

A Novel Artificial Intelligence (AI) Ensembled Method to Identify Alzheimer's Disease Progression

Md Mehedi Hasan¹, Senjuti Rahman¹, Harshit Parmar², and Suman K. Chowdhury^{1*}

¹Department of Industrial, Manufacturing and Systems Engineering, Texas Tech University, Lubbock, Texas, USA

²Neuroimaging Institute, Texas Tech University, Lubbock, Texas, USA

*All correspondence should be addressed to—

Dr. Suman K. Chowdhury, Assistant Professor, Department of Industrial, Manufacturing and Systems Engineering, Texas Tech University

905 Canton Ave, Lubbock 79409-3061, Texas, USA.

E-mail: suman.chowdhury@ttu.edu

Acknowledgments

This work was primarily supported by the National Science Foundation (2239110).

Abstract

Purpose: This study aimed to develop an AI-ensembled network to identify five stages of Alzheimer's disease (AD) progression—normal cognition (NC), early mild cognitive impairment (EMCI), mild cognitive impairment (MCI), late MCI (LMCI), and AD—using brain features and regions.

Methods: T1-weighted MRI data of 1000 participants were retrieved from Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The dataset was preprocessed, and brain volume was parcellated into 170 regions of interest (ROIs) using automated anatomical labeling-3 (AAL-3) atlas. In each ROI, volumes of gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) was estimated to generate 13 brain features. Data outliers were identified, and the dataset was divided into training, testing, and validation in a 70:15:15 ratio. An AI-ensembled network comprising a random forest (RF) model and 3D ResNet-18 was tested using four combinations of features and regions. Hyperparameters were tuned via five-fold cross-validation.

Results: The RF model identified GM-to-WM, WM-to-CSF, and GM-to-CSF, volumetric ratios as top predictors of AD progression. Thalamus, amygdala, and hippocampus brain regions were consistently affected across all stages. The ResNet-18 network performed best with combination-1 (RF selected three features and 60 regions) input, achieving 66% F-1 score, 76% sensitivity, and 93.5% specificity. Five-fold cross-validation confirmed 60.02% accuracy for combination-1.

Conclusion: The proposed AI-ensembled network, first-of-its-kind, can effectively identify the AD continuum, particularly the EMCI stage. Its implication in clinical settings can assist in obtaining disease-modifying targeted therapeutic interventions, extending patient's life expectancy. Further enhancements could be achieved with expanded training data and transfer learning techniques.

Keywords: Alzheimer's disease, random forest, brain features, brain regions, progression pattern.

Statements and Declarations

No funding was received for conducting this study.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder resulting in brain cell death and brain tissue shrinkage [1]. People with AD can progressively experience various adverse effects, including breathing difficulties, heart failure, and, ultimately, death [2]. It is also the most prevalent form of dementia, which collectively indicates various forms of cognitive impairments (CI), such as memory loss, communication difficulty, and decline in thinking and reasoning skills [3]. According to the 2020 World Health Organization (WHO) report, approximately 55.2 million individuals have dementia and AD worldwide, and this number is predicted almost to double every 20 years, reaching 78 million in 2030 and 139 million by 2050 [4]. In the United States alone, about 6.9 million people aged 65 or older were living with AD in 2024. Approximately \$360 billion was spent on the long-term medical care of Alzheimer's patients in 2024, demonstrating the significant economic impact that AD can have on the society [5].

The pathophysiological process of AD begins years before its clinical symptoms appear. The mild cognitive impairment (MCI) condition is considered as the risk state [6], and only 6–15% of patients with MCI progress to AD [7]. However, not all MCI cases progress to AD; some remain stable or even improve over time with appropriate treatments [8]. In research and clinical settings [9], the progression of MCI is divided into two stages: early MCI (EMCI), and late MCI (LMCI). This classification distinguishes between individuals at the initial stage of clinical symptom (EMCI) and those at a later stage (LMCI), progressing directly towards the development of AD [10]. As there is a lack of effective cure for AD, early and accurate detection is crucial [11]. Identifying the early stages of Alzheimer's Disease (AD), such as EMCI and MCI through imaging modalities would facilitate more effective treatment strategies for clinicians [12].

Previous studies have reported various MCI-associated brain degenerations, such as shrinking of the hippocampus, significant atrophy in the medial temporal lobe, and loss of gray matter volume in the frontal and parietal areas [13-15] using different imaging technologies like computed tomography (CT), PET (positron emission tomography), and MRI [16, 17]. MRI is the most commonly employed imaging technique for predicting stages of AD [18]. In order to reveal hidden information and eliminate noise inherent in complex MRI sequences, the extraction of important features becomes imperative [19]. Numerous studies have utilized critical input features, such as the measurement of brain volume, cortical thickness, brain surface area, texture, and white matter lesion (WML) volume [20, 21] extracted from MRI data to optimize performance [22-25]. Putcha et al. [26] introduced a new approach to measure structural brain changes by using the gray matter to white matter ratio, which emerged as a promising feature for detecting progression from earlier stages to the most prevalent stage (i.e., AD). Dubois et al. [27] discussed the development of *vivo* features that shifted the diagnosis of AD from the later dementia stages of the disease towards the earlier stages. They showed structural MRI reveals atrophy, or the loss of volume of gray matter, which denotes neurodegeneration in individuals with AD. Khagi et al. [28] revealed that a decrease in gray matter (GM) and brain volume, as well as a minor increase in cerebrospinal fluid (CSF), can be an effective feature for the detection of AD. Most of these features were derived by considering the whole brain instead of its various parcels [10, 29], which implied that there remains the need for a method capable of detecting AD-induced localized brain lesions using a minimal amount of brain features and brain ROIs. ROI-based techniques frequently utilize the

use of 3D volume and shape properties of specific brain regions [29, 30]. An atlas is a map of brain's anatomy and structure that can aid in partitioning the brain into anatomically distinct regions or ROIs. In recent research papers, automated anatomical labelling-3 (AAL3) atlas was utilized to divide the brain into 170 ROIs [31, 32].

The field of artificial intelligence, particularly machine learning (ML) and deep learning (DL), is continually evolving and playing an important role in analyzing MRI data to classify Alzheimer's Disease (AD) stages [33-35]. Among the various ML techniques, random forest has proven to be one of the most effective models for identifying minimal brain features and regions of interest (ROIs). This is due to its resilience to noise and its ability to handle complex, multimodal data [36, 37]. Deep learning methods have been utilized to predict the progression of AD using features derived from structural MRI data [30, 35, 38, 39]. The Convolutional Neural Network (CNN), a widely used deep learning network, has been employed for the classification and prediction of AD [40-43]. In CNN, there was limited interaction between the feature maps from different layers. The advent of Residual Network (ResNet) marked a significant milestone in the evolution of deep learning methods. It introduced a residual unit to connect the current layer to the previous one, known as skip-connection, addressing the degradation problem [44]. This architecture allowed ResNet to become deeper, and achieve better performance than conventional CNN [45].

3D CNNs have been used in predicting Alzheimer's disease mostly as they extract intricate features from 3D data and enhance model comprehensibility [44, 45]. Long et al. [32] proposed a 3D densely connected convolutional neural network (CNN) with a connection-wise attention mechanism to learn the multi-level features of brain MR images for AD classification. They used MRI of 968 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database to discriminate (1) AD versus healthy subjects, (2) MCI converters versus healthy subjects, and (3) MCI converters versus non-converters. Their proposed method achieved 97.35% accuracy for distinguishing AD patients from healthy control, 87.82% for MCI converters against healthy control, and 78.79% for MCI converters against non-converters. In the training process, they adjusted the parameters of the deep CNN model, including the number of layers, the size, and the number of kernels in each layer; nevertheless, network convergence was still challenging. Furthermore, the study didn't consider the multi-classification among the categories and considered the whole brain instead of its various parcels. Folego et al. [46] developed an end-to-end deep 3D CNN (namely, LeNet-5, VGG, GoogLeNet, and ResNet-18) for the multiclass AD biomarker identification task, using the whole image volume as input. It was composed of three main steps: brain extraction and normalization, 3D CNN processing, and domain adaptation to classify subjects into AD, MCI, or NC groups. They utilized the ADNI dataset and reached 52.3% accuracy in the testing set. They found the degradation problem, where traditional models similar to VGG stopped improving performance after a certain number of layers and even started getting worse afterward. To overcome this problem, they proposed the residual function, which was the basic building block of a Residual Network (ResNet). In their approach, the training was stopped after 50 epochs without further improvement in average TPF (true positive rate) over the validation set and did not consider disease specific features, such as hippocampal volume, demographic information. Khagi et al. [47] used SPM tool to get a 3D image of gray matter and fed into 3D CNN. Once trained, an untested MRI can

be passed through CNN to determine whether it is a healthy control (HC), or Mild Cognitive Impairment (MCI) due to AD (mAD) or AD dementia (ADD). They collected a dataset from the Gwangju Alzheimer's Disease and Related Dementia (GARD) center. The dataset consisted of 42 Alzheimer's disease dementia (ADD), 42 HC, and 39 MCI due to AD (mAD). They achieved an accuracy of around 40% for mutual information. The detailed feature extraction and analysis were still under study, and the overall classification result was not very high, which was due to the use of limited training materials. Chen et al. [48] introduced an ensemble deep learning model for AD classification, which incorporated Soft-NMS (Non-Maximum Suppression) into the Faster R-CNN architecture to classify the three categories, i.e. 115 AD patients, 106 subjects with MCI, and 185 NC subjects. Using a validation accuracy of about 50% and fine-tuning the ADNI dataset, they were able to reach an accuracy of 84.37% for the 3-way classification. They did not consider the five-way classification, and recommended using a larger dataset, key features and regions as identification of early biomarkers for future work. We considered their recommendations for the proposed methodology by using important features and regions. Only a few studies [49, 50] have addressed the classification of the four stages of AD. However, to our knowledge, no study has explored the utilization of a five-way categorization scheme, coupled with the implementation of minimal features and regions using an ensembled 3D deep learning network, for effectively categorizing the cognitive impairment conditions associated with Alzheimer's disease.

