Enhancing Alzheimer's Disease Classification using 3D Convolutional Neural Network and Multilayer Perceptron Model with Attention Network

Enoch A. Frimpong[1,2], Zhiguang Qin[1*], Regina E. Turkson[3], Bernard M. Cobbinah[4], Edward Y. Baagyere[5], and Edwin K. Tenagyei[6]

[1] School of Information and Software Engineering, University of Electronic Science and Technology of China,
No.4, Section 2, North Jianshe Road, 610054, Chengdu, P.R. China.
[e-mail: frimpongadjeienoch@ymail.com, qinzg@uestc.edu.cn]
[2] Department of Computer Science and Technology, Cape Coast Technical University,
Box DL 50, Cape Coast, Ghana.
[3] Department of Computer Science and Information Technology, University of Cape Coast,
PMB University Post Office, Cape Coast, Ghana.
[e-mail: rturkson@ucc.edu.gh]
[4] School of Computer Science and Engineering, University of Electronic Science and Technology of China,
Chengdu, 611731, China
[e-mail: cobbinahben@std.uestc.edu.cn]
[5] Department of Computer Science, C. K. Tedam University of Technology and Applied Sciences,
P.O. Box 24, Navrongo, Ghana.
[e-mail: ebaagyere@cktutas.edu.gh]
[6] School of Engineering and Built Environment, Griffith University, QLD 4111, Australia.
[e-mail: edwinkwadwo.tenagyei@griffithuni.edu.au]
[*]Corresponding author: Qin Zhiguang

Abstract

Alzheimer's disease (AD) is a neurological condition that is recognized as one of the primary causes of memory loss. AD currently has no cure. Therefore, the need to develop an efficient model with high precision for timely detection of the disease is very essential. When AD is detected early, treatment would be most likely successful. The most often utilized indicators for AD identification are the Mini-mental state examination (MMSE), and the clinical dementia. However, the use of these indicators as ground truth marking could be imprecise for AD detection. Researchers have proposed several computer-aided frameworks and lately, the supervised model is mostly used. In this study, we propose a novel 3D Convolutional Neural Network Multilayer Perceptron (3D CNN-MLP) based model for AD classification. The model uses Attention Mechanism to automatically extract relevant features from Magnetic Resonance Images (MRI) to generate probability maps which serves as input for the MLP classifier. Three MRI scan categories were considered, thus AD dementia patients, Mild Cognitive Impairment patients (MCI), and Normal Control (NC) or healthy patients. The performance of the model is assessed by comparing basic CNN, VGG16, DenseNet models,

and other state of the art works. The models were adjusted to fit the 3D images before the comparison was done. Our model exhibited excellent classification performance, with an accuracy of 91.27% for AD and NC, 80.85% for MCI and NC, and 87.34% for AD and MCI.

1. Introduction

Dementia is defined as memory impairment and other cognitive functions. People with dementia may find it difficult to remember or think well to execute simple daily tasks such as eating or getting dressed [1]. Alzheimer's disease (AD) is a neurodegenerative disorder known to be one of the common causes of dementia in the elderly population [2]. About 24.3 million people between the ages of 60 and 84 suffer from AD [3] and an estimated 40 million patients globally [4]. In the United States, AD is said to be the sixth leading cause of death [5], yet getting a cure still eludes researchers even today. The destructive combination of plagues and tangles works simultaneously to break down the structure of the brain. Degeneration begins in the patient's brain cortex, triggering abrupt shrinkage in the hippocampal region responsible for the formation of memories. This is why the first symptom of AD is mostly short-term memory loss. As the disease progresses, other parts of the brain may be invaded which creates a unique transformation that indicates the progression phases of AD [6] which is shown in **Fig. 1**.
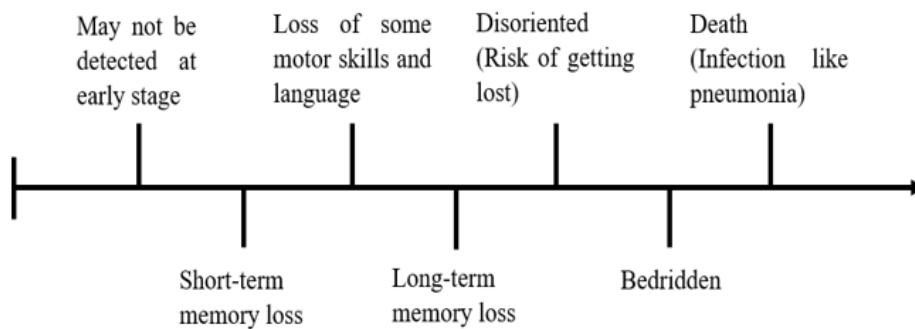

Fig. 1. Progressive symptoms of AD.

At the early stage, people with AD are considered as having mild cognitive impairment (MCI), however, not all MCIs develop into AD. MCI is the intermediate state from healthy to AD [6]. Diagnosing AD depends on clinical assessment [7, 8, 9] and the application of neuroimaging methodologies like as positron emission tomography (PET), functional magnetic resonance, and structural magnetic resonance imaging (sMRI) [10-20]. The type of neuroimaging technique used depends on the severity of the disease stage; for instance, other modalities such

as single photon emission computed tomography (SPECT) or PET could be used when MRI is unable to detect cerebrum modifications. Currently, AD has no remedy. Therefore, the need to develop effective model for early detection is very important, as this could help health practitioners give treatments to delay its progression. Mini-mental state examination (MMSE) and clinical dementia classification are most widely used indicators for AD detection. Employing these indicators as ground truth marking for AD detection could be imprecise, as the reported clinical conclusion on the correctness of AD compared with subsequent findings is within 70%-90% range [7]. There have been several computer-aided frameworks developed by researchers such as rule-based and recently supervised models [17]. In the case of supervised models, features of the task-related images are extracted by human expects for model training. This is tedious, time consuming and may also require a lot of funds [7].

Deep learning (DL) models have proven to be efficient in feature extraction from task-related images with no human intervention. This also reduces human errors and increases models' efficiency on disease diagnosis. DL is a machine learning method used to train computers to execute tasks like humans, for text, sound, or image classifications. Also, it is a powerful technique for attaining state-of-the-art accuracy because of its ability to learn every feature of the task related data [19, 20].

2. Related Works

Conventionally, neuro-experts, assisted by psychologists perform several mental and physical examination. Some of these examinations are Mini-Mental State Examination (MMSE) [21], neurological evaluations [22], Screen tests and physical assessments [23]. The conventional methods such as clinical dementia rating MMSE have lower accuracy rates.

The magnetic resonance imaging, otherwise known as MRI, is a mostly used method to obtain detailed tissue by tissue information about the nervous structure as compared to the conventional methods. It has been successfully used to diagnose disorders such as tumours and cancers. Additionally, efficient brain image processing may show significant biomarkers long before an individual gets Alzheimer's disease [24] which the conventional procedures have failed to do by only studying tissue changes caused by complicated pixel patterns.

Artificial Neural Network and Deep Neural Network approaches are becoming more prevalent in AD diagnosis. The primary justification for its appeal is its capacity to pick up information about the best features in a dataset and use that knowledge to enhance its prediction accuracy. In this section, we discuss some recent state-of-the-art publications in respect to AD detection. Existing frameworks predominantly rely on 2D images and supervised models, which may not capture the full scope of information present in 3D images such as MRI scans. 2D CNNs predict the segmentation map for each slice using 2D convolutional kernels. Predictions for segmentation maps for full volume are made for a slice at a time. To create predictions, the 2D convolutional kernels may leverage context throughout the width and height of the slice. However, since 2D CNNs only accept one slice as input, they are incapable of leveraging context from subsequent slices. Voxel information from close by slices could be useful as well for prediction of segmentation map. 3D CNNs overcome this issue by predicting segmentation for a volumetric patch of a scan using 3D convolutional kernels. However, the capacity to exploit inter-slice context to boost performance comes at a computational expense because of the increasing number of parameters employed by these convolutional neural networks. By transitioning to a 3D CNN-Multi Layer Perceptron architecture, we enhance the model's ability to exploit the full depth of these 3D images which 2D CNNs are incapable of. An attention mechanism is also added to the framework to extract more complex and relevant

features for accurate classification. This is because a model's ability to accurately predict an outcome largely depends on its capacity to extract the best features in the dataset which is lacking in most 2D and 3D existing models. The limitations of the methods which motivated our study are summarized in **Table 1**.

