Convolutional Neural Network Based Prediction of Conversion from Mild Cognitive Impairment to Alzheimer's Disease: A Technique using Hippocampus Extracted from MRI

Gulshan MUKHTAR, Saima FARHAN
Lahore College for Women University, Lahore, Pakistan
gulshan.mukhtar@lcwu.edu.pk

***Abstract*—Alzheimer's disease (AD) is an irreversible neurodegenerative disorder. Mild Cognitive Impairment (MCI) is a prodromal stage of AD and its identification is very crucial for early treatment. MCI to AD conversion is of imperative concern in current Alzheimer's research. In this study, we have investigated the conversion from MCI to AD using different types of features. The impact of structural changes in entire brain tissues captured through MRI, genetics, neuropsychological assessment scores and their combination are investigated. Computational cost can be significantly reduced by examining only the hippocampi region, atrophy of which is visible in the earliest stages of the disease. We proposed a CNN based deep learning approach for the prediction of conversion from MCI to AD using above mentioned features. Highest accuracy is achieved when left hippocampus is used as a region of interest (ROI). The proposed technique outperforms the other state of the art methods, while maintaining a low computational cost. The main contribution of the research lies in the fact that only a single slice based small region of MRI is used resulting in an outstanding performance. The accuracy, sensitivity and specificity achieved are 94%, 92% and 96% respectively.**

***Index Terms*—artificial neural networks, computer aided diagnosis, image analysis, image classification, pattern recognition.**

I. INTRODUCTION

Alzheimer's disease (AD) is a neurodegenerative disorder characterized by continuous loss of brain tissues that leads to severe memory decline, other cognitive disabilities and ultimately death. AD is the 6th leading cause of death in United States. By 2018, it is estimated that 5.7 million people of all ages are living with AD in America. By the year 2025, the population of age 65 and older with AD is expected to reach 7.1 million – a rise of nearly 29% [1].

The progressive nature of brain atrophy in AD, demands for developing automated techniques in order to detect the cognitive decline at an early stage [2]. Mild Cognitive Impairment (MCI), a prodromal stage of AD, is characterized as the intermediary period between Cognitive Normal (CN) and the expected AD. No medical treatments have been established so far for AD. Hence, the development of neuropsychological rehabilitations has attracted tremendous consideration. These rehabilitation therapies can have a better control when the diagnosis of MCI is done at an initial stage [3]. A patient with MCI is at a greater risk of progression towards Alzheimer's and other dementias. The research in the identification and detection of AD and MCI is now effectively moving towards the prediction of MCI conversion to AD. Fig. 1 shows sample MRIs of CN, MCI and AD patients.

Modern multimodal neuroimaging techniques are receiving great attention in current dementia research. These modalities have enabled researchers to analyze and quantify brain structures and functions. Structural Magnetic Resonance Imaging (sMRI) [4], functional Magnetic Resonance Imaging (fMRI) [5], Positron Emission Tomography (PET) [6] and Single-Photon Emission Computed Tomography (SPECT) [7] have successfully instigated for early detection of brain tissue degeneration and the prediction of MCI progression towards AD.

Laplace Eigen maps model is designed to improve the clinical trial of MCI to AD conversion [6]. A Computer-Aided Diagnosis (CAD) system has been designed for the prediction of MCI conversion in one to three year time before medical interpretation is made [8]. PET and SPECT are common neuroimaging modalities that use nuclear medicine injected in patient's body to track the disease. These procedures are time consuming, expensive and require latest technology. Moreover, the imaging tests that use nuclear medicine cannot be repeated frequently to monitor the progression of MCI to AD.

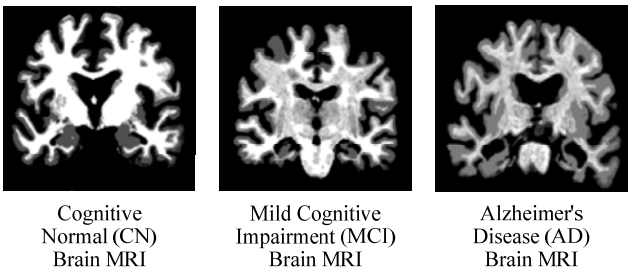

Figure 1. Sample brain MRIs of CN, MCI and AD patients

Recently, researches have revealed that up to 20% of population above 65 years of age might have MCI [8]. Numerous machine learning techniques like Support Vector Machine (SVM) [9, 10], Naïve Bayes [11], Multilayer Perceptron [12] and Decision Trees [4] are being used for the efficient classification of AD, MCI and CN. A number of different features and modalities have been investigated for prediction of MCI and AD. An ensemble of classifiers based approach for AD prediction has been proposed using

Digital Object Identifier 10.4316/AECE.2020.02013

structural features from MRI [4]. Another approach using ensemble of classifiers for prediction of AD based on fusion of volumetric, textural and hemodynamic features from *fMRI* has been used effectively [13]. The prediction of conversion from MCI to AD using MRI and structural features has also been investigated [14]. Researchers have proposed short term prediction of MCI to AD conversion using longitudinal studies of brain MRIs [9]. As hippocampus is known to be effected in the earliest stage of Alzheimer's, a vast number of researchers have focused to segment it from MRI and perform volumetric studies on it [15-17]. A CAD technique has been proposed performing grey matter atrophy classification for AD detection [18].

Another prognostic model has been designed for the prediction of conversion form MCI to AD and the prediction of most probable progression window using conformal prediction (CP) approach is done [11, 19]. A number of studies determined the effectiveness of combined use of sMRI and cognitive methods to observe and track MCI and its progression [15, 20]. Clinicians and researchers mostly use Mini-Mental State Examination (MMSE), Clinical Dementia Rating (CDR), Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) and other clinical assessments to measure the cognitive dysfunction of brain at different stages of dementia [9, 19]. MMSE helps diagnose dementia, its progression and severity [21]. CDR is used to characterize multiple aspects of intellectual and functional abilities that help quantify the severity of AD and dementia [18]. The consequences of these tests being used individually result in inappropriate classification of dementia stage [22]. The ability of MMSE to distinguish subjects with MCI from CN aging is not highly reliable [23].The genetic aspects cannot be ignored as it plays a vital role in the onset and progression of AD. Apolipoprotein E (APOE) with allele 'ε4' has been consistently associated with this disease. APOE- ε4 is the strongest risk factor and key predictor of the development of AD, specifically in MCI subjects [24].

Machine learning techniques are being used for MCI and AD detection but majority of the work is concerned with classification instead of prediction. Although multimodal neuroimaging techniques are being used for accurate classification but these techniques are computationally and economically very expensive. AD classification from handcrafted features is not reproducible and reliable. With advancements in technology and increase in CPU power, artificial neural networks (ANN) started to be used more efficiently and effectively [25]. ANNs have been widely used for decision making in medical applications [26]. Various modeling approaches have been introduced in literature for diverse application e.g. modeling derived from Bayesian filtering [27] and surrogate model based optimization of traffic lights [28]. Another modeling approach based on ANN is deep learning .The use of deep learning, which learns the features itself by analysis of data, has overcome the limitations of other modeling approaches.

