation coefficients data is provided [26]. Figure 9(a) shows the ROC curves for the time series input of the R-fMRI series. The ROC encoder and input correlation curve are shown in Figure 9(b).

Figure 10 shows the results of AD diagnosis based on ADNI Datasets.It noticed that the R-fMRI time series provides lower predictive precision than the correlation coefficients for a certain classification model. In addition, the deep neural network exists with the same training features in another traditional method. The boundary of decision is a hybrid dividing the input data into 2 or more classes when a deep neural network is used in classification [27], [28]. The value of the input characteristics may be related to the decision limit. By selecting weight, it can determine associations which can play a major role in predicting AD between certain brain areas. Different node weights reflect the value of different MCI nodes. Figure 11(a) shows the sagittal and radial view of all brain nodes. The edge weight shows the value of every connection in the input R-fMRI network between certain brain nodes. Each item shows the average importance of the corresponding edge after average and reconstruction in the normal matrix. The sagittal and axial views on the important edges in the first hidden layer showed



(a) Accuracy of the time-series R-fMRI data (b) Accuracy of the correlation coefficient data

FIGURE 8. (a) Accuracy of the time-series R-fMRI data. (b) Accuracy of the correlation coefficient data.



(a) ROC curve of the time-series R-fMRI data (b) ROC curve of the correlation coefficient data

FIGURE 9. (a) ROC curve of the time-series R-fMRI. (b) ROC curve of the correlation coefficient data.

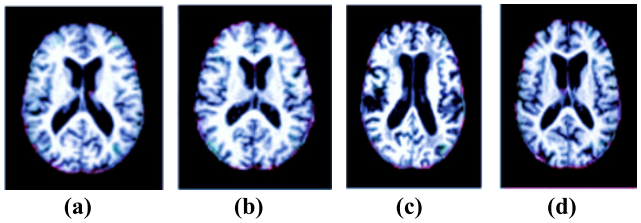


FIGURE 10. Results of AD diagnosis based on ADNI datasets.

that important functional interrelationships between the brain node are important in the detection of MCI (Figure11(b)).

Sensitivity and specificities are also used to further evaluate the classification results. The sensitivity in AD is the ability of subjects with MCI to detect correctly and subjects without the MCI description. In order to further assess classification performance, sensitivity and specificity are used. In the case of AD prediction, the sensitivity can resolve MCI detection, and the accuracy can identify subjects correctly without MCI.

In order to assess the efficiency of the MCI detector model, sensitivity and specificity (Table 1 and Table 2) are measured. The autoencoder is 94.6%, which is much greater than that of DPM, SPECT, SICA, VHFE methods. The highest improvement for a certain system in the evaluation of switching effectiveness from R-fMRI to correlation coefficients data is the input of a correlation coefficient.

The MCI model proposed with brain network characteristics and autoencoding outperforms conventional techniques.

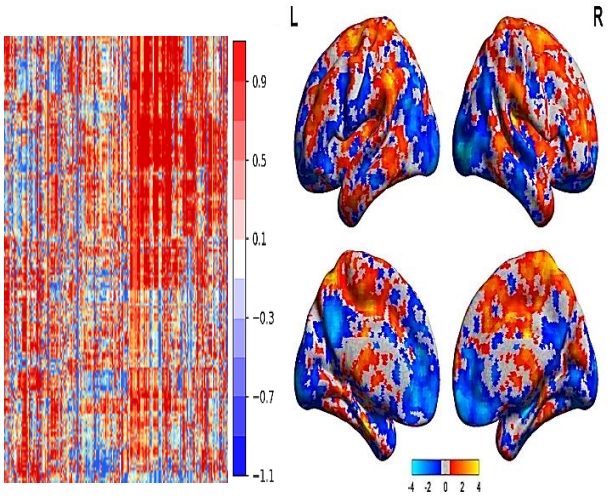


FIGURE 11. (a) Axial views of all nodes (b) Sagittal and important edge axial views.