Table 1. Classification methods and their limitation(s)

Methods	Limitation(s)
Traditional (Clinical dementia rating, MMSE)	Low accuracy
2D Convolutional Neural Networks discussed in related works	Inability to explore full depth of 3D images
	Manual feature selection
	Low accuracy
3D Convolutional Neural Networks discussed in related works	Manual feature selection
	Low accuracy

Huang et al [25] introduced VGG-like set of multi-modality Alzheimer's disease classifiers which employ fluorodeoxyglucose positron emission tomography as input data for prediction. The model achieved 84.82% average performance score. As alternatives to MRI classification, authors in [26] proposed 2D CNN models. Using various fusion strategies, the methods apply 2D CNN models to 3D image data. Their method was able to achieve a good computational time. Gunawardena et al [27] employed numerous 2D CNNs comprising of two C-layers, one S-layer, and an FC layer for all their experiments on 2D slices of coronal view of the MRI images. Aderghal et al [28] employed a 2D CNN architecture with two C-layers over 2D sagittal planes to analyze MRI data from the hippocampal region. To diagnose AD, Luo et al [29] applied 7 2D CNNs to 7 sections, numbered 25 to 31, each of which contained five neighborhood of 3D MRI brain image slices. Wang et al [30] applied 8-layered 2D convolutional neural network on sagittal view of MRI data for AD detection. Three distinct pooling functions thus Sigmoid, ReLu, and leaky ReLu were experimented. Their best results were at achieved at the combination of max-pooling and leaky ReLu activation function for AD classification. Liu et al [31] also proposed a method based on the blend of 2D convolutional neural network and recurrent neural network (RNN). The framework can learn both the intra and inter slice attributes of the dataset after the 3D positron emission tomography are decomposed into sequence of 2D slices. The 2D CNNs were designed to collect the attributes of the image slices, while the RNN's gated recurrent unit (GRU) is cascaded to acquire and incorporate inter-slice information for classification of images. Marwa et al [32] proposed a deep learning-based system for precise detection and categorization of Alzheimer's disease stages. The proposed analytic method makes use of basic Convolutional Neural Network (CNN) architecture and 2-dimensional MR brain images. Zhang et al [33] proposed a 3D Residual attention network model which employs self-attention residual procedure to study long-range dependencies and a gradient-based localization class activation mapping to provide visual interpretations of AD predictions made.

Although several research have been conducted for brain image analysis and classification and its related diseases, we believe further research could be done to attain near perfect deep learning model for accurate feature extraction from input images which could also increase model's overall efficiency.

In our research, we developed a novel 3D Convolutional Neural Networks (CNNs) Multilayer Perceptron (MLP)-based pipeline for classifying AD using MRI scans. The fundamental problem we address in our study is the inaccurate diagnosis of Alzheimer's disease (AD) using the existing indicators such as the Mini-mental state examination (MMSE) and clinical dementia rating. These indicators, though widely utilized, are often imprecise when it comes

to accurately identifying AD. Moreover, the need for an accurate diagnostic method becomes crucial considering that the earlier AD is detected, the more successful the treatment can be. Again, we also address 2D CNNs inability to fully explore in depth 3D MRI images as well as low accuracy performance presented by some existing state of the art 3D models. The proposed model innovatively integrates the Attention Mechanism with a standard machine learning model to enhance the precision and automatic extraction of relevant features from Magnetic Resonance Images (MRI). These extracted features are then used to generate probability maps, which in turn serve as the input for the MLP classifier. The innovative aspect lies in the model design as most existing frameworks typically rely on 2D images and supervised models, which may not capture the full scope of information present in 3D images such as MRI scans. By transitioning to a 3D CNN-MLP architecture, we can exploit the full depth of these 3D images to extract more complex features for accurate classification. This approach aids in achieving a more accurate classification of AD patients compared to traditional methods.

Three MRI scan groups were considered (AD dementia patients, MCI, and NC) and designed a novel diagnostic model for classification using 3D CNNs MLP- based classifier for three-binary AD classification task (AD vs. NC, MCI vs. NC, and AD vs. MCI). We implemented multi-class pipeline for classifying complex dynamic brain activities. Specifically, a raw MRI data undergoes a pre-processing to obtain the regions of interests (ROI). The preprocessed MRI scans are then used as input for the 3D CNN to extract discriminative features which are riche in multi-modal feature information. To learn the hierarchy of features, several convolutional layers are placed on each other, and each convolutional layer is employed to extract feature maps from its preceding layer. Finally, the extracted multi-modal traits from the 3D CNN are used to generate the CNN probability maps which in turn serves as the input to the MLP for classification and prediction of AD. This research primarily aims to investigate how classification features and classifier models impact classification accuracy. We experimented the robustness of our proposed model using our preprocessed MRI scans. Moreover, we used only those MRI slices of subjects (as shown in **Fig. 2**) having the most amount of relevant information in it by comparing their respective entropies, hence strengthening the model's overall robustness.
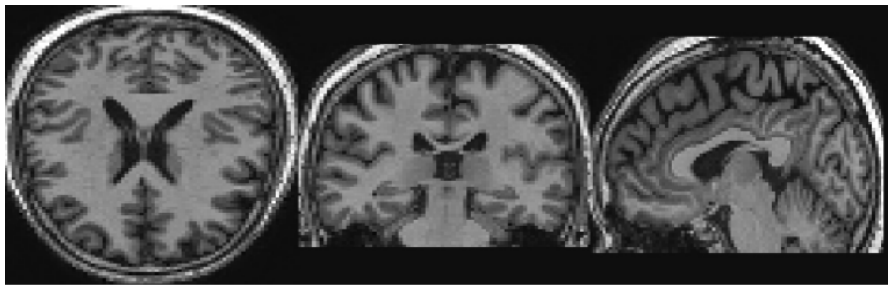

Fig. 2. Slices of an MRI scan of AD patient (from left to right, in axial view, coronal view, and sagittal view)

The rest of the paper is arranged as follows: In Section 3, we provide the details of the proposed method. The experimental results are presented in Section 4. The discussion of experimental result implication is presented in section 5 and finally, the conclusion is drawn in Section 6.