In this study, we present a Convolutional Neural Network (CNN) based approach for the prediction of conversion from MCI to AD using Volume of Interest (VOI) of left hippocampus. MRI scans have been pre-processed, segmented and the area of left hippocampus has been extracted using ROI mask mapping technique. The prediction of MCI convertors (MCIc) and MCI non-convertors (MCInc) is performed using CNN. For comparison, we have also extracted handcrafted features from same MRI scans. Volumes of Grey Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF) are the best known features to classify AD [4, 8]. We have also used MMSE and CDR scores as features for classification using traditional machine learning algorithms. The proposed approach is computationally and economically inexpensive and provides promising results.

Rest of the paper is organized as follows: section 2 provides statistics of data used in this work and the methodology to design CNN network. The results and discussions of the proposed technique are presented in section 3. The paper ends with conclusion and future work in section 4.

II. Materials and Methods

Dataset: In this research, T1-weighted MRI volumes of the brain obtained from ADNI database (adni.loni.usc.edu) are used. T1-weighted MRI volumes present better contrast of white and grey matter as compared to T2-weighted MRI. A total of 200 MRI scans have been downloaded from ADNI with NIFTI file format including MRIs of MCI and AD subjects in the age range of 56-90 years. Out of these, 100 subjects are categorized as MCIc based on MMSE and CDR scores. A sample of 100 subjects who remained stable (MCInc) are also included in the research. APOE genotype data is collected for the same subjects provided by the ADNI database. Subject attributes are shown in Table I.

Table I. Demographics of Subjects Involved

Group	MCIc	MCInc
Number of subjects (n)	100	100
Male	52	70
Female	48	30
Age (mean ± S.D)	76.36 ± 6.6	76.17 ± 7.6
Age Range	56-91	56-90
MMSE (mean ± S.D)	24 ± 1.84	26.52 ± 1.77
CDR (mean ± S.D)	0.69 ± 0.25	0.5 ± 0

Proposed Approach: The proposed approach is a longitudinal study for prediction of conversion from MCI to AD. It involves the classification of left hippocampus segmented images extracted from brain MRIs using CNN. For comparison with deep learning techniques, volumetric features are also extracted from preprocessed MRI and classification is performed using traditional machine learning methods. The traditional techniques are enriched with variety of features including the most well-known image features, neuropsychological scores and genetic data with a target to achieve best possible results. The purpose of comparison is to validate the effectiveness of CNN in the domain of Alzheimer's research. The proposed approach comprises of preprocessing, segmentation, feature set construction and classification. The MRIs obtained from ADNI are preprocessed images that have gone through

various types of corrections. Preprocessing performed on these scans include Gradwarp, B1 non-uniformity correction and N3 bias field correction as shown in Fig. 2.

Gradwarp is an automatic correction of MRI geometry misinterpretation due to stochastic gradient descent. It is important to perform gradient non-linearity correction for decreasing faults in longitudinal behavior response procedures. If it is ignored, gradient non-linearity (GNL) manifests as image geometric distortion [29, 30].

B1 non-uniformity correction technique applies the B1 calibration scans to correct the image contrast non-uniformities. N3 reduces intensity contrast due to the magnetic consequence at 3T. This technique is applied to all images to reduce residual intensity inhomogeneity [31].

Brain Surface Extraction (BSE) and Bias Field Correction (BFC) have been applied to preprocessed and corrected images. Second phase segments the image into different regions and extracts the most useful features from segmented image using Partial Volume Classifier (PVC). Deep learning methods are effective when dataset size is large, resulting in greater training time as compared to conventional methods. Scaling down the images, in order to reduce its size and ultimately the computational time, results in information loss. The effected region in the brain during early development of AD is hippocampi, which is a proven biomarker of AD progression. Therefore, we extracted left hippocampus as ROI and used it for classification. It has been observed that the left hippocampus is more vulnerable than right one to AD pathology [32] and it is affected in early stages of dementia [33, 34]. In the last phase, classification between MCIc and MCInc has been done using CNN. A detailed methodology of the proposed approach is shown in Fig. 3.

1. Preprocessing: Brain Surface Extraction (BSE) is a preprocessing phase in which non-brain tissues i.e., skull and scalp are eliminated from the MR images. Skull stripping is done using a combination of anisotropic diffusion filtering, edge detection and a series of mathematical morphological operators [35-37]. Parameters in this process are set as: automated iterations = 5, diffusion iterations = 3, diffusion constant = 25, edge contrast = 0.64 and erosion size = 1. MR images are also processed to trim spinal cord in BSE in order to get brain only structures.

The obtained image from BSE is processed for Bias Field Correction (BFC) [37]. This is the evaluation of local gain alterations by investigation of local ROIs which are spread all over the MRI scan. Within each ROI, a partial volume measurement model is fitted to the histogram of the ROI. One module in this model is gain alteration, thus the fitting method produced an approximation of the gain alterations for the ROI. A tri-cubic B-spline is then equipped to the robust set of local approximations to produce a correction field for the whole brain volume, which is then eliminated from the image to produce non-uniformity corrected image. Parameters applied for non-uniformity correction are: histogram radius = 12, sample spacing = 16, control point spacing = 64, spline stiffness = 0.0001 and cuboid shape is selected for ROI.

2. Segmentation and Feature Set Construction: Segmentation is performed with the help of PVC. The partial volume measurement model is used in this classification under the hypothesis that the gain is uniform, and is combined with a spatial prior 0.1 that models the largely contiguous nature of brain tissue types. PVC provides labels of multiple classes of tissues including GM, WM and CSF volumes along with their corresponding combinations [37].

Hippocampus is extracted from MRI for classification of subjects using CNN. Hippocampus is primarily associated with memory; particularly the long-term memory. It is also linked with spatial memory that enables navigation. The healthy function of the hippocampus can be affected by AD in its very preliminary stages. Left hippocampus is affected in initial phase of the disease as compared to right hippocampus [33, 34]. The extraction of left hippocampus is done using ROI mask mapping technique [4] as shown in Fig. 4.

Two cognitive assessment scores i.e. MMSE and CDR have also been incorporated in this research. These cognitive biomarkers play important role in the classification of AD, MCI and CN along with the prediction of conversion from MCI to AD. The ADNI Biomarker Core provides genetic information of contributing subjects, which are determined from subject's blood test at the time of baseline scans. The APOE gene is an individual categorical variable of an individual which is strongly associated with the risk of AD development. It comes in one of the six possible combinations i.e. (ε2, ε2), (ε2, ε3), (ε2, ε4), (ε3, ε3), (ε3, ε4) and (ε4, ε4). The subjects with ε4 have higher risk of AD development. The genetic information makes it possible to predict the conversion well before time.

