



Automated Alzheimer's disease diagnosis using radiomics feature extraction on magnetic resonance images

T. Manochandar^{1*}



P. Kumaraguru
Diderot²

^{1,2}Department of ECE, Hindustan Institute of Technology and Sciences,
Padur, Chennai, India.

¹Email: manochandarresearchscholar@gmail.com

²Email: pkguru@hindustanuniv.ac.in



(+ Corresponding author)

ABSTRACT

Article History

Received: 1 August 2023

Revised: 7 November 2023

Accepted: 20 November 2023

Published: 22 December 2023

Keywords

ADNI dataset
Alzheimer's disease
Deep learning
MRI images
Neuroimaging data.

The current study investigates automated Alzheimer disease diagnosis using radiomic feature extraction on magnetic resonance images. Alzheimer's disease (AD) is a progressive, incurable neurological brain disorder. Early diagnosis of AD might prevent brain tissue damage and assist with proper treatment. Researchers examine numerous statistical and machine learning (ML) techniques for the detection of AD. Analyzing magnetic resonance imaging (MRI) is a traditional way of analyzing for AD in clinical examination. Diagnosis of AD is challenging because of the similarity in MRI statistics and standard healthy MRI information for elderly people. The application of DL to earlier diagnosis and automatic classification of AD is gaining immense popularity in recent times, as rapid development in neuroimaging techniques has generated large-scale multimodal neuroimaging datasets. This study develops a new Automated Alzheimer's Disease Diagnosis using Deep Learning Model (AADD-DLM) on MRI images. The presented AADD-DLM technique examines the MRI images to assist in the AD diagnostic process. In the presented AADD-DLM technique, three major processes are involved, namely skull stripping, segmentation, and feature extraction. Initially, the AADD-DLM technique uses the U-Net model for the skull stripping process, which enables the removal of the skull regions in the brain MRI. Next, the QuickNAT model is utilized for an effective brain MRI segmentation process. Moreover, the radiomics feature extraction approach is used to generate a useful set of feature vectors. For exhibiting the promising performance of the AADD-DLM technique, widespread experimentation analysis is made on the ADNI database. The optimized model achieves 99.6% accuracy in the ADNI database. The simulation outcomes revealed the improved effectiveness of the AADD-DLM technique over other recent approaches.

Contribution/Originality: The primary contribution of this study is using DL models to create a reliable MRI image analysis model for diagnosing AD. Present a U-Net-based skull stripping approach to remove the skull area from a brain MRI. Then a QuickNAT model was proposed for efficient MRI brain segmentation. To generate feature vectors, we introduce radiomics features and verify them using the ADNI benchmark dataset.

1. INTRODUCTION

MRI can be exploited to evaluate the anatomical structure of the brain owing to its ability to contrast soft tissues and its high spatial resolution [1]. Generally, MRI was linked with more health hazards than other modalities like PET and CT [2]. Understanding brain structure with an MRI and assessing strokes have both made tremendous strides in recent years. Brain-related disorders like multiple sclerosis and Alzheimer's disease

(AD) can be diagnosed with MRI [3]. The segmenting process of brain MRI captured at distinct times can also be utilized for measuring structural variations in the brain. When trying to diagnose something like Alzheimer's disease, it becomes increasingly important to correctly categorize and distinguish harmful tissues and the surrounding healthy architecture [4].

An enormous volume of data is needed for more precise diagnoses. However it becomes a challenge for doctors who evaluate complex and large MRI datasets and abstract imperative data manually [5]. In addition, there are other issues with intra- or inter-operator heterogeneity that make manual reading of a brain MRI tedious and error-prone [6]. Therefore, it can be essential to formulate an automatic segmentation technique to offer precise outcomes with great confidence. Computerized methods for registration, MRI segmentation, and visualization were utilized on large-scale datasets to assist medical practitioners in making qualitative diagnoses [7]. Deep learning (DL) is an advanced method; techniques like convolutional neural networks (CNNs), stacked AE auto-encoders (SAE), and deep belief networks (DBNs) can mechanically construct further abstract high-level representations of learning mechanisms by compiling lower-level features embedded under the dataset. The CNN technique is being commonly exploited for object detection, classification, and segmentation, having many advantages [8]: CNN is able to receive images directly as input, use spatial data embedding from neighboring pixels, and efficiently lessen the model parameter count with the use of subsampling, local receptive domains, and weight sharing [9]. If a CNN technique can be trained with MRI scanner, image features can be mechanically retrieved, eradicating the necessity of selecting the features manually for learning [10]. In the meantime, ensemble learning benefited from its robustness and performance by compiling many learning systems, which were implemented in MRI.