3. Methodology and Implementation

We constructed a robust deep-learning model for AD classification. A 3D CNN MLP-based architecture was designed which takes raw MRI data and pre-processes the MRI data to obtain the ROI. We then used a 3D CNN to extract discriminative features. The feature maps obtained from the 3D CNN were used to generate a probability map to ensure the probability that the features are above or below some specific threshold. The probability map generated then serves as the input for the MLP for classification and prediction of the three-binary AD classification. Thus, the proposed approach has a deep residual block of convolution that concentrates on significant features. Additionally, to counter the effect of cell-scale variations, a feature-based probability map attention mechanism was developed to enhance feature information and improve prediction precision. The proposed method's framework is illustrated in **Fig. 3**, consisting of the following components: raw MRI, pre-processed MRI (ROI), a 3D-CNN, a probability map and an MLP.
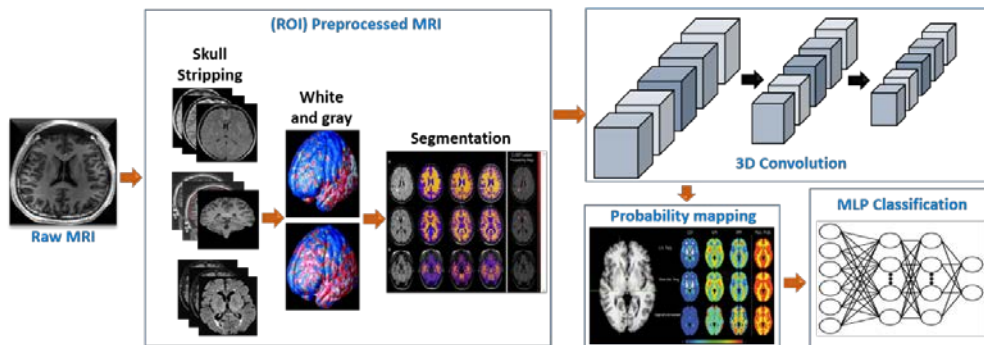

Fig. 3. Proposed model's architecture.

3.1 MRI Data Acquisition and Pre-processing

1.5T and 3T scanners were used to acquire T1-weighted structural MRI images. The subsequent sequences were obtained: (1) 1.5 T T1-weighted; acquisition criteria were repetition time (TR)=2400 ms, flip angle=8° or 9°, minimum full echo time (TE), inversion time (TI)=1000 ms, field-of-view (FOV)=240mm, thickness of slice = 2 mm, acquisition matrix with differing x, y, and z dimensions.

MRI scans are shown as three-dimensional Nifti volumes. The original T1 and T2 weighted structural magnetic resonance imaging (sMRI) images were acquired from 1.5 T or 3 T scanning machines utilizing Siemens MRI scanners, Philips Medical Systems, and General Electric (GE) Medical System at all locations. For the same brain substructures to be registered at the same image coordinates for all participants, we pre-processed the original MRI scans to change the original multicentre brain scans into normal image space.

To prepare the raw MRI scans for analysis, we implemented a standard data pre-processing pipeline. This process included linear image registration, skull-stripping, segmentation, intensity normalization, background removal, and outlier clipping. All MRI scans were in NIFTI format and were registered using the MNI152 template, except for 3-T MRI scans of the same subjects, which were co-registered with the registered 1.5-T MRI scans. The FLIRT tool from the FSL package was used for this purpose. Upon manual review, the automatic registration was found to be adequate for the majority cases. For cases that were not adequately registered, we applied affine transformations for manual registration. After normalization, we

performed outlier clipping by adjusting voxel intensity to the range of [-1, 2.5], setting any voxels with intensity below -1 to -1 and those above 2.5 to 2.5. Finally, we removed background voxels by setting their intensity to -1 to ensure uniform background intensity.

The best discriminator set consisted of the Entorhinal Area, Hippocampus, Amygdala, and Posterior Cingulate Cortex. These four (4) ROIs were utilized for categorization. All data received were standardized to zero mean and unit variance for every feature after pre-processing, using a common scalar function. Thus, given the data matrix A, in which rows indicate subjects and columns indicate features, and the normalized matrix with elements $a(i, j)$ is computed in (1) as:

$$A_n(i, j) = \frac{a(i, j) - mean(a_j)}{\text{std}(a_j)} \tag{1}$$

Where the jth column of the matrix (A) is denoted as $A_{i \cdot j}$.

3.2 AD MRI Feature Extraction with Deep 3D Convolution Neural Networks (CNNs)

The utilization of 3D data as an input generally enhances the speed, despite a rise in memory costs and computational complexity due to high number of factors. As a result, we employed a deep 3D CNN. As opposed to 2D convolutions, the kernel of the 3D CNN slides in three dimensions. It is usually applied to MRI 3D image data for feature extraction and examination of the brain and other internal organs. The deep 3D CNN architecture effectively utilizes the network's capabilities by reusing features, resulting in compact models that are simple to train and excel at extracting high-quality AD features. We then train the deep 3D convolutional neural network. After training the network, we utilized the encoder's weights and biases from the learned architecture of the 3D CNN layer [34]. The 3D CNN topology improved on general feed-forward training by using spatial relationships to reduce the number of parameters that must be learned. The DenseNet121 architecture [35] was modified as three-dimensional layers for the deep convolutional network used in this work. It is known for its dense connections, which allow the network to learn rich feature representations from input data. The DenseNet121 architecture consists of series of convolutional and pooling layers, followed by a series of dense blocks. Each dense block contains a series of convolutional layers, where the output of each layer is concatenated with the input of the subsequent layer. This creates a dense connectivity pattern, where each layer receives direct connections from all previous layers. The dense blocks are separated by transition layers, which perform convolutional operations and down sampling to reduce the spatial dimensions of the feature maps as shown in **Fig. 4**.

Fig. 4. Examples of convolutions with the 3D deep densenet121.

The conventional DenseNet model consists of multiple layers, denoted as l, where z_l represents the output of the lth layer. Each layer implements a non-linear transformation, denoted as $M_l(.)$, with l indexing the layer. To address the issue of vanishing gradients and enhance information flow during network training, DenseNet employs connections from a layer to all subsequent layers.

In this study, we extend the concept of dense connectivity to the task of 3D volumetric image classification. Specifically, l is defined in (2) as:

$$z_l = M_l([\mathrm{z}_0, \mathrm{z}_1, \mathrm{z}_2, ..., \mathrm{z}_{l-1}]) \tag{2}$$

Here, $z_0, z_1, \dots, z_{l-1}$ represent 3D feature volumes generated in previous layers, and [...] denotes the concatenation function. The function $M_l(.)$, encompasses three main operations: a batch normalization (BN) layer to reduce internal covariate shift, spatial convolution employing k 3×3×3 convolution kernels to generate 3D feature volumes, and a rectified linear unit (ReLU) to expedite the training phase.

A dense unit within the DenseNet framework consists of a single layer within a dense block, where each layer in the block is connected to all subsequent layers. This dense connectivity facilitates efficient utilization of features and limits feature growth for each layer compared to traditional convolutional neural networks (CNNs). Consequently, the DenseNet models are more compact and possess fewer parameters than other network architectures. Although DenseNet can be configured with different depths, for this research, we employed the 121-depth configuration, excluding the last linear layer and the classification layer from the original configuration.

Our 3D DenseNet121 architecture has a total of 121 layers, including 121 3D convolutional layers and 121 dense blocks. It has a receptive field size of $121 \times 121 \times 121$ pixels, and it is trained on the AD 3D MRI data. The kernel size and stride size of the convolutional layers are $7 \times 7 \times 7$ and $2 \times 2 \times 2$ respectively. For the Max pooling layers, the kernel size and stride size are $3 \times 3 \times 3$ and $2 \times 2 \times 2$ respectively. The DenseNet121 has been used for tasks like object detection, semantic segmentation, and image classification. We train the network with Adam optimizer with a learning rate of 0.005.

3.3 Feature Probability Map with Attention Network

The discriminative AD features extracted from the deep 3D Convolution Neural Networks are further passed through an attention network to learn very discriminative feature probability map for the MRI features. Specifically, a 3D CNN attention model is trained on the output of the deep 3D convolutional network, and the resulting features are used to create a probability map that enhances the ability to discriminate between different AD classes. The learned probability map is then utilized as the input for the MLP for classification and prediction purpose. The network was built using a method of visual attention to ensure its effectiveness. The attention mechanism targets the dataset's most pertinent characteristic that is required for classification. Each input may be given an attention weight that indicates the relative relevance of that input element compared to the other input elements. This instructs the model to focus on specific inputs that are more crucial for carrying out the task at hand.