TABLE 1. Sensitivity ratio.

Number of Experiments	DPM	SPECT	SICA	VHFE	Autoencoder Network
20	65.5	65.8	65.9	66.3	66.7
40	72.8	76.9	79.5	82.5	85.4
60	60.9	65.4	69.8	73.6	78.7
80	76.5	79.4	80.3	85.9	87.6
100	80.6	84.7	86.2	87.5	94.6

TABLE 2. Specificity ratio.

Number of Experiments	DPM	SPECT	SICA	VHFE	Autoencoder Network
20	56.8	59.7	63.5	66.5	70.5
40	65.7	69.3	74.8	79.8	83.7
60	74.6	78.9	81.7	85.8	87.3
80	60.9	64.3	68.5	73.6	79.8
100	80.7	85.3	89.8	92.5	96.7

When MCI is to be identified than the characteristics of each region, the relation between the brain structure regions is more significant. This research will improve disease prediction accuracy and reliability.

VI. CONCLUSION

In this paper, the resting-state fMRI based earlier detection framework has been suggested for alzheimer's disease based on deep neural networks and different medical data. The model training and data classification uses all FMRI images and texts, including age, sex and genetics. Based on the R-fMRI signal correlation, functional intellectual networks are built and then used for improved neural network formation as correlation coefficient information. The methodology proposed increases diagnostic accuracy by approximately 25% compared with traditional approaches, which means combining the brain with improved deep learning is an excellent way to diagnose neurological disorders early. More or similar methodologies for able to diagnose other neurological disease, which is a basis for continuous diagnosis, can be used in this research.

REFERENCES

[1] Alzheimer's Association, W. Thies, and L. Bleiler, "Alzheimer's disease facts and figures," *Alzheimer's Dementia*, vol. 9, no. 2, pp. 208–245, 2013.

[2] I. E. Jansen, J. E. Savage, K. Watanabe, J. Bryois, D. M. Williams, S. Steinberg, J. Sealock, I. K. Karlsson, S. Hägg, L. Athanasiu, and N. Voyle, "Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk," *Nature Genet.*, vol. 51, no. 3, pp. 404–413, 2019.

[3] GBD 2016 Dementia Collaborators, "Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: A systematic analysis for the global burden of disease study 2016," *LANCET Neurol.*, vol. 18, no. 1, pp. 88–106, Jan. 2019.

[4] F. V. Chirila, T. K. Khan, and D. L. Alkon, "Fibroblast aggregation rate converges with validated peripheral biomarkers for Alzheimer's disease," *J. Alzheimer's Disease*, vol. 42, no. 4, pp. 1279–1294, 2014.

[5] D. P. Veitch, M. W. Weiner, P. S. Aisen, L. A. Beckett, N. J. Cairns, R. C. Green, D. Harvey, C. R. Jack, W. Jagustm, J. C. Morris, R. C. Petersen A. J. Saykin, L. M. Shaw, A. W. Toga, J. Q. Trojanowski, and Alzheimer's Disease Neuroimaging Initiative, "Understanding disease progression and improving Alzheimer's disease clinical trials: Recent highlights from the Alzheimer's disease neuroimaging initiative," *Alzheimer's Dementia*, vol. 15, no. 1, pp. 106–152, 2019.

[6] S. S. Dominy, C. Lynch, F. Ermini, M. Benedyk, A. Marczyk, A. Konradi, M. Nguyen, U. Haditsch, D. Raha, C. Griffin, and L. J. Holsinger, "Porphyromonas gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors," *Sci. Adv.*, vol. 5, no. 1, 2019, Art. no. eaau3333.

[7] S. Salloway, R. Sperling, N. C. Fox, K. Blennow, W. Klunk, M. Sabbagh, L. S. Honig, A. P. Porsteinsson, S. Ferris, and M. Reichert, "Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease," *New England J. Med.*, vol. 370, no. 4, pp. 322–333, 2014.