This study develops a new AADD-DLM on MRI images. There are primarily three steps in the described AADD-DLM method: skull stripping, segmentation, and feature extraction. Initially, the AADD-DLM technique uses the U-Net model for the skull stripping process, which enables the removal of the skull regions in the brain MRI. Next, the QuickNAT model is utilized for an effective brain MRI segmentation process. Moreover, radiomics feature extraction approach is used. Obtain a relevant collection of feature vectors of information. For exhibiting the promising performance of the AADD-DLM technique, widespread experimentation analysis is made on the ADNI dataset. In a nutshell, the paper's main achievements are outlined below:

- Employ an effective MRI image analysis model for AD diagnosis using DL models.
- Present a U-Net-based skull stripping process to remove the skull region from the brain MRI
- Propose a QuickNAT model for an accurate brain MRI segmentation process.
- Introduce radiomics features for the generation of feature vectors.
- Validate the results on the benchmark ADNI dataset.

2. RELATED WORKS

Kong, et al. [11] modelled an image fusion technique to merge PET images and MRIs from AD patients. Moreover, the author uses 3D-CNN to assess the efficacy of this image fusion technique in both multi-classification and dichotomous errands. The 3-D convolutions of the fused images are employed for extracting the data from the features, leading to richer multi-modal feature data. At last, the derived multi-modal traits were predicted and classified through a fully connected NN. In Alhassan [12], the authors presented an EFEHO for OTSU segmentation, called EFEHO-OTSU. Firstly, exploiting EFEHO in the suggested approach to seek the optimum segmentation threshold for the OTSU approach. Secondly, DA-MIDL was suggested for prompt analysis of AD and its prodromic stage, MCI.

Ajagbe, et al. [13] are meant for advancing AD image classifications with DCNN involving CNN and TL utilizing MRI and extended estimation metrics, as the capacity and limitations of methods cannot be exposed by some metrics. The purpose of this study was to use numerous assessment indicators to assess neurologists' ability to categorize AD images into four established categories. This study utilized computer methods, primarily TL and

DCNN, for classifying AD. Marwa, et al. [14] developed a DL-related pipeline for precise stratification and diagnosis of AD phases. The presented analysis pipeline uses 2D T1-weighted MRI brain images and shallow CNN structures. The presented pipeline presents an accurate and fast AD diagnosis module as well as local and global classifiers.

Fang, et al. [15] devised a new structure that ensembles 3 existing DCNNs with multi-modality imageries for an AD classifier. Moreover, to reject low discrimination slices of category probability, a ‘dropout’ system was presented. Then average reserved slices of all subjects were needed as new features. Liu, et al. [16] modelled a multi-model DL structure related to CNN for AD classification through structural MRI data and joint automatic hippocampal segmentation. Primarily, a multi-task DCNN technique was built for disease classifiers and jointly learning hippocampal segmentation. Afterward, the author framed 3D Dense Net to study of features of 3-D patches derived on the basis of the hippocampal segmentation outcomes for the classifier task.

3. THE PROPOSED MODEL

In this study, we have presented a novel AADD-DLM technique for automated AD diagnosis using brain MRIs. In the presented AADD-DLM technique, three major processes are involved, namely:

- U-Net-based skull stripping.
- QuickNAT-based segmentation.
- Radiomics feature extraction.