This study leverages on attention mechanism that enables the network to selectively examine specific feature maps in every layer, as depicted in **Fig. 5**.
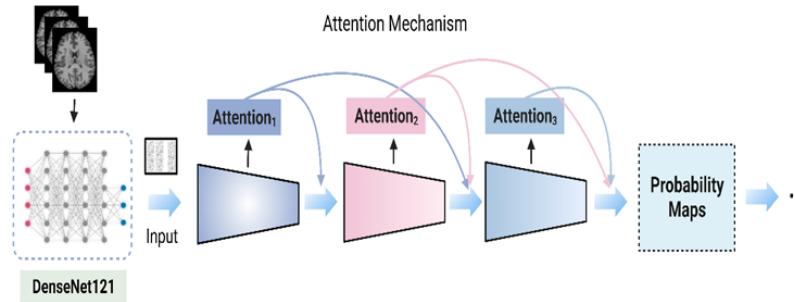
Fig. 5. The AD probability map extraction mechanism with attention network.

Instead of combining all previous layers' features like in DenseNet, the proposed attention mechanism uses a probability map algorithm to selectively integrate features from previous layers through a weighted sum, where the weights are learned during training. This approach simplifies and improves the efficiency of the network by emphasizing the most important information. The network is composed of three types of layers: input, where images are fed in, convolutional, where filters are applied to produce feature maps, and the third one is the attention layer which use the connection-wise attention model to direct the focus on specific feature maps in each layer. The probability map was built with three attention modules. Each attention module consists of three convolutional layers preceded by a max pooling layer and a SoftMax activation layer. The convolution filters used were $1\times1\times1$ in size and, the third element of the architecture is the pooling layer where the max pooling method is employed. This technique decreases the size of feature maps in the spatial dimensions by retaining only the maximum value from each cube, thus preserving the most relevant features for image classification. In this research, we used the max pooling technique on $2\times2\times2$ regions and employed the SoftMax function as the activation function in these layers due to its effectiveness in CNNs. The final operation involved the use of input and output neurons that created a linear combination of all neurons from the previous layer, which were then passed through a nonlinearity. The fully connected layers' inputs and outputs were represented as 1D vectors and no longer retained their spatial location. The proposed model featured two unique characteristics: first, it employed the ResNet connection technique, which combined both identity function and output of the SoftMax through summation, resulting in a reduction of parameters during training. Also, it uses the feature maps of all previous layers to obtain the feature maps of the current layer, where a self-learned weighting coefficient is assigned to each layer to focus on more informative features. The attention probability map is carried out in each channel of the network layers, and the attention weight calculation of each region is processed in (3):

$$\begin{cases} \text{key}_i = \text{features}_i \cdot W^{(k)} \\ \text{value}_i = \text{features}_i \cdot W^{(v)} \\ \text{query}_i = \text{features}_i \cdot W^{(q)} \end{cases} \tag{3}$$

The features represented by i is the initial output of the MRI features. The convolution layers are represented by W^k, W^v, and W^q. We updated the output using the attention mechanism in (4) and (5):

$$\text{Weight}_{ij} = \text{softmax}\left(\text{query}_i \odot \text{key}_j\right) \tag{4}$$

$$\text{probability\_map}_i = \sum_j^r \left(\text{weight}_{ij} \odot \text{value}_j\right) \tag{5}$$

3.4 Classification layer with Multilayer Perceptron

Multilayer Perceptron is a popular pixel-based machine learning algorithm which maps set of input data to a set of outputs in a feedforward way. MLP network is a structure of interconnected nodes in multiple layers, including input, hidden, and output layers, all connected to their preceding and succeeding layers. The outputs of the individual node are weighted units that can distinguish non-linearly separable data [36]. It is a neural network classifier used to learn nonlinear spectral features at the pixel level regardless of statistical properties [37]. The algorithm of MLP is a simple model architecture with a shallow classifier of one or two feature representation levels. In essence, the MLP is a pixel-based classifier with a shallow structure (one or two layers) [38], where it predicts the membership association of a pixel for each class.

In this work, the features extracted from the 3D attention probability map are used as an input to the MLP for the three-binary AD classification tasks. The MLP classifier is capable of providing a different feature representation with strong complementary and thus, help in enhancing the final classification performance.

The cross-entropy loss function was utilized in the classification model. This loss function was chosen due to its effectiveness in both weakly-supervised discriminative localization and multi-task regression tasks. The loss function L is represented in (6) as:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i\log\left(\hat{y}_i\right) + (1 - y_i)\log\left(1 - \hat{y}_i\right)\right] \tag{6}$$

Where N is the number of samples, y_i is the true label, and $\hat{y}_i$ is the predicted probability of the positive class.

4. Experiments

4.1 Dataset

The training data employed in this research was acquired from the ADNI database [39]. The ADNI is a public initiative that provides verifiable clinical and imaging datasets for Alzheimer's disease researchers. In this research, 2,160 participants were recruited from the ADNI database. The dataset contains participants with threshold T1 and T2 weighted scans from ADNI-1, ADNI-2, and ADNI-GO. The dataset was separated into three groups. The groups were: cognitively healthy individuals (Mini-Mental State Examination (MMSE) score greater than 24, Clinical Dementia Rating (CDR) score of 0, and no depression), Alzheimer's disease patients (MMSE score between 26, CDR score greater than 0.5), and mild cognitive impairment patients (MMSE score greater than 24, CDR score of 0.5, with objective memory loss). **Table 2** shows the dataset's demographic information.

Table 2. Dataset Demographics and Characteristics

		AD	MCI (with objective memory loss)	NC (non-depressed)
Sex	Male	300	410	481
	Female	310	340	319
Age (y)		75.2 ± 7.2	71.4 ± 6.2	65.2 ± 4.0
CDR		1.2 ± 0.5	0.5 ± 0.2	0 ± 0
MMSE		19.09 ± 5.55	27.40 ± 1.70	29.10 ± 1.0

To facilitate comparison across subjects, the original MRI images were transformed into standard space employing affine transformation and re-sampled to a uniform size of $224 \times 224 \times 224$. This resulted in a total of 2,260 records for training and testing, including 710 in AD category, 750 in MCI category, and 800 in NC category.

4.2 Experimental Setup

In our study, we randomly divided each group into training, validation, and testing sets, with 80% (1808 MRI scans), 10% (226 MRI scans), and 10% (226) of subjects respectively. To enhance the robustness of the 3D Magnetic Resonance Imaging (MRI) training dataset and to mitigate the risk of overfitting, data augmentation was applied. This methodology facilitates the creation of additional images, consequently expanding the breadth of training examples available to the proposed deep learning model, thereby curtailing the likelihood of overfitting. As for data augmentation, we use a series of techniques to increase the volume and diversity of our training dataset. These techniques include: (1) Rotation: We randomly rotate the MRI images along the three axes to simulate the varying angles at which scans can be captured in a real-world scenario, (2) Flipping: We perform random flipping in the 3D space to create additional variability in the dataset, (3) Scaling: We also perform random scaling of the MRI images to simulate the varying sizes of brains in the population. With a rigorous application of these augmentation techniques, the dataset corresponding to each of the three subject groups was magnified by a factor of three. Consequently, this process culminated in a comprehensive training dataset composed of 5424 MRI images. Specifically, 1704 training MRI images for AD, 1800 for MCI, and 1920 for NC images after augmenting the initial training MRI training set.

The classification method was implemented using Pytorch in Python 3.7 and was run on a PC with NVIDIA RTX 3080 GPU in an Ubuntu 20.04 environment. We used the Adam optimizer with an initial learning rate of 0.005, momentum of 0.9, and batch size of 4. To prevent overfitting, we employed dropout regularization as well as L1 regularization in our network. The entire framework was trained for 100 epochs. We evaluated the performance of the proposed algorithm in three classification tasks: AD vs. NC, MCI vs. NC, and MCI vs. AD.