[8] K. G. Mawuenyega, W. Sigurdson, V. Ovod, L. Munsell, T. Kasten, J. C. Morris, K. E. Yarasheski, and R. J. Bateman, "Decreased clearance of CNS β-amyloid in Alzheimer's disease," *Science*, vol. 330, no. 6012, p. 17774, 2010.

[9] J. Cummings, G. Lee, A. Ritter, M. Sabbagh, and K. Zhong, "Alzheimer's disease drug development pipeline: 2019," *Alzheimer's Dementia, Transl. Res. Clin. Intervent.*, vol. 5, no. 1, pp. 272–293, 2019.

[10] T. Jonsson, J. K. Atwal, S. Steinberg, J. Snaedal, P. V. Jonsson, S. Bjornsson, H. Stefansson, P. Sulem, D. Gudbjartsson, J. Maloney, and K. Hoyte, "A mutation in APP protects against Alzheimer's disease and age-related cognitive decline," *Nature*, vol. 488, no. 7409, pp. 96–99, 2012.

[11] B. W. Kunkle, B. Grenier-Boley, R. Sims, J. C. Bis, V. Damotte, A. C. Naj, C. Bellenguez, A. Boland, M. Vronskaya, S. J. Van Der Lee, and A. Amlie-Wolf, "Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing," *Nature Genet.*, vol. 51, no. 3, p. 414, 2019.

[12] M. Madmoli, Y. Modheji, A. Rafi, R. Feyzi, P. Darabiyan, and A. AfsharNia, "Diabetes and its predictive role in the incidence of Alzheimer's disease," *Med. Sci.*, vol. 23, no. 95, pp. 30–34, 2019.

[13] Z. S. Khachaturian and T. S. Radebaugh, "Synthesis of critical topics in Alzheimer's disease," in *Alzheimer's Disease*. Boca Raton, FL, USA: CRC Press, 2019, pp. 3–12.

[14] E. P. Moreno-Jiménez, M. Flor-García, J. Terreros-Roncal, A. Rábano, F. Cafini, N. Pallas-Bazarra, J. Ávila, and M. Llorens-Martín, "Adult hippocampal neurogenesis is abundant in neurologically healthy subjects and drops sharply in patients with Alzheimer's disease," *Nature Med.*, vol. 25, no. 4, pp. 554–560, 2019.

[15] M. T. Heneka, M. P. Kummer, A. Stutz, A. Delekate, S. Schwartz, A. Vieira-Saecker, A. Griep, D. Axt, A. Remus, T. C. Tzeng, and E. Gelpi, "NLRP3 is activated in Alzheimer's disease and contributes to pathology in APP/PS1 mice," *Nature*, vol. 493, no. 7434, pp. 674–678, 2013.

[16] M. Lorenzi, M. Filippone, G. B. Frisoni, D. C. Alexander, S. Ourselin, and Alzheimer's Disease Neuroimaging Initiative, "Probabilistic disease progression modeling to characterize diagnostic uncertainty: Application to staging and prediction in Alzheimer's disease," *NeuroImage*, vol. 190, pp. 56–68, Apr. 2019.

