Received: 17 July 2020 | Revised: 11 January 2021 | Accepted: 17 January 2021

DOI: 10.1002/ima.22553

RESEARCH ARTICLE

WILEY

Deep learning based Alzheimer's disease early diagnosis using T2w segmented gray matter MRI

Shaik Basheera[1] | **M Satya Sai Ram**[2]

[1]Department of ECE, Acharya Nagarjuna Uniersity College of Engineering, ANU, Guntur, Andhra Pradesh, India

[2]Department of ECE, RV R & JC College of Engineering, Guntur, Andhra Pradesh, India

Correspondence
Shaik Basheera, Department of ECE, Acharya Nagarjuna Uniersity College of Engineering, ANU, Guntur, Andhra Pradesh, India.
Email: shaikbphd@gmail.com

Abstract

Diagnosing Alzheimer's disease at early stage required an effective classification mechanism to differentiate mild cognitive impairment from cognitive normal and AD. In this paper, we used data set collected from ADNI and OASIS. Instead of using the whole volume of the MRI, high informative slices are selected using entropy. The selected slices are pre-processed by removing unwanted tissues using skull stripping algorithm and extracted gray matter using EICA. In this work, we used CNN model with inception blocks to extract deep features from the GM slices used to predict AD at early stage. The model avoids data leakage by considering all the slices of an MRI as a unit. The model trained with 80% of ADNI subject MRI volumes and tested with the remaining 20% subject MRI volumes, to provide great variance in training and testing data, the model further tested with OASIS data sets. 10-fold cross-validation is used to test the model without biasing. The model performance is evaluated using accuracy. The model achieves 98.73%, 100%, 93.72%, and 95.6% of accuracy for differentiating CN-MCI, CN-AD, AD- MCI and CN-MCI and AD. At 10-fold cross-validation it gives $92.92 \pm 3\%$, $98 \pm 2\%$, $90 \pm 4\%$ and $94.9 \pm 2\%$ accuracy to differentiate CN-MCI, CN-AD, AD-MCI, and CN-MCI-AD using ADNI. We further tested the model with 135 MRI volumes selected from OASIS data set, we achieved 92%, 91.76%, 88.23%, 81.48% of accuracy with CN-AD, MCI-AD, CN-MCI, and three-way classification. The model gives good accuracy and sensitivity of early AD Diagnosis.

KEYWORDS

Alzheimer's disease, classification, CNN, deep learning, gray matter

1 | INTRODUCTION

Alzheimer's disease (AD) is a progressive dementia. Alzheimer's degrade the memory, change in personality and day-to-day life of the person. AD is caused due to the loss of brain cells and reduces brain volume over a time.[1] In India there were around 3.7 million Indians are projected to AD and increases to 7.6 million by 2030.[2] As the AD is progressive dementia and no cure or treatment for the AD, accurate and early diagnosis of disease will give better medication to the person.[3] AD is caused due to the two proteins called A-beta and Tau, Tau protein accumulate neurofibillary tangles (NFT), and amyloid-β (Aβ) those cause neuron death.[4] Amyloid-β plaques are observed by diagnosing the CSF. Lumbar puncture with fluid collection technique is used to collect CSF from the subject but is leads some medical issues in the subject. Later mini mental state examination (MMSE) is used to diagnosis the subject dementia based on the score.[5]

Medical image modalities such as computer tomography (CT), positron emission tomography (PET) and magnetic resonance imaging (MRI) has enabled noninvasive

in vivo investigations to diagnosis the anatomical changes the Brain. MRI is one among them that gives better tissue intensity and contrast.[6] MRI is used for diagnosis majority of Brain disorders.

There are different MRI sequences named T1w and T2w MRIs. Gray matter atrophy and White matter atrophy are used as features to make the classification of AD. More tissue level atrophy changes are evaluated using T2w MRI than compared with the T1w MRI, brain atrophy regions are identified better in T2w MRI, whereas in T1w MRI NFT and deposit of plaque is not identified. By observing the Probability density of T2w MRI, GM, and WM are skewed to each other, discrimination of tissues is possible and used for diagnosis of neurodegenerative disorders such as AD.[7]

Atrophy changes are observed in the MRI image as the disease progress. Tissues such as Gray matter, hippocampus volume, and total brain volume are get change as the disorder progress. AD is characterized by the accumulation of amyloid-β (Aβ or A-beta) forming neurotic plaques, neurofibrillary tangles, and eventually progressive loss of neurons within the brain of cerebral. At early stage, histopathology changes are monitored effectively using T2w MRI. Atrophy changes in the brain tissues at different stages of single subject are shown in Figure 1. It observes the changes in the brain region as the AD progresses. We observe that the tissues are changes over a time as AD Progresses. At early stage the atrophy changes are negligible compared with MCI, as it progress to AD we observe lot of change in texture of the brain region are shown in Figure 1.

In recent days, computer aided diagnosis (CADx) techniques used Medical Image[8] along with machine learning approach aims to provide facilitating medical diagnosis.[9] A large number of promising machine learning applications have used MRI for AD prediction.[10] Voxel based morphometric is used as a feature for classification of AD.[11] Features such as brain volume, atrophy changes,[12] hippocampus volume,[13] used for Diagnosis of AD. CADx has used machine learning approaches such as KNN, random forest,[14] Cortical thickness from MRI is used as feature and performed classification using SVM,[15] boosting algorithms,[16] and Feed forward neural network also used to make the classification of AD from mild cognitive impairments (MCI) and cognitive normal (CN) those used to assist the physician to diagnosis the AD, where these classifiers work on handy crafter features from the images. Handy picked features collected from the region of interest (ROI) are typically prone to introduce error and consume time.[17]

At early stage of AD Blood flow in brain cortex get reduce it effects the decay in glucose to the brain. Gray matter atrophy changes are used to diagnosis the stage of the AD, as the gray matter atrophy increases AD get progress.[18] GM morphological changes in structural magnetic resonance imaging (sMRI) increases the accuracy of early diagnosis of AD.[19]

Accurate prediction of MCI is named early detection of AD. In this study morphological change in T2w MRI gray matter is used as input to the CNN model with inception block used to perform classification of AD. The inception block was developed with the inspiration of Inception block used in GoogleNet and ResNet. In this paper, instead of using entire volumes, we selected high entropy slices from the whole volume. The approach improves early detection of AD. The Model differentiates MCI with high accuracy. Features of small atrophy changes are effectively extracted using the inception block, those features makes better prediction at early stage of AD. The atrophy changes in gray matter are used for analysis of the AD.

The paper is further organized as, section 2 briefly discuss the related work of performing classification of AD using deep learning, section 3 deals with preprocessing on MRI data by removing the skull, extracting gray matter using segmentation algorithms, the preprocessed slices are used to train the model and test the model with independent data sets is discussed in section 4, section 5 briefly discuss about the results and compare the performance with the existing work. Finally concluded the work with a research line.

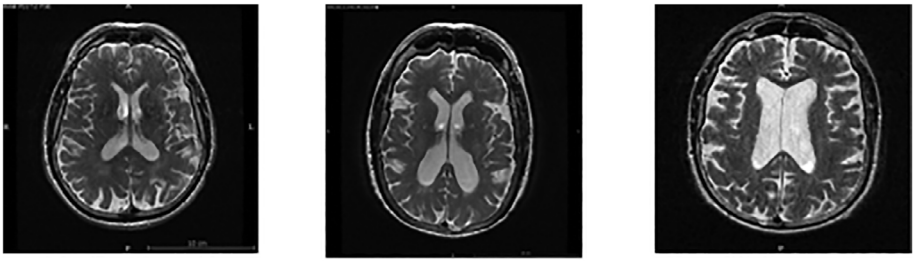

FIGURE 1 Atrophy changes in the MRI at different stages

2 | RELATED WORK

In recent days, Deep learning algorithms work as state of art of machine learning algorithm, used to achieve optimal results in medical image analysis and classification, detection and segmentation of the tissues.[20,21] Gupta et al,[22] used sparse auto encoder (SAE) to learn the bases from a set of natural images and then applied convolutional network for AD classification, they performed Binary and three-way classification of AD from MCI and CN. S. Liu[23] used SAE long with Softmax layer, Auto encoder extract deep features and Softmax layer perform early prediction AD and its prodromal stage, MCI. For early diagnosis of AD timely prediction of MCI is essential, Shi et al[24] proposed a multi modal neuroimaging data such as MRI and PET are used to perform the diagnosis of AD, they used Deep polynomial network to perform the prediction of AD, they stacked the DPN in hierarchical format one on the another the initial stages are generating high level features to perform the classification of AD. F. Li et al,[25] used Multi model convolution network is used where first model is used for transferring the MRI into more compact high-level features and multi scale CNN model is used for extracting further features. For the small sample problem of medical images, transfer learning was used on the classification of AD. Danni Cheng et al,[26] used multiple CNN models ensemble together and performed CN vs AD Classification.