3.1. Skull Stripping Process

The areas of the skull are masked in MRIs of the brain by the U-Net method. The U-Net method comprises an encoder-decoder architecture with skip connections [17]. In all the encoding and decoding blocks, two replications of convolution layers are used with the same kernel size $K = (3,3,3)$. Every convolution layer is then non-linear activation and BN with ReLU. The primary count of feature maps, following the initial convolution layer, was set to $F = 32$ for each model, and the feature map count is doubled (halved) after every encoding (decoder) block. The dropout layer is used afterward in the encoding and decoding layers to implement dropout sampling for uncertainty quantification and to avoid over-fitting. Max-pooling afterward, every encoding block halves the feature map dimension. Similarly, up-sampling with transpose convolution afterward doubles the feature mapping size, and lastly, restores the primary dimension to the outcome. The number of max-pooling functions determines the depth D of U-Net framework that is set for $D = 4$ for each trained model. The bottleneck blocks restrict data flow from the encoding to decoding block and comprise 2 convolution layers, each afterward rectified by linear activation and batch normalization. Unlike the encoding and decoding blocks, we don’t apply a dropout layer to the bottleneck block. Due to memory limitations, the input brain volume is typically 2 million voxels in size, evenly dispersed across the imaging dimension. Then, train the network with the help of the Adam optimization using primary learning rates of 0.001. In the process of training, a set of random transformations is applied, for example, rotations and translations, to the volume for data augmentation.

The mixture of Dice scores is summed over each class label, and the categorical cross-entropy function:

$$\mathcal{L} = - \sum_{r \in \mathcal{S}} \left(\frac{2 \sum_x Q_r(x) T_r(x)}{\sum_x Q_r(x) + T_r(x)} - \sum_x T_r(x) \log(Q_r(x)) \right). \quad (1)$$

Now, $Q_r(x)$ indicates the softmax outcome of networks at voxel location x and $T_r(x)$ denotes the ground truth at a similar location. Theoretically, the impact of Dice loss and cross-entropy is further weighted, however, it is found small impact on performance and thereby neglect further weighting.

3.2. Image Segmentation Process

In this work, the Quick NAT model is exploited for the active brain MRI segmentation process, which recognizes various regions. Quick NAT [18] includes three 2D FCNN architectures, which segment the input image slice-wise alongside axial, sagittal, and coronal axes. After that, the view aggregation phase takes place, where the three segmentations generated are integrated to provide a concluding segmentation.

Every 2D F-CNN models have an encoding-decoding-based structure, with four encoding and decoding blocks detached by the bottleneck blocks. The dense connection was added within all the encoding and decoding blocks to promote learning of better representations and promote feature re-usability. A skip connection exists between all the encoding and decoding blocks, just as in U Net. Dice degradation and gravity nonlinear losses were optimized during model training. Quick NAT is trained by simultaneously augmenting 2 loss functions: (1) the multi-class Dice loss and (2) the weighted logistic loss. The logistic loss gives the pixel-wise probabilistic estimation of resemblance between the manually annotated and estimated labels.

The Dice loss has been stimulated by the Dice overlap ratio, which evaluates resemblance between the manually annotated and estimated labels. Initially, it was presented for two-class classification and extended to multi-class classification. Assuming the expected probability $Q(k)$ of k pixel belongs to class l and the actual class $g_l(k)$,

$$\mathcal{L} = \underbrace{-\sum_k \omega(k)g_l(k)\log(Q(k))}_{\text{LogisticLoss}} - \underbrace{\frac{2\sum_k Q_l(k)g_l(k)}{\sum_k Q_l^2(k) + \sum_x g_l^2(k)}}_{\text{DiceLoss}} \quad (2)$$

Multiclass logistic loss and Dice loss are two types of loss functions. Then, present a weight $\omega(k)$ that balances the relative significance of the pixels from the loss.

Next, apply weight to resolve 2 problems: (i) errors in segmentation at the anatomical boundary and (ii) class imbalance. Assuming the frequency f_l of classes l from the trained dataset, viz., the 2D gradient operator ∇ , the class prior probability, the trained segmentation S , and the indicator function \mathbb{I} , the weighted are determined by

$$\omega(x) = \sum_l I(S(x) = l) \frac{