This study aimed to develop an AI-ensembled network using important brain features and brain regions in order to identify five stages of AD progression—NC, EMCI, MCI, LMCI, and AD—and link their associated brain lesions effectively. The efficacy of our model was validated using a five-fold cross-validation approach, statistical analysis as well as existing literature data from similar works.

Materials and Methods

The methodological approach of this work was divided into four major steps, comprising of 3D MRI data retrieved from ADNI database (available at <http://adni.loni.usc.edu>), pre-processing, feature extraction, and development of an ensembled AI network to predict CI condition using a minimal set of brain features. The entire process of the suggested methodology was thoroughly introduced in this portion of the paper.

Data Acquisition

In this study, We retrieved T1-weighted MRI data of 1000 participants from the widely-recognized ADNI database, including 200 AD patients, 200 LMCI subjects (progressing within 18 months), 200 MCI subjects (not progressing within 5 years), 200 EMCI subjects (showing early signs of AD), and 200 cognitively normal controls (NC) maintaining stability over 3 years [51]. Demographic and cognitive test data for each group, such as gender, age, clinical dementia rating (CDR), and mini-mental state examination (MMSE) scores (decreasing with Alzheimer's progression)—were presented in Table 1. The CDR score, indicates dementia presence and severity, ranged from 0 (no dementia) to 3 (severe dementia) [52] [53].

Table 1 Demographic and cognitive examination scores of the retrieved ADNI data.¹

Category	#Subjects	Gender		Age average $\pm$ SD	MMSE average $\pm$ SD	CDR average $\pm$ SD
		F	M			
AD	200	100	100	74.40 $\pm$ 5.78	20.16 $\pm$ 6.60	2.98 $\pm$ 0.60
LMCI	200	100	100	73.78 $\pm$ 7.29	25.82 $\pm$ 2.98	1.97 $\pm$ 0.61
MCI	200	100	100	69.65 $\pm$ 7.18	27.87 $\pm$ 4.63	0.99 $\pm$ 0.22
EMCI	200	100	100	72.078 $\pm$ 8.05	28.86 $\pm$ 1.99	0.435 $\pm$ 0.20
NC	200	100	100	76.735 $\pm$ 3.40	29.93 $\pm$ 1.79	0.06 $\pm$ 0.19

Preprocessing

Data preprocessing steps subsequently consisted of brain extraction, bias field correction, noise reduction, image registration, normalization, segmentation, and parcellation of 170 regions using an AAL-3 brain atlas. Statistical parametric mapping (SPM12) toolbox was utilized in preprocessing data due to its widespread acceptance in recent studies [54, 55]. Brain extraction distinguished brain voxels structures like the brainstem, cerebrospinal fluid (CSF), gray matter (GM), white matter (WM), and subcortical regions, from the non-brain voxels, including scalp, eyes, bones, and other tissues [56]. Bias field correction adjusted image contrast affected by magnetic field inhomogeneity, crucial for analysis at magnetic field strengths of 1.5 T, 3 T, or higher [57]. Additionally, noise reduction techniques were employed to mitigate Rician noise, improving overall performance of the proposed methodology. Image registration, in inter and intra-patient forms, was used to align images to common anatomical spaces [58], resulting in each MRI having dimensions of 79x95x79 mm³. The images were then uniformly scaled through normalization and segmentation, generating the volumes of gray matter, white matter, and CSF based on input modalities, as shown in Fig. 1. The automated anatomical labeling (AAL) atlas, specifically the widely used AAL3 version [59], was applied to create 170 parcellated brain region masks. Outliers were identified and reduced using the interquartile range (IQR) method [60], leading to the exclusion of approximately 40 subjects from the dataset.

Feature Extraction

Using the AAL3 atlas, we created region masks and computed the volumes of gray matter (GMV), white matter (WMV), and CSF (CSFV) in each region, as shown in Equation 1.

$$Volume_{GM/WM/CSF} = \sum_{i=1}^N Voxel_i \quad (1)$$

Where, $Voxel_i$ denoted the volume of each voxel within the GM/WM/CSF region and N was the total number of ROIs.

The first three features are the ratios of GM to CSF (GCR), WM to CSF (WCR), and GM to WM (GWR), with GWR being a novel feature [26]. These ratios were calculated using the equations in Table 2.

¹ NC = normal cognition, EMCI = early mild cognitive impairment, MCI = mild cognitive impairment, LMCI = late mild cognitive impairment, AD = Alzheimer's disease, F = female, M = male, SD = standard deviation, MMSE = mini-mental state examination, CDR = clinical dementia rating

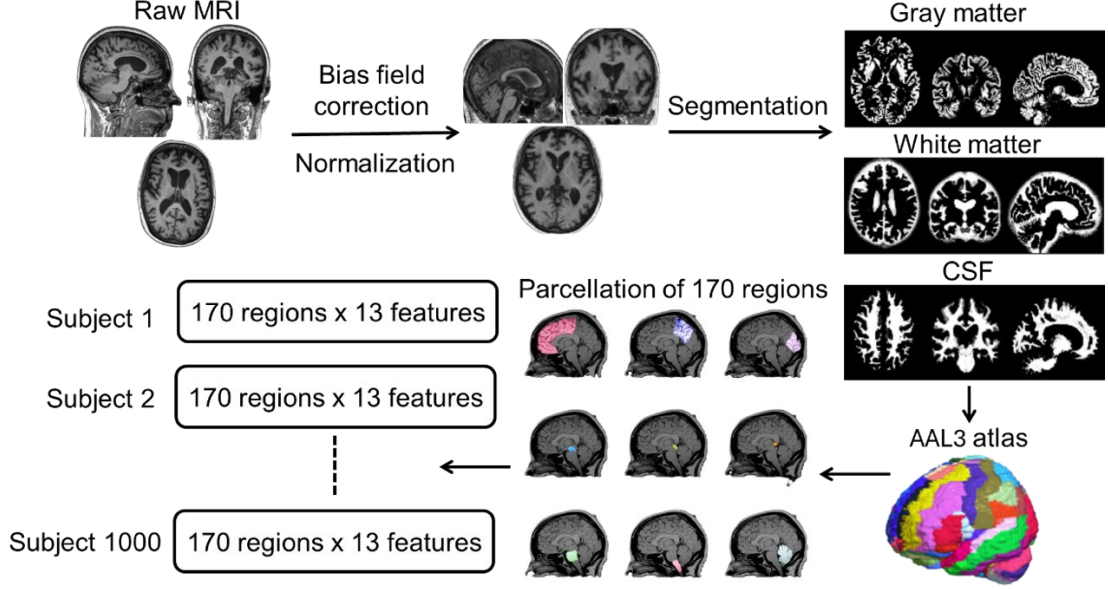


Fig. 1 Pre-processing steps of MRI with respect to automated anatomical labelling atlas-3 (AAL-3). Here, MRI = magnetic resonance imaging, CSF = cerebrospinal fluid, AAL-3 = automated anatomical labeling atlas-3.

Utilizing three separate Montreal Neurological Institute (MNI) template as the ground truth [14] for gray matter, white matter, and CSF, resulted in three additional features. We evaluated the overlap between the volume of each ROIs (A) and respected ground truth volume (B) to assess the Dice similarity coefficient (DSC) — a statistical measure used to measure the similarity between two sets [61]— generating three features of GM, WM and CSF.

$$Dice - Coefficients = \frac{2|A \cap B|}{|A| + |B|} \quad (2)$$

Where A = GM/ WM/ CSF volume, B = respected ground truth of GM/ WM/ CSF volume.