3. Classification: The ROIs extracted from MRIs have been classified using CNN. The extraction of features, from MRI volumes, is computationally expensive and creates very large datasets for training of classifiers. The advantage of using CNN is better training than a typical neural network and this leads to noise reduction and improved performance of classification task. The CNN uses deep learning approach to perform image classification [38]. CNNs have been widely proven to be an effective class of networks for many applications including image and video recognition.

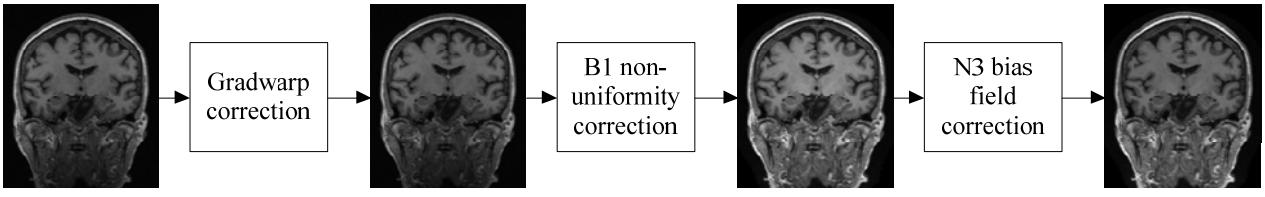

Figure 2. Gradwarp, B1 and N3 Corrections on a sample MR image

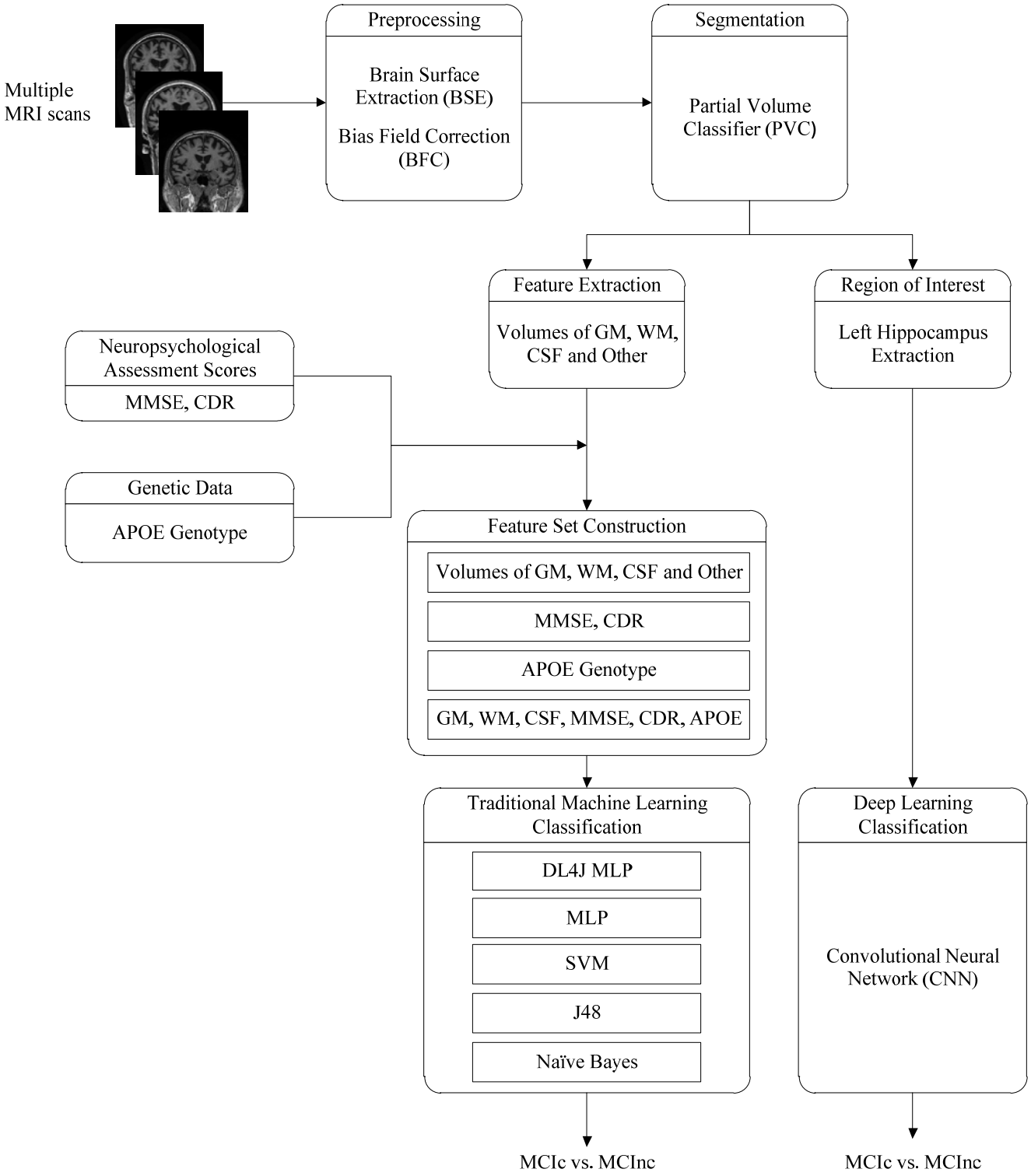

Figure 3. Methodology of the proposed approach

The CNN designed for this work comprises of seven layers as shown in Fig. 5. Each layer demonstrates some operators as convolution and subsampling. The convolution operators calculate the outcome of neurons to be reconnected to local regions in the input, by computing a dot product between their weights and a small region to link it to the input image.

The proposed CNN extracts most discriminative features from the image to classify it. The inputs are the hippocampus images extracted from the MR images for both MCI and AD classes. The resolution of the input images is 150×150×1, representing height, width and number of channels.

In convolutional layer, we have applied 32 filters each of size 5×5, showing neurons that connects to the same region of the output. The main function of this layer is to detect local combinations of features from preceding layer that automatically map their presence to a feature map. As a result of convolution, the image is divided into a number of perceptrons. For a N×N square neuron layer, a m×m filter w is used. The output of convolutional layer is of size (N-m+1) × (N-m+1).