If the data set is small transfer learning is used for classification of the images, where pre trained models are used for classification, LeNet-5,[27] AlexNet,[28] OXFORD_VGG,[29] GoogleNet,[30] ResNet,[31] trained with ImageNet data set and the weights of the filters are preserved used for the required application, classification is performed using fully connected network with the required number of layers. B. Khagi et al,[32] works with Alexnet to perform classification with 2D slices collected from the 3D MRI volume. Ashkan Shakarami et al,[33] used transfer learning with AlexNet, replaced its last three layers with SVM classifier they work on PET 2D slices are used to perform AD-CN Classification. Rachna Jain et al,[34] used VGG − 16 network trained with Image Net data set used to perform classification of CN-AD. Ciprian et al,[35] modified the 16 layers VGGNet for the three ways classification of AD, MCI, and NC based on the Alzheimer's disease neuroimaging initiative (ADNI) data set and achieved an overall accuracy 91.85%. Farooq et al,[36] work with GoogleNet Inception transfer learning to perform classification of AD. Aly Valliani et al,[37] used ResNet with Data augmentation to perform the AD-CN and three-way classification of the AD. M. Hon et al,[38] worked with transfer learning using VGG 16 and Inception V4 with 2D slices.

The slices are selected from 3D MRI volume using entropy of the MRI. VGG − 16 transfer learning generates accuracy of 92.3% and inception V4 gives the accuracy of 96.25%. Morphological changes in the Brain tissues used for analysis of such as hippocampal size get reduced as the dementia get progresses, ventricles get enlarge, atrophy changes in gray and white matter, texture changes are observed as the AD progresses. Fan Li et al,[39] developed a hybrid model of CNN and RNN for Early diagnosis of AD using Hippocampal 3D image patch for early diagnosis of AD. Atrophy changes in white matter and GM slices are used to perform early diagnosis of AD. Andrés Ortiz et al,[40] work with GM of the brain have been split into 3D patches using the regions labeled by AAL atlas and these patches are used to train different deep belief networks. Final prediction done by ensemble of majority voting. Ji, H,[41] used ensemble CNN classifier by taking white matter and GM as input to the CNN to perform the classification of AD. They performed binary class classification of 97.65% for AD/MCI and 88.37% of accuracy for MCI and CN. Shaik Basheera et al,[42] used T2w MRI's to perform classification of AD using CNN model trained from scratch the CNN model have five convolution layers, five zero padding layers, five max pooling layers and drop out layers are connected in sequence followed by six fully connected layers perform binary and multi class classification where they compare both T1w MRI and T2w MRI images to train the CNN and segmented gray matter of T2w MRI used to train the model achieves 86.7% of accuracy in multi class classification, 100% of accuracy in classification of AD-CN, 96.2% of accuracy in classifying AD-MCI and 98.0% of accuracy in classifying CN-MCI. Shaik Basheera et al,[43] used segmented T2w Gray matter to train the CNN model with and the model tested with clinical information performs multi class classification achieves 90.47% of accuracy, 86.66% of recall and 92.59% of precision.

The summary of Primary contribution of the early detection of AD is shown in Table 1.

3 | MATERIAL AND METHODS

3.1 | Material

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu) and Open Access Series of Imaging Studies (OASIS). 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 MRI, PET, other biological markers, and clinical and

TABLE 1 Comparison of conventional Alzheimer's early detection techniques using deep learning

Algorithm	Content	Key approach	Classification	Modalities	Data set
Gupta et al,[22]	Full brain volume	SAE learned features from natural images	CNN	sMRI	ADNI
S. Liu[23]	Full brain volume	Auto encoder	Softmax	sMRI	ADNI
Shi et al,[24]	Full brain volume	Stacked CNN	CNN	sMRI+PET	ADNI
F. Li et al,[25]	Full brain volume	Multi Model CNN	CNN	sMRI	ADNI
Cheng and Liu 2017[26]	Full brain volume	Multi Model CNN with Ensemble	CNN	sMRI	ADNI
B. Khagi et al,[32]	Slices from 3D MRI	Transfer learning	AlexNet	sMRI	ADNI
Ashkan Shakarami et al,[33]	Slices from 3D MRI	Transfer learning	AlexNet+SVM	sMRI	ADNI
Rachna Jain et al,[34]	Slices from 3D MRI	Transfer learning	VGG-16	sMRI	ADNI
Ciprian et al,[35]	Slices from 3D MRI	Transfer learning	VGG-16	sMRI	ADNI
Farooq et al,[36]	Slices from 3D MRI	Transfer learning	Inception-V3	sMRI	ADNI
Aly Valliani et al,[37]	Slices from 3D MRI	Transfer learning	ResNet-50	sMRI	ADNI
M. Hon et al[38]	Slices from 3D MRI	Transfer learning	Inception-V3, VGG-16	sMRI	OASIS
Fan Li et.al,[39]	Hippocampal volume	Measuring hippocampal volume for early detection	CNN + RNN	sMRI	ADNI
Andrés Ortiz et al,[40]	Gray matter patch	Ensemble	CNN	sMRI	ADNI
Ji, H[41]	White matter and gray matter as input	Ensemble of VGG-16	CNN	sMRI	ADNI
Shaik Basheera[42]	Gray matter as input	Stacked CNN	CNN	sMRI	ADNI
Shaik Basheera[43]	Segmented gray matter	CNN	CNN	sMRI	ADNI

neuropsychological assessment can be combined to measure the progression of MCI and early AD. Where we collected 1820 number of T2w MRI volumes, and generated 18 017 numbers of 2D Slices, those are used to make the classification of AD from MCI and CN. We are interested in heterogeneity of data so, we collected MRI of both male and female of different age group with different MMSE score, global deterioration (GD) Scale, global clinical dementia rating (CDR), and neuropsychiatric inventory questionnaire (NPI-Q) total score. The MRI's acquired five to six times in 3 years, they includes 6th, 12th, 18th, 24th, and 36th month. An average five to six MRI's are collected from each subject over a period of 3 years. The data set we collected from 349 subjects at different visiting periods on an average of five to six MRI's are collected from each subject total 1820 number of MRI volumes are collected. 1820 MRI volumes are represented as 635 AD, 637 CN, and 548 MCI. The corresponding demographic representation of subjects is shown in Table 2.

We further used 135 T2w MRI volumes from OASIS, of which 50 AD, 50 CN, and 35 MCI, these volumes are used to test the model work with independent data set. The demography of OASIS data set is shown in Table 3.

3.2 | Methodology

In this work, we used T2 weighted MRI image to perform binary and multi class classification of AD from MCI and CN using segmented gray matter (GM) extracted from the slice images.

MRI volumes are collected from ADNI data set is used for both training and testing and 10-fold cross validation. The schematic diagram of the proposed approach shown in the Figure 2. Initially the model trained with 80% patient MRI volumes collected from ADNI and tested the model with remaining 20% patient MRI volumes and 135 OASIS MRI Volumes. Slices are selected using Slice selection algorithm, unwanted tissues are removed from the slices using skull stripping algorithm, Gray matter is extracted using EICA. In this paper MRI is taken as a unit and performed training. Independence is carried at both training and testing data sets to overcome data leakage

TABLE 2 Demographic representation of MRI volumes collected from ADNI Online respiratory

Research Group	Number of Subjects	Age	Number of MRI Volumes	MMSE Score	NPI-Q Total	GD Score	Global CDR	Number of Slices
CN	117	71-96	637	29-30	0-10	0-10	-1 to 0.5	6000
MCI	112	61-96	548	27-29	0-9	0-9	−1 to 0.5	6000
AD	120	55-93	635	21-27	0-20	0-20	0.5 to 2	6017

TABLE 3 Demographic representation of MRI volumes collected from OASIS online respiratory

Research group	Number of subjects	Age	Number of MRI volumes	MMSE score	Global CDR	Number of slices
CN	10	42-94	50	27-30	0	509
MCI	8	69-82	35	27-29	<0.5	350
AD	10	66-85	50	10-30	0.5-2	503