We calculated the tenth feature by averaging the intensity values in each ROI. The final three features were a quantitative measure of the surface area. To derive the final three features, we applied erosion to the ROI volume, then subtracted the eroded volume from the original. The process of extracting these 13 features from the brain atlas of 170 regions was visually depicted in Fig. 1. After extracting features, we utilized the brain atlas to assign feature values to each region, forming a 4D matrix with dimensions $79 \times 95 \times 79 \times l$, where l denoted either all features or RF selected features.

Ensemble RF Network for Feature and Region Selection

To identify CI categories effectively, a feature selection algorithm was required for selecting optimal features. Studies showed that reducing features not only speeds up computation but also enhances classification performance [62, 63]. In this study, random forest, an ensemble machine learning technique was used for the selection of most important features and mostly affected regions.

The parameters for the random forest were chosen as outlined in Table 3 for both feature selection and identifying the most affected regions. The selected features and regions were presented in the results section.

Table 2 Names and equations of the features.²

Feature number	Name of the features	Equations for the features
Feature 1	Gray matter to CSF ratio (GCR)	$GCR = \frac{100 \times (GMV - CSFV)}{0.5(GMV + CSFV)}$
Feature 2	White matter to CSF ratio (WCR)	$WCR = \frac{100 \times (WMV - CSFV)}{0.5(WMV + CSFV)}$
Feature 3	Gray matter to white matter ratio (GWR)	$GWR = \frac{100 \times (GMV - WMV)}{0.5(GMV + WMV)}$
Feature 4	Gray matter to gray matter ground truth ratio (GGR)	$GGR = \frac{100 \times (GMV - GGMV)}{0.5(GMV + GGMV)}$
Feature 5	White matter to white matter ground truth ratio (WGR)	$WGR = \frac{100 \times (WMV - GWMV)}{0.5(WMV + GWMV)}$
Feature 6	CSF to CSF ground truth ratio (CGR)	$CGR = \frac{100 \times (CSFV - GCSFV)}{0.5(CSFV + GCSFV)}$
Feature 7	Dice coefficient of gray matter	$Dice - coefficient = \frac{2 GMV \cap GGMV }{ GMV + GGMV }$
Feature 8	Dice coefficient of white matter	$Dice - coefficient = \frac{2 WMV \cap GWMV }{ WMV + GWMV }$
Feature 9	Dice coefficient of CSF	$Dice - coefficient = \frac{2 CSFV \cap GCSFV }{ CSFV + GCSFV }$
Feature 10	Average intensity	$Average\ intensity = \frac{1}{N} \sum_{i=1}^N Intensity_i$
Feature 11	Surface area of gray matter	$Surface\ area = GMV - erosion\ of\ GMV$
Feature 12	Surface area of white matter	$Surface\ area = WMV - erosion\ of\ WMV$
Feature 13	Surface area of CSF	$Surface\ area = CSFV - erosion\ of\ CSFV$

Residual Neural Network Architecture

The evolution of CNNs into 3D CNNs [64] enhances the analysis by capturing spatial, temporal, and depth data especially valuable in medical imaging (CT, MRI, ultrasound). To address the vanishing gradient issue encountered in training deep convolutional networks, the ResNet (residual network) was introduced [65].

The 3D ResNet-18 architecture employed in this study was designed to process volumetric data shown in Fig. 2. The input layer accepted volumetric brain data, formatted as 3D tensors. The initial convolutional layer applied a 3D convolution with a kernel size of 7×7×7, a stride of 2×2×2, and padding of 3×3×3, producing a set of feature maps. Following this, a 3D batch normalization layer was applied to standardize the outputs from the convolutional layer. A ReLU activation function was introduced to enable non-linearity. Subsequent to the initial convolution and normalization steps, a max pooling layer with a kernel size of 3×3×3 and a stride of 2×2×2 was added to down sample the feature maps.

² GMV: volume of gray matter, WMV: volume of white matter, CSFV: volume of CSF, GGMV = ground truth of GMV, GWMV = ground truth of GWMV, GCSFV = ground truth of CSFV, N: total number of regions, Intensity_i: the intensity value of the i-th region

Table 3 Selected parameters for the random forest model used to select important brain features and regions in this study.

Parameter name	Description	Chosen value
n_estimators	The number of decision trees	100
max_features	The largest number of features to consider when branching	5
max_depth	The maximum depth of a single tree	25
min_samples_split	The minimum number of samples required to split an internal node	6
min_samples_leaf	The minimum number of samples required to be at a leaf node.	1

The core of the ResNet-18 architecture consisted of two types of residual blocks: Residual Block 1, and Residual Block 2. Residual Block 1 contained two convolutional layers but did not incorporate skip connections. Each convolutional layer within this block used a 3×3×3 kernel with padding to maintain spatial dimensions. Residual Block 2 also incorporated two convolutional layers, but, unlike Residual Block 1, it included skip connections. These connections were important in addressing the vanishing gradient problem and enabling the construction of deeper networks. Each convolutional layer in Residual Block 1 and 2 used a 3×3×3 kernel with padding, followed by batch normalization and ReLU activation.

After the series of residual blocks, the model employed an average pooling layer, which reduced each feature map to a single value. Following the average pooling layer, the architecture included three fully connected layers. These layers further processed the pooled features, allowing for more complex representations to be learned. The final fully connected layer mapped the processed features to the output classes, corresponding to the five impairment conditions (NC, EMCI, MCI, LMCI, and AD) in Alzheimer's disease. This final layer used a Softmax activation function to produce probabilities for each class, enabling the classification task.

Methodology of AD Progression Classification

The proposed methodology followed a thorough process of feature and region selection, coupled with deep learning classification shown in Fig. 3. The RF algorithm played a critical role in identifying the most relevant features and brain regions from the expansive initial datasets. Subsequently, the ResNet-18 model processed brain features and regions through four distinct combinations, each aimed at classifying subjects into the five stages of Alzheimer's disease progression: normal cognition, early mild cognitive impairment, mild cognitive impairment, late mild cognitive impairment, and Alzheimer's disease. This approach ensured that the classification was based on the most important data, effectively combining the strengths of both traditional machine learning and advanced deep learning techniques.

hyperparameters were selected at this stage shown in Table 4. For two or three-category classifications, we didn't repeat the procedure outlined in Fig. 4, as our primary focus centered on categorizing five categories. We only determined the classification accuracy for two or three categories to facilitate comparison with other models.

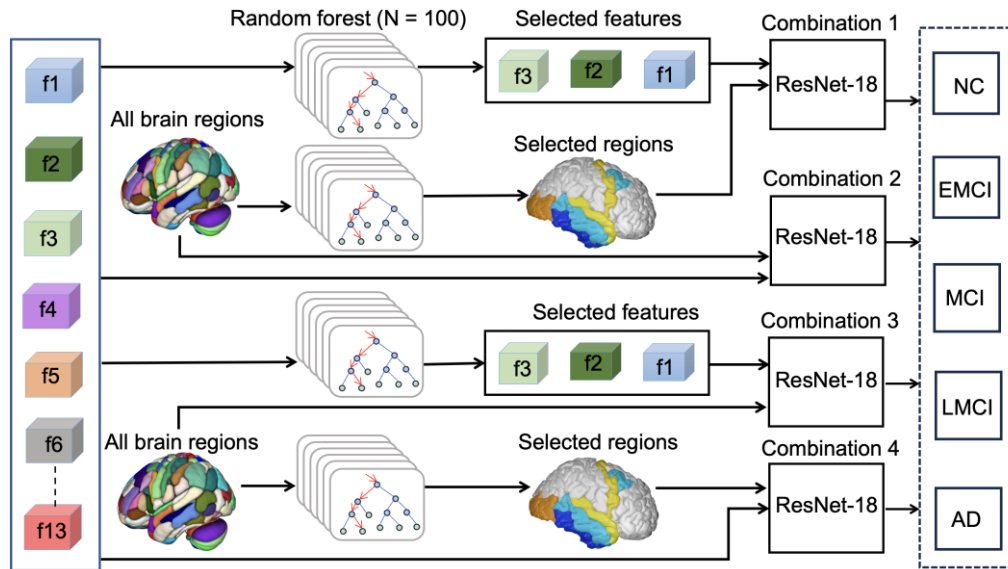


Fig. 3 An overview of the methodology to develop an AI-ensembled network to classify Alzheimer's disease progression. Here, NC = normal cognition, EMCI = early mild cognitive impairment, MCI = mild cognitive impairment, LMCI = late mild cognitive impairment, AD = Alzheimer's disease, f1 = gray matter to CSF ratio (GCR), f2 = white matter to CSF ratio (WCR), f3 = gray matter to white matter ratio (GWR), f4 = gray matter to gray matter ground truth ratio (GGR), f5 = White matter to white matter ground truth ratio (WGR), f6 = CSF to CSF ground truth ratio (CGR), f13 = Surface area of CSF, N = number of decision trees.