When evaluating the proposed model, we compared it against three popular networks: a simple CNN, VGG16, ResNet 110 and DenseNet models. The basic CNN had 6 3D convolutional layers with a kernel size of $2 \times 2 \times 2$, stride of $2 \times 2 \times 2$, 4 max pooling layers with a size of $2 \times 2 \times 2$, and 2 fully connected layers, Dropout was employed as a regularization technique to avoid overfitting and enhance model robustness. We modified the popular 2D convolutional VGG16 and Densenet121 to its respective 3D convolutional version to fit with the 3D MRI images. The parameter settings for all three networks were the same as those for the proposed algorithm.

4.3 Evaluation of the proposed model and comparison with other CNNs using classification performance metrics.

The proposed model demonstrated excellent classification performance, with an accuracy of 91.27% between discriminating AD and NC, 80.85% for discriminating MCI and NC, and 87.34% for discriminating AD against MCI. When evaluating the proposed method's performance, we compared it to other well-known convolutional neural networks such as the basic CNN, VGG16, ResNet110, and DenseNet using the same datasets. Our method showed superior results in all group classifications.

In all three classification tasks, the proposed method with attention probability map outperformed popular networks such as VGG16 and DenseNet. In the AD vs. NC task, the basic CNN had a classification performance (measured by sensitivity (SEN), specificity (SPE), and accuracy (ACC)) lower than 88%. The performance of our model and the aforementioned models were evaluated by calculating their respective accuracies (ACC), sensitivities (SEN), specificities (SPE) in (7), (8), and (9) respectively.

$$Accuracy = \frac{TN + TP}{TP + TN + FP + FN} \tag{7}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

$$Specificity = \frac{TN}{TN + FP} \tag{9}$$

Where TN, TP, FP, and FN denote, true negatives, true positives, false positives and false negatives respectively.

The Densenet improved this accuracy to 88.90%, a 6.12% improvement over the basic CNN for the AD vs. NC classification task. The DenseNet, which enhances feature propagation and reuse, further improved accuracy to 82.45% for AD vs. MCI classification task. However, our proposed model attained the highest accuracy of 91.27%, 80.85%, and 87.34% for the classification tasks AD vs. NC, MCI vs. NC and AD vs. MCI respectively. As shown in **Fig. 6a**, **Fig. 6b**, and **Fig. 6c**, the area under curve (AUC) of all the tasks for our proposed technique was higher than 85%.
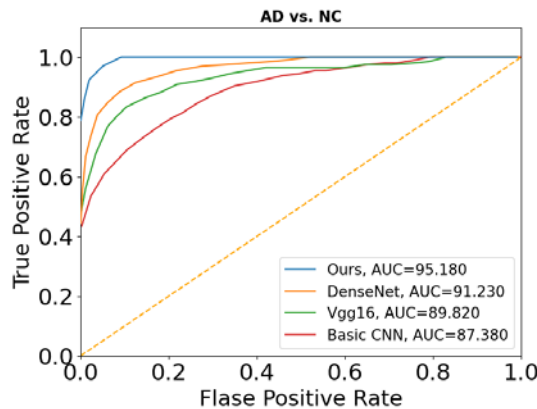

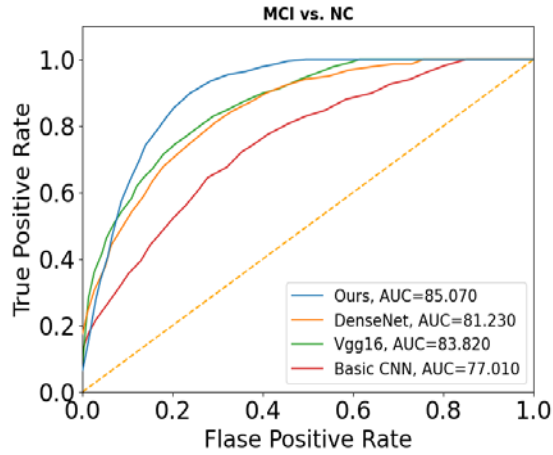**Fig. 6a.** Classification accuracy comparison for AD vs. NC.

2936

Frimpong et al.: Enhancing Alzheimer's Disease Classification using 3D Convolutional
Neural Network and Multilayer Perceptron Model with Attention Network

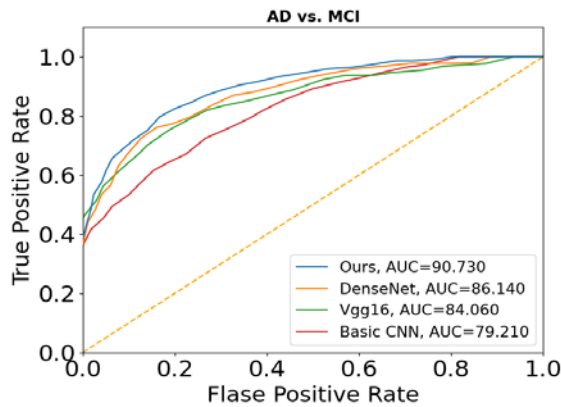**Fig. 6b.** Classification accuracy comparison for MCI vs. NC.

Fig. 6c. Classification accuracy comparison for AD vs. MCI

In the MCI vs. NC task, the basic CNN had an accuracy lower than 71%. Both the DenseNet and proposed method performed well, with accuracies of 88.90% and 91.27 % respectively, for AD vs. NC. In the MCI vs. NC task, the classification accuracies of the four competing methods were between 71%, 76%, 75%, and 74%. The basic CNN had an accuracy of 71.06%, with sensitivity and specificity of 70.66% and 70.00%. The VGG16 and DenseNet had accuracies of 76.10% and 75.12% respectively. Our proposed method again demonstrated the highest accuracy for all the AD classification tasks. The results are presented in **Table 3**. Furthermore, in **Table 4**, we compare the average performance of the proposed model with other existing models in [40-42]. Our model had the highest average performance.

Table 3. Comparison of the proposed network's classification results with other common networks on distinguishing AD vs. NC, MCI vs. NC, and AD vs. MCI.

Methods		Basic CNN	Vgg 16	DenseNet	ResNet 110	Ours
AD vs. NC	ACC	82.78	87.89	88.90	88	91.27
	SEN	83.09	84.50	87.32	85.91	88.73
	SPE	82.50	91.25	90.00	90.00	93.75
	AUC	87.38	89.82	91.23	90.12	95.12
MCI vs. NC	ACC	70.49	76.10	75.12	74.23	80.85
	SEN	70.66	76.00	74.66	73.33	77.33
	SPE	70.00	76.25	75.00	75.00	86.25
	AUC	77.01	83.92	81.83	81.74	85.07
AD vs. MCI	ACC	74.44	81.83	82.45	80.43	87.34
	SEN	73.23	80.28	81.69	78.87	84.50
	SPE	76.00	86.11	82.66	81.33	89.62
	AUC	79.21	84.06	86.14	85.63	90.73

Table 4. Average performance comparison with some state-of-the-art AD classification methods.

Authors	Dataset	Average Performance
Zhang et al [33]	ADNI	84.82%
Zhao et al [40]	ADNI	77.39%
Xin Bi et al [41]	ADNI	83.27%
Lian et al [42]	ADNI	82.63%
Our Model (Frimpong et al)	ADNI	91.27%

Comparing the confusion matrices in **Fig. 7**, **Fig. 8**, and **Fig. 9**, it is evident that our proposed method achieved the highest number of true positives (TP) and true negatives (TN) among all methods for all classification tasks (AD vs. NC, MCI vs. NC and AD vs. MCI). In **Fig. 7**, our proposed model demonstrated superior performance in correctly classifying both AD and NC samples. However, it also had a slightly higher number of false negatives (FN) compared to the other methods. This means that it generally makes fewer mistakes of both types: false positives and false negatives. Densenet exhibited a similar trend, with a relatively higher number of TP and TN and a slightly higher number of FN compared to VGG16 and Basic CNN. The VGG16 model displayed balanced performance regarding TP, TN, FN, and FP. Basic CNN showed a higher number of false positives (FP) and false negatives (FN) compared to the other methods, indicating potential challenges in correctly classifying both AD and NC samples. Densenet and vgg16 have comparable performances, with Densenet showing slightly better accuracy and precision but marginally less sensitivity. The model performance trend using the confusion matrix follows through for all the other classification tasks. The Basic CNN performs the worst among the four models in terms of accuracy, specificity, and sensitivity.

