Alzheimer's & Dementia: Translational Research & Clinical Interventions 5 (2019) 974-986

Featured Article

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

Shaik Basheera*, M Satya Sai Ram

Department of ECE, Acharya Nagarjuna University College of Engineering and Technology, Guntur, India

Abstract In recent times, accurate and early diagnosis of Alzheimer's disease (AD) plays a vital role in patient care and further treatment. Predicting AD from mild cognitive impairment (MCI) and cognitive normal (CN) has become popular. Neuroimaging and computer-aided diagnosis techniques are used for classification of AD by physicians in the early stage. Most of the previous machine learning techniques work on handpicked features. In the recent days, deep learning has been applied for many medical image applications. Existing deep learning systems work on raw magnetic resonance imaging (MRI) images and cortical surface as an input to the convolution neural network (CNN) to perform classification of AD. AD affects the brain volume and changes the gray matter texture. In our work, we used 1820 T2-weighted brain magnetic resonance volumes including 635 AD MRIs, 548 MCI MRIs, and 637 CN MRIs, sliced into 18,017 voxels. We proposed an approach to extract the gray matter from brain voxels and perform the classification using the CNN. A Gaussian filter is used to enhance the voxels, and skull stripping algorithm is used to remove the irrelevant tissues from enhanced voxels. Then, those voxels are segmented by hybrid enhanced independent component analysis. Segmented gray matter is used as an input to the CNN. We performed clinical valuation using our proposed approach and achieved 90.47% accuracy, 86.66% of recall, and 92.59% precision.

Keywords: Alzheimer's disease; Independent component analysis; Gaussian Mixture model; CNN; Clinical evaluation

1. Introduction

Alzheimer's disease (AD) is a progressive dementia, which causes a loss of connection between nerve cells in elders. Owing to AD, the brain shrinks, hippocampal size decreases, and the brain ventricles enlarge. As AD progresses, it debases memory, thinking ability, and the person's expressions to the problem in day-to-day activities. Understanding AD, mild cognitive impairment (MCI), and cognitive normal (CN) manifestation is one of the most challenging tasks faced by neurologists from the past few years. Physicians are using different clinical methodologies to perform classification of AD. Clinically, cerebrospinal fluid (CSF) concentration deals with AD. The level of norepinephrine increases in the CSF as the disease progresses. The CSF is collected using a ventricular puncture; the physician makes a hole in the skull and collects the CSF directly from one of the brain ventricles [1]. It is a laborious procedure, and it may have a risk of bleeding in the brain. With the development of medical imaging techniques, neuroimaging plays a major role in the diagnosis of structural and functional changes in the brain and encompasses computer tomography, magnetic resonance imaging (MRI), positron emission tomography, functional MRI, and single-photon emission CT. MRI is used to

*Corresponding author. Tel.: 9059735487.
E-mail address: shaikbphd@gmail.com

https://doi.org/10.1016/j.trci.2019.10.001

analyze structural changes caused by AD, CN, and MCI manifestation because of its ease of accessibility. The most common MRI sequences are T1-weighted and T2-weighted scans. T2-weighted scans are used in this work. Neuroimaging techniques help visualize the anatomical changes in the brain. In Fig. 1, change in the hippocampal size and enlargement of ventricles are observed in the MRI image AD brain having cortical atrophy compared with the MRI image of CN brain and MCI brain.

It is evident that the texture of the brain changes as the disease progresses from CN to MCI to AD. Shape transformation in the brain is used as a morphological signature of the brain structure. Morphological changes in the brain texture, structure, and volume are used to classify the healthy brain from a diseased brain [2,3]. AD is caused by degeneration of brain cells and changes in the brain volume. The early effect of AD is observed based on changes in the hippocampus, the size of which is used to classify the AD stage [4]. Change in the white matter (WM) is estimated to analyze the area of the brain affected due to AD [5]. The gray matter (GM) is used to analyze AD [6]. AD is classified using the volume of interest [7]. Image volume has more number of voxels and high dimensionality. The huge information is reduced via wavelet transformation where the classification is carried out in a voxel-by-voxel manner instead of classifying the entire data [8]; selecting an appropriate voxel and relevant area will result in good specificity and sensitivity [9]; voxel-based features are used to classify AD stages [10].

1.1. Related work

Since the last few years, computer-aided diagnosis (CAD) is used to assist and give a second opinion to the physicians. Many researchers are developing different CAD systems to diagnose AD. Most physicians use physical tests and the Mini-Mental State Examination [2,11] to verify the stage of AD. Clinically AD classification is performed by collecting different parameters and by developing biomarkers to test the AD stage. A 5-stage route map was developed for CSF-based diagnosis of preclinical AD using Aβ ratios rather than Aβ42 [12]. Recent CAD systems use machine learning as a computational technique to analyze patterns of medical data. Different machine learning approaches such as regression, classification, and clustering are used in the CAD system. Machine learning approach gives better classification accuracy based on the features that are extracted from the images; to detect the structural and textural changes in the brain MRI, single modalities and multimodalities are used as features [13]. Brain volume, shape, voxel intensity, CSF measurement, and genetic information are used as features to perform the classification of AD, using random forest [14]; those features are correlated using PCA, the dimensionality of the features is reduced, and they are classified using support vector machine (SVM) and particle swarm optimization [14,15]. As AD progresses, it affects brain tissues such as the WM, GM, and hippocampus. The WM and GM are segmented from brain MRI using learning vector quantization, an unsupervised approach, and classification was performed using SVM. Texture changes in the WM and GM are used to differentiate AD from MCI and CN; texture changes are measured using first-order statistical parameters that are extracted from the histogram, and then the second-order statistical features are extracted from GLCM and Gabor filters [16]—these features are used to differentiate AD from CN using KNN [16] and SVM [17] classifiers. Hybrid features generated by combining texture and volume information, such as texture features along with GM volume, are used to perform the classification of AD using SVM-random Fourier expression (SVM-RFE) [18]. Hybrid features extracted from segmented brain image and clinical data are used for multiclass classification of AD from MCI and CN [19]. As the features are more, the classification accuracy increases, but it makes accurate training of a classifier more complicated; greedy score is used to select the important features, and kernel-based discriminative method is used to perform feature selection of complex features [20]. Hippocampal volume is used to differentiate AD and MCI [21]. Hippocampal volume is verified patchwise [22]; patch-based image features are selected by professional and medical experts with knowledge in medical segmentation. Texture features are extracted patchwise using Gabor filter, and classification is performed using a weak classifier [23]. In all the aforementioned approaches, features are extracted manually, and it requires expert knowledge in selecting the features.