We trained our deep learning models on an Alienware Aurora R15 Windows 11 enterprise- 64 bit) machine with 13th Gen Intel® Core i9 3.00 GHz CPUs (32 CPUs), 65536MB of memory, and a 64GB NVIDIA GeForce RTX 4090 GPU. The computation time depends on the features and regions, the combination of more features and regions led to more converging time whereas the selected three features and 60 regions maintained the convergence within or less than 3 hours.

Table 4 Description of hyperparameters tuned using five-fold cross-validation approach among four input combinations.

Model	Three features-60 regions	13 features-60 regions	13 features-all regions	Three features-all regions
Training Optimization function	Adam	Adam	Adam	Adam
Mini-batch size	35	35	35	35
Maximum Epoch	50	50	50	50
Learning rate	5e-05	1e-05	1e-05	5e-05
Drop Factor	0.75	0.50	0.50	0.75
L2-Regularization	0.05	0.1	0.1	0.05

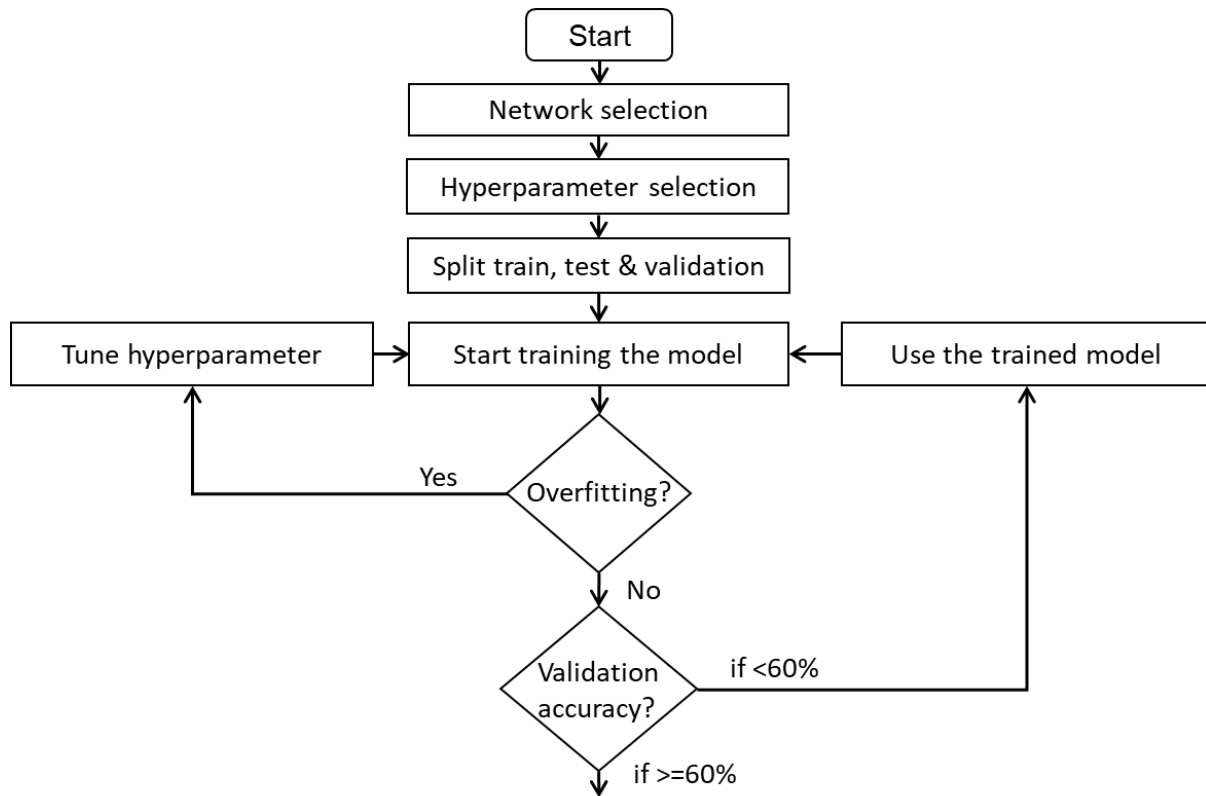


Fig. 4 The methodology of training, testing and validating the proposed model.

Validation

We followed a five-fold cross-validation strategy to evaluate the classification performance on the proposed model for selected features and brain lesions (combination-1). In the experiment, we split the data into five non-overlap folds, where one-fold was used as the testing data and the remaining four folds are used for training at each time. We repeated the whole process five times to avoid any possible bias caused by dataset partition. The final classification accuracy was reported by averaging the classification results from cross-validations. To validate the significance of the most important regions across five distinct categories, ANOVA tests were performed on an equal number of category instances within the most important three features. This methodological approach aimed to ascertain the pivotal role of these features in delineating the most crucial region.

Results

Experimental Analysis of the Important Features and Most Affected Regions

The RF model respectively identified GM-to-WM, WM-to-CSF, and GM-to-CSF, volumetric ratios as first, second, and third most important features with a threshold level of 1 illustrated in Fig. 5. Subsequently, the RF model was utilized to select the important brain regions among 170 parcellated ROIs which were mostly affected during the progression of AD. To reduce time and computational complexity, we selected the top 60 regions using the same threshold level of 1. Out of 60 RF-selected brain regions, ventral posterolateral, pulvinar lateral, ventral lateral, and superior temporal gyrus regions of brain thalamus, left and right amygdala regions, and all hippocampus regions were identified as the most affected brain regions, indicating that any clinical manifestation (cognitive changes) in these regions during EMCI and MCI stage could potentially be used as AD biomarkers. Fig. 6 illustrates a few most affected

regions of a patient classified under the MCI category, displayed in coronal, sagittal, axial, and render planes, with red indicating the highly affected areas, based on the selected features. The anatomical description of these affected regions was presented in Table 5.

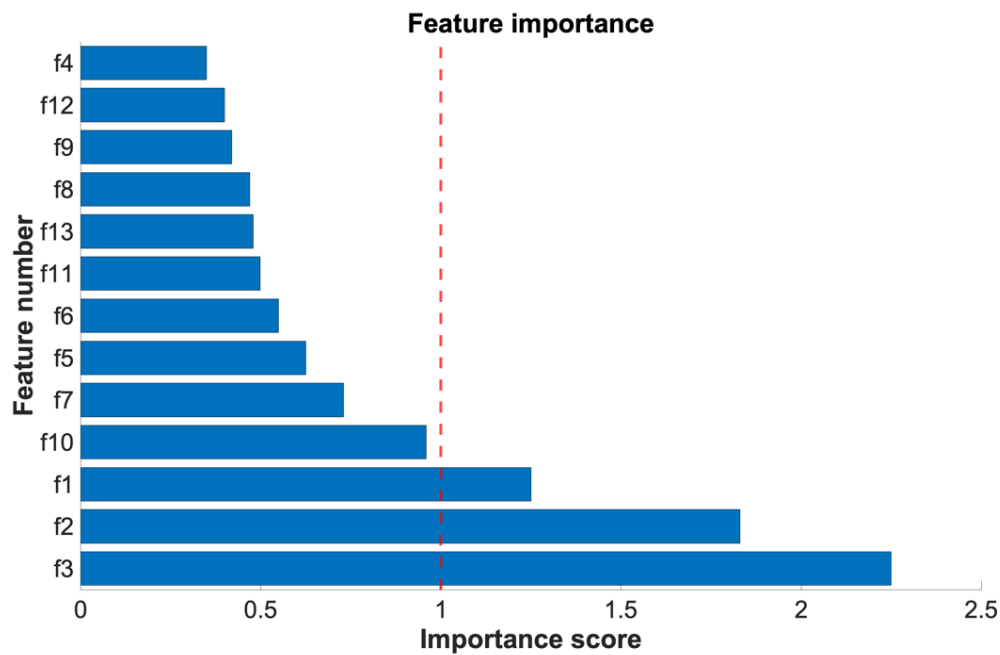


Fig. 5 Feature importance as determined by random forest model. Threshold level 1 identified the most important features: gray matter to white matter ratio (f3), white matter to CSF ratio (f2), gray matter to CSF ratio (f1).