Overall, the confusion matrices provide insights into the performance of the baseline methods for the AD classification tasks. Our proposed method achieved a relatively high accuracy in

both AD and NC classification, followed closely by the Densenet and VGG16 methods. However, the Basic CNN method exhibited a higher misclassification rate, indicating a potential need for improvement. These findings can guide further research and development of classification models for Alzheimer's disease prediction.
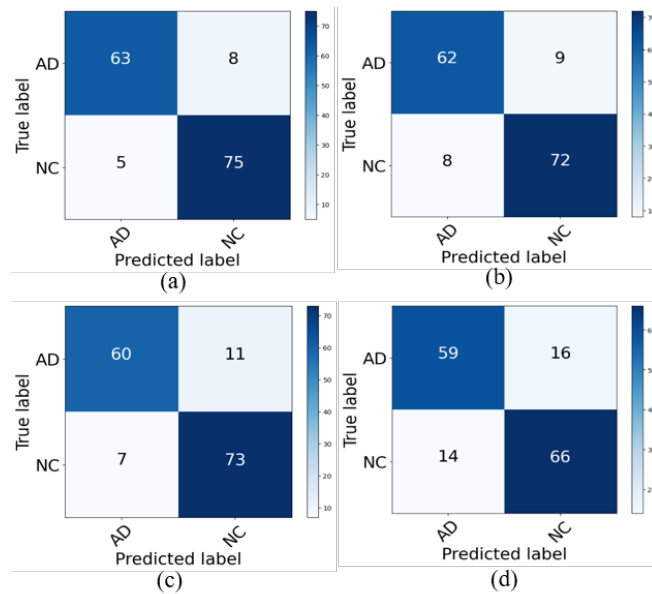

Fig. 7. Confusion matrix for all competing methods on AD vs. NC test data set: (a) Our (DenseNet + Attention Mechanism) (b) DenseNet, (c) Vgg16, (d) Basic CNN.
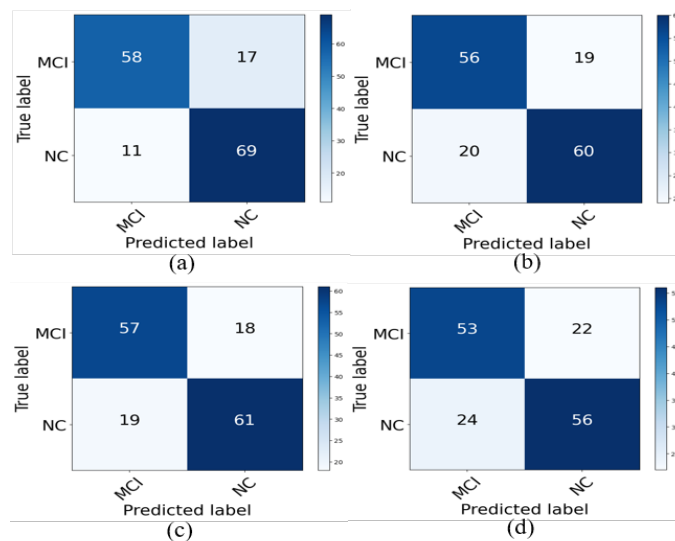

Fig. 8. Confusion matrix for all competing methods on MCI vs. NC test data set: (a) Our (DenseNet + Attention Mechanism), (b) DenseNet, (c) Vgg16, (d) Basic CNN.
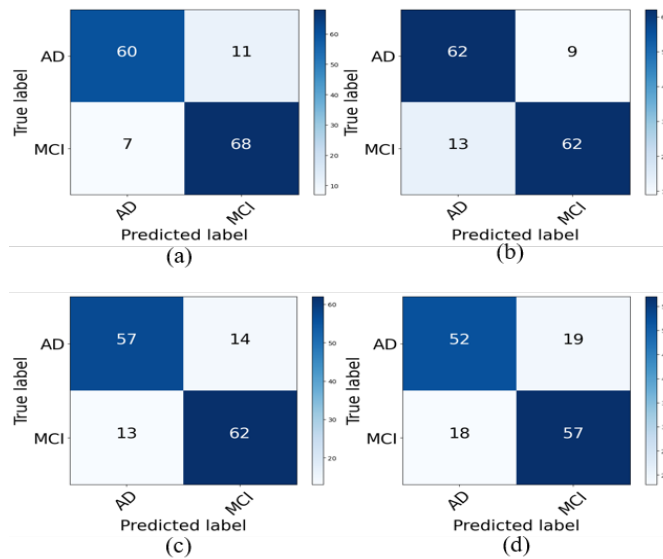
Fig. 9. Confusion matrix for all competing methods on AD vs. MCI test data set: (a) Our (DenseNet + Attention Mechanism), (b) DenseNet, (c) Vgg16, (d) Basic CNN.

5. Discussion

Early diagnosis of AD is crucial for effective disease management and intervention. Researchers are striving to develop computer-assisted systems that can identify Alzheimer's disease (AD) in its early stages. Our proposed approach, a densely connected convolutional network with a connection-wise attention map probability technique, represents a shift away from conventional, manually crafted feature-based methods. It aims to identify and predict AD or Mild Cognitive Impairment (MCI) using MRI scans of the brain.

The utilization of convolutional neural networks (CNNs) for image-based diagnosis and disease categorization has become popular. Although utilizing deeper and more intricate networks can result in higher classification accuracy, it also brings problems like gradient vanishing or explosion. Conventional CNN designs pass feature information in a one-way, one-time flow, resulting in restricted interaction between feature maps from different layers. Our experiments showed that the basic CNN method had a classification accuracy of 82.78%, 70.49%, and 74.44% for AD vs.NC, MCI vs. NC, and AD vs. MCI classification, which was not as good as the performance of other networks. The introduction of the Vgg16 network marked a significant advancement in deep learning methods, as it introduced a robust architecture in the AD classification tasks. The DenseNet took the idea of residual networks a step further by introducing a new structure with direct connections between any two layers. This architecture allows each layer to receive input from the outputs of all preceding layers and sends the feature map created by the current layer as input to all subsequent layers. This design effectively addresses the issue of gradient vanishing, promotes feature reuse, and greatly reduces the number of parameters. As a result, the DenseNet achieved a classification accuracy of 88.90%, 75.12%, and 82.45% for the classification tasks of AD vs. NC, MCI vs. NC, and AD vs. MCI, respectively.

In this work, we introduce a robust network architecture that incorporates connection-wise attention probability mapping for the purpose of AD classification. Traditional attention mechanisms can be grouped into two main categories: spatial attention, which concentrates on

regions of an image containing crucial information, and channel attention, which enables the network to focus on specific filters for improved classification accuracy. Our proposed attention mechanism diverges from these as it centers on the feature maps between different layers of the network. While the DenseNet merges features from different layers, the contribution of each layer to classification performance may differ, and treating all information as equally important can result in redundancy and increased computation. By utilizing these weighting parameters, the network can automatically adjust the contribution of each layer to enhance classification efficiency. Our experimental results demonstrate that the proposed method outperforms Basic CNN and DenseNet architectures in terms of classification accuracy.