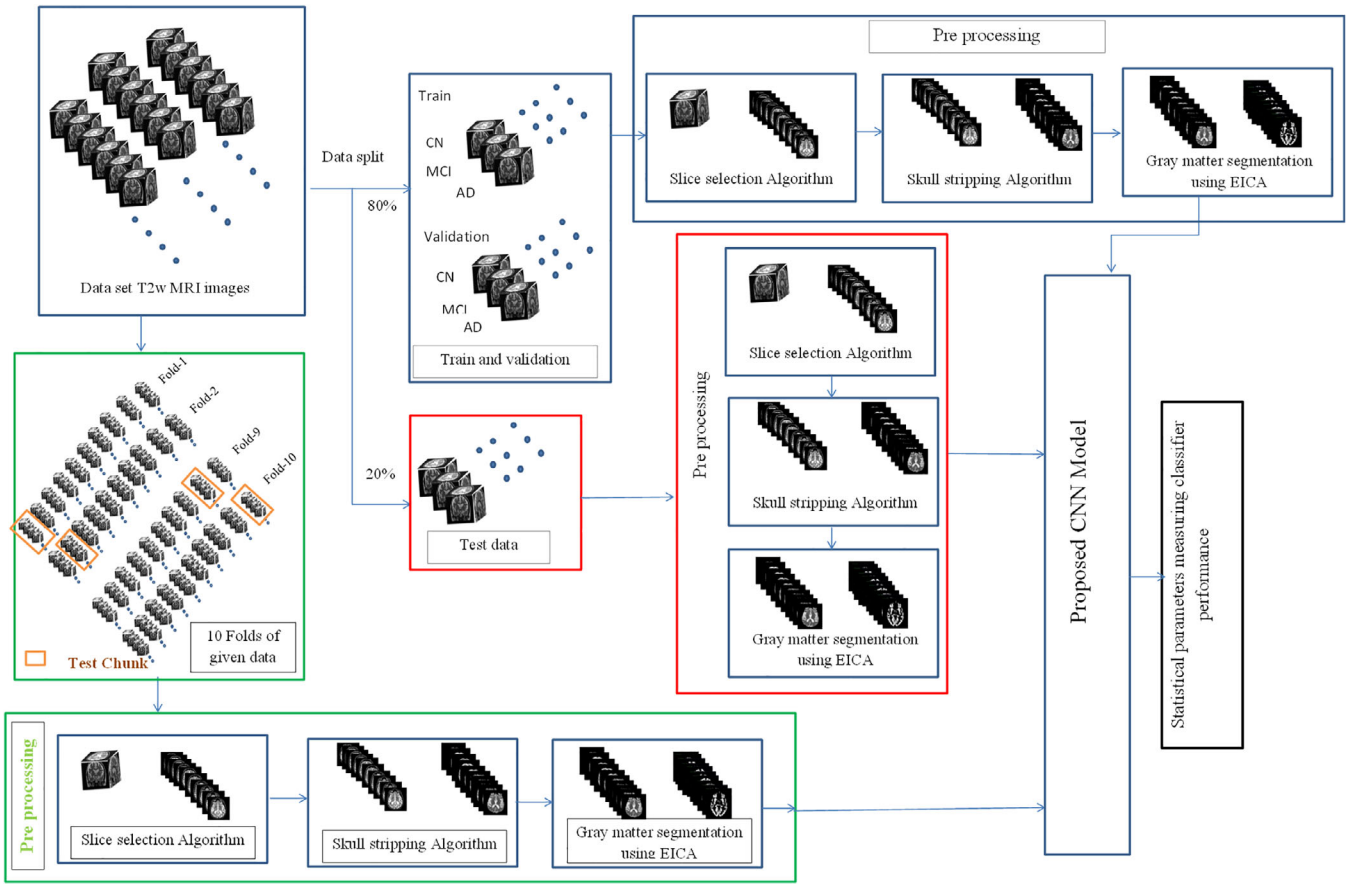

FIGURE 2 Proposed algorithm block diagram [Color figure can be viewed at wileyonlinelibrary.com]

and biasing. To perform the classification T2w MRI volumes are initially sliced into 2D images and pre-processed the slices by reducing noise using Gaussian Filter, irrelevant tissues such as skull, dura, Eyes and Ears are stripped off from the MRI image using morphological operation along with active contour using skull stripping algorithm, and GM of the MRI image is extracted using Enhanced ICA algorithm8. GM Extracted from T2w MRI is considered as a region of interest to perform the classification of AD using the proposed CNN model.

The collected T2w MRI having the size of $256 \times 256 \times 182$ to $256 \times 256 \times 170$ some are having the

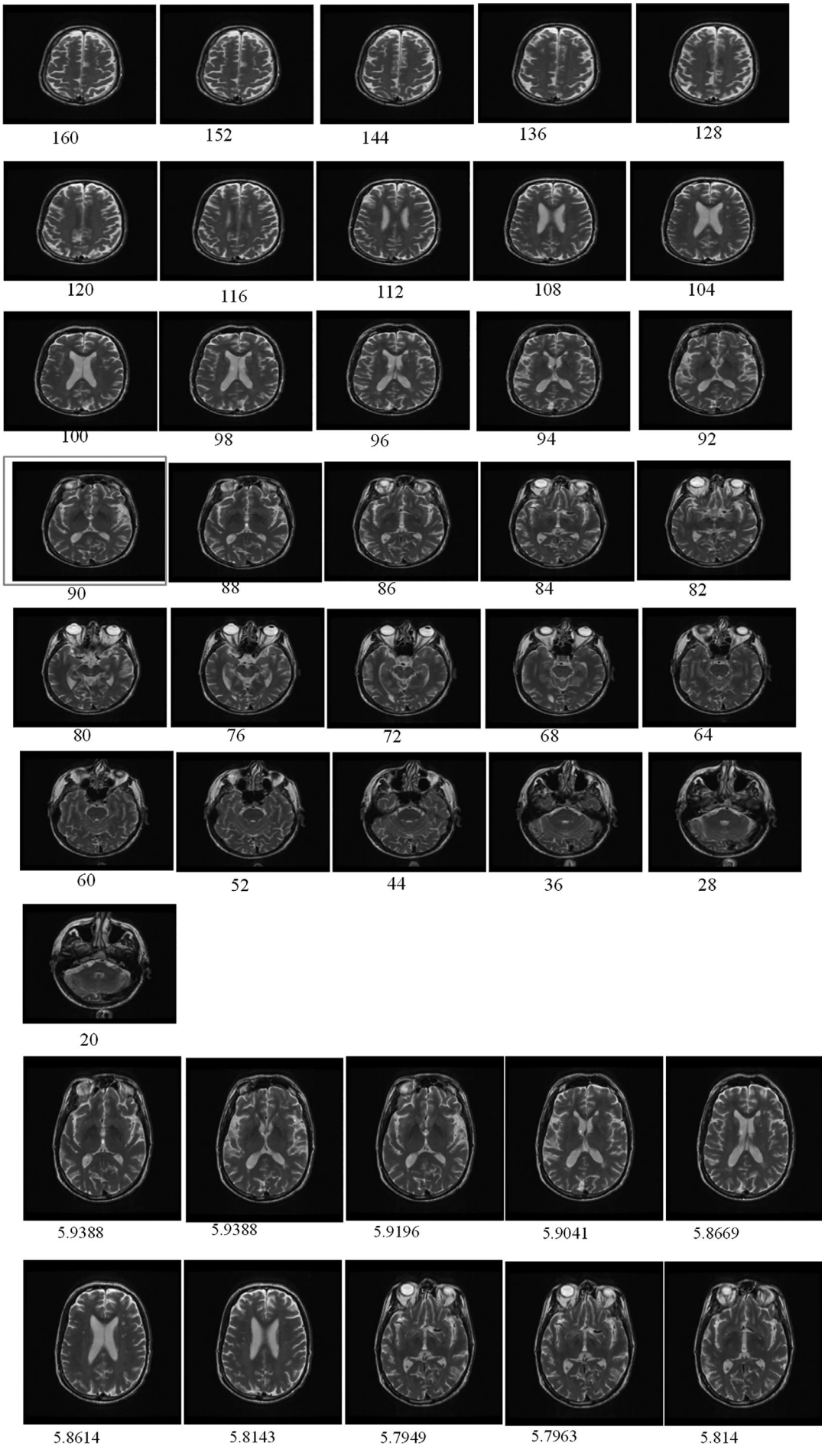

FIGURE 3 Selected Slices from a MRI 3(A) slices selected at first phase of algorithm, 3(B) Top images having high entropy

size of $256 \times 256 \times 170$, with the slice thickness of 1 mm × 1 mm × 1.2 mm. The entire volume of the MRI is not required for analysis as some of the slices have more back ground than the brain so we had selected the few slices based on a fixed procedure.

Procedure of selecting Slices:

1. 40 slices are ignoring the top and bottom slices as they are not having required brain tissues
2. Let assume there are M slices remain
3. Select the center slice of the MRI volume in axial direction, let the center slices is $\frac{M}{2}$
4. Initially we selected every second slice from the $\frac{M}{2}$ toward top direction and bottom direction up to five each direction
5. Later selected every fourth slice toward top and bottom direction, total five on top direction and five toward bottom direction are selected
6. Next every eighth slice is selected in top and bottom directions, collected five in each direction
7. We had selected total 25 to 31 slices from each MRI image based on the number of slices that a MRI is having
8. We calculated entropy of all the slices and collected approximately top 10 slices having high entropy

Figure 3A shows a marked slice, it is the center slice from a volume of MRI's, the slices are selected using 1 to 6 steps formulated. For better diagnosis of AD at early stage more informative slices are required. Entropy of an image used to measure the information of the image, based on the entropy value the slices are sorted and selected top 10 Slices those have high entropy. The individual entropy of the slice is calculated using

$$E = -sum(P_i \log_2(P_i))$$

E represents the entropy of the image, P_i represents the probability of ith pixel, P_i is calculated using normalized histogram values of individual image pixels. It is observed that the image with higher entropy is having more information. The slices at center have higher entropy than the slices at edge. Set of top MRI slices collected along with respective entropy are shown in the Figure 3B.