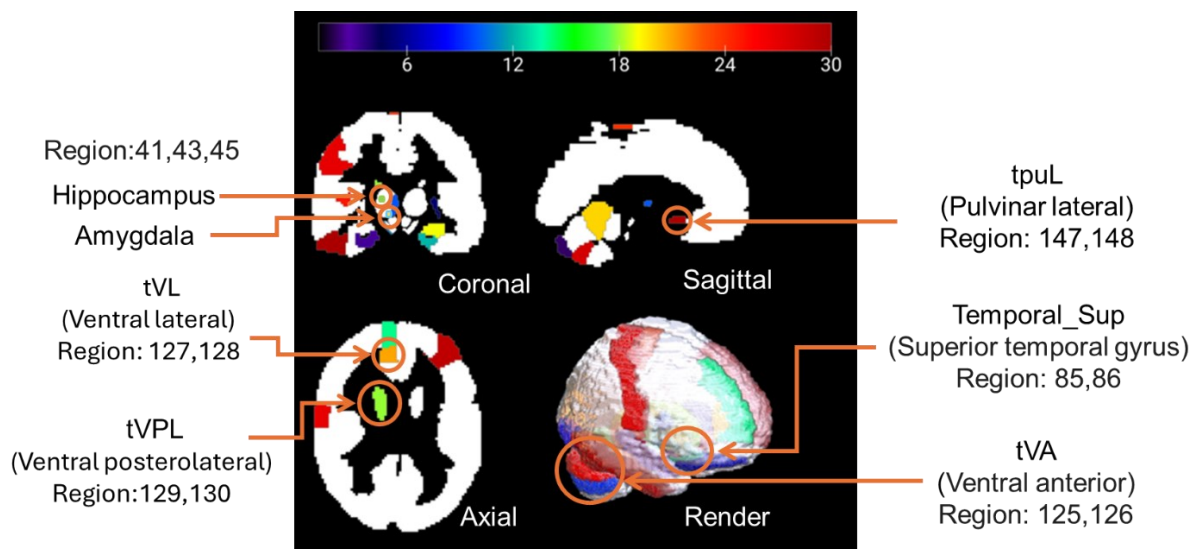


Fig. 6 Most affected brain regions selected by random forest for a subject of MCI Category.

Table 5 Anatomical description of the random forest selected most affected brain regions during AD progression [31].

Region	Anatomical description	Label
41,42	Hippocampus left and right	Hippocampus
43,44	Left and right para-hippocampus	Para-Hippocampal
45,46	Left and right-amygdala	Amygdala
85,86	Superior temporal gyrus	Temporal_Sup
125,126	Ventral anterior	Thal_VA
127,128	Ventral lateral	Thal_VL
129,130	Ventral posterolateral	Thal_VPL
147,148	Pulvinar lateral	Thal_PuL

To validate the importance of brain regions influenced by the three most significant features (GM-to-WM, WM-to-CSF, and GM-to-CSF volumetric ratios), we performed one-way ANOVA tests. Moreover, we explored the significance across five AD stages through post hoc multivariate testing, employing Tukey's Honestly Significant Difference (HSD) method. Table 6 shows the top five significant brain regions and their significance in categorizing AD conditions.

Feature-1 (GM-to-CSF volumetric ratio) significantly affected five brain regions, such as amygdala, ventral posterolateral, pulvinar lateral, ventral lateral, and superior temporal gyrus. The most significant regions were the superior temporal gyrus ($4.91e-13$) and ventral lateral ($6.30e-10$). Among these, the pulvinar lateral and ventral lateral regions were severely affected in the earlier stages of AD progression (NC, EMCI, and MCI), while the amygdala and ventral posterolateral regions were greatly impacted in the AD stage. Feature-2 (WM-to-CSF volumetric ratio) identified the pulvinar lateral, superior temporal gyrus, ventral posterolateral, hippocampus, and ventral lateral as the most affected brain regions. The most significant changes were observed in the ventral lateral ($p = 2.5e-10$) and hippocampus ($8.89e-09$) regions. The superior temporal gyrus, hippocampus, and ventral lateral were primarily affected during the EMCI, MCI, and LMCI stages. Pulvinar lateral and ventral posterolateral regions exhibited significant changes that differentiated patients from the early stages to the final stage of AD prominently. However, feature-3 (GM-to-WM volumetric ratio) highlighted the pulvinar lateral, ventral posterolateral, hippocampus, ventral lateral, and ventral anterior as the most affected brain regions. The ventral anterior region showed the highest significance ($8.32e-12$), particularly between EMCI and MCI conditions, while the hippocampus ($3.74e-09$) was another critically affected region with significant changes across disease progression stages. Except for the pulvinar lateral region, analyzing the other four regions for feature-3 could be helpful in detecting AD progression at early stages.

This analysis highlighted the hippocampus and ventral lateral as consistently significant regions across all three features for detecting the disease progression at early stages. GM-to-WM volumetric ratio demonstrated its potential utility in categorizing AD at earlier stages, as four out of the five regions showed significant changes.

Table 6 Statistical analysis of the affected brain regions based on random forest selected features and their category-wise significance.³

Features	Most affected five regions		Category-wise significance (p-value)
	Region name	Significance	
Feature 1	Amygdala	6.82e-06	NC vs. AD (<1.07e-05) EMCI vs. AD (<6.97e-05) MCI vs. AD (<0.0423) LMCI vs. AD (<0.0076)
	Ventral posterolateral	1.66e-06	NC vs. MCI (<0.0128) NC vs. AD (<1.32e-06) EMCI vs. AD (<0.0003) LMCI vs. AD (<0.0041)
	Pulvinar lateral	4.90e-08	NC vs. MCI (<0.0279) EMCI Vs. LMCI (<5.08e-10) MCI vs. AD (<0.0026)
	Ventral lateral	6.30e-10	NC vs. MCI (<0.0279) MCI vs. AD (<5.08e-10) EMCI Vs. MCI (< 1.35e-06)
	Superior temporal gyrus	4.91e-13	NC vs. MCI (<0.0145) NC vs. LMCI (< 0.0175) LMCI vs. AD (<2.36e-13)
Feature 2	Pulvinar lateral	0.0004	NC vs. AD (<0.0002) EMCI vs. AD (<0.0047)
	Superior temporal gyrus	0.0002	EMCI vs. MCI (<0.0161) MCI vs. LMCI (<8.62e-05) MCI vs. AD (<0.0086)
	Ventral posterolateral	1.56e-06	NC vs. AD (<0.0226) EMCI vs. AD (<0.0001)
	Hippocampus	8.89e-09	NC vs. LMCI (<0.0210) EMCI vs. LMCI (<8.75e-09) MCI vs. AD (<6.79e-06)
	Ventral lateral	2.59e-10	EMCI vs. MCI (<8.85e-05) MCI vs. LMCI (<7.34e-08) MCI vs. AD (<6.54e-06)
Feature 3	Pulvinar lateral	0.0002	NC vs. AD (<0.0001) MCI vs AD (<0.0096)
	Ventral posterolateral	6.63e-07	EMCI vs. LMCI (<0.0002)
	Hippocampus	3.74e-09	EMCI vs MCI (<0.0002) MCI vs LMCI (<0.0001)
	Ventral lateral	2.62e-09	EMCI vs. MCI (<1.61e-06) MCI vs. LMCI (<0.0010) LMCI vs. AD (<0.0002)
	Ventral anterior	8.32e-12	NC vs. MCI (<0.0001) EMCI vs. MCI (<1.97e-08) MCI vs. LMCI (<0.0001)

³ NC = normal cognition, EMCI = early mild cognitive impairment, MCI = mild cognitive impairment, LMCI = late mild cognitive impairment, AD = Alzheimer's disease

Experimental Analysis of the Ensembled Network and Validation

The performance outcomes of the proposed models encompassing run-1 through 4, with comparisons drawn among them using sensitivity, specificity, and F1-score metrics was presented in Table 7. It was found that 3D CNN (ResNet-18) network with three features and 60 brain regions model (run-1) outperformed the same models for other combinations (run-2, run-3 and run-4). The training accuracies for the three features with 60 regions (run-1) and for the three features with all 170 regions (run-4) were reported to be 99.99% and 99.96%, respectively. In contrast, the training accuracies for run-2 (13 features and 60 regions) and run-3 (13 features and all regions) were 77.29% and 80.10%, respectively, which were lower than those for run-1 and run-4.

The sensitivity, F-1 score, and specificity values were reported in Table 7 for four runs which demonstrated that the combination of selected features (three features) and selected regions (60 regions) yielded the most promising results across performance metrics. To validate the model, we also utilized the proposed model for two and three AD categories based on the selected three features and most affected 60 regions. The performance parameters for the two and three categories are shown in Table 8. The highest training accuracy was reported for AD vs. NC (93.33%). Through five-fold cross-validation, the proposed model incorporated three features and 60 regions (run-1) demonstrated a validation accuracy of approximately 60.02%.

Table 7 Performance parameters of the proposed approaches (for five AD categories).

Run number	Regions and features	Sensitivity	F-1 score	Specificity
Run-1	3 features – 60 regions	0.760	0.660	0.935
Run-2	13 features – 60 regions	0.600	0.439	0.883
Run-3	13 features – all regions	0.630	0.470	0.840
Run-4	3 features – all regions	0.733	0.650	0.941

Experimental Analysis of the Progression Pattern

In neuroimaging research, the absolute mean Z-score serves as a common metric for quantifying the level of atrophy across various brain regions, particularly in Alzheimer's disease (AD) and Mild Cognitive Impairment (MCI), where a higher absolute mean Z-score usually signifies a greater degree of atrophy [68]. The progression pattern of volumetric changes in different brain regions among three RF selected features was presented in Fig. 7. To determine the Z-score of all categories (EMCI, MCI, LMCI, AD), we utilized the mean and standard deviation of NC subjects. The absolute mean Z-score value revealed a noticeable alteration in regional volume throughout the progression from the early MCI stage to the final stage of AD disease. From the observed progression pattern, it became apparent that the absolute mean Z-score value exhibited an increasing trend (indicative of higher atrophy) during the transition from the EMCI stage to AD.