There are limitations to the proposed technique. Although we minimized the size of the input data rather than using whole brain images, the method is still computationally expensive for training. Additionally, altering the deep CNN model's parameters, like the number of layers and the size and number of kernels in each layer, did not fully resolve the convergence issue. Reducing the patch size to 4 of the current size may help improve convergence and reduce training time, but this also requires a trade-off between classification efficiency and effectiveness.

6. Conclusion

In this study, we present a method for predicting AD based on MRI data. The proposed model employs the benefits of both Convolutional Neural Networks and Multilayer Perceptron. Our model uses a 3D CNN to effectively extract features from 3D MRI scans, which is an enhancement over traditional 2D CNNs used in similar research. Furthermore, our model integrates an attention mechanism within the 3D CNN, which helps it automatically focus on the most relevant features in the MRI scans. This is a notable departure from many models that rely on manual feature selection and is particularly beneficial in medical imaging where the correct feature selection can be challenging. Following feature extraction, the generated probability maps serve as input for an MLP classifier, which handles the final classification task. The MLP allows for robust decision boundaries, making it adept at distinguishing between the AD, MCI, and NC categories.

We achieved classification accuracies of 91.27 %, 80.85 %, and 87.34 % for distinguishing AD from NC, MCI from NC, and MCI from non- AD, respectively. Our method outperforms general neural networks on the same dataset, demonstrating top-ranked classification accuracy for the detection of the AD classification task. However, there are some limitations thus, the proposed method been computationally expensive for training, and not fully resolving convergence issues altering the parameters. These issues may be resolved by reducing the patch size, but there might be a trade-off with respective to classification efficiency. In future work, we will be addressing convergence issues, and ways to reduce computational cost of training, without compromising on models' efficiency.

Acknowledgment

References

[1] F. Ammarah, S. M. Anwar, M. Awais, and S. Rehman, "A deep CNN based multi-class classification of Alzheimer's disease using MRI," in *Proc. of IEEE International Conference on Imaging systems and techniques (IST)*, pp. 1-6, 2017. Article (CrossRef Link)

[2] B. Ron, E. Johnson, K. Ziegler-Graham, and H. M. Arrighi, "Forecasting the global burden of Alzheimer's disease," *Alzheimer's & dementia*, vol. 3, no. 3, pp. 186-191, 1 July 2007. Article (CrossRef Link)

[3] S. J. Dennis, and J. Hardy, "The amyloid hypothesis of Alzheimer's disease at 25 years," *EMBO molecular medicine*, vol. 8, no. 6, pp.595-608, Jun 2016. Article (CrossRef Link)

[4] Alzheimer Association, "Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, pp.327-406, 2021. Article (CrossRef Link)

[5] F. P. Cleusa, M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, M. Ganguli, K. Hall, "Global prevalence of dementia: a Delphi consensus study," *The lancet*, vol. 366, no. 9503, pp.2112-2117, 17 Dec 2005. Article (CrossRef Link)

[6] Osmosis. *Alzheimer's disease – plaques, tangles, causes, symptoms & pathology* [Video]. YouTube. (2016, March 22). [Online]. Available: https://www.youtube.com/watch?v=v5gdH_Hydes&t=31s

[7] S. Monika, S. Ahuja, S. Rani, D. Koundal, A. Zaguia, and W. Enbeyle, "An Exploration: Alzheimer's Disease Classification Based on Convolutional Neural Network," *BioMed Research International*, Jan 2022. Article (CrossRef Link)

[8] G. Emilie, G. Chételat, M. Chupin, R. Cuingnet, B. Desgranges, H. Kim, M. Niethammer, "Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging," *Neuroimage*, vol. 47, no. 4, pp.1476-1486, 1 Oct 2009. Article (CrossRef Link)

[9] K. Stefan, C. M. Stonnington, C. Chu, B. Draganski, R. I. Scahill, J. D. Rohrer, N. C. Fox, C. R. Jack Jr, J. Ashburner, and R. S. J. Frackowiak, "Automatic classification of MR scans in Alzheimer's disease," *Brain*, vol. 131, no. 3, pp.681-689, 1 March 2008. Article (CrossRef Link)

[10] F. Montenegro, J. Manuel, B. Villarini, A. Angelopoulou, E. Kapetanios, J. Garcia-Rodriguez, and V. Argyriou, "A survey of alzheimer's disease early diagnosis methods for cognitive assessment," *Sensors*, vol 20, no. 24, p.7292, 18 Dec 2020. Article (CrossRef Link)

[11] E. M. Amir, S. Luo, and R. Chiong, "Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review," *Computer methods and programs in biomedicine*, vol. 187, p.105242, 1 April 2020. Article (CrossRef Link)

[12] L. P. Jason, J. Pruessner, A. P. Zijdenbos, D. L. Collins, S. J. Teipel, H. Hampel, and A. C. Evans, "Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls," *Neurobiology of aging*, vol. 29, no. 1, pp.23-30, 1 Jan2008. Article (CrossRef Link)

[13] S. Alessia, A. Cerasa, and A. Quattrone, "Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review," *Frontiers in aging neuroscience*, vol. 9, p.329, 6 Oct 2017. Article (CrossRef Link)

[14] L. Manhua, D. Cheng, K. Wang, and Y. Wang, "Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis," *Neuroinformatics*, vol. 16, no. 3, pp.295-308, Oct 2018. Article (CrossRef Link)

[15] I. Jyoti, and Y. Zhang, "Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks," *Brain informatics*, vol. 5, no. 2, pp. 1-14, Dec 2018. Article (CrossRef Link)

[16] I. Jyoti, and Y. Zhang, "An ensemble of deep convolutional neural networks for Alzheimer's disease detection and classification," in *Proc. of NIPS 2017 Workshop on Machine Learning for Health*, 2 Dec 2017. Article (CrossRef Link)

[17] M. R. Krishn, S. Urolagin, J. A. A. Jothi, A. S. Neogi, and N. Nawaz, "Deep learning-based sentiment analysis and topic modeling on tourism during Covid-19 pandemic," *Frontiers in Computer Science*, vol. 3, p.775368, 5 Nov 2021. Article (CrossRef Link)

[18] I. Jyoti, and Y. Zhang, "A novel deep learning based multi-class classification method for Alzheimer's disease detection using brain MRI data," in *Proc. of International conference on brain informatics*, Springer, Cham, pp. 213-222, 16 Nov 2017. Article (CrossRef Link)

[19] H. Ehsan, R. Keynton, and A. El-Baz, "Alzheimer's disease diagnostics by adaptation of 3D convolutional network," in *Proc. of IEEE international conference on image processing (ICIP)*, IEEE, pp. 126-130, 25 Sep 2016. Article (CrossRef Link)

[20] F. Feng, P. Wang, K. Zhao, B. Zhou, H. Yao, Q. Meng, L. Wang, "Radiomic features of hippocampal subregions in Alzheimer's disease and amnestic mild cognitive impairment," *Frontiers in aging neuroscience*, vol. 10, p. 290, 25 Sep 2018. Article (CrossRef Link)

[21] I. Arevalo-Rodriguez, N. Smailagic, M. i Figuls, A. Ciapponi, E. Sanchez-Perez, A. Giannakou, O. Pedraza, X. Bonfill Cosp, S. Cullum, "Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI)," *Cochrane database of systematic reviews*, vol. 3, 2015. Article (CrossRef Link)

[22] F. J. Huff, F. Boller, F. Lucchelli, R. Querriera, J. Beyer, S. Belle, "The neurologic examination in patients with probable Alzheimer's disease," *Archives of Neurology*, vol. 9, no. 44, pp. 929-932, 1987. Article (CrossRef Link)

[23] Mayo Clinic Staff. Learn How Alzheimer's Is Diagnosed. [Online]. Available: https://www.mayoclinic.org/diseasesconditions/alzheimers-disease/in-depth/alzheimers/art-20048075 2019