3.2.1 | Skull stripped algorithm

To get accurate classification unwanted tissues are removed from the MRI slices using the proposed skull stripped algorithm given in our previous paper.[42]

3.2.2 | Gray matter segmentation

Atrophy changes in gray matter are used to estimate the stage of AD. Early prediction of AD is achieved using segmented gray matter of the T2w MRI slices. Atrophy changes are estimated using segmented tissues. Segmentation of brain tissues is a challenging task as the tissues are overlap and takes smooth transmission of the tissues. In this paper, we used Enhanced Independent component analysis (EICA) is used to perform the segmentation of the gray matter.[44] To work with EICA, it assumes that the tissues are independent to each, we used Mixture model to perform the segmentation of the tissues. In this paper, unsupervised segmentation approach is used to extract tissues from brain by considering all the tissues are Gaussian in nature, modifying Gaussian mixture model (GMM) with Expected maximization perform segmentation of the Brain Tissues, spatial dependence is achieved using Hidden markov random fields (HMRF), Gaussian density function weights are mapped using Gibbs density function to achieve better spatial relation. The resultant images are shown in the Figure 4.

3.2.3 | Classifier

To perform the prediction of AD at early stage, we proposed a CNN model with the inspiration of using inception type of block to extract deep features from the segmented MRI images. The extracted features are much useful for early diagnosis of the AD. In general a CNN model is having three types of layers named Convolution layer, Zero padding layer and Max pooling layer. In our model we used all the above said layers not only in sequential model also non sequential model also added with the inspiration of inception and residual network way of representation is carried in the network.

Proposed CNN model is shown in the Figure 5 is having three sections named initial stage, intermediate stage and final stage.

At initial stage, $224 \times 224 \times 3$ gray matter segmented slice is given to a Convolution layer, the convolution layer is having 32 number of (3×3) kernels with stride 1, followed by zero padding of (3, 3) and performed max pooling with (3, 3) window. The initial stage followed by intermediate stage.

Intermediate stage are having two types of blocks are cascaded, first one is convolution block and second one is designed with the inspiration of inception and residual mode of formation.

a. *Convolution Block*: It is having Convolution layer with different number of filters with different sizes at

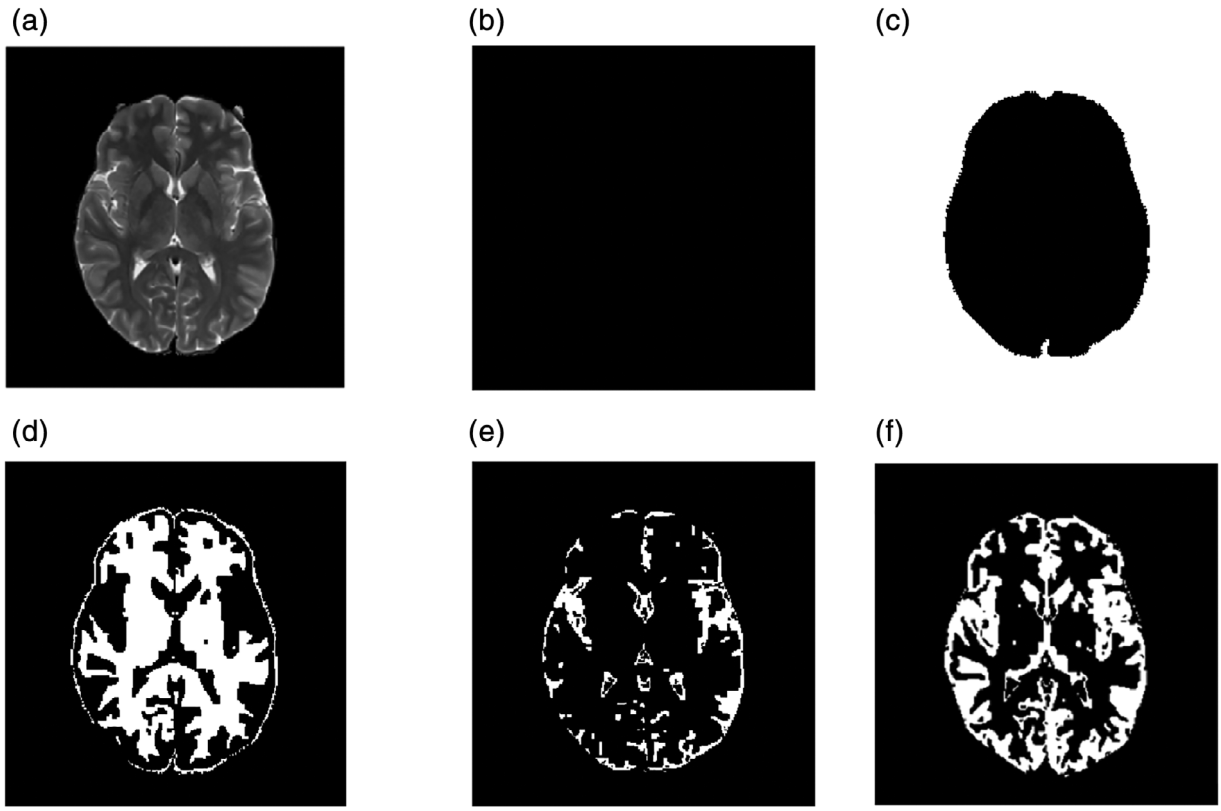

FIGURE 4 Gray mater segmentation, A, Skull stripped image, B, Back ground, C, Brain region selecting mask, D, White matter and dura, E, Cerebral spinal fluid, F, Gray matter

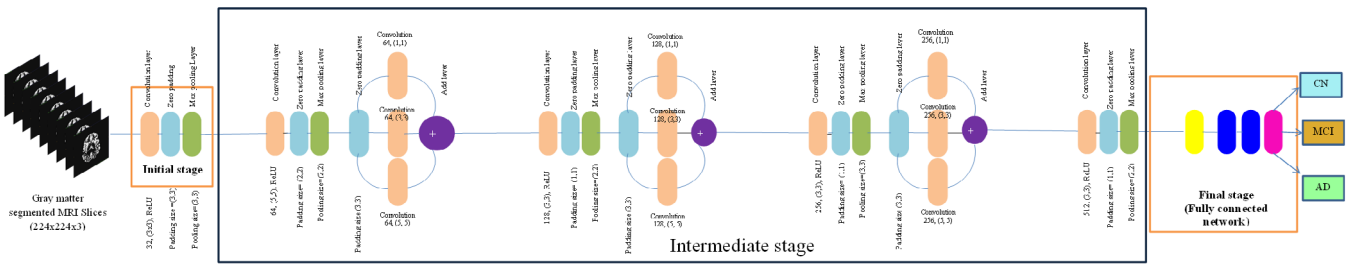

FIGURE 5 Proposed CNN Model [Color figure can be viewed at wileyonlinelibrary.com]

different locations, with ReLU as activation function connected to zero padding layer with the size of (2, 2) followed by Max pooling layer of (2, 2) size.

b. *Inception Block*: The second block is having the zero padding layer of (3, 3) size, combination of Inception mode of representation are performed by adding the response of 3 filters of (1, 1), (3, 3), and (5, 5), 8 number of filters run on the same layer with ReLU activation used to extract multi-level features from the same module. Early atrophy changes in the MRI are observed at local changes in gray matter of the T2w MRI images and global change in Gray matter is used to describe the worsen AD. Filter size (1, 1) used to extract local features, (3, 3) used to extract medium features and (5, 5) used to extract global features. For early diagnosis local features, global features are keen to make prediction of AD. Perfect selection of kernel is much required to extract the required features from the gray matter, small kernel size is used for extracting local information, and large kernels are used for extract global information from gray matter of the T2w MRI.

To extract the features the network should be deep but it leads over fitting and hard to update the gradient of entire neural network. To overcome this multiple and different sized filters are stacked in parallel at the same level.

The proposed inception model is like a naïve inception model but we are not introduced maxpooling and instead of cascading layer in this model residual connections are performed where add output of the entire convolution layer out comes with same dimensions. The model over comes the vanishing of gradient, increase overall accuracy of the model.

c. *Final stage*: At final stage fully connected network is used with three fully connected layers. The outcome of Convolution network is flatten and converted to 1D array and connected to fully connected layer having three fully connected layers. First fully connected layer has 1024 number of neurons, second fully connected layer has 1024 number of neurons and third layer is having three neurons for multi class classification and two neurons for binary class classification. In this model first and second fully connected layer are activated using ReLU and last layer is activated using Softmax for Multi class classification and Sigmoid when binary class classification is performed.