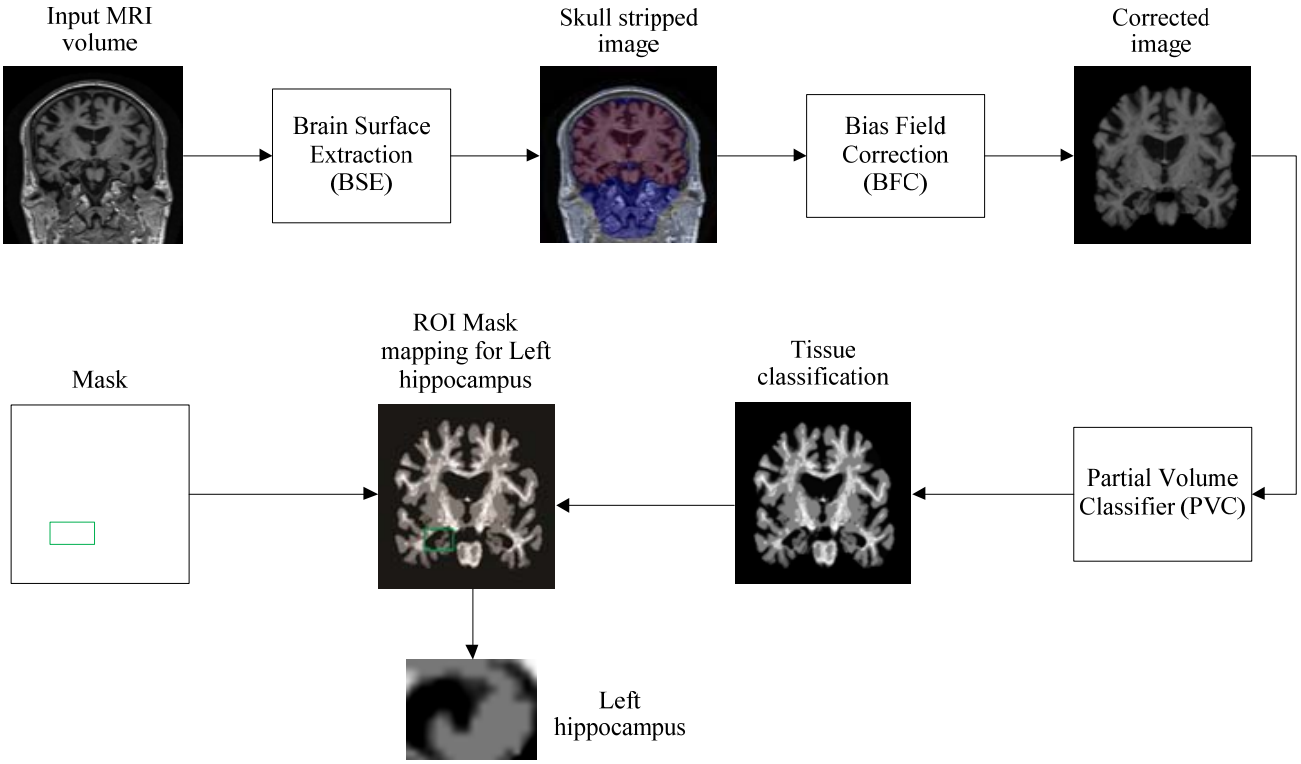

Figure 4. Proposed approach for left hippocampus ROI extraction

To compute the pre-nonlinearity input to the unit X^l_{ij} in the layer, we summed up the contributions from previous layer cells as shown in equation (1).

$$y_{ij}^{p} = \sum_{a=0}^{k-1}\sum_{b=0}^{k-1} x_{ab}\, z_{(i+a)(j+b)}^{p-1} \tag{1}$$

The nonlinearity of convolutional layer is shown in equation (2).

$$z_{ij}^{p} = \sigma(y_{ij}^{p}) \tag{2}$$

The convolutional layer is followed by a nonlinear activation Rectified Linear Units (ReLU) layer. These units are a specific implementation that associates non-linearity and rectification in CNNs. A ReLU is a piecewise linear function as defined in equation (3).

$$z_i^{(p)} = \max(0, z_i^{(p-1)}) \tag{3}$$

The pooling layer p is responsible for reducing the spatial size of activation maps. The pooling size is 2×2 with max pooling operator, resulting in faster convergence and better performance. Stride size is also set to 2, so pools do not overlap. The hyper parameters used here are spatial extent of the filter f(p) and the stride s(p). It takes an input volume of size $n_1^{(p-1)}\times n_2^{(p-1)}\times n_3^{(p-1)}$ and results in an output volume of size $n_1^{(p)}\times n_2^{(p)}\times n_3^{(p)}$ where the values of $n_1^{(p)}$, $n_2^{(p)}$ and $n_3^{(p)}$ are computed as shown in equation (4), (5) and (6) respectively.

$$n_1^{(p)} = n_1^{(p-1)} \tag{4}$$

$$n_2^{(p)} = (n_2^{(p-1)} - f^{(p)})/s^{(p)} + 1 \tag{5}$$

$$n_3^{(p)} = (n_3^{(p-1)} - f^{(p)})/s^{(p)} + 1 \tag{6}$$

The next layer in our proposed CNN is fully-connected layer in which the neurons are connected to all the neurons in the preceding layer. The fully-connected layer is usually used to project all the features across the image to identify larger patterns. The fully connected layers in a CNN are particularly multi-layer perceptrons (MLP) that maps the $n_1^{(p-1)}\times n_2^{(p-1)}\times n_3^{(p-1)}$ activation volume. Here, these activation volumes are gained from the combination of preceding layers into a class probability distribution. Theoretically, a fully connected layer is defined as in equation (7).

$$z_i^{(p)} = f(w_i^{(p)}) \tag{7}$$

where

$$w_i^{(p)} = \sum_{i=1}^{n_1^{(p-1)}} x_{i,j}^{(p)}\, z_i^{(p-1)} \tag{8}$$

The fully-connected layer is followed by softmax layer. The softmax activation function standardizes the outcome of the previous layer and is defined in equation (9).

$$softmax(z)_i = \frac{\exp(w_i)}{\sum_i \exp(w_i)} \tag{9}$$

The outcome of softmax layer contains positive numbers, which are used as classification probabilities by the next layer. The final layer is the classification layer. It uses the probabilities provided by the softmax layer for each input in order to allocate it to one of the mutually exclusive classes. It also computes the loss through the loss function (cross entropy), evaluated on the outcome value of softmax layer.

Accuracy, sensitivity and specificity are used as performance evaluation measures for our proposed approach in order to compare it with existing approaches. These performance evaluation measures have been used widely in research and are presented in equations (10), (11) and (12).