Table 8 Performance parameters of the proposed approaches (for two and three AD categories).⁴

Run number	Category	Sensitivity	F1 score	Specificity
Run-1	AD Vs. NC	0.933	0.912	0.932
Run-2	EMCI vs. MCI	0.866	0.787	0.860
Run-3	LMCI Vs. AD	0.832	0.819	0.830
Run-4	MCI Vs. AD	0.833	0.793	0.830
Run-5	MCI Vs. LMCI	0.801	0.768	0.800
Run-6	MCI vs. NC	0.702	0.700	0.702
Run-7	NC Vs. MCI Vs. AD	0.800	0.727	0.900
Run-8	NC Vs. LMCI Vs. AD	0.866	0.833	0.9667
Run-9	EMCI Vs. LMCI Vs. AD	0.690	0.689	0.867
Run-10	NC Vs. MCI Vs. LMCI	0.800	0.645	0.967

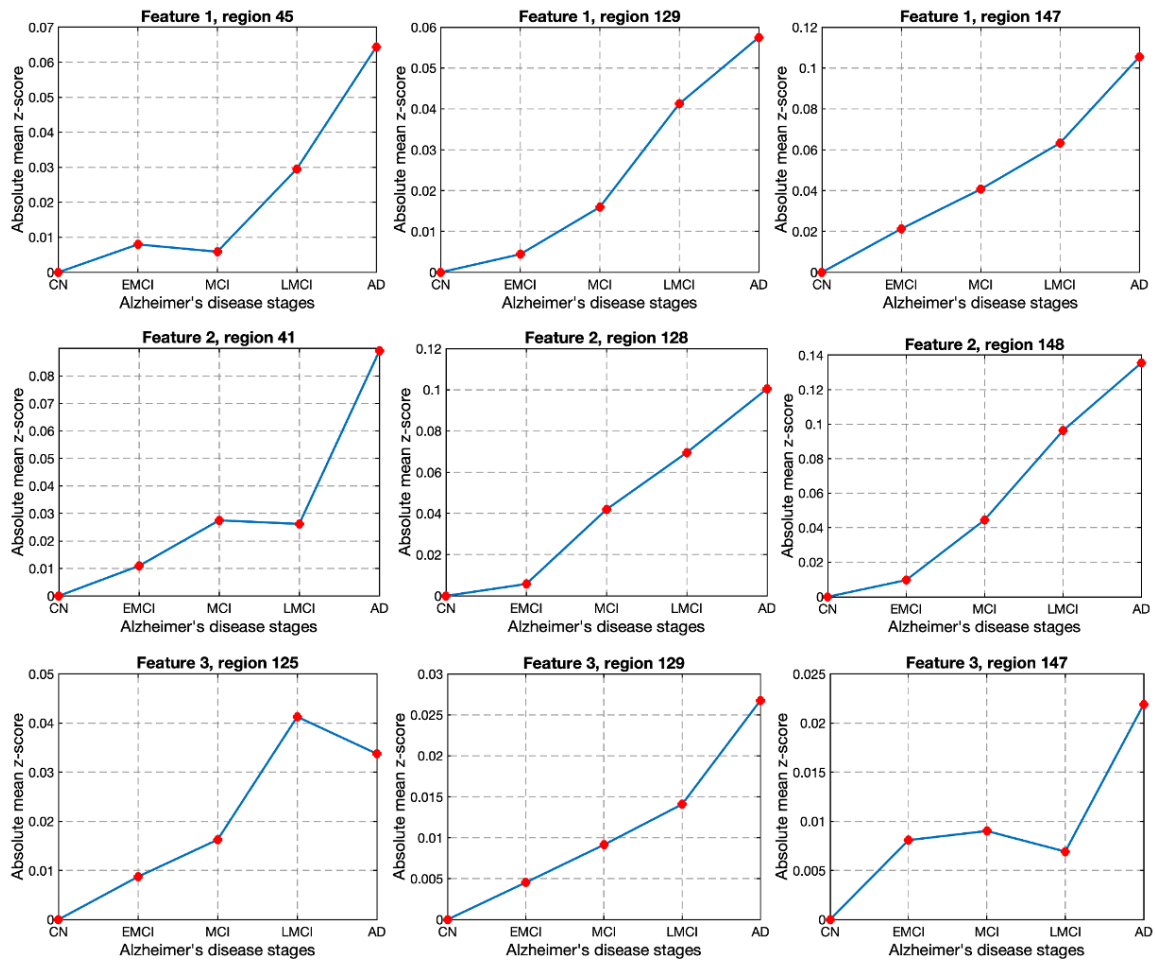


Fig. 7 The Progression trend among five AD stages based on absolute mean Z-score value. Here, NC = normal cognition, EMCI = early mild cognitive impairment, MCI = mild cognitive impairment, LMCI = late mild cognitive impairment, AD = Alzheimer's disease.

⁴ NC = normal cognition, EMCI = early mild cognitive impairment, MCI = mild cognitive impairment, LMCI = late mild cognitive impairment, AD = Alzheimer's disease

Discussion

In comparison to previous AI models in the literature (developed based on ADNI datasets), our model provided higher values across all performance metrics, which could be attributed to our systematic hyperparameter tuning process, the presence of skip connections, and residual blocks in the ResNet-18 network. In this study, the aim was to classify the five categories of Alzheimer's disease and predict the progression pattern through EMCI to AD stage. The progression pattern would help the clinicians to detect the symptoms in earlier stages (e.g. EMCI, MCI) and in obtaining disease-modifying targeted therapeutic interventions to assist potential AD patients, which can eventually extend their life expectancy. Notably, combination-1 (utilizing RF selected three features and 60 regions) yielded the highest performance parameters which could help in detecting early stages of AD with less computational time and cost. This combination of selected features and brain regions achieved the highest accuracy as they were significantly contributing to classifying the five categories of the AD.

To validate the significance of our results, we conducted statistical analysis (ANOVA) on RF selected three features to evaluate the most important regions. Our findings indicated that the Amygdala, Hippocampus, Ventral posterolateral, Pulvinar lateral, Ventral lateral, and Superior temporal gyrus regions were among the most affected regions, showing significance across different disease categories of cognitive impairment condition. The analysis was aimed to delineate the progression pattern for categorizing the early stages of cognitive impairment, achieved by leveraging the absolute mean Z-score. This pattern revealed an upward trend when progressing from the NC stage to AD. To show the novelty of the work, we compared our findings with the existing literature presented in Table 9. Most previous studies focused on analyzing two or three categories of Alzheimer's disease. Therefore, we extended our analysis among five AD stages to validate our novel approach using the proposed ensemble AI network. To ensure fair comparison, we utilized common platforms, including MRI datasets, activation functions, classifiers, categories, validation approaches, and performance parameters.

Although our proposed network provided a validation accuracy of nearly 60% using RF selected three important features and most affected 60 brain regions, the result could be improved by applying the transfer learning algorithm, expanding training data, and/or further analyzing the voxels as 2D slices. The total number of MRI samples was 1000, which could be increased to reduce the effects of overfitting. Different features (selection of different weights/coefficients and different combination) and atlases could be utilized to localize the brain regions and determine the progression rate for each combination of feature and atlas. Furthermore, depending upon the regions of hippocampus and amygdala, the biomarkers could be developed for different specific features.