4 | EXPERIMENT

We perform the experiment for early diagnosis of AD using T2w gray matter segmented MRI images. We used 1820 MRI Volumes in our experiment collected from 349 subjects, out of these 80% subject and their MRI's of ADNI Volumes are used for train, validation and reaming 20% subject volumes are used to test the model. We also test the model OASIS Data set. In this experiment we have three stages First is preprocessing, second is train the model and third is test the model with two independent data sets, also verified the model with 10-fold cross validation with ADNI data set.

4.1 | Experiment setup

We used R2015a, MATLAB software to perform preprocessing, where we initially performed slice selection from the MRI volume, removed the unwanted tissues from the MRI image using skull stripping algorithm and segmented the Gray matter using Enhanced ICA. To train our convolution neural network, we used Google Colab with open source python 3.0 version.

In this experiment GPU: 1xTesla K80, having 2496 CUDA cores, compute 3.7, 24GB (23.439GB Usable) GDDR5, VRAM to perform training and validation. Our classifier is designed using Keras, Tensorflow Modules. Our network is having five convolution layers followed by zero padding and Max pooling In between each such segment a incept block is added further to extract in depth features, this network is followed by three fully connected layers to make the classification of the AD from MCI and CN, the model is shown in Figure 3.

Our proposed network is trained with gray matter segmented MRI slices collected from 80% patients are split into train and validation. We trained the model with 128 batch size, augmented the MRI images by rescaling, zooming, shear and horizontal flipping to get huge value of image data to train the model. We used Adam optimizer with learning rate of 0.001, fit the model with 25 train generator epochs, 10 valid steps and the training process carried with 200 epochs the model generates 26 985 794 trained parameters.

4.2 | Creating training set and test set

We used 18 017 number of gray matter segmented MRI slices in our experiment. To avoid the biasing we split the data set into 80:20 at subject level. We perform binary classification and multi class classification. The data split is shown in the Table 4. We trained and validate the model with gray matter segmented slices collected from 80% subjective MRI's of ADNI and remaining patient MRI's are used for testing the model. While training 80% Patient MRI's are further divided into two subsets as training and validation. For training 80% of subject MRI's are used and for validation 20% patient MRI's are used.

5 | RESULT

We trained the model using the data collected from the ADNI data set, test with 20% of subject MRI's from each modality and remaining 80% subjects are used for training and validation. In our experiment skull stripped, gray matter segmented MRI slices are used as input to the model to perform the classification of the AD from MCI and CN. We initially performed binary classification of CN-MCI, CN-AD, and AD-MCI, Later we performed multi class classification of AD from MCI and CN.

From the Figure 6, we observed that Train and validation curves are followed one with another the error occurred while validation is also less and the model get trained well and give better results.

In this experiment, we tested the trained CNN model with 20% of independent subject ADNI MRI volumes and 135 MRI volumes collected from publically available OASIS data set. The model performance is evaluated using different statistical parameters such as accuracy, precision, recall, and F1 score and specificity using confusion matrix. The parameters are described as in Equations (1) to (5). Where all the slices of one MRI are considered as a sample.

TABLE 4 Training set and test set sizes

Classification type	Class label	Total MRI	Train MRI	Test MRI	Train slices	Test slices	Total slices
3 level classification	CN-MCI-AD	1820	1455	365	14 414	3603	18 017
Binary class classification	CN-MCI	1185	947	238	9600	2400	12 000
	CN-AD	1272	1017	255	9614	2403	12 017
	AD-MCI	1183	946	237	9614	2403	12 017

a. *Accuracy*: it is defined as how many samples are perfectly classified out of total samples

$$Accuracy = \frac{TP + TN}{Total\ Number\ of\ samples} \quad (1)$$

b. *Precision*: it is defined as the number of samples correctly predicted out of total predicted value of the particular class

$$Precision\ (or) Positive\ Prediction\ Value = \frac{TP}{TN + FP} \quad (2)$$

c. *Sensitivity (or) recall (or) true positive rate*: It is defined as the numbers of samples are correctly classified out of total actual samples of the class

$$Sensitivity\ (or) Recall\ (or)\ True\ positive\ Rate = \frac{TP}{TP + FN} \quad (3)$$

d. *Specificity (or) selectivity (or) true negative rate:* It is defined as the ratio of number of True negative classes to number of true negative and false positive samples

$$Specificity\ (or)\ selectivity\ or\ true\ negative\ rate\ (TNR) = \frac{TN}{TN + FP} \quad (4)$$

e. *F1 score*: it is defined as the Harmonic mean of the Precision and recall

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (5)$$

Table 5 describes the confusion matrix on test the model with ADNI T2w MRI volumes; we evaluated the model performance both at binary and multi class classification. The evaluated parameters of the ADNI test data is shown in Table 6.

To verify the model performance at different data set we used OASIS data set to evaluate the performance of the model at high variance data set quite different from the trained data set with the proposed approach. We used 135 number of T2w MRI volumes collected from the 20 subjects. The confusion matrix is shown in the Table 7, the evaluated parameters of the OASIS data set is given in Table 8.

We further evaluated the accuracy, average sensitivity, and average specificity all the classes on testing both the data sets with the trained model the resultant is given in the Table 9.

Accuracy, average sensitivity, and average specificity of the ADNI data set, we observed that the model work well when it tested with ADNI data set. The model differentiates AD and CN more accurately, whereas MCI also differentiated from AD and CN that leads early detection of AD. Even at multi-stage also MCI MRI's are differentiated with remarkable accuracy that leads better diagnosis of AD at early stage.

Accuracy, average sensitivity, and average specificity of the OASIS data set, we observed that the model work well in differentiating MCI from AD, whereas separating MCI from CN gives less accuracy remaining Binary class Classification. The results are slightly less than the model tested with ADNI data set. Even without having OASIS Data set in training the model gives significant results.

5.1 | Ten-fold cross validation

To perform 10-fold cross validation, we had taken subject as a unit and performed the training with the MRI's. To perform 10-fold cross validation instead of taking random selecting of Images we selected the Subject randomly to train and test the model. In 10-fold cross validation total subjects are split into 10 chunks. The model trained with nine chunks and test with the remaining one chunk. We avoid the random dependence of slices. We evaluate the