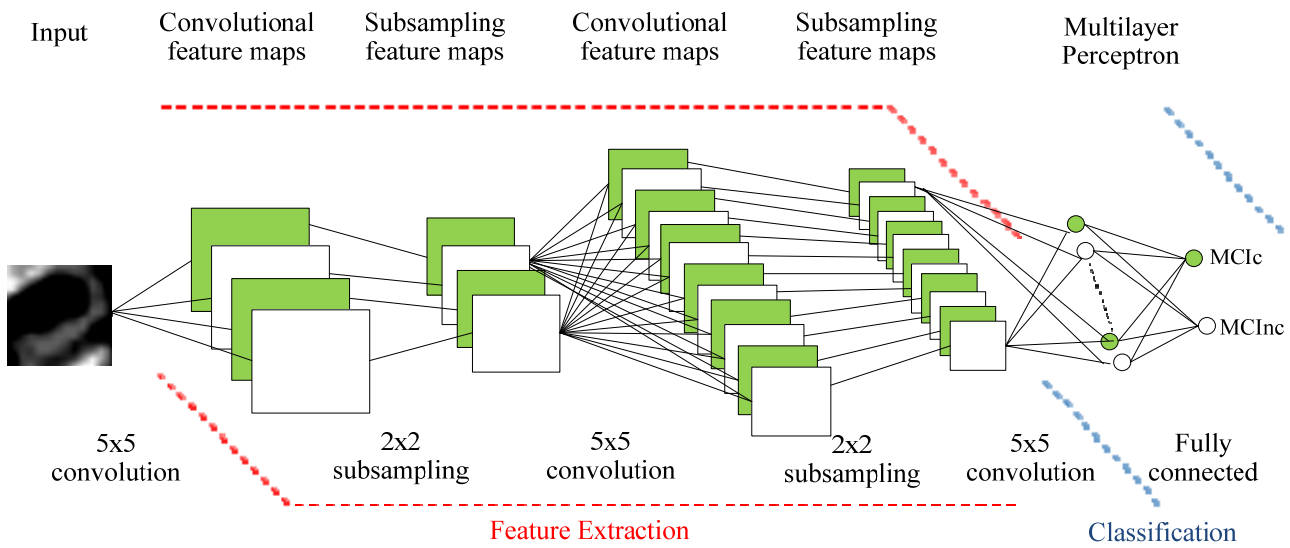

Figure 5. Basic structure of CNN architecture designed for the classification of MR images

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (10)$$

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (11)$$

$$Specificity = \frac{TN}{(TN + FP)} \quad (12)$$

Here,
TP is True Positives i.e. number of subjects with AD that are correctly classified as MCIc,
TN is True Negatives i.e. number of subjects with MCI that are correctly classified as MCInc,
FN is False Negatives i.e. number of subjects with AD that are incorrectly classified as MCInc,
FP is False Positives i.e. number of subjects with MCI that are incorrectly classified as MCIc.

III. RESULTS AND DISCUSSION

The CNN approach of deep learning is used to extract most relevant features from images and classify them. The segmented ROI images are smaller in size and, therefore, their classification is computationally less expensive. The entire dataset (200 images) is split into training (150 images) and test (50 images) set. The CNN model is trained on 150 images from training dataset. The trained CNN model is then tested with 50 images from test set and accuracy is reported. A total of 15 epochs has been used in CNN model. Initial training accuracy is 35% (1st epoch), which converges to 99% in last epoch (15th epoch), resulting in an overall training accuracy of 99% and validation accuracy of 94%. The validation specificity is 96% and validation sensitivity is 92%.

In order to validate the significance of this work, a number of traditional machine learning methods are also used to classify the handcrafted features of MRIs for the prediction of MCI conversion to AD. WEKA tool has been used to classify image features (volume of GM, WM, CSF), neuropsychological assessment scores (MMSE, CDR) and the genetic data (APOE) individually as well as collectively. The classifiers used include DL4J, MLP, SVM, J48 and Naïve Bayes.

It is observed that volumetric image features presented higher accuracy as compared to neuropsychological assessment scores and genetic data. It is also evident that results obtained from the collection of different feature types are slightly better but still lower than the results achieved from volumetric image features. The accuracy, sensitivity and specificity of different feature sets are shown in Table II.

TABLE II. RESULTS OF IMAGE FEATURES, MMSE + CDR, APOE AND COMBINED FEATURES PRESENTED IN PERCENTAGE

Feature Set	DL4J			MLP			SVM			J48			Naïve Bayes		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
GM+WM+CSF	86	83	88	84	82	85	86	83	88	75	70	80	86	83	88
MMSE+CDR	76	73	78	74	72	75	76	73	78	70	61	79	76	73	78
APOE	70	65	75	66	62	69	68	62	73	46	46	46	68	62	73
All above features	81	72	90	83	82	83	84	80	88	78	80	76	77	82	72

The task of this research is to investigate whether neuropsychological test scores and genetic data along with volumetric image features improve the classification accuracy. The answer is 'No' as is evident from the results. Moreover it is examined that whether ROI based image data using CNN produces better results as compared to volumetric image features. In this case the answer is 'Yes'. The hippocampal image data, trained and tested on CNN, provides better classification results than volumetric image features. The hippocampus extracted from MR images are smaller in size and computationally less expensive and provides higher accuracy (94%) as compared to the accuracy obtained from volumetric image features (highest 86%). CNN algorithms have higher AD prediction accuracy due to the fact that multiple hidden layers are present which lead the classifier for taking better decisions.

Most of the literature investigates the classification or prediction of AD and NC subjects. Machine learning as well as deep learning techniques has been exploited effectively in this regard. The new research trends are now focusing on the pre-clinical stage of MCI. For early diagnosis of AD, timely prediction of MCI is vital in order to stop or at least slow down its progression. A number of longitudinal studies have been performed that predict the conversion of MCI to AD. Although a number of different modalities have been used, but for comparison purposes we only consider MRI.

Among various machine learning algorithms, SVM has been used most widely for binary classification of neuroimaging data in AD diagnosis. Feature ranking and classification error is investigated for AD and NC classification, achieving an accuracy of 92.48% [21]. Similarly, an ensemble classification method for AD, MCI and NC is proposed [39]. The best accuracy resulted for the classification of AD and NC i.e. 98.83%. Classification of AD and MCI resulted in 91.66% accuracy, whereas, 90.83% accuracy is achieved for MCI and NC classification.

Recently, deep learning models, especially CNNs, have gained popularity in the domain of image classification. CNNs are effectively investigated for the classification of single modality and multi-modality neuroimaging data as well. Multi-CNN model for classification of AD and NC has been exploited for diagnosis of AD with an accuracy of 87.15% [40]. Spectral CNN is used for classification of AD, MCI and NC [41]. Accuracies of 91.07%, 87.72% and 85.45% have been reported for binary classification AD and NC, AD and MCI, and, MCI and NC, respectively. Considering multi-modality, MRI and PET scans, have been studied for binary as well as multi-class classification using stacked autoencoders [42]. For binary classification of AD and NC, the reported accuracy is 91.40%, whereas, for MCI and NC, it is 82.10%. For multi-class classification of AD, MCI and NC, the accuracy obtained is 53.79%.

The focus of our research work is the classification of MCIc and MCInc because of its importance in early diagnosis of AD. This task is more complicated due to the fact that there exists least discriminating patterns in the brains of the two groups. For comparison with our research work, we have studied a number of research works that perform the classification of MCIc and MCInc.