Table 9 Comparison of the proposed model with the existing literature.⁵

Networks	Layers	Classifier	Categories	Dataset and validation approach	Accuracy (%)
3D CNN and CNN-LSTM [69]	12 (3D CNN) and 23 layers (CNN-LSTM)	Softmax	NC Vs. AD	894 MRI scans from OASIS; 10-fold cross-validation	0.53 (3D CNN) 0.63 (CNN-LSTM)
3D CNN [46]	LeNet-5, VGG, Google Net, and ResNet	Softmax	AD Vs. MCI Vs. NC	MRI scan from ADNI; leave-one-out cross-validation	Highest accuracy :0.52
3D CNN (VGG16, ResNet-50 architecture) [70]	16 layers	Softmax	Binary and Multiple Classification	MRI dataset (EDLB, ADNI); 10-fold cross-validation	AD Vs. NC: 0.89 4-way classification: 0.66
3D CNN [47]	21 layers	Softmax	AD Vs. MCI Vs. NC	MRI from Gwangju Alzheimer's disease and Related Dementias (GARD) center.	Highest Accuracy: 0.4211
3D Connection-wise-attention-model (CAM-CNN) [45]	20 layers	Softmax	AD Vs. MCI Vs. NC	MRI from ADNI Validation: 15% for validation	0.973, 0.878, and 0.787 accuracy for mild AD, MCI and stable MCI against NC
3D CNN Model-SFENet with GCN (graph convolutional network) [71]	11 Blocks (9 convolution and 2 fully connected layer)	ReLU	EMCI Vs. LMCI Vs. NC	DTI of 298 subjects from the ADNI	accuracy: 0.833
3D CNN Classifier [72]	27 layers	Softmax	AD Vs. MCI Vs. NC	1230 PET scans from ADNI database	AD Vs. NC; accuracy: 0.887
DenseCNN2, a lightweight 3D deep convolutional network model [73]	23 layers	Softmax	AD Vs. NC	MRI dataset from ADNI	Average accuracy of 0.925
ResNet-18 Network [74]	72 layers	Softmax	EMCI, MCI, LMCI, and AD (Binary Classification)	413 MRI scans from ADNI	MCI vs. EMCI: 0.896
3D deep learning approach (3D-SENet) [75]	15 layers	Softmax	AD Vs. HC Vs. MCI, Binary Classification	509 MRI scans from ADNI	AD vs. HC: 0.84, MCIc Vs HC: 0.79
3D DenseNets [67]	121 layers	Softmax	EMCI Vs. LMCI Vs. AD Vs. NC	480 MRI scans from ADNI	AD, EMCI, LMCI, NC: 0.667
3D Deep Network (DenseNet and ResNet) [76]	55 layers	Softmax	AD Vs. MCI Vs. NC	449 MRI scans from ADNI	AD vs. NC: 0.889, MCI Vs NC: 0.760
CNN - LeNet [77]	--	--	AD, MCI and NC; Binary Classification	815 MRI scans from ADNI	AD Vs. NC: 0.868, AD Vs. MCI: 0.714, MCI Vs. NC: 0.698
Proposed Work: ResNet-18 - RF (selected features and regions)	72 layers	Softmax	AD, LMCI, MCI, EMCI and NC (2-, 3- and 5-way classification)	1000 MRI scans from ADNI; 5-fold cross-validation	AD Vs. NC: 0.933 AD Vs. MCI: 0.783 MCI Vs. NC: 0.700 AD Vs. NC Vs. MCI: 0.644 AD, LMCI, MCI, EMCI and NC: 0.663

⁵ NC = normal cognition, EMCI = early mild cognitive impairment, MCI = mild cognitive impairment, LMCI = late mild cognitive impairment, AD = Alzheimer's disease

Conflict of Interest

The authors have no competing interests to declare that are relevant to the content of this article.

Code Availability

The codes used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Muhammed Raees PC, Thomas V. Automated detection of Alzheimer's Disease using Deep Learning in MRI. *Journal of Physics Conference Series*. 1921. 2021:012024.
2. Sharma S, Guleria K, Tiwari S, Kumar S. A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer Disease using MRI scans. *Measurement: Sensors* 2022;24:100506.
3. Health Nif, Excellence C. Dementia: assessment, management and support for people living with dementia and their carers. National Institute for Health and Clinical Excellence; 2018.
4. Global status report on the public health response to dementia: executive summary. World Health Organization 2021; 2021.
5. 2024 Alzheimer's Disease Facts and Figures Report: Executive Summary. Alzheimer's Association; 2024.
6. Mild cognitive impairment (MCI). MAYO CLINIC; 2024.
7. Chen Y-X, Liang N, Li X-L, Yang S-H, Wang Y-P, Shi N-N. Diagnosis and treatment for mild cognitive impairment: a systematic review of clinical practice guidelines and consensus statements. *Frontiers in Neurology* 2021;12:719849.
8. Wang Y, Gao R, Wei T, Johnston L, Yuan X, Zhang Y, et al. Predicting long-term progression of Alzheimer's disease using a multimodal deep learning model incorporating interaction effects. *Journal of Translational Medicine* 2024;22(1):265.
9. Zhang T, Zhao Z, Zhang C, Zhang J, Jin Z, Li L. Classification of early and late mild cognitive impairment using functional brain network of resting-state fMRI. *Frontiers in Psychiatry* 2019;10:572.
10. 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(3):681-9.
11. Kyaga ZNVGS. The Early Diagnosis of Alzheimer's Disease: A Patient-Centred Conversation with the Care Team. 2022.
12. Mosconi L, Brys M, Glodzik-Sobanska L, De Santi S, Rusinek H, De Leon MJ. Early detection of Alzheimer's disease using neuroimaging. *Experimental gerontology* 2007;42(1-2):129-38.
13. Dang M, Yang C, Chen K, Lu P, Li H, Zhang Z, Beijing Aging Brain Rejuvenation Initiative ftAsDNI. Hippocampus-centred grey matter covariance networks predict the development and reversion of mild cognitive impairment. *Alzheimer's Research & Therapy* 2023;15(1):27.
14. Wu Z, Peng Y, Hong M, Zhang Y. Gray matter deterioration pattern during Alzheimer's disease progression: a regions-of-interest based surface morphometry study. *Frontiers in Aging Neuroscience* 2021;13:593898.

15. Yao Z, Zhang Y, Lin L, Zhou Y, Xu C, Jiang T, Initiative AsDN. Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease. *PLoS computational biology* 2010;6(11):e1001006.
16. Schwarz CG. Uses of human MR and PET imaging in research of neurodegenerative brain diseases. *Neurotherapeutics* 2021;18(2):661-72.
17. Yen C, Lin C-L, Chiang M-C. Exploring the frontiers of neuroimaging: a review of recent advances in understanding brain functioning and disorders. *Life* 2023;13(7):1472.
18. Kubi P, Islam A, Zaher MAHB, Ripon SH. A Deep Learning-Based Technique to Determine Various Stages of Alzheimer's Disease from 3D Brain MRI Images. *International Conference on Information Integration and Web Intelligence*. Springer; 2023:162-75.
19. Chakraborty D, Zhuang Z, Xue H, Fiecas MB, Shen X, Pan W. Deep learning-based feature extraction with MRI data in neuroimaging genetics for Alzheimer's disease. *Genes* 2023;14(3):626.
20. Stephen R, Liu Y, Ngandu T, Antikainen R, Hulkkonen J, Koikkalainen J, et al. Brain volumes and cortical thickness on MRI in the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER). *Alzheimer's research & therapy* 2019;11:1-10.
21. Seo K, Pan R, Lee D, Thiyyagura P, Chen K. Visualizing Alzheimer's disease progression in low dimensional manifolds. *Heliyon* 2019;5(8).
22. Diciotti S, Ginestroni A, Bessi V, Giannelli M, Tessa C, Bracco L, et al. Identification of mild Alzheimer's disease through automated classification of structural MRI features. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE; 2012:428-31.
23. Zheng Y, Guo H, Zhang L, Wu J, Li Q, Lv F. Machine learning-based framework for differential diagnosis between vascular dementia and Alzheimer's disease using structural MRI features. *Frontiers in neurology* 2019;10:456891.
24. Sierra H, Cordova M, Chen C-SJ, Rajadhyaksha M. Confocal imaging-guided laser ablation of basal cell carcinomas: an ex vivo study. *The Journal of investigative dermatology* 2015;135(2):612.
25. 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. *Journal of healthcare engineering* 2017;2017.
26. Putcha D, Katsumi Y, Brickhouse M, Flaherty R, Salat DH, Touroutoglou A, Dickerson BC. Gray to white matter signal ratio as a novel biomarker of neurodegeneration in Alzheimer's disease. *NeuroImage: Clinical* 2023;37:103303.
27. Dubois B, von Arnim CA, Burnie N, Bozeat S, Cummings J. Biomarkers in Alzheimer's disease: role in early and differential diagnosis and recognition of atypical variants. *Alzheimer's Research & Therapy* 2023;15(1):175.
28. Khagi B, Kwon G-R. 3D CNN based Alzheimer's diseases classification using segmented Grey matter extracted from whole-brain MRI. *JOIV: International Journal on Informatics Visualization* 2021;5(2):200-5.
29. Long X, Wyatt C. An automatic unsupervised classification of MR images in Alzheimer's disease. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE; 2010:2910-7.
30. Zhang D, Wang Y, Zhou L, Yuan H, Shen D, Initiative AsDN. Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* 2011;55(3):856-67.