In recent years, deep learning framework has achieved greater success in many fields. An artificial neural network has more influence on the development of deep learning architecture. There are many machine learning approaches adopted to perform classification of medical images using CAD. The advantage of neural networks is that the CAD system used in recent days [24]. In deep convolutional neural networks, hierarchical layers are connected and have the advantage over artificial neural networks. Deep learning achieves good performance in medical image analysis [25]. Deep radiomic features are extracted from the three-dimensional MRI image using entropy convolution neural network (CNN) to perform AD classification [26]. Multimodal three-dimensional CNN is used to extract the features and perform AD classification [27]. Features from stacked autoencoders and low-level features in combination help to build the classification model [28]. Extracting texture from the center slices of the MRI image and using those as input for performing AD classification using bootstrap algorithm as the region of interest is used to collect the features from MRI [29]. Transfer learning using the VGG-16 pretrain model is used to perform the classification of AD-NC-MCI

[30]. Hippocampal volume patches are used to perform the classification using the hybrid classifier CNN and recurrent neural network [22].

In this article, we propose a CNN classifier for automatic classification of AD from MCI and CN using GM. We evaluated the architecture performance using T2-weighted MR images collected from a standardized data set, Alzheimer's Disease Neuroimaging Initiative (ADNI). The contribution of the article is summarized as follows:

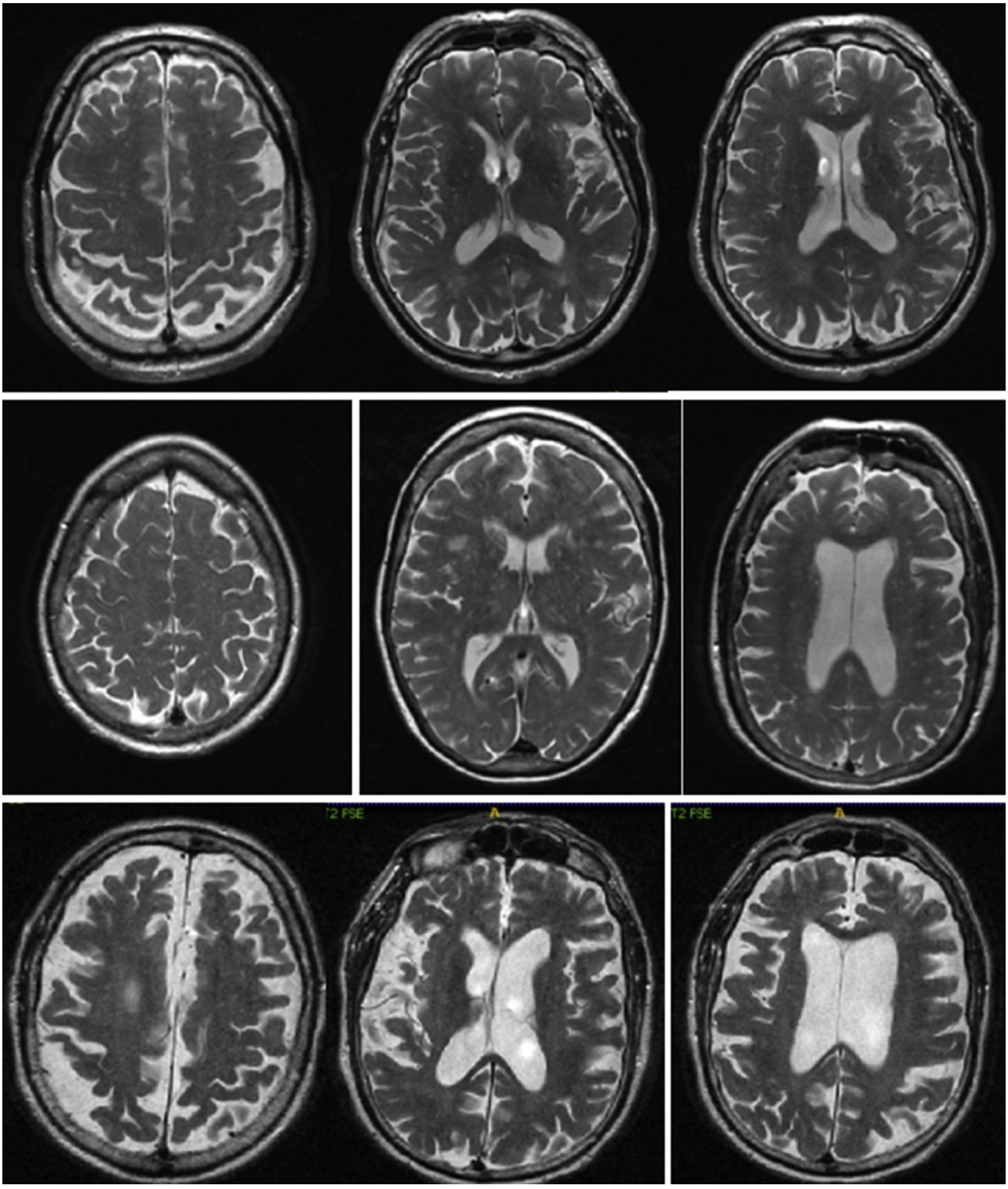

Fig. 1. Cross sections from MRI images of CN (the top row), MCI (the middle row), and AD (the bottom row). Abbreviations: MRI, magnetic resonance imaging; CN, cognitive normal; MCI, mild cognitive impairment; AD, Alzheimer's disease.

Table 1
Demographic representation of MRI images

Date source	Research group	Number of subjects	Sex		Age (years)	Number of MRI volumes	Image slices	Imaging protocol
			M	F				
ADNI	AD	120	59	61	55–93	635	6017	Axial, 2D, 1.5 Tesla field strength
	CN	117	50	67	71–96	637	6000	
	MCI	112	66	66	61–96	548	6000	

Abbreviations: AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; CN, cognitive normal; MCI, mild cognitive impairment; MRI, magnetic resonance imaging; M, male; F, female.

a) Threshold and morphological operations are performed to remove the unwanted tissues from the voxels.
b) We specifically used GM for atrophy detection. In this article, GM tissues are segmented using hybrid enhanced independent component analysis (ICA).
c) CNN architecture is trained using segmented GM voxels.
d) The trained CNN is evaluated using independent MRI voxels collected from a local MRI center and correlated with clinical information, which achieves remarkable accuracy.

The aim of this work is to develop a computer-based diagnosis system that provides additional support for the medical staff to support their diagnosis evidence.

2. Materials

In our work, a total of 1820 MRI images are obtained from the ADNI database (adni.loni.usc.edu). We used 1.5-Tesla, T2-weighted MRI volumes, which are of 420 × 462 × 32 voxels. We collected AD, MCI, and CN MRIs of individuals of different age groups, both male and female. MRIs and their demographic representation are shown in Table 1. Overall, we collected 635 AD MRIs, 548 MCI MRIs, and 637 CN MRIs.