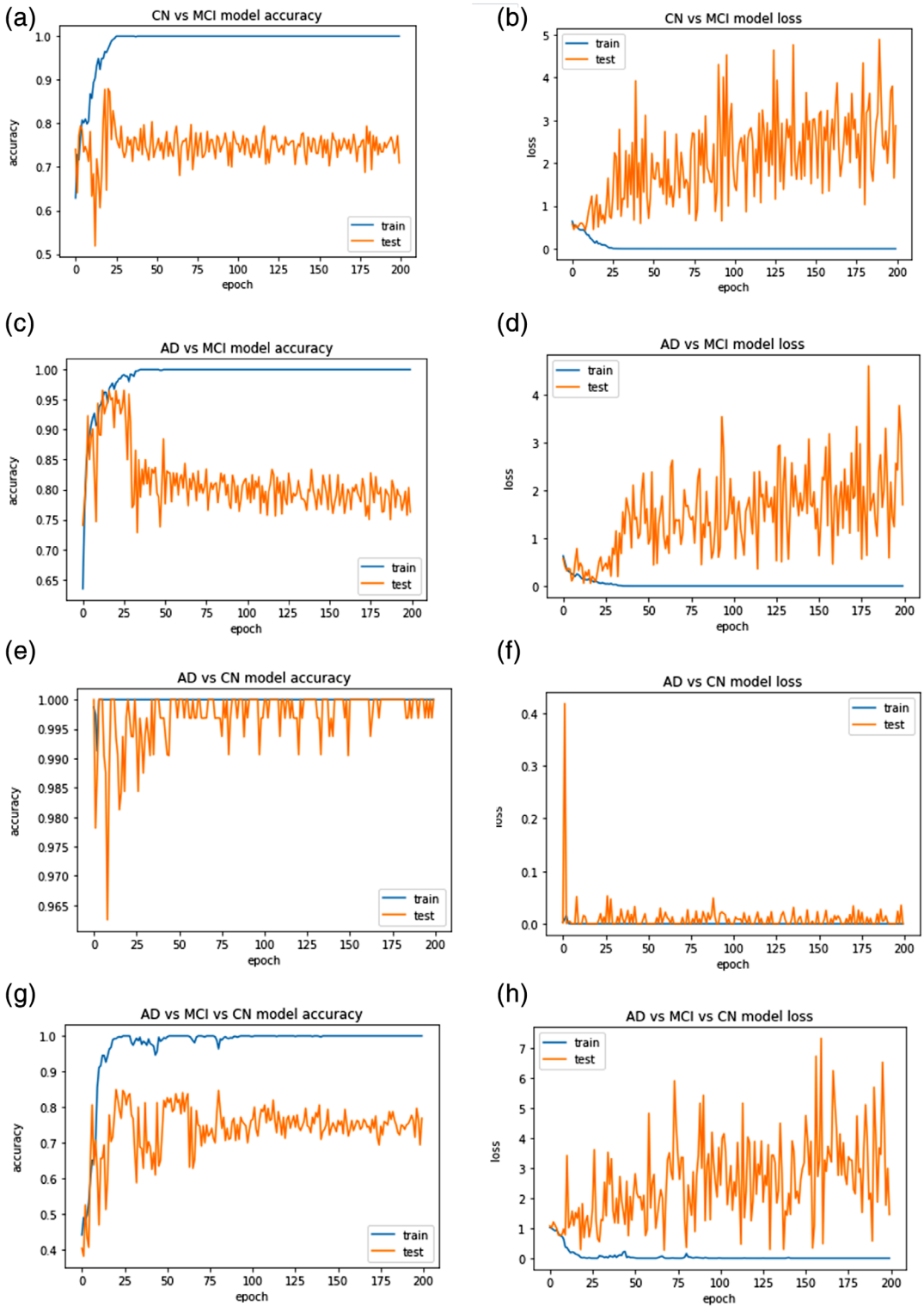

FIGURE 6 Accuracy and loss curves of Binary class classification, multi-class classification, during training and validation of GM extracted T2w MRI, A, CN-MCI accuracy Curve, B, CN-MCI Loss Curve, C, AD-MCI Accuracy curve, D, AD-MCI Loss Curve, E, AD-CN Accuracy Curve, F, AD-CN Loss Curve, G, AD-CN-MCI Accuracy curve and, H, AD-CN-MCI Loss curve [Color figure can be viewed at wileyonlinelibrary.com]

model performance using Accuracy, sensitivity and specificity. The system performance is shown in the given table for both binary and multi class classification the result is shown in Table 10.

TABLE 5 Confusion matrix binary and multi class classifications tested with ADNI data set

		Predicted label		
Multi class classification CN-MCI-AD		**AD**	**CN**	**MCI**
Actual label	AD	127	0	0
	CN	0	128	0
	MCI	13	3	94
Binary class classification AD-CN		**Predicted label**		
		AD	CN	
Actual label	AD	127	0	
	CN	0	128	
Binary class classification AD-MCI		**Predicted label**		
		AD	MCI	
Actual label	AD	127	0	
	MCI	16	94	
Binary class classification CN-MCI		**Predicted label**		
		CN	MCI	
Actual label	CN	128	0	
	MCI	3	107	

TABLE 6 Precession, recall and F1 Score for different classifications using ADNI data set

Classification Type	Class label	Precision (%)	Sensitivity (or) Recall (%)	F1 Score (%)	Specificity (%)
CN-MCI-AD	AD	90.71	100	95.131	94.46
	CN	97.7	100	98.84	98.66
	MCI	100	85.5	92.15	100
CN-AD	AD	100	100	100	100
	CN	100	100	100	100
MCI-AD	AD	88.88	100	94.07	85.45
	MCI	100	85.45	92.15	100
CN-MCI	CN	97.70	100	98.84	97.27
	MCI	100	97.27	98.61	100

The model gives $92.92 \pm 3\%$ of accuracy and make classification of CN-MCI, $98 \pm 2\%$ of accuracy achieved on performing classification of CN-AD, $90 \pm 4\%$ of accuracy when AD-MCI Binary classification had performed, $92.9 \pm 2\%$ of accuracy when 3 way classifications with 10-fold cross validation.

We compare the proposed approach with different frame works discussed in literature shown in Table 11.

5.2 | Discussion

Early, effective and accurate diagnosis of AD is very much required to provide medication to the AD patients. Many researchers had developed computer aided diagnosis techniques for early diagnosis of AD with the accuracy vary from the range of 74.12% to 99.3% at different stages.[25,26,33-35,37-40] All the literatures work based on T1w structural MRI (sMRI) except[33] they used PET for early diagnosis of AD. Except[38] remaining all the literature work with ADNI data set,[38] work on OASIS Data set. Entire volumes are used for early detection of AD by Fan Li et al,[25] they used multi model features, they used two sections one with 3D CNN and second with CNN model to perform the classification of AD from CN with the accuracy of 88.31%. Cheng et al,[26] work with multiple CNN models and ensemble the results, they used skull stripped T1w MRI images. Instead of whole MRI they taken patch from the MRI and performed the classification of AD from CN and achieve accuracy of 87.15%.

TABLE 7 Confusion matrix binary and multi class classifications tested with OASIS data set

		Predicted label		
Multi class classification CN-MCI-AD		**AD**	**CN**	**MCI**
Actual label	AD	50	0	0
	CN	8	42	0
	MCI	7	10	18
Binary class classification AD-CN		**Predicted label**		
		AD	CN	
Actual label	AD	50	0	
	CN	8	42	
Binary class classification AD-MCI		**Predicted label**		
		AD	MCI	
Actual label	AD	50	0	
	MCI	7	28	
Binary class classification CN-MCI		**Predicted label**		
		CN	MCI	
Actual label	CN	50	0	
	MCI	10	25	

TABLE 8 Precession, recall and F1 Score for different classifications using OASIS data set

Classification Type	Class label	Precision (%)	Sensitivity (or) Recall (%)	F1 Score (%)	Specificity (%)
CN-MCI-AD	AD	76.92	100	86.95	80
	CN	80.76	84.0	82.35	87.17
	MCI	100	51.42	67.92	100
CN-AD	AD	86.20	100	92.59	84.0
	CN	100	84.0	91.30	100
MCI-AD	AD	87.71	100	93.45	80.0
	MCI	100	80	88.88	100
CN-MCI	CN	83.33	100	90.90	71.42
	MCI	100	71.42	83.33	100

TABLE 9 Accuracy, average sensitivity and average specificity on test with ADNI and OASIS data set

	ADNI data set				OASIS data set			
Parameter	**CN-AD**	**MCI-AD**	**CN-MCI**	**CN-MCI-AD**	**CN-AD**	**MCI-AD**	**CN-MCI**	**CN-MCI-AD**
Accuracy	100	93.24	98.73	95.61	92.00	91.76	88.23	81.48
Average Sensitivity	100	92.72	98.63	95.15	92.00	90.0	85.71	78.47
Average Specificity	100	92.72	98.63	97.77	92.00	90.0	85.71	89.05

Using 3D MRI is computationally cost and required more memory to train a CNN model. 2D Slices collected from the 3D MRI are used for prediction the stage of AD by different researchers. Ashkan Shakarami et al,[33] they used transfer learning to diagnosis the AD from CN using AlexNet as feature extracted from PET 2D Slices, and performed classification using SVM they achieved 96.3% of accuracy. Transfer learning approaches are adopted in AD diagnosis by different researchers C.D. Billones et al,[35] used VGG16 to perform three-way classification

TABLE 10 Ten-fold cross validation result

Classification	Accuracy	Sensitivity	Specificity
CN-MCI	92.92 ± 3%	92.11 ± 1	95 ± 4%
CN-AD	98 ± 2%	96 ± 3%	97 ± 2%
AD-MCI	90 ± 4%	90 ± 7%	95 ± 3%
CN-MCI-AD	92.9 ± 2%	92.01 ± 3%	92 ± 4%

of CN-MCI and AD achieves 91.85% of accuracy. Valliani.A et al,[37] proposed ResNet-18 to perform classification of AD from CN and achieves accuracy of 81.3%. Marcia Hon et al,[38] work with VGG-16 model and Google Inception V4 Net, they developed VGG16 from scratch and performed classification achieves 74.12% of accuracy, as the model is trained from scratch so it achieves remarkably less accuracy, at the same they work with OXFORD-VGG-16 pre trained model to perform the classification of AD from CN they achieves 92.3% of accuracy the limitation of VGG is large number of trained parameters, they are overcome by Inception V4 Net, this net not only reduce the learning parameters but also feedback the error at intermediate positions and achieves 96.25% of accuracy in Differentiating AD from CN they used OASIS data set. Jain et al,[34] adopted Mathematical modeling in VGG-16 and performed Binary and 3 way classification and achieves 99.14% of accuracy at AD-CN, 99.3% accuracy at AD-MCI, 99.2% of accuracy at CN-MCI and 95.73% accuracy at 3 way classification. Ji, H,[40] used multiple VGG-16 pre trained model and ensemble the results of the each model and select the highest voting as the result they achieves 97.62% of accuracy in differentiating MCI from AD and 88.37% of accuracy in differentiating CN and MCI.

A study on classification of early AD is using the brain tissues such as hippocampus, cerebral cortex, gray matter structural changes. Fan Li et al,[38] they used hippocampus volume as a feature to perform the early diagnosis of AD by differentiating AD from CN, CN from MCI and pMCI from sMCI, they achieved 91.00% of accuracy to differentiate AD from CN, 75.8% of accuracy at CN-MCI and 74.60% of accuracy in differentiating pMCI from sMCI, they used CNN with RNN monitoring the hippocampus volume changes.