31. Rolls ET, Huang C-C, Lin C-P, Feng J, Joliot M. Automated anatomical labelling atlas 3. *Neuroimage* 2020;206:116189.
32. Long Z, Li J, Fan J, Li B, Du Y, Qiu S, et al. Identifying Alzheimer's disease and mild cognitive impairment with atlas-based multi-modal metrics. *Frontiers in Aging Neuroscience* 2023;15.
33. Li H, Habes M, Wolk DA, Fan Y, Initiative AsDN. A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data. *Alzheimer's & Dementia* 2019;15(8):1059-70.
34. Islam J, Zhang Y. Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. *Brain informatics* 2018;5:1-14.
35. Pan D, Huang Y, Zeng A, Jia L, Song X, Initiative AsDN. Early diagnosis of Alzheimer's disease based on deep learning and GWAS. *Human Brain and Artificial Intelligence: First International Workshop, HBAI 2019, Held in Conjunction with IJCAI 2019, Macao, China, August 12, 2019, Revised Selected Papers 1*. Springer; 2019:52-68.
36. Zukotynski K, Gaudet V, Kuo PH, Adamo S, Goubran M, Scott CJ, et al. The use of random forests to identify brain regions on amyloid and FDG PET associated with MoCA score. *Clinical nuclear medicine* 2020;45(6):427-33.
37. Sarica A, Cerasa A, Quattrone A. Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. *Frontiers in aging neuroscience* 2017;9:329.
38. Ebrahimighahnavieh MA, Luo S, Chiong R. Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review. *Computer methods and programs in biomedicine* 2020;187:105242.
39. Liu S, Liu S, Cai W, Che H, Pujol S, Kikinis R, et al. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE transactions on biomedical engineering* 2014;62(4):1132-40.
40. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data* 2021;8:1-74.
41. Jain R, Jain N, Aggarwal A, Hemanth DJ. Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research* 2019;57:147-59.
42. Lin W, Tong T, Gao Q, Guo D, Du X, Yang Y, et al. Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment. *Frontiers in neuroscience* 2018;12:777.
43. Huang Y, Xu J, Zhou Y, Tong T, Zhuang X, Initiative AsDN. Diagnosis of Alzheimer's disease via multi-modality 3D convolutional neural network. *Frontiers in neuroscience* 2019;13:509.
44. Hu Y, Tang H, Pan G. Spiking deep residual networks. *IEEE Transactions on Neural Networks and Learning Systems* 2021.
45. Zhang J, Zheng B, Gao A, Feng X, Liang D, Long X. A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification. *Magnetic Resonance Imaging* 2021;78:119-26.
46. Folego G, Weiler M, Casseb RF, Pires R, Rocha A. Alzheimer's disease detection through whole-brain 3D-CNN MRI. *Frontiers in bioengineering and biotechnology* 2020;8:534592.

47. Khagi B, Kwon G-R. 3D CNN based Alzheimer's diseases classification using segmented Grey matter extracted from whole-brain MRI. *JOIV: International Journal on Informatics Visualization* 2021;5(2):200-5.
48. Chen Y, Wang L, Ding B, Shi J, Wen T, Huang J, Ye Y. Automated Alzheimer's disease classification using deep learning models with Soft-NMS and improved ResNet50 integration. *Journal of Radiation Research and Applied Sciences* 2024;17(1):100782.
49. Pasnoori N, Flores-Garcia T, Barkana BD. Histogram-based features track Alzheimer's progression in brain MRI. *Scientific Reports* 2024;14(1):257.
50. Raza N, Naseer A, Tamoor M, Zafar K. Alzheimer disease classification through transfer learning approach. *Diagnostics* 2023;13(4):801.
51. Wyman BT, Harvey DJ, Crawford K, Bernstein MA, Carmichael O, Cole PE, et al. Standardization of analysis sets for reporting results from ADNI MRI data. *Alzheimer's & Dementia* 2013;9(3):332-7.
52. Morris JC. The Clinical Dementia Rating (CDR) current version and scoring rules. *Neurology* 1993;43(11):2412--a.
53. Tariot PN, Boada M, Lanctôt KL, Hahn-Pedersen J, Dabbous F, Udayachalerm S, et al. Relationships of change in Clinical Dementia Rating (CDR) on patient outcomes and probability of progression: observational analysis. *Alzheimer's Research & Therapy* 2024;16(1):36.
54. Parmar H, Nutter B, Long R, Antani S, Mitra S. Spatiotemporal feature extraction and classification of Alzheimer's disease using deep learning 3D-CNN for fMRI data. *Journal of Medical Imaging* 2020;7(5):056001-.
55. Huang Y, Li W. Resizer Swin transformer-based classification using sMRI for Alzheimer's disease. *Applied Sciences* 2023;13(16):9310.
56. Duta N, Sonka M. Segmentation and interpretation of MR brain images. An improved active shape model. *IEEE transactions on medical imaging* 1998;17(6):1049-62.
57. Kanakaraj P, Yao T, Cai LY, Lee HH, Newlin NR, Kim ME, et al. DeepN4: Learning N4ITK Bias Field Correction for T1-weighted Images. *Neuroinformatics* 2024:1-13.
58. Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang M-C, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 2009;46(3):786-802.
59. Long Z, Li J, Liao H, Deng L, Du Y, Fan J, et al. A multi-modal and multi-atlas integrated framework for identification of mild cognitive impairment. *Brain Sciences* 2022;12(6):751.
60. Vinutha H, Poornima B, Sagar B. Detection of outliers using interquartile range technique from intrusion dataset. *Information and decision sciences: Proceedings of the 6th international conference on ficta*. Springer; 2018:511-8.
61. Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26(3):297-302.
62. Dai Z, Yan C, Wang Z, Wang J, Xia M, Li K, He Y. Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3). *Neuroimage* 2012;59(3):2187-95.
63. Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, et al. Prediction of individual brain maturity using fMRI. *Science* 2010;329(5997):1358-61.

64. Tiwari S, Jain G, Shetty DK, Sudhi M, Balakrishnan JM, Bhatta SR. A Comprehensive Review on the Application of 3D Convolutional Neural Networks in Medical Imaging. *Engineering Proceedings* 2023;59(1):3.
65. Borawar L, Kaur R. ResNet: Solving vanishing gradient in deep networks. *Proceedings of International Conference on Recent Trends in Computing: ICRTC 2022*. Springer; 2023:235-47.
66. Zhang F, Petersen M, Johnson L, Hall J, O'Bryant SE. Accelerating hyperparameter tuning in machine learning for Alzheimer's disease with high performance computing. *Frontiers in Artificial Intelligence* 2021;4:798962.
67. Ruiz J, Mahmud M, Modasshir M, Shamim Kaiser M, Alzheimer's Disease Neuroimaging Initiative ft. 3D DenseNet ensemble in 4-way classification of Alzheimer's disease. *Brain Informatics: 13th International Conference, BI 2020, Padua, Italy, September 19, 2020, Proceedings* 13. Springer; 2020:85-96.
68. Odano I, Maeyatsu F, Hosoya T, Asari M, Oba K, Taki Y. Diagnostic approach with Z-score mapping to reduce artifacts caused by cerebral atrophy in regional CBF assessment of mild cognitive impairment (MCI) and Alzheimer's disease by [99mTc]-ECD and SPECT. *Japanese Journal of Radiology* 2024:1-11.
69. Liu W, Zhang J, Zhao Y. A Comparison of Deep Learning and Traditional Machine Learning Approaches in Detecting Cognitive Impairment Using MRI Scans. *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE; 2022:998-1001.
70. Soliman A, Chang JR, Etminani K, Byttner S, Davidsson A, Martínez-Sanchis B, et al. Adopting transfer learning for neuroimaging: a comparative analysis with a custom 3D convolution neural network model. *BMC medical informatics and decision making* 2022;22(Suppl 6):318.
71. Kong H, Pan J, Shen Y, Wang S. Adversarial Learning Based Structural Brain-Network Generative Model for Analyzing Mild Cognitive Impairment. *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*. Springer; 2022:361-75.
72. Islam J, Zhang Y. Understanding 3D CNN behavior for Alzheimer's disease diagnosis from brain PET scan. *arXiv preprint arXiv:191204563* 2019.
73. Katabathula S, Wang Q, Xu R. Predict Alzheimer's disease using hippocampus MRI data: a lightweight 3D deep convolutional network model with visual and global shape representations. *Alzheimer's Research & Therapy* 2021;13(1):1-9.
74. Nicholas PJ, To A, Tanglay O, Young IM, Sughrue ME, Doyen S. Using a ResNet-18 Network to Detect Features of Alzheimer's Disease on Functional Magnetic Resonance Imaging: A Failed Replication. *Comment on Odusami et al. Analysis of Features of Alzheimer's Disease: Detection of Early Stage from Functional Brain Changes in Magnetic Resonance Images Using a Finetuned ResNet18 Network*. *Diagnostics* 2021, 11, 1071. *Diagnostics* 2022;12(5):1094.
75. Pan D, Zeng A, Jia L, Huang Y, Frizzell T, Song X. Early detection of Alzheimer's disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning. *Frontiers in neuroscience* 2020;14:259.
76. Liu M, Li F, Yan H, Wang K, Ma Y, Shen L, et al. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. *Neuroimage* 2020;208:116459.

77. Aderghal K, Afdel K, Benois-Pineau J, Catheline G. Improving Alzheimer's stage categorization with Convolutional Neural Network using transfer learning and different magnetic resonance imaging modalities. *Heliyon* 2020;6(12).