3. Methodology

In our work, MRIs are initially sliced into voxels. These voxels are preprocessed to correct for geometric distortion and reduce noise using Gaussian filter. Nonbrain tissues are removed from the voxels using the skull stripping algorithm. Structural and texturual changes in the brain are used to differentiate healthy and diseased tissues, and enhanced ICA is used to perform segmentation of the brain into the WM, GM, and CSF. Brain tissue atrophy is used to detect AD stage. AD is a progressive disease, in which the brain experiences changes in GM and WM texture and volume, as well as expansion of ventricles. In our work, we classify the AD based on GM atrophy. We used CNN as a classifier which is used in different computer vision techniques, since the past couple of years. Our classification approach has 3 major sections: (1) preprocessing; (2) train, test, and validation of the classifier; and (3) perform clinical valuation—as shown in Fig. 2.

3.1. Preprocessing

3.1.1. Skull stripping algorithm

Skull stripping is the most important preprocessing technique. For accurate classification of images, unwanted and nonbrain tissues are initially removed from the voxels and the brain tissues are left. The proposed skull stripping algorithm has a sequence of steps. Before applying skull stripping, the image is enhanced using a Gaussian filter, and detailed information and noise are reduced.

The enhanced voxels are convolved by a 3 × 3 filter as given in equation (1).

$$k(x,y) = \begin{bmatrix} 1 & -1 & 1/2 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (1)$$

The filtered image is segmented into brain and nonbrain tissues using thresholding technique; selection of threshold value is crucial to generate the initial binary image. Morphological operation is performed on the binary image to remove the unwanted regions. Active contour is applied on the binary image to separate foreground from background and generate a final binary mask. On multiplying the final binary mask with the original voxel, the skull is stripped and unwanted tissues such as skull, scalp, dura, eyes, and so forth are removed from the voxel. The skull stripping algorithm steps are as follows.

$I(x, y)$ is the input image
$k(x, y)$ 3 × 3 Kernal
I_5 is the output image

Step1: *I = input image*
Step2: *I_1 = enhanced image*
Step3: Convolution of input image and kernel $I_2 = I_1$ *conv* k
Step4: Threshold the image
for $i = 1$: *row length of* I_2
for $i = 1$: *column length of* I_2
if $I_2(x, y) \leq 254$
$I_2(x, y) = 0$
else
$I_2(x, y) = 1$
end if
end if
end if
Step5: Erode: I_3 = erode I_2 image
Step6: Active contour: I_4 = Perform active contour on I_3
Step7: Skull stripped image $I_5 = I \times I_4$

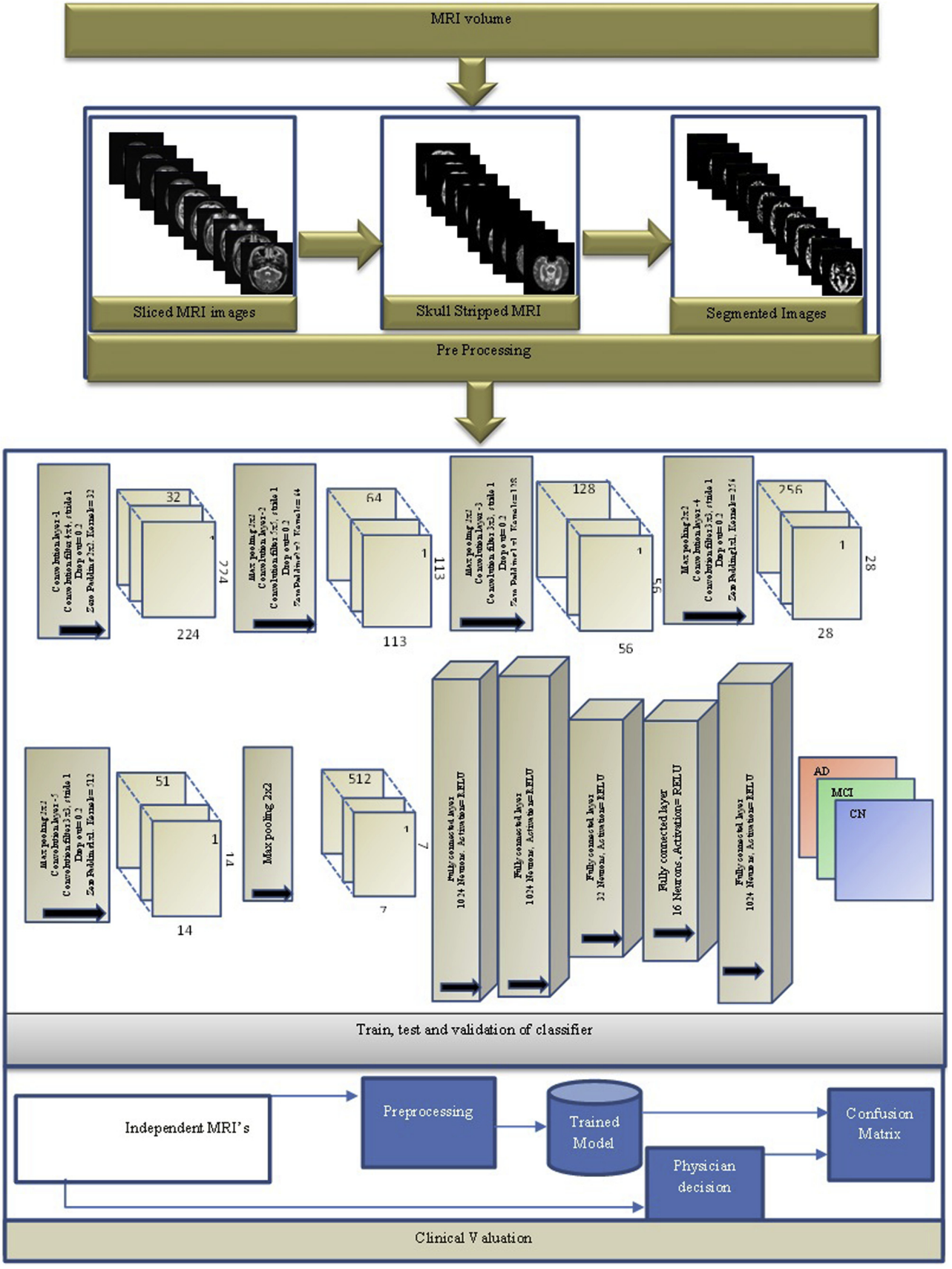

Fig. 2. Proposed framework with CNN. Abbreviation: CNN, convolution neural network.

3.1.2. Segmented algorithm

Blind separation of brain tissues in the MRI is carried out by using an unsupervised segmentation approach. In our work, we have used hybrid enhanced ICA. K-means and expected maximization (EM) are combined to form a hybrid strategy to cluster the brain tissues in the MRI. This combination achieves the capability of providing clusters for well-distributed image pixels and compactness through EM.