Traditional machine learning algorithms have been applied to predict conversion of MCI to AD. Beheshti et al. achieved an accuracy of 93.01% for classification of AD and NC, and 75% accuracy for the classification of static MCI and progressive MCI [8]. Lotjonen et al. showed that different groups of training and validating sets progress to a classification accuracy ranging between 53% and 77% for the prediction of MCI conversion [43]. Linear discriminant analysis (LDA) with MRI resulted in an accuracy of 73.95% for classification of MCIc and MCInc [15]. Similarly, an accuracy of 73.91% is reported when most widely used SVM is applied for predicting conversion of MCI to AD [44].

Deep learning models are recently exploited for prediction of conversion from MCI to AD. Multitask deep learning framework is investigated in this regard using MRI and PET data [45]. CNN based MRI image analysis for MCIc and MCInc classification resulted in 79.90% accuracy [46]. Deep belief networks and SVM are combinely studied for a number of binary classification tasks using MRI and PET scans [47]. Among others, the accuracy of MCIc and MCInc turned out to be 78%. Sparse regression models and 2D CNN are investigated for predicting the conversion of MCI to AD [48]. Their results reported an accuracy of 74.82%.

In this research, demographic data of the participants is not included as a feature for conversion prediction. However, this data might present integral information to improve the performance. Age has been proven as a risk factor of AD [49]. Older adult subjects are more probable to develop AD than younger ones, demonstrating that age is a significant forecaster of AD. However, normal aging has comparable shrinkage effects on brain's specific regions as AD which would limit to find the disease related alterations for classification task. In recent studies [50, 51], the aging effect was dropped out before the classification so that the particular disease related modifications can be used to train and validate the classifiers.

Latest approaches are now making use of pre-trained deep learning models known as transfer learning for neuroimaging data. These methods avoid the pipeline phases of feature extraction by using previously trained models. ResNet and VGGNet pre-trained models have been used for classification of AD and NC subjects using MRI, resulting in an accuracy of 88% [52].

It is evident from the studied literature, that predicting conversion from MCI to AD, or in other words, classification of MCIc and MCInc, is the most challenging task. The proposed approach is novel in the sense that it is less computationally expensive by using only a single region of brain (hippocampus) from a single slice of MRI, yet producing better accuracy. The comparison of our proposed approach with a number of existing approaches is presented in Table III.

IV. CONCLUSION

Timely diagnosis of Alzheimer's disease is one of the challenging tasks in medical field. Early detection may prevent severe dementia or at least slow down its progression. Now a days, deep learning and in particular CNN, has shown state of the art performance in many computer vision fields including medical sciences. In this research, a CNN framework is proposed for improved

prediction of MCI conversion to AD. Two experiments are performed on the MRI dataset obtained from ADNI. In the first experiment, entire 3D MRI images are pre-processed for brain surface extraction, skull stripping, non-uniformity correction and partial volume tissue classification. The segmentation of pre-processed images results in volumes of GM, WM and CSF as features, which are used in combination with neuropsychological assessment scores i.e. MMSE and CDR as well as genetic data i.e. APOE ε4 to train and test a number of machine learning algorithms. It is observed that handcrafted volumetric image features alone outperform other features.

In the second experiment, hippocampus is extracted from MRI images using ROI mask mapping technique. The CNN is trained on hippocampus segmented images. The outcomes of this experiment are quite promising as compared to the outcomes obtained through volumetric image feature set because the hidden layers of CNN lead the classifier for making better decisions. We achieved the prediction accuracy of 94%, which is better than state of the art techniques. From these results, it is concluded that hippocampus extracted images, which are smaller in size and computationally less expensive, provide better outcomes than the volumetric features from the entire MRI. The experimental results indicate that the CNN approach is an effective technique for the prediction of MCI conversion and early detection of AD. The proposed technique may be beneficial in the domain of medical sciences for assisting medical practitioners in making timely and accurate prognosis.

TABLE III. COMPARISON OF PROPOSED APPROACH WITH EXISTING APPROACHES, N/A: NOT AVAILABLE

Approach	Image Modality	Classifier	Classification Group	ACC (%)	SEN (%)	SPE (%)
[8]	MRI	SVM	AD vs NC MCIc vs MCInc	93.01 75.00	89.13 76.92	96.80 73.23
[15]	MRI	LDA	MCIc vs MCInc	73.95	74.14	73.77
[21]	MRI	SVM	AD vs NC	92.48	91.07	93.89
[39]	MRI	SVM	AD vs NC AD vs MCI MCI vs NC	98.83 91.66 90.83	100 92 88	96 89 91
[40]	MRI	3D CNN	AD vs NC	87.15	86.36	85.93
[41]	MRI	Spectral CNN	AD vs NC AD vs MCI MCI vs NC	91.07 87.72 85.45	88.24 84.38 92.86	95.45 92.00 77.78
[42]	MRI , PET	Stacked autoencoders	AD vs NC MCI vs NC AD vs MCI vs NC	91.40 82.10 53.79	92.32 60.00 52.14	90.42 92.32 86.98
[44]	MRI	SVM	MCIc vs MCInc	73.91	70.5	77.1
[45]	MRI, PET	Multitask Deep Learning framework	AD vs MCI MCIc vs MCInc	70.1 57.4	n/a	n/a
[46]	MRI	CNN	MCIc vs MCInc	79.9	84	74.8
[47]	MRI, PET	Deep Belief Networks, SVM	AD vs NC AD vs MCInc MCInc vs NC MCIc vs NC MCIc vs MCInc	90 84 80 83 78	86 79 60 67 61	94 89 90 95 88
[48]	MRI	Sparse regression models and 2D CNN	AD vs NC MCIc vs MCInc	91.02 74.82	92.72 70.93	89.94 78.82
[52]	MRI	3D CNN (ResNet, VGGNet)	AD vs NC	88	n/a	n/a
[53]	FDG PET	CNN	AD vs NC MCIc vs MCInc	96.00 84.2	93.5 81.00	97.8 87.00
Proposed Approach	MRI	CNN	MCIc vs MCInc	94	92	96

REFERENCES

[1] Alzheimer's Association, "2018 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 14, pp. 367-429, 2018. doi:10.1016/j.jalz.2018.02.001

[2] R. A. Sperling, P. S. Aisen, L. A. Beckett, D. A. Bennett, S. Craft, et al., "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimer's & Dementia, vol. 7, pp. 280-292, 2011. doi:10.1016/j.jalz.2011.03.003

[3] C. Cabral, P. M. Morgado, D. C. Costa, M. Silveira, and ADNI, "Predicting conversion from MCI to AD with FDG-PET brain images at different prodromal stages," Computers in Biology and Medicine, vol. 58, pp. 101-109, 2015. doi:10.1016/j.compbiomed.2015.01.003

[4] S. Farhan, M. A. Fahiem, and H. Tauseef, "An Ensemble-of-Classifiers Based Approach for Early Diagnosis of Alzheimer's Disease: Classification Using Structural Features of Brain Images," Computational and Mathematical Methods in Medicine, vol. 2014, pp. 11, September 2014. doi:10.1155/2014/862307

[5] A. Khazaee, A. Ebrahimzadeh, A. Babajani-Feremi, and ADNI, "Classification of patients with MCI and AD from healthy controls using directed graph measures of resting-state fMRI," Behavioural Brain Research, vol. 322, pp. 339-350, 2017. doi:10.1016/j.bbr.2016.06.043

[6] J. Ding and Q. Huang, "Prediction of MCI to AD conversion using Laplace Eigenmaps learned from FDG and MRI images of AD patients and healthy controls," 2nd International Conference on Image, Vision and Computing (ICIVC), 2017, pp. 660-664.