In Reference 42, we had used T2w MRI segmented gray matter to perform classification of AD stages with stacked CNN model. It gives 100% accuracy in differentiating AD from CN, 93.0% of accuracy in differentiating MCI from CN, 96.2% of accuracy at AD-MCI, 85.7% accuracy in three-way classifications. It is observed that the model face problem in early diagnosis of AD. In Reference 43 we had used T2w MRI segmented MRI slices to train a CNN model and tested the model with MRI images collected from local hospital and tested the performance of the model and achieved 90.47% accuracy, 86.66% of recall, and 92.59%.

There are few limitations in the existing frame works, they are facing data leakage and not provided the high variance in training data and test data. For early diagnosis of AD perfect estimation of MCI is the key factor. Most of the literature focused on differentiating AD from CN. They are not focusing on neuropathological changes in the MRI. The main advantage of the proposed approach is producing high accuracy as well as specificity in classifying the stage of the AD. The disease stages may vary from subject to subject as normal, MCI and AD. We have used neuropathological changes in Gray matter as phenotypes used to predict the stage of AD. The proposed CNN added with inception Block help in extracting the deep features from GM of T2w MRI to discriminate the stages of the AD. In this paper, we focused on high variance in train data and test data. The proposed approach achieves remarkable results on testing the model with both ADNI and OASIS data sets. There is no single early imaging biomarker for AD is available that deals with neuropathological changes in the Brain tissues till date. Early detection of anatomical atrophy changes help in providing the required medication to restrict disease progression.

5.3 | Contribution of the work

1. Slices are selected from the sMRI using slice selection algorithm
2. Performed skull stripping using morphological and threshold, removed the unwanted tissues from the MRI slices.
3. Extracted gray matter from the MRI slices using segmentation algorithm
4. Trained the proposed deep CNN model using segmented gray matter of the sMRI slice.
5. The performance of the model is tested using independent both ADNI.[45] Open access series of imaging studies (OASIS)[46] data sets.
6. The model work on high variant data set and provides remarkable performance in accurate prediction of MCI.

5.4 | Advantages of our experiment

1. We used a CNN model with an Incepted block that extracts deep features to perform early diagnosis of AD by accurately Predicting MCI.
2. We provided independence in both training and testing data, we trained the model with 80% subject ADNI

TABLE 11 Comparing the different frame works with our proposed algorithm

Authors(Years)	Data set	Sequence	Processes	Classifier	Modalities	Accuracy	Sensitivity (or) recall	Specificity	Precision
Fan Li (2017)[25]	ADNI	sMRI (T1w)	Multi-Model Features	CNN	AD-HC	88.31	91.4	84.42	NA
Cheng and Liu 2017[26]	ADNI	sMRI (T1w)	Multi-Mode CNN	CNN	CN-AD	87.15	86.36	85.93	NA
Ashkan Shakarami et.al., (2020)[32]	ADNI	PET	Extracting features from AlexNet and Classify using SVM	CNN + SVM	AD-CN	96.3	95	97.78	NA
Jain (2019)[34]	ADNI	sMRI	Mathematical Model ($P_E S_E C_{TL}$)	Transfer learning VGG-16	AD-CN	99.14	99.5	NA	99.5
					AD-MCI	99.3	99.5		99.5
					CN-MCI	99.2	99.5		99.5
					CN-MCI-AD	95.73	96.0		96.33
C.D. Billones (2016)[35]	ADNI	sMRI (T1w)	Transfer learning	VGG-16	CN-MCI-AD	91.85	NA	NA	NA
Valliani, A., & Soni, A[37]	ADNI	sMRI (T1w)	Transfer Learning	ResNet-18	CN-AD	81.3	NA	NA	NA
					AN-MCI-AD	56.8			
Marcia Hon, Khan [2017][38]	OASIS	sMRI (T1w)	Scratch model	VGG-16	CN-AD	74.12	NA	NA	NA
			Transfer Learning	VGG-16 Transfer learning	CN-AD	92.3			NA
				Inception V4 (Transfer learning)	CN-AD	96.25			NA
Fan Li et al (2019)[39]	ADNI	sMRI (T1w)	Hippocampus	CNN + RNN	AD-CN	91.00	NA	NA	
					MCI-CN	75.80			
					pMCI-sMCI	74.60			
Andrés Ortiz et al,[40]	ADNI	sMRI (T1w)	Gray matter	CNN	AD-CN	90	86.0	94.0	NA
					MCI-AD	84	79	89	NA
					CN-MCI	83	67	95	NA
Ji, H (2019)[41]	ADNI	sMRI	Gray matter and White matter Ensemble of 3 VGG-16 CNN Models	CNN	AD-MCI	97.62	96.00	100	NA
					CN-MCI	88.37	80.56	94.00	NA
					AD-CN	98.59	97.22	100	NA
Shaik Basheera et. all.,[42]	ADNI	sMRI	CNN Trained with T2w MRI segmented gray matter	CNN	CN-AD	100	100	100	100

(Continues)

TABLE 11 (Continued)

Authors(Years)	Data set	Sequence	Processes	Classifier	Modalities	Accuracy	Sensitivity (or) recall	Specificity	Precision
					CN-MCI	98.0	96.0	100	97.45
					AD-MCI	96.2	93.0	100	95.7
					CN-MCI-AD	86.7	89.6	86.61	88.1
Shaik Basheera et. all.,[43]	ADNI	sMRI	Trained with T2w gray matter	CNN	CN-MCI-AD	90.47	86.66	92.59	89.59
Proposed approach	ADNI	sMRI	T2w MRI segmented Gray Matter	CNN with inception block	CN-AD	100	100[a]	100[b]	100[c]
					CN-MCI	98.73	98.63[a]	98.63[b]	98.85[c]
					AD-MCI	93.24	92.72[a]	92.72[b]	94.44[c]
					CN-MCI-AD	95.61	95.15[a]	97.77[b]	96.136[c]
	OASIS	sMRI	T2w MRI segmented gray matter	CNN with inception block	CN-AD	92.00	92.00[a]	92.00[b]	93.1[c]
					CN-MCI	88.23	85.71[a]	85.71[b]	91.665[c]
					AD-MCI	91.76	90.00[a]	90.00[b]	93.85[c]
					CN -MCI-AD	81.48	78.47[a]	89.05[b]	85.89[c]

[a]Average sensitivity of all the classes.
[b]Average specificity of all the classes.
[c]Average precision of all the classes.

data set and tested with remaining 20% subject ADNI data set and another publically available independent data set OASIS, provide independence between training and testing data, it over comes data leakage.

5.5 | Limitations of the experiment

1. We are not taking clinical parameters into consideration.
2. Our model is not to replace a physician but it strengthens the decision of the physician and gives second opinion.

5.6 | Research line

1. By adding different clinical evaluations to the model it become strong and gives better opinion to the physician.
2. By training the model with different neurological disorders it is possible to diagnose more disorders.

6 | CONCLUSION

In this paper we used CNN model with incepted block for early diagnosis AD by accurately differentiating MCI from AD and CN. In this approach we used gray matter segmented T2w MRI image to perform multi-class and binary class classification. Our model achieves accuracy of 96.78% in case of multi-class classification, 100% with AD-CN, 93.24% with AD-MCI and 98.73% of accuracy in case of CN-MCI, on testing with OASIS data set with the trained model it achieves 92.00% of accuracy with CN-AD, 88.23% of accuracy with CN-MCI, 91.76% of accuracy with AD-MCI and 8148% of accuracy with three-way classification. Change in gray matter was advantageous to differentiate MCI accurately. That gives promising results that provide required assistance to the physician for early diagnosis of the disease.

CONFLICT OF INTEREST

Authors not have any conflict of interest.

AUTHOR CONTRIBUTIONS

Shaik Basheera, presents the idea, theoretical designed, performed the experiment, computations, analytical methods and investigate the work, contributed in data collection, He wrote the manuscript, performed calculations and the numerical simulations. M Satya Sai Ram, contributed in experiment design and data collection. He provided the feedback in preparing manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from ADNI and OASIS. Restrictions apply to the availability of these data, which were used under license for this study. Data are available http://adni.loni.usc.edu/ and https://www.oasis-brains.org/#access with the permission of ADNI and OASIS.