In our proposed hybrid enhanced ICA, the concept of mixture model is introduced and it is characterized into mutually exclusive classes. In the modified GMM approach, spatial information is added to GMM using Markov random field (MRF) and takes spatial dependency into account. In EM, the expected step is computed using log likelihood with mean and variance calculated using modified K-means and latent variable calculated through Gibbs density function. The aforementioned parameters are used as input parameters to hybrid enhanced ICA to perform the segmentation of the brain MRI voxels. The algorithm is further explained using the following steps:

3.2. Classifier

The CNN is used for classification. In our article, we used 224 × 224-sized gray segmented images as input to the CNN. The performance of the CNN depends on the network architecture and weights that are set. The architecture of the CNN depends on the specific task, and the requirements of the data for the network need to be known. The size of the MRI slice, filter size, number of kernels, padding, and strides determine the particular convolution layer size. Our classifier has 5 convolution layers with 32, 64, 128, 256, and 512 filters with different sizes (4,4), (5,5), (3,3), (3,3), and (3,3), respectively, at different stages of stride 1, padding, followed by max pooling layers used to extract features and 6 fully connected layers used to perform classification. The network is trained using back-to-back propagation with 200 epochs; we used Adam optimization. Equations (2), (3), (4), (5), (6), (7), and (8) show the layerwise parameter calculation and activation functions used at convolution layer and fully connected layers.

$g(x, y)$ *is the input image* into k identically independent GMMs with parameters $\theta_k = \{\mu_k, E_k\}$

Step1: Represent g(x,y) in vector$\{g_i: i = 1, 2, 3, 4\ldots N\}$

Step2: Modified K-means to find the prior information of the Gaussian mixture model such as mean and covariance

{Mixing Coeficient π_k: $i = 1, 2, 3, 4 \ldots N\}$, g_i *is the gray level.*

a. Partition of N pixels into K equal sets
b. Center of each set as a centroid $c1, c2, c3, \ldots c_k$
c. Find the distance between Euclidean distance between g(x, y) and the cluster centers
d. Find the centroid that is close to the particular g(x, y)
e. Recalculate the centroids of each clusters
f. Repeat the steps from c to e
g. If the distance between g(x,y) and new cluster center is less than or equal to the previous distance, then g(x,y) will be in the same cluster otherwise it moves to another cluster based on the distance.
h. The process continues until the clusters are convergence
i. Collect mean and covariance of the clusters $\theta_k = \{\mu_k, E_k\}$

Step2: $\theta_k = \{\mu_k, E_k\}$

for i = 1: pixels

for k = 1: number of k

Probability density of the mixture model is considered as $p(x_i|\pi,\theta)=\sum_{k=1}^{c}\pi_i^k\ p\ (x_i|\theta_k)$

end

end

Step 3: log likelihood of the density function is calculated to find the probability of pixels that belong to the particular Gaussian ln $p(=p(X|\theta)=\sum_{i=1}^{N}\ln p(x_i|\theta)$, $\theta = \{z, \mu, \sigma\}$, z, latent variable and calculated using expectation and maximization

Step 4: E step for I:

$Q_{(i)}\ (Z^{(i)}) : = P(z^i|x^i,\theta)$

Step 5: M Step for all z

$\theta:=argmax_\theta \sum_i \sum_{Z^i} Q_i(z^i)\ log\ \frac{p(x^i,z^i;\theta)}{Q_i(z^i)}$, Q_i *the posterior distribution of* (z^i)

Step 6: Prior distribution of π is given by MRF model through Gibbs density function $p(\pi)=\text{expv}\frac{-\beta\sum_{i=1}^{N}V_{Ni}(\pi)}{a}$, a is a normalizing constant, $V_{Ni}\ (\pi)$ *is the clique potential function*

Step 7: Process stops when $\left\|\theta^{new} - \theta^{old}\right\| \leq error$

$$Convolution_{width} = \frac{Input\ MRI\ Slice\ Size_{width} - Filter_{width} + (2XPadding)}{Strides_{Width}} + 1 \quad (2)$$

$$Convolution_{Height} = \frac{Input\ MRI\ Slice\ Size_{Height} - Filter_{Height} + (2XPadding)}{Strides_{Height}} + 1 \quad (3)$$

$$No.\ of\ Neurons\ in\ Convolution\ layer\ =\ Convolution_{width} \times Convolution_{Heigh} \times Number\ of\ Filters \quad (4)$$

$$Max\ Pooling\ resultant\ image_{size} = \frac{Convolution_{width}}{2} \quad (5)$$

$$Fully\ Connected\ layer_{parameters} = No.\ of\ parameters\ form\ previous\ stage \times No.\ of\ nodes\ in\ the\ present\ layer \quad (6)$$

$$\text{Rectified linear unit used as activation function. } f(x) = \begin{cases} 0\ for\ x < 0 \\ x\ for\ x \geq 0 \end{cases} \quad (7)$$

$$soft\ max\ function = \frac{e^{xi}}{\sum_{j=1}^{n} e^{xj}}\ for\ \ i = 1, 2, 3, \ldots n \quad (8)$$

3.3. Model development and training

In our work, we used MATLAB R2015b to perform slicing, skull stripping, and segmentation of the image. To train deep neural network data, parallel processing is needed, so we used an open-source software package Python, version 3.0, Google Colab to perform the training and validation of the classifier (GPU: 1xTesla K80, having 2496 CUDA cores, compute 3.7, 12 GB [11.439 GB useable] GDDR5 VRAM). We used Keras library over TensorFlow modules to design our proposed model.

3.4. Creating training and test set

Our total data set has 18,017 GM segmented images. We shuffled and split the data set in the ratio 80:20 as training and test data sets. We used this data set for multiclass classification and binary classification; the data set is summarized in Table 2.

Table 2
Training set, validation set, and test set sizes

Classification type	Class label	Training set	Test set	Total images
Multiclass classification	AD	4814	1203	6017
	MCI	4800	1200	6000
	CN	4800	1200	6000
Binary class classification	AD-MCI	9614	2403	12,017
	AD-CN	9614	2403	12,017
	CN-MCI	9600	2400	12,000

Abbreviations: AD, Alzheimer's disease; CN, cognitive normal; MCI, mild cognitive impairment.

Our classifier has 5 convolution layers with 32, 64, 128, 256, and 512 filters with different sizes (4,4), (5,5), (3,3), (3,3), and (3,3) , respectively, at various stages of stride 1, padding, followed by max pooling of the feature extractor followed by 6 fully connected layers. The networked is trained by Adam optimization using back-to-back propagation with 200 epochs.