[17] P. Padilla, M. Lopez, J. M. Gorriz, J. Ramirez, D. Salas-Gonzalez, and I. Alvarez, "NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease," *IEEE Trans. Med. Imag.*, vol. 31, no. 2, pp. 207–216, Feb. 2012.
[18] M. H. Ahmad, M. Fatima, and A. C. Mondal, "Influence of microglia and astrocyte activation in the neuroinflammatory pathogenesis of Alzheimer's disease: Rational insights for the therapeutic approaches," *J. Clin. Neurosci.*, vol. 59, pp. 6–11, Jan. 2019.
[19] R. Cui and M. Liu, "Hippocampus analysis by combination of 3-D DenseNet and shapes for Alzheimer's disease diagnosis," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 5, pp. 2099–2107, Sep. 2019.
[20] L. Yue, X. Gong, J. Li, H. Ji, M. Li, and A. K. Nandi, "Hierarchical feature extraction for early Alzheimer's disease diagnosis," *IEEE Access*, vol. 7, pp. 93752–93760, 2019.
[21] V. T. Chan, Z. Sun, S. Tang, L. J. Chen, A. Wong, C. C. Tham, T. Y. Wong, C. Chen, M. K. Ikram, H. E. Whitson, and E. M. Lad, "Spectral-domain OCT measurements in Alzheimer's disease: A systematic review and meta-analysis," *Ophthalmology*, vol. 126, no. 4, pp. 497–510, 2019.
[22] E. F. Fang, Y. Hou, K. Palikaras, B. A. Adriaanse, J. S. Kerr, B. Yang, S. Lautrup, M. M. Hasan-Olive, D. Caponio, X. Dan, P. Rocktäschel, D. L. Croteau, M. Akbari, N. H. Greig, T. Fladby, H. Nilsen, M. Z. Cader, M. P. Mattson, N. Tavernarakis, and V. A. Bohr, "Mitophagy inhibits amyloid-β and tau pathology and reverses cognitive deficits in models of Alzheimer's disease," *Nature Neurosci.*, vol. 22, no. 3, pp. 401–412, Mar. 2019.
[23] H. Fuse, K. Oishi, N. Maikusa, T. Fukami, and Japanese Alzheimer's Disease Neuroimaging Initiative, "Detection of Alzheimer's disease with shape analysis of MRI images," in *Proc. Joint 10th Int. Conf. Soft Comput. Intell. Syst. (SCIS) 19th Int. Symp. Adv. Intell. Syst. (ISIS)*, Dec. 2018, pp. 1031–1034.
[24] H. Li and Y. Fan, "Early prediction of Alzheimer's disease dementia based on baseline hippocampal MRI and 1-year follow-up cognitive measures using deep recurrent neural networks," 2019, *arXiv:1901.01451*. [Online]. Available: http://arxiv.org/abs/1901.01451
[25] T. Tong, Q. Gao, R. Guerrero, C. Ledig, L. Chen, D. Rueckert, and Alzheimer's Disease Neuroimaging Initiative, "A novel grading biomarker for the prediction of conversion from mild cognitive impairment to Alzheimer's disease," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 155–165, Jan. 2017.
[26] K. Oishi, H. Fuse, N. Maikusa, T. Fukami, and Japanese Alzheimer's Disease Neuroimaging Initiative, "Classification of patients with Alzheimer's disease and healthy subjects from MRI brain images using the existence probability of tissue types," in *Proc. Joint 10th Int. Conf. Soft Comput. Intell. Syst. (SCIS) 19th Int. Symp. Adv. Intell. Syst. (ISIS)*, Dec. 2018, pp. 1035–1038.
[27] H. M. T. Ullah, Z. Onik, R. Islam, and D. Nandi, "Alzheimer's disease and dementia detection from 3D brain MRI data using deep convolutional neural networks," in *Proc. 3rd Int. Conf. Converg. Technol. (I2CT)*, Apr. 2018, pp. 1–3.
[28] T. A. Song, S. R. Chowdhury, F. Yang, H. Jacobs, G. E. Fakhri, Q. Li, K. Jhonson, and J. Dutta, "Graph convolutional neural networks for Alzheimer's disease classification," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 414–417.

HAIBING GUO received the Ph.D. degree in statistics. He is currently an Associate Professor with Jiangsu Ocean University, China. His research interests include semi-parametric model and data mining.

YONGJIN ZHANG received the bachelor's degree from Nanjing Medical University. He is currently the Associate Chief Physician with The Lianyungang First People's Hospital, China. His research interests include neurointervention and neurodegenerative disease.