[24] C. Ledig, A. Schuh, R. Guerrero, R. A. Heckemann, D. Rueckert, "Structural brain imaging in Alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database," *Scientific reports*, vol. 1, no. 8, p.11258, 2018. Article (CrossRef Link)

[25] Y. Huang, J. Xu, Y. Zhou, T. Tong, X. Zhuang, & Alzheimer's Disease Neuroimaging Initiative (ADNI), "Diagnosis of Alzheimer's disease via multi-modality 3D convolutional neural network," *Frontiers in neuroscience*, vol. 13, p. 509, 2019. Article (CrossRef Link)

[26] G. Liang, X. Xing, L. Liu, Y. Zhang, Q. Ying, A. Lin, N. Jacobs, "Alzheimer's disease classification using 2d convolutional neural networks," in *Proc. of 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3008-3012, 2021. Article (CrossRef Link)

[27] K. Gunawardena, R. Rajapakse, N. Kodikara, "Applying convolutional neural networks for pre-detection of alzheimer's disease from structural MRI data," in *Proc. of 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, pp. 1-7, November 2017. Article (CrossRef Link)

[28] K. Aderghal, M. Boissenin, J. Benois-Pineau, G. Catheline, K. Afdel, "Classification of sMRI for AD Diagnosis with Convolutional Neuronal Networks: A Pilot 2-D+ Study on ADNI," in *Proc. of MultiMedia Modeling: 23rd International Conference, MMM 2017, Proceedings, Part I*, pp. 690-701, 2017. Article (CrossRef Link)

[29] S. Luo, X. Li, and J. Li, "Automatic Alzheimer's disease recognition from MRI data using deep learning method," *Journal of Applied Mathematics and Physics*, vol. 5, no. 9, pp. 1892–1898, 2017. Article (CrossRef Link)

[30] S.-H. Wang, P. Phillips, Y. Sui, B. Liu, M. Yang, and H. Cheng, "Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling," *Journal of Medical Systems*, vol. 42, no. 5, p. 85, 2018. Article (CrossRef Link)

[31] M. Liu, D. Cheng, W. Yan, & Alzheimer's Disease Neuroimaging Initiative, "Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images," *Frontiers in neuroinformatics*, vol. 12, p. 35, 2018. Article (CrossRef Link)

[32] E. G. Marwa, H. Moustafa, F. Khalifa, H. Khater, E. AbdElhalim, "An MRI-based deep learning approach for accurate detection of Alzheimer's disease," *Alexandria Engineering Journal*, vol. 63, pp. 211-221, 2023. Article (CrossRef Link)

[33] X. Zhang, L. Han, W. Zhu, L. Sun, D. Zhang, "An explainable 3D residual self-attention deep neural network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI," *IEEE journal of biomedical and health informatics*, vol. 11, no. 26, pp. 5289-5297, 2022. Article (CrossRef Link)

[34] A. Ardekani, A. Bachman, "Model-based automatic detection of the anterior and posterior commissures on MRI scans," *Neuroimage*, vol. 46, no. 3, pp. 677-682, 1 Jul 2009. Article (CrossRef Link)

[35] J. Mark, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp.825-841, 1 Oct 2002. Article (CrossRef Link)

[36] L. Christian, R. Wolz, P. Aljabar, J. Lötjönen, R. A. Heckemann, A. Hammers, D. Rueckert, "Multi-class brain segmentation using atlas propagation and EM-based refinement," in *Proc. of 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pp.896-899, 2 May 2012. Article (CrossRef Link)

[37] K. Zhaokai, M. Zhang, W. Zhu, Y. Yi, T. Wang, and B. Zhang, "Multi-modal data Alzheimer's disease detection based on 3D convolution," *Biomedical Signal Processing and Control*, vol. 75, p.103565, 1 May 2022. Article (CrossRef Link)

[38] S.R. Braulio, R. Villalón-Fonseca, and G. Marín-Raventós, "Alzheimer's disease early detection using a low cost three-dimensional densenet-121 architecture," in *Proc. of International conference on smart homes and health telematics*, pp. 3-15, 2020. Article (CrossRef Link)

[39] J. R. Clifford, M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 27, no. 4, pp.685-691, Apr 2008. Article (CrossRef Link)

[40] Y. Zhao, B. Ma, P. Jiang, D. Zeng, X. Wang, S. Li, "Prediction of Alzheimer's disease progression with multi-information generative adversarial network," *IEEE Journal of Biomedical and Health Informatics*, vol. 3, no. 25, pp. 711-719, 2020. Article (CrossRef Link)

[41] X. Bi, X. Zhao, H. Huang, D. Chen, Y. Ma, "Functional brain network classification for Alzheimer's disease detection with deep features and extreme learning machine," *Cognitive Computation*, vol. 12, pp. 513-527, 2020. Article (CrossRef Link)

[42] C. Lian, M. Liu, J. Zhang, D. Shen, "Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI," *IEEE transactions on pattern analysis and machine intelligence*, vol. 4, no. 42, pp. 880-893, 2020. Article (CrossRef Link)

ENOCH A. FRIMPONG completed his M.Eng. Degree in Computer Science and Engineering at the University of Electronic Science and Technology of China (UESTC), in 2017, where he is currently pursuing the Ph.D. in software engineering. He is currently a Lecturer/researcher at the Cape Coast Technical University, Ghana. His research interests include machine learning, artificial intelligence and IoT.

QIN ZHIGUANG is a full professor at the School of Information and Software Engineering in the University of Electronic Science and Technology of China (UESTC). He is a recipient of special national allowance, a member of the Computer Science and Technology Group of the Sixth and Seventh Discipline Evaluation Group of the State Council, a member of the second and third decision-making advisory committees of the Sichuan Provincial Party Committee and Provincial Government, and chairman of the Sichuan Software Industry Association. He is a famous teacher in Sichuan Province and an academic and technical leader in Sichuan Province. He has won one second-class national teaching achievement award; three second-class provincial and ministerial scientific and technological progress awards; and one first-class teaching achievement award in Sichuan Province. He has undertaken several major national and key projects.

REGINA E. TURKSON completed her Ph.D. in computer science at the University of Electronic Science and Technology of China, she is currently a Lecturer/researcher at the University of Cape Coast, Ghana. Her research interests include machine learning, artificial intelligence, and computer security.

COBBINAH M. BERNARD completed his master's degree in computer science at the University of Electronic Science and Technology of China and is currently pursuing a PhD in the same field at the same university. His research focuses on areas such as federated learning, data stream mining, and neuroimaging.

EDWARD Y. BAAGYERE is a Computer Scientist and an Associate professor in the Department of Computer Science, at C. K. Tedam University of Technology and Applied Sciences (CKT-UTAS). He received the B.Sc. degree (Hons.) in Computer Science from the University for Development Studies (UDS), Tamale, Ghana, in 2006, the M.Phil. degree in Computer Engineering from the Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana, in 2011, and the Doctor of Engineering (DEng) degree in Computer Science and Technology from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 2016. Prof. Baagyere was a Postdoctoral Fellow at the Network and Data Security Key Laboratory of Sichuan Province at UESTC, Chengdu, China between December 2016, and August 2021. Prof. Baagyere currently serves as the Dean of the School of Computing and Information Sciences at C. K. Tedam University of Technology and Applied Sciences, Navrongo, Ghana. His current research interests include Deep/Machine learning, mobile sensor networks, cryptography, and social networks.

Edwin K. Tenagyei is currently a PhD student in Griffith University under the school of Engineering and Built Environment. He received M.Eng. and B.Eng. Degree in software engineering from the University of Electronic Science and Technology of China in 2021 and 2019 respectively. His current research interests include deep learning, and computer vision.