Table 3
Summary of the architecture of CNN

Layer 1	Kernel size	Feature map
Input image	224 × 224	—
Convolution layer 1	4 × 4	221 × 221 × 32
Dropout layer 1	20%	—
Zero padding layer 1	3 × 3	227 × 227 × 32
Max pooling 1	2 × 2	113 × 113 × 32
Convolution layer 2	5 × 5	109 × 109 × 64
Dropout layer 2	20%	—
Zero padding layer 2	2 × 2	113 × 113 × 64
Max pooling 2	2 × 2	56 × 56 × 64
Convolution layer 3	3 × 3	54 × 54 × 128
Dropout layer 3	20%	—
Zero padding layer 3	1 × 1	56 × 56 × 128
Max pooling 3	2 × 2	28 × 28 × 128
Convolution layer 4	3 × 3	26 × 26 × 256
Dropout layer 4	20%	—
Zero padding layer 4	1 × 1	28 × 28 × 256
Max pooling 4	2 × 2	14 × 14 × 256
Convolution layer 5	3 × 3	12 × 12 × 512
Dropout layer 5	20%	—
Zero padding layer 5	1 × 1	14 × 14 × 512
Max pooling 5	2 × 2	7 × 7 × 512
Fully connected layer 1	1024	—
Fully connected layer 2	1024	—
Fully connected layer 3	32	—
Fully connected layer 4	16	—
Fully connected layer 5	1024	—

Abbreviation: CNN, convolution neural network.

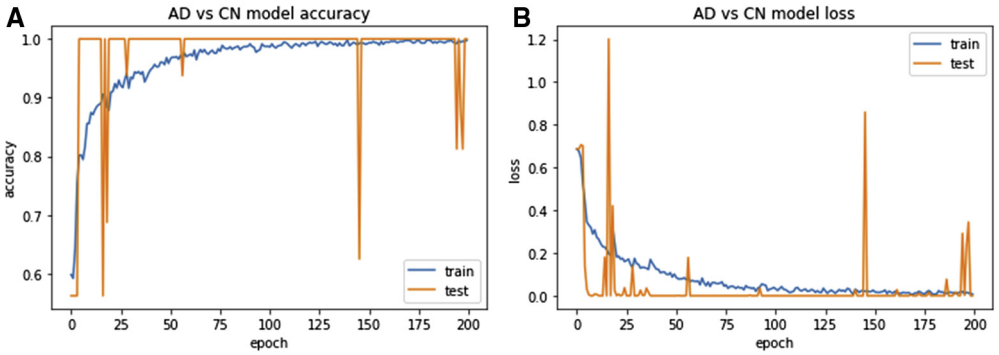

Fig. 3. Accuracy and loss calculation of AD-CN during training and testing. (A) AD-CN accuracy calculation. (B) AD-CN loss calculation. Abbreviations: AD, Alzheimer's disease; CN, cognitive normal.

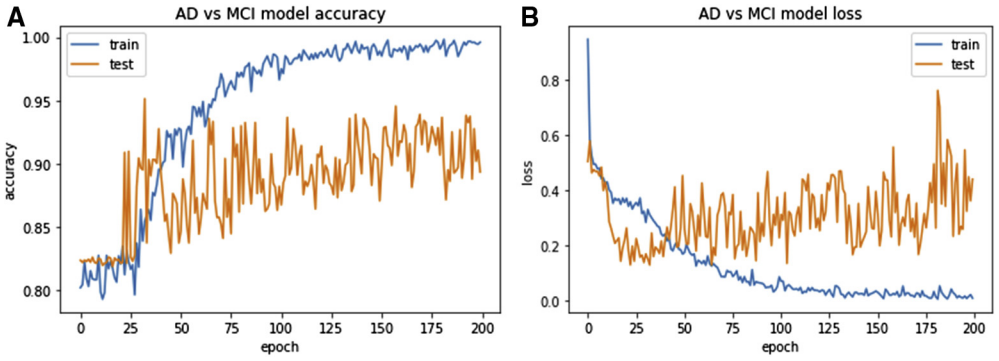

Fig. 4. Accuracy and loss calculation of AD-MCI during training and testing. (A) AD-MCI accuracy calculation. (B) AD-MCI loss calculation. Abbreviations: AD, Alzheimer's disease; MCI, mild cognitive impairment.

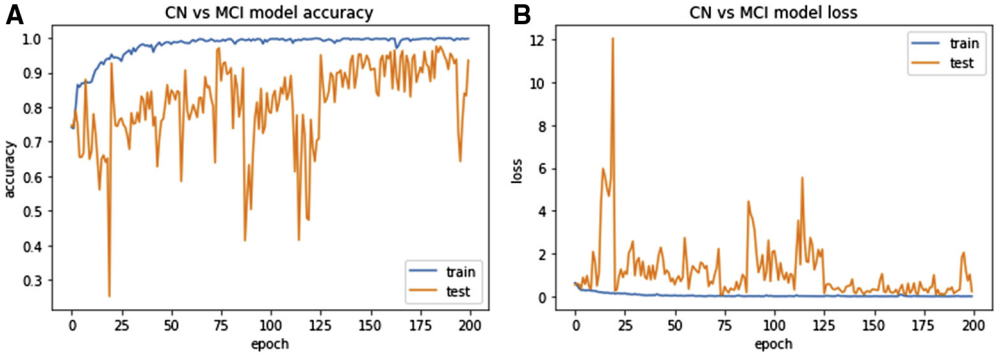

Fig. 5. Accuracy and loss calculation of CN-MCI during training and testing. (A) CN-MCI accuracy calculation. (B) CN-MCI loss calculation. Abbreviations: CN, cognitive normal; MCI, mild cognitive impairment.

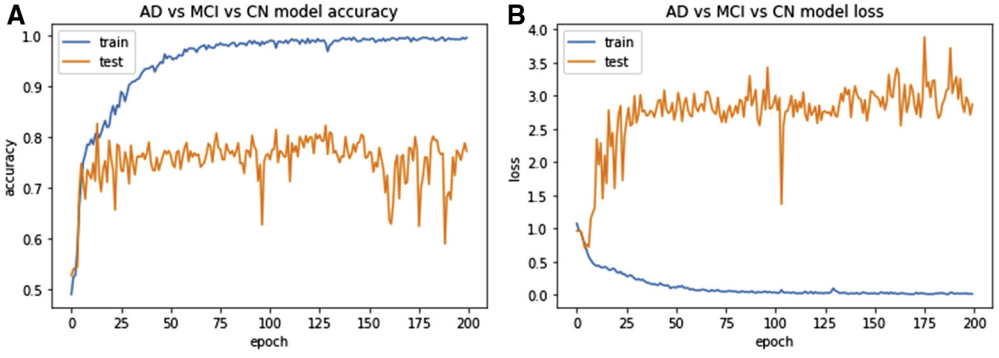

Fig. 6. Accuracy and loss calculation of AD-CN-MCI during training and testing. (A) AD-CN-MCI accuracy calculation. (B) AD-CN-MCI loss calculation. Abbreviations: AD, Alzheimer's disease; CN, cognitive normal; MCI, mild cognitive impairment.