[7] L. D. Declercq, R. Vandenberghe, K. V. Laere, A. Verbruggen, and G. Bormans, "Drug development in Alzheimer's disease: the contribution of PET and SPECT," Frontiers in Pharmacology, vol. 7, pp. 88, 2016. doi:10.3389/fphar.2016.00088

[8] I. Beheshti, H. Demirel, H. Matsuda, and ADNI, "Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm," Computers in Biology and Medicine, pp. 109-119, 2017. doi:10.1016/j.compbiomed.2017.02.011

[9] J. E. Arco, J. Ramirez, C. G. Puntonet, J. M. Gorriz, and M. Ruz, "Improving short-term prediction from MCI to AD by applying searchlight analysis," IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016, pp. 10-13. doi:10.1109/ISBI.2016.7493199

[10] F. Previtali, P. Bertolazzi, G. Felici, and E. Weitschek, "A novel method and software for automatically classifying Alzheimer's disease patients by magnetic resonance imaging analysis," Computer Methods and Programs in Biomedicine, vol. 143, pp. 89-95, 2017. doi:10.1016/j.cmpb.2017.03.006

[11] T. Pereira, A. Mendon, F. Ferreira, S. Madeira, and M. Guerreiro, "Towards a reliable prediction of conversion from mild cognitive impairment to Alzheimer's disease: stepwise learning using time windows," in Medical Informatics and Healthcare, 2017, pp. 19-26.

[12] S.-H. Wang, Y. Zhang, Y.-J. Li, W.-J. Jia, F.-Y. Liu, et al., "Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization," Multimedia Tools and Applications, vol. 77, pp. 10393-10417, 2018. doi:10.1007/s11042-016-4222-4

[13] F. Malik, S. Farhan, and M. A. Fahiem, "An Ensemble of Classifiers based Approach for Prediction of Alzheimer's Disease using fMRI Images based on Fusion of Volumetric, Textural and Hemodynamic Features," Advances in Electrical and Computer Engineering, vol. 18, pp. 61-71, 2018. doi:10.4316/AECE.2018.01008

[14] A. Ayub, S. Farhan, M. A. Fahiem, and H. Tauseef, "A novel approach for the prediction of conversion from mild cognitive impairment to Alzheimer's disease using MRI images," Advances in Electrical and Computer Engineering, vol. 17, pp. 113-123, 2017. doi:10.4316/AECE.2017.02015

[15] J. E. Arco, J. Ramirez, J. M. Gorriz, C. G. Puntonet, and M. Ruz, "Short-term prediction of MCI to AD conversion based on longitudinal MRI analysis and neuropsychological tests," in Innovation in Medicine and Healthcare 2015, ed: Springer, 2016, pp. 385-394. doi:10.1007/978-3-319-23024-5_35

[16] C. Misra, Y. Fan, and C. Davatzikos, "Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI," Neuroimage, vol. 44, pp. 1415-1422, 2009. doi:10.1016/j.neuroimage.2008.10.031

[17] A. M. Mubeen, A. Asaei, A. H. Bachman, J. J. Sidtis, B. A. Ardekani, et al., "A six-month longitudinal evaluation significantly improves accuracy of predicting incipient Alzheimer's disease in mild cognitive impairment," Journal of Neuroradiology, vol. 44, pp. 381-387, 2017. doi:10.1016/j.neurad.2017.05.008

[18] I. Beheshti, H. Demirel, and ADNI, "Probability distribution function-based classification of structural MRI for the detection of Alzheimer's disease," Computers in Biology and Medicine, vol. 64, pp. 208-216, 2015. doi:10.1016/j.compbiomed.2015.07.006

[19] S. Tangaro, A. Fanizzi, N. Amoroso, R. Bellotti, and ADNI, "A fuzzy-based system reveals Alzheimer's disease onset in subjects with Mild Cognitive Impairment," Physica Medica, vol. 38, pp. 36-44, 2017. doi:10.1016/j.ejmp.2017.04.027

[20] S. Minhas, A. Khanum, F. Riaz, S. A. Khan, and A. Alvi, "Trajectory based predictive modeling of conversion from mild cognitive impairment to Alzheimer's disease," in 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), 2017, pp. 385-388. doi:10.1109/BHI.2017.7897286

[21] I. Beheshti, H. Demirel, F. Farokhian, C. Yang, H. Matsuda, et al., "Structural MRI-based detection of Alzheimer's disease using feature ranking and classification error," Computer Methods and Programs in Biomedicine, vol. 137, pp. 177-193, 2016. doi:10.1016/j.cmpb.2016.09.019

[22] K. R. Chapman, H. Bing-Canar, M. L. Alosco, E. G. Steinberg, B. Martin, et al., "Mini Mental State Examination and Logical Memory scores for entry into Alzheimer's disease trials," Alzheimer's research & therapy, vol. 8, pp. 9, 2016. doi:10.1186/s13195-016-0176-z

[23] M. J. Baek, K. Kim, Y. H. Park, and S. Kim, "The validity and reliability of the mini-mental state examination-2 for detecting mild cognitive impairment and Alzheimer's disease in a Korean population," PloS One, vol. 11, pp. e0163792, 2016. doi:10.1371/journal.pone.0163792

[24] L. Zheng, X. Kong, Y. Cui, Y. Wei, J. Zhang, et al., "Conversion from MCI to AD in patients with the APOEϵ4 genotype: prediction by plasma HCY and serum BDNF," Neuroscience Letters, vol. 626, pp. 19-24, 2016. doi:10.1016/j.neulet.2016.05.018

[25] K. K. Sarma, "Neural network based feature extraction for assamese character and numeral recognition," International Journal of Artificial Intelligence, vol. 2, pp. 37-56, 2009

[26] A. Albu, R.-E. Precup and T.-A. Teban, Results and Challenges of Artificial Neural Networks Used for Decision-Making in Medical Applications, Facta Universitatis, Series: Mechanical Engineering (University of Nis), vol. 17, no 4, pp. 285-308, 2019