ORCID

Shaik Basheera https://orcid.org/0000-0001-7815-5203

REFERENCES

1. A. Payan and G. Montana, Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. arXiv Preprint arXiv:1502.02506, 2015
2. According to the Dementia India Report 2010 By the Alzheimer's and Related Disorders Society of India (ARDSI)
3. Bron EE, Smits M, van der Flier WM, et al. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge. *Neuroimage*. 2015;111:562-579.
4. Sevigny J, Chiao P, Bussière T, et al. The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. *Nature*. 2016;537 (7618):50-56. https://doi.org/10.1038/nature19323.
5. Klekociuk SZ, Summers JJ, Vickers JC, et al. Reducing false positive diagnoses in mild cognitive impairment: the importance of comprehensive neuropsychological assessment. *European Journal of Neurology*. 2014;21(10):1330e83. https://doi.org/10.1111/ene.12488.
6. Zhang Y, Wang S, Preetha P, et al. Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC. *Biomed. Signal Process. Control*. 2015;21:58-73. https://doi.org/10.1016/j.bspc.2015.05.014.
7. Diaz-de-Grenu LZ, Acosta-Cabronero J, Pereira JMS, Pengas G, Williams GB, Nestor PJ. MRI detection of tissue pathology beyond atrophy in Alzheimer's disease: Introducing T2-VBM. *NeuroImage*. 2011;56(4):1946-1953. https://doi.org/10.1016/j.neuroimage.2011.03.082.
8. Evgin Gocer, Biomedical information technology: image based computer aided diagnosis systems. Papr prsented at International Conference on Advanced Technologies. Antalya, Turkey. 132, 2018
9. Cao P, Liu X, Yang J, Zhao D, et al. Nonlinearity-aware based dimensionality reduction and over-sampling for AD/MCI classification from MRI measures. *Comput. Biol. Med*. 2017;91: 21-37. https://doi.org/10.1016/j.compbiomed.2017.10.002.
10. Mateos-Pérez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, Evans AC. Structural neuroimaging as clinical predictor: a review of machine learning applications. *NeuroImage Clin*. 2018;20:506-522. https://doi.org/10.1016/j.nicl.2018.08.019.
11. Tanveer M, Richhariya B, et al. Machine learning techniques for the diagnosis of Alzheimer's disease: a review. *ACM Trans. Multimedia Comput. Commun. Appl*. 2020;16(1):35. https://doi.org/10.1145/3344998.
12. Herrup K. Commentary on "Recommendations from the National Institute on Aging-Alzheimer's Association

workgroups on diagnostic guidelines for Alzheimer's disease." Addressing the challenge of Alzheimer's disease in the 21st century. *Alzheimer's & Dementia*. 2011;7(3):335-337. https://doi.org/10.1016/j.jalz.2011.04.002.

13. Uysal G, Ozturk M. Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods. *Journal of Neuroscience Methods*. 2020;337:108669. https://doi.org/10.1016/j.jneumeth.2020.108669.
14. Kulkarni N, Bairagi V. Chapter five - classification algorithms in diagnosis of Alzheimer's disease. In: Kulkarni N, Bairagi V, eds. *EEG-Based Diagnosis of Alzheimer Disease*. United Kingdom: Academic Press; 2018:61-71. https://doi.org/10.1016/B978-0-12-815392-5.00005-8.
15. Evgin Gocer, Automated measurement of changes in cortical thickness from MR images, Paper prsented at: International Conference on Applied Analysis and Mathematical Modeling (ICAAMM2018), Istanbul, Turkey, pg. 78, 2018
16. Hinrichs C, Singh V, Mukherjee L, Xu G, Chung MK, Johnson SC. Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. *Neuroimage*. 2009;48:138-149. https://doi.org/10.1016/j.neuroimage.2009.05.056.
17. Li F, Liu M, Alzheimer's Disease and Neuroimaging Initiative. Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks. *Computer Med Imag Graph*. 2018;70:101-110. https://doi.org/10.1016/j.compmedimag.2018.09.009.
18. Anderson VM, Schott JM, Bartlett JW, Leung KK, Miller DH, Fox NC. Gray matter atrophy rate as a marker of disease progression in AD. *Neurobiol Aging*. 2012;33(7):1194-1202. https://doi.org/10.1016/j.neurobiolaging.2010.11.001.
19. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270-279. https://doi.org/10.1016/j.jalz.2011.03.008.
20. Carneiro G, Nascimento JC, Freitas A. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. *IEEE Trans Image Process*. 2012;12(3):968-982. https://doi.org/10.1109/TIP.2011.2169273.
21. Shin HC, Orton MR, Collins DJ, et al. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. *IEEE Trans Pattern Anal Mach Intell*. 2013;35(8):1930-1943. https://doi.org/10.1109/TPAMI.2012.277.
22. A. Gupta, M. Se, C. Ayhan, and A. Maida. 2013. Natural image bases to represent neuroimaging data. In ICML'13. JMLR, Atlanta, USA, 987–994
23. S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis and D. Feng, Early diagnosis of Alzheimer's disease with deep learning. Paper presented at: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, 2014, pp. 1015-1018. https://doi.org/10.1109/ISBI.2014.6868045.
24. Shi J, Zheng X, Li Y. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. *IEEE Journal of Biomedical and Health Informatics*. 2017, 2017;22(1):173-183. https://doi.org/10.1109/JBHI.2017.2655720.
25. F. Li, D. Cheng and M. Liu, Alzheimer's disease classification based on combination of multi-model convolutional networks. Paper presented at: 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, 2017, pp. 1–5. https://doi.org/10.1109/IST.2017.8261566.
26. Danni C, Manhua L, Jianliang F, et al. Classification of MR brain images by combination of multi-CNNs for AD diagnosis. Paper presented at: Proceedings of SPIE 10420, Ninth International Conference on Digital Image Processing (ICDIP 2017), 1042042. https://doi.org/10.1117/12.2281808.
27. LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput*. 1989;1(4):541-551.
28. Alex Krizhevsky Sutskever, Ilya; Hinton, Geoffrey Eet al., Geoffrey Hinton University of Toronto
29. Simonyan, Karen, and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
30. Szegedy C, Wei L, et al. Going deeper with convolutions. Paper presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.
31. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Paper presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
32. B. Khasgi, C. G. Lee and G. Kwon, Alzheimer's disease Classification from Brain MRI based on transfer learning from CNN. Paper presented at: 2018 11th Biomedical Engineering International Conference (BMEiCON), Chiang Mai, 2018, pp. 1-4. https://doi.org/10.1109/BMEiCON.2018.8609974.
33. Shakarami A, Tarrah H, Mahdavi-Hormat A. A CAD system for diagnosing Alzheimer's disease using 2D slices and an improved AlexNet-SVM method. *Optik*. 2020;212:164237. http://dx.doi.org/10.1016/j.ijleo.2020.164237.
34. Jain R, Jain N, Aggarwal A, et al. Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research*. 2019;57:147-159. https://doi.org/10.1016/j.cogsys.2018.12.015.
35. Billones CD, Demetria D, and Hostallero D. 2016. *DemNet: A Convolutional Neural Network for the Detection of Alzheimer's Disease and Mild Cognitive Impairment*. In TENCON. IEEE, Singapore, 3724–3727. https://doi.org/10.1109/TENCON.2016.7848755.
36. A. Farooq, S. Anwar, M. Awais and S. Rehman, "A deep CNN based multi-class classification of Alzheimer's disease using MRI. Paper presented at: 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, 2017, pp. 1-6. https://doi.org/10.1109/IST.2017.8261460.
37. Aly Valliani, Ameet Soni, Deep residual nets for improved Alzheimer's diagnosis. Paper presented at: 8th ACM International Conference, 2017, Boston, MA, USA, 2017. https://doi.org/10.1145/3107411.3108224
38. M. Hon and N. M. Khan, "Towards Alzheimer's disease classification through transfer learning. Paper presented at: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, 2017, pp. 1166-1169. https://doi.org/10.1109/BIBM.2017.8217822
39. Li F, Liu M. A hybrid convolutional and recurrent neural network for hippocampus analysis in alzheimer's disease. *Journal of Neuroscience Methods*. 2019;323:108-118. https://doi.org/10.1016/j.jneumeth.2019.05.006.

40. Ortiz A, Munilla J, Górriz JM, Ramírez J. Ensembles of deep learning architectures for the early diagnosis of the alzheimer's disease. *International Journal of Neural Systems*. 2016;26(07): 1650025. https://doi.org/10.1142/s0129065716500258.
41. Ji, H., Liu, Z., Yan, W. Q., and Klette, R. (2019). Early diagnosis of alzheimer's disease using deep learning. Paper presented at: Proceedings of the 2nd International Conference on Control and Computer Vision (New York, NY: Association for Computing Machinery), 87–91. https://doi.org/10.1145/3341016.3341024
42. Basheera S, Satya Sai Ram M. A novel CNN based Alzheimer's disease classification using hybrid enhanced ICA segmented gray matter of MRI. *Computerized Medical Imaging and Graphics*. 2020;81:101713. http://dx.doi.org/10.1016/j.compmedimag.2020.101713.
43. Basheera S, Sai Ram MS. Convolution neural network–based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*. 2019;5(1):974-986. https://doi.org/10.1016/j.trci.2019.10.001.
44. Basheera S, Ram MSS. Gray matter segmentation of brain mri using hybrid enhanced independent component analysis in noisy and noise free environment. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*. 2020;47:75–103. http://dx.doi.org/10.4028/www.scientific.net/jbbbe.47.75.
45. Alzheimer's Disease Neuroimaging Initiative, http://adni.loni.usc.edu/
46. Open Access Series of Imaging Studies, http://www.oasis-brains.org.

How to cite this article: Basheera S, Ram MSS. Deep learning based Alzheimer's disease early diagnosis using T2w segmented gray matter MRI. *Int J Imaging Syst Technol*. 2021;1–19. https://doi.org/10.1002/ima.22553