Table 4
Comparing the proposed approach with previous frameworks

Author (year)	Resources	Processing and training	Classification	Modalities	Accuracy	Sensitivity	Specificity	AUC
Fung and Stoeckel (2007) [9]	SPECT	Relevant area and selection of voxels	SVM	AD-HC	–	84.40%	90.90%	–
Escudero et al. (2011) [21]	MRI	Volumetric and cortical thickness of the hippocampus	SVM	AD-HC	89.20%	–	–	–
				AD-MCI	72.70%			
Suk and Shen (2013) [28]	MRI, PET	SAE	Multikernel SVM	AD versus HC	95.50%	–	–	–
				MCI versus HC	85.00%			
				MCI_C versus MCI_{NC}	75.80%			
Adaszewski et al. (2013) [4]	MRI	Hippocampal temporoparietal atrophy	SVM	HC	80.30%	–	–	–
				AD	73.50%			
				cMCI	63.70%			
				ncMCI	69.00%			
Yang et al. (2013) [15]	MRI	Volume and shape	PCA + SVM	AD-NC(Vol)	82.35%	–	–	–
				MCI-NC(Vol)	77.72%			
				AD-NC(Sha.)	94.12%			
				MCI-NC (Sha.)	88.89%			
Gray et al. (2013) [14]	PET	MRI volumes, voxel-based FDG-PET signal intensities, CSF biomarker measures, and categorical genetic information	Random Forest	AC-HC	89%	–	–	–
				AD-MCI	75%			
Ortiz et al. (2013) [31]	MRI	Tissue information	SVM	AD-CN	90%	95%	–	–
Li et al. (2017) [27]	MRI	Multimodel features	CNN	AD-HC	88.31	91.4	84.42	92.73
Lama et al. (2017) [20]	MRI	Cortical thickness, folding index	10-fold CV	SVM	AD-CN	60.1	74.63	88.81
				IVM		59.5	62.3	62.85
				RELM		77.3	62.12	79.85
			LOO CV	SVM	AD-CN	78.01	75.81	79.12
				IVM		73.36	70.97	75.95
				RELM		75.66	72.13	77.22
Altaf et al. (2018) [19]	MRI	GLCM, SIFT, LBP, HoG's, Clinical Data	SVM	AD versus CN	97.80%	100%	95.65%	–
				AD versus MCI	85.30%	75.00%	94.29%	
				CN versus MCI	91.80%	90.00%	93.33%	
Hett et al. (2018) [23]	MRI	Gray Matter + Gabor Filter	Weak classifier	Intensity-based grading histo				
				CN versus AD		93.5	95.5	82.7
				CN versus pMCI		90	81.8	81.4
				AD versus sMCI		81.1	78.5	68.3
				sMCI versus pMCI		74.9	77.6	67.2
				Texture-based grading histo				
				CN versus AD		94.6	94.2	86.6
				CN versus pMCI		92	92.5	81.2
				AD versus sMCI		82.6	77.6	72.6
				sMCI versus pMCI		76.1	74.9	70.2

(*Continued*)

Table 4
Comparing the proposed approach with previous frameworks (*Continued*)

Author (year)	Resources	Processing and training	Classification	Modalities	Accuracy	Sensitivity	Specificity	AUC
Chaddad et al. (2018) [26]	MRI	Selecting MRI based on entropy, intensity, texture, shape	CNN	AD-HC	–	–	–	92.58%
Jain (2019) [30]	MRI	Mathematical model $P_FS_EC_{TL}$	Transfer learning VGG-16	AD-CN-MCI	95.73	–	–	–
				AD-CN	99.14			
				AD-MCI	99.3			
				CN_MCI	99.22			
Kim et al. (2019) [32]	MRI	Cortical thickness	Hierarchical approach	CN-dementia	86.10%	87.00%	85.40%	0.917
				AD versus FTD	90.80%	87.50%	92.00%	0.955
				bvFTD versus PPA	86.90%	92.10%	77.10%	0.865
				nfvPPA versus svPPA	92.10%	97.40%	88.00%	0.955
Vaithinathan et al. (2019) [29]	MRI	Region of interest	SVM + bootstrapped	AD-CN	–	89.58	85.82	–
Li et al. (2019) [22]	MRI	Hippocampus	CNN + RNN	AD-NC	–	–	–	91.00%
				MCI-NC				75.80%
				pMCI-sMCI				74.60%
Basaia et al. (2019) [33]	MRI	No feature engineering	CNN	AD-NC	99.2	98.9	99.5	–
				AD-MCI	75.4	74.5	76.4	
				MCI-NC	87.1	87.8	86.5	
Proposed Method	MRI	Segmented gray matter using enhanced ICA	CNN	AD-CN-MCI	86.7	89.6	86.61	88.50
				AD-CN	100	100	100	100
				AD-MCI	96.2	93.0	100	98.72
				CN-MCI	98.0	96.0	100	99.87

Abbreviations: CNN, convolution neural network; PET, positron emission tomography; ICA, independent component analysis; SPECT, single-photon emission computed tomography; AD, Alzheimer's disease; CN, cognitive normal; RNN, recurrent neural network; MCI, mild cognitive impairment.

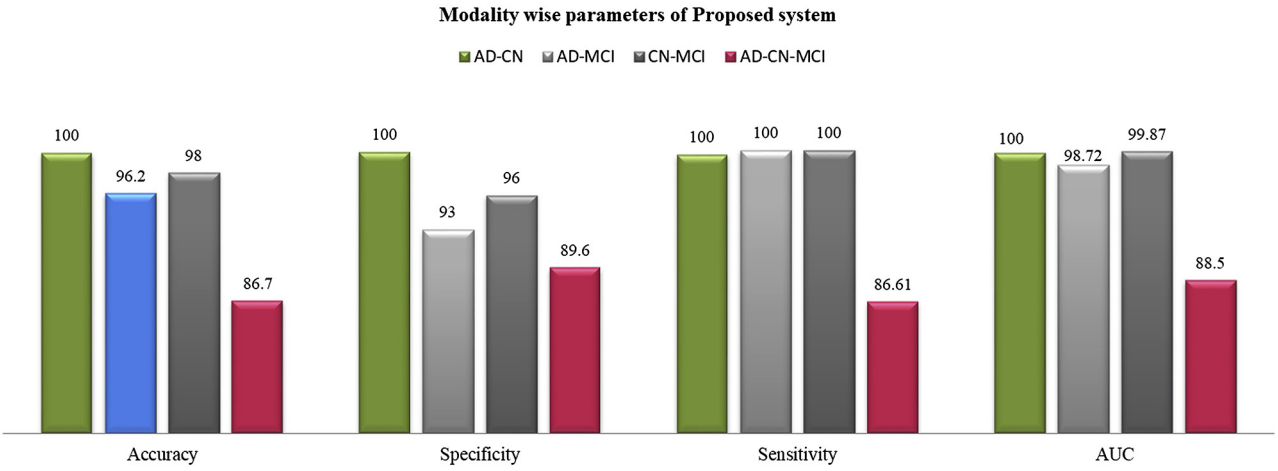

Fig. 7. Proposed system parameters modality wise.