[27] C. Pozna, R.-E. Precup, J. K. Tar, I. Skrjanc, and S. Preitl, "New results in modelling derived from Bayesian filtering," Knowledge-Based Systems, vol. 23, pp. 182-194, 2010. doi:10.1016/j.knosys.2009.11.015

[28] R. A. Gil, Z. C. Johanyak, and T. Kovacs, "Surrogate model based optimization of traffic lights cycles and green period ratios using microscopic simulation and fuzzy rule interpolation," Int. J. Artif. Intell, vol. 16, pp. 20-40, 2018

[29] D. C. Newitt, E. T. Tan, L. J. Wilmes, T. L. Chenevert, J. Kornak, et al., "Gradient nonlinearity correction to improve apparent diffusion coefficient accuracy and standardization in the american college of radiology imaging network 6698 breast cancer trial," Journal of Magnetic Resonance Imaging, vol. 42, pp. 908-919, 2015. doi:10.1002/jmri.24883

[30] J. Jovicich, S. Czanner, D. Greve, E. Haley, A. van Der Kouwe, et al., "Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data," Neuroimage, vol. 30, pp. 436-443, 2006. doi:10.1016/j.neuroimage.2005.09.046

[31] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," IEEE transactions on medical imaging, vol. 17, pp. 87-97, 1998. doi:10.1109/42.668698

[32] M. J. Müller, D. Greverus, P. R. Dellani, C. Weibrich, P. R. Wille, et al., "Functional implications of hippocampal volume and diffusivity in mild cognitive impairment," Neuroimage, vol. 28, pp. 1033-1042, 2005. doi:10.1016/j.neuroimage.2005.06.029

[33] P. M. Thompson, K. M. Hayashi, G. I. de Zubicaray, A. L. Janke, S. E. Rose, et al., "Mapping hippocampal and ventricular change in Alzheimer disease," Neuroimage, vol. 22, pp. 1754-1766, 2004. doi:10.1016/j.neuroimage.2004.03.040

[34] J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, C. Avedissian, et al., "Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls," Neuroimage, vol. 45, pp. S3-S15, 2009. doi:10.1016/j.neuroimage.2008.10.043

[35] V. Dill, P. C. Klein, A. R. Franco, and M. S. Pinho, "Atlas selection for hippocampus segmentation: Relevance evaluation of three meta-information parameters," Computers in Biology and Medicine, vol. 95, pp. 90-98, 2018. doi:10.1016/j.compbiomed.2018.02.005

[36] S. Sandor and R. Leahy, "Surface-based labeling of cortical anatomy using a deformable atlas," Medical Imaging, IEEE Transactions on, vol. 16, pp. 41-54, 1997. doi:10.1109/42.552054

[37] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," NeuroImage, vol. 13, pp. 856-876, 2001. doi:10.1006/nimg.2000.0730

[38] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, pp. 2278-2324, 1998. doi:10.1109/5.726791

[39] M. A. Iftikhar and A. Idris, "An ensemble classification approach for automated diagnosis of Alzheimer's disease and mild cognitive impairment," in 2016 International Conference on Open Source Systems & Technologies (ICOSST), 2016, pp. 78-83. doi:10.1109/ICOSST.2016.7838581

[40] D. Cheng, M. Liu, J. Fu, and Y. Wang, "Classification of MR brain images by combination of multi-CNNs for AD diagnosis," in *Ninth International Conference on Digital Image Processing (ICDIP 2017)*, 2017, p. 1042042. doi:10.1117/12.2281808

[41] X. Li, Y. Li, and X. Li, "Predicting clinical outcomes of alzheimer's disease from complex brain networks," in International Conference on Advanced Data Mining and Applications, 2017, pp. 519-525. doi:10.1007/978-3-319-69179-4_36

[42] S. Liu, S. Liu, W. Cai, H. Che, S. Pujol, et al., "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease," IEEE Transactions on Biomedical Engineering, vol. 62, pp. 1132-1140, 2014. doi:10.1109/TBME.2014.2372011

[43] J. Lotjonen, R. Wolz, J. Koikkalainen, V. Julkunen, L. Thurfjell, et al., "Fast and robust extraction of hippocampus from MR images for diagnostics of Alzheimer's disease," Neuroimage, vol. 56, pp. 185-196, 2011. doi:10.1016/j.neuroimage.2011.01.062

[44] R. Wei, C. Li, N. Fogelson, and L. Li, "Prediction of conversion from mild cognitive impairment to Alzheimer's Disease using MRI and structural network features," Frontiers in Aging Neuroscience, vol. 8, pp. 76, 2016. doi:10.3389/fnagi.2016.00076

[45] F. Li, L. Tran, K.-H. Thung, S. Ji, D. Shen, et al., "A robust deep model for improved classification of AD/MCI patients," IEEE Journal of Biomedical and Health Informatics, vol. 19, pp. 1610-1616, 2015.doi:10.1109/JBHI.2015.2429556

[46] W. Lin, T. Tong, Q. Gao, D. Guo, X. Du, et al., "Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment," Frontiers in neuroscience, vol. 12, p. 777, 2018. doi:10.3389/fnins.2018.00777

[47] A. Ortiz, J. Munilla, J. M. Gorriz, and J. Ramirez, "Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease," International journal of neural systems, vol. 26, p. 1650025, 2016. doi:10.1142/S0129065716500258

[48] H.-I. Suk, S.-W. Lee, D. Shen, and A. s. D. N. Initiative, "Deep ensemble learning of sparse regression models for brain disease diagnosis," Medical image analysis, vol. 37, pp. 101-113, 2017. doi:10.1016/j.media.2017.01.008

[49] J. Lindsay, D. Laurin, R. Verreault, R. Hebert, B. Helliwell, et al., "Risk factors for Alzheimer's disease: a prospective analysis from the Canadian Study of Health and Aging," American Journal of Epidemiology, vol. 156, pp. 445-453, 2002. doi:10.1093/aje/kwf074

[50] E. E. Bron, M. Smits, W. M. Van Der Flier, H. Vrenken, F. Barkhof, et al., "Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge," NeuroImage, vol. 111, pp. 562-579, 2015. doi:10.1016/j.neuroimage.2015.01.048

[51] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, J. Tohka, et al., "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," Neuroimage, vol. 104, pp. 398-412, 2015. ,doi:10.1016/j.neuroimage.2014.10.002

[52] S. Korolev, A. Safiullin, M. Belyaev, and Y. Dodonova, "Residual and plain convolutional neural networks for 3D brain MRI classification," in 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), 2017, pp. 835-838. doi:10.1109/ISBI.2017.7950647

[53] H. Choi, K. H. Jin, and ADNI, "Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging," Behavioural Brain Research, vol. 344, pp. 103-109, 2018. doi:10.1016/j.bbr.2018.02.017