4. Results

In this work, a total of 18,016 MRI axial slices are used; these are generated from 1820 T1-weighted MRIs collected from the standard AD data set, the ADNI. All the voxels are preprocessed by enhancing them using rotationally invariant Gaussian filters. Irrelevant tissues such as scalp, skull, ears, dura, and eyes are removed from the MRI voxels using skull stripping algorithm. We focused on the GM for atrophy detection. GM tissues are segmented using hybrid enhanced ICA. Atrophy in the GM is used to differentiate AD from MCI and CN.

Our proposed CNN model has 5 convolution layers and 5 fully connected layers; each convolution layer is followed by dropout, padding, and max pooling layers with ReLu as an activation function. Preprocessed images are augmented to increase the sample size and train the CNN model. We used Keras with TensorFlow to build the proposed CNN model using Python. Our proposed approach performs binary classification and multiclass classification to fit the model in a batch size of 128 in 200 epochs using Google Colab. It takes around 7 hours to train the model. The total architecture is summarized in Table.3.

We train the model by Adam optimization with a learning rate of 0.001, beta1 of 0.9, and beta2 of 0.999. Our classifier is trained with 200 epochs. We used it to perform binary classification and multiclass classification. During binary classification, we first trained the classifier with AD-CN segmented images and the model resulted in 99.75% training accuracy. Then, we trained the classifier with segmented AD-MCI voxels, which achieved 98.72% training accuracy. Later, it was trained with segmented CN-MCI images which resulted in 99.87% training accuracy. Multiclass classification was performed by the trained classifier, using segmented AD-MCI-CN images which achieved 99.50% training accuracy. Training, testing accuracy, and loss graphs are shown in Fig. 3, Fig. 4, Fig. 5, and Fig. 6. It is required that our proposed framework is trained and the prediction is made with utmost accuracy.

We compared the performance of the proposed system with that of different models discussed in literature review as shown in Table 4. It is observed that our classifier achieves remarkable performance both in binary classification and multiclass classification. Fig. 7 shows parameters of both binary and multiclass classifications.

We further performed the clinical evaluation using our proposed approach on 21 independent MRI slices collected from Poorvi MRI Center, Chirala and compared the predicted results generated by our system with results diagnosed by the physician. The confusion matrix of clinical analysis is shown in Fig. 8.

We calculated some important performance measurement parameters as follows using Equations (9), (10), (11), and (12):

$$Accuracy = \frac{\mathbf{TN+TP}}{\mathbf{TN+FP+FN+TP}} \tag{9}$$

$$Recall = \frac{\mathbf{TP}}{\mathbf{FN+TP}} \tag{10}$$

Actual/ Predicted	AD	MCI	CN
AD	9	0	0
MCI	0	3	2
CN	0	0	7

Fig. 8. Confusion matrix for clinical analysis of images.

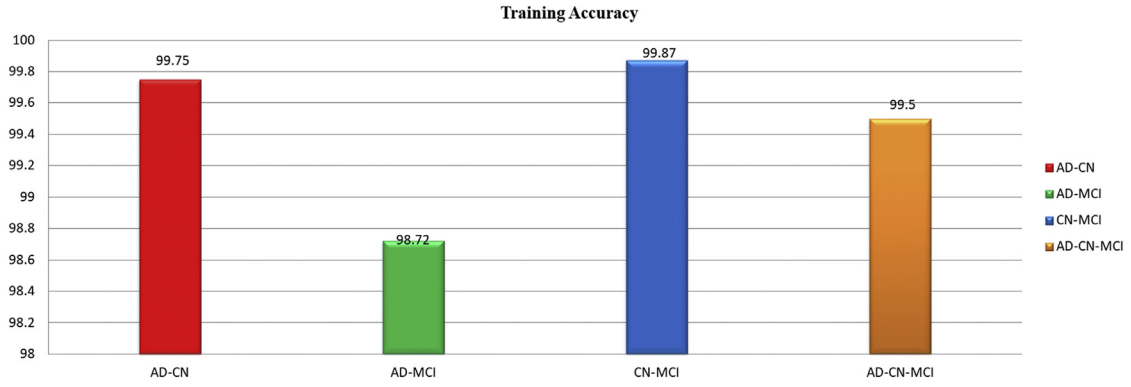

Fig. 9. Clinical evaluation of proposed system.

$$True\ \ Negative\ \ Rate = \frac{\mathbf{FP}}{\mathbf{TN}+\mathbf{FP}} \tag{11}$$

$$Precision = \frac{\mathbf{TP}}{\mathbf{FP}+\mathbf{TP}} \tag{12}$$

We achieve 90.47% accuracy, 86.66% recall, and 92.59% precision in comparison of our system with physician decision. Bar diagram of the clinical evaluation is given in Fig. 9.

5. Conclusion

Effective diagnosis of AD helps the patient to get a featured treatment. Many researchers are focusing on this challenging task; they had developed many CAD systems to perform the diagnosis of AD. In our workflow, we developed a deep learning approach to perform the classification based on GM segment, using hybrid enhanced ICA.

Our proposed framework has more strengths than the previous techniques. We use heterogeneous MRI volumes of different age groups and gender. In our experiment, we used T2-weighted MRI to perform the classification. The GM has neuron cell bodies and non-neuron brain cells called glial cells. The GM undergoes development and growth throughout childhood and adolescence; it is used to carry glucose to the brain, and changes in this affect the memory, speech, and motor controls. In our work, we mainly focused on the use of the GM to classify AD. In our work, we observed that the framework is not affected with noise and data augmentation.

Our deep learning model got trained and was validated and tested on the MRI collected from the database, and we performed binary classification such as AD-MCI, AD-CN, and MCI-CN and multiclass classification such as AD-MCI and CN. We further compared the classifier performance with the physician's decision and achieved good results. No other framework performed the comparison of the system with the physician's decision. Our system is recommend not to replace but to support the physician decision.

RESEARCH IN CONTEXT

1. Systematic review: In our work we had used AD data collected from online repository to train the model and test the model using images collected from local MRI center. We used 21 MRI slices of different age groups of 60 to 92 years both male and female, and compared the test result of model with a physician decision based on MSME score, to evaluate the system accuracy.
2. Future directions: System is further improved by adopting multiple image data such as T1, T2 and meta data along with the proposed system to improve evaluation of AD at clinical level.

References

[1] Elrod R, Peskind ER, DiGiacomo L, Brodkin KI, Veith RC, Raskind MA. Effects of Alzheimer's disease severity on cerebrospinal fluid norepinephrine concentration. Am J Psychiatry 1997;154:25–30.
[2] Weissberger GH, Strong JV, Stefanidis KB, Summers MJ, Bondi MW, Stricker NH. Diagnostic accuracy of memory measures in Alzheimer's dementia and mild cognitive impairment: a systematic review and meta-analysis. Neuropsychol Rev 2017;27:354–88.
[3] Lorio S, Kherif F, Ruef A, Melie-Garcia L, Frackowiak R, Ashburner J, et al. Neurobiological origin of spurious brain morphological changes: A quantitative MRI study. Hum Brain Mapp 2016; 37:1801–15.
[4] Adaszewski S, Dukart J, Kherif F, Frackowiak R, Draganski B, Alzheimer's Disease Neuroimaging Initiative. How early can we predict Alzheimer's disease using computational anatomy? Neurobiol Aging 2013;34:2815–26.
[5] Maggipinto T, Bellotti R, Amoroso N, Diacono D, Donvito G, Lella E, et al. DTI measurements for Alzheimer's classification. Phys Med Biol 2017;62:2361–75.
[6] Klöppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, et al. Automatic classification of MR scans in Alzheimer's disease. Brain 2008;131:681–9.

[7] Liu Y, Teverovskiy L, Carmichael O, Kikinis R, Shenton M, Carter CS, et al. Discriminative MR image feature analysis for automatic schizophrenia and Alzheimer's disease classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2004. p. 393–401.
[8] Lao Z, Shen D, Xue Z, Karacali B, Resnick SM, Davatzikos C. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. NeuroImage 2004; 21:46–57.
[9] Fung G, Stoeckel J. SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information. Knowl Inf Syst 2007;11:243–58.
[10] Suk HI, Lee SW, Shen D, Alzheimer's Disease Neuroimaging Initiative. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. NeuroImage 2014;101: 569–682.
[11] Klekociuk SZ, Summers JJ, Vickers JC, Summers MJ. Reducing false positive diagnoses in mild cognitive impairment: the importance of comprehensive neuropsychological assessment. Eur J Neurol 2014; 21:1330–6. e82-e83.
[12] Adamczuk K, Schaeverbeke J, Vanderstichele HM, Lilja J, Nelissen N, Van Laere K, et al. Diagnostic value of cerebrospinal fluid aβ ratios in preclinical Alzheimer's disease. Alzheimers Res Ther 2015;7:75.
[13] Shen D, Wee CY, Zhang D, Zhou L, Yap PT. Machine Learning Techniques for AD/MCI Diagnosis and Prognosis. In: Dua S, Acharya U, Dua P, eds. Machine Learning in Healthcare Informatics. Intelligent Systems Reference Library. Berlin, Heidelberg: Springer; 2014.
[14] Gray KR, Aljabar P, Heckemann RA, Rueckert D. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. NeuroImage 2013;65:167–75.
[15] Yang S-T, Lee J-D, Chang T-C, Huang C-H, Wang J-J, Hsu W-C, et al. Discrimination between Alzheimer's disease and mild cognitive impairment using SOM and PSO-SVM. Comput Math Methods Med 2013;2013:253670.
[16] Kadhim W. Alzheimer Disease Diagnosis using the K-means, GLCM and K_NN. J Univ Babylon 2017;26:57–65.
[17] Reghu A, John R. GLCM Based Feature Extraction of Neurodegenerative Disease for Regional Brain Patterns. Int J Eng Res Gen Sci 2016;4:319–23.
[18] Xiao Z, Ding Y, Lan T, Zhang C, Luo C, Qin Z. Brain MR Image Classification for Alzheimer's Disease Diagnosis Based on Multifeature Fusion. Comput Math Methods Med 2017;2017:1952373.
[19] Altaf T, Anwar SM, Gul N, Majeed MN, Majid M. Multi-class Alzheimer's disease classification using image and clinical features. Biomed Signal Process Control 2018;43:64–74.
[20] Lama RK, Gwak J, Park J-S, Lee S-W. Diagnosis of Alzheimer's disease based on structural MRI images using a regularized extreme learning machine and PCA features. J Healthc Eng 2017; 2017:5485080.
[21] Escudero J, Zajicek JP, Ifeachor E. Machine Learning classification of MRI features of Alzheimer's disease and mild cognitive impairment subjects to reduce the sample size in clinical trials. Conf Proc IEEE Eng Med Biol Soc 2011;2011:7957–60.
[22] Li F, Liu M, Alzheimer's Disease Neuroimaging Initiative. A hybrid Convolutional and Recurrent Neural Network for Hippocampus Analysis in Alzheimer's Disease. J Neurosci Methods 2019;323:108–18.
[23] Hett K, Ta VT, Manjón JV, Coupé P, Alzheimer's Disease Neuroimaging Initiative. Adaptive fusion of texture-based grading for Alzheimer's disease classification. Comput Med Imaging Graph 2018;70:8–16.
[24] Suk HI, Wee CY, Lee SW, Shen D. State-space model with deep learning for functional dynamics estimation in resting-state fMRI. Neuroimage 2016;129:292–307.
[25] Shen D, Wu G, Suk H. Deep learning in medical image analysis. Annu Rev Biomed Eng 2017;19:221–48.
[26] Chaddad A, Desrosiers C, Niazi T. Deep Radiomic Analysis of MRI Related to Alzheimer's Disease. IEEE Access 2018;6:58213–21.
[27] Li F, Cheng D, Liu M. Alzheimer's disease classification based on combination of multi-model convolutional networks, 2017 IEEE International Conference on Imaging Systems and Techniques. Beijing: IST; 2017. p. 1–5.
[28] Suk HI, Shen D. Deep learning-based feature representation for AD/MCI classification. Med Image Comput Comput Assist Interv 2013; 16:583–90.
[29] Vaithinathan K, Parthiban L, Alzheimer's Disease Neuroimaging Initiative. A Novel Texture Extraction Technique with T1 Weighted MRI for the Classification of Alzheimer's Disease. J Neurosci Methods 2019;318:84–99.
[30] Jain R, Jain N, Aggarwal A, Jude Hemanth D. Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. Cogn Syst Res 2019;57:147–59.
[31] Ortiz A, Górriz JM, Ramírez J, Martínez-Murcia FJ. LVQ-SVM based CAD tool applied to structural MRI for the diagnosis of the Alzheimer's disease. Pattern Recognit Lett 2013;34:1725–33.
[32] Kim JP, Kim J, Park YH, Park SB, Lee JS, Yoo S, et al. Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. NeuroImage: Clin 2019;23:101811.
[33] Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. NeuroImage Clin 2019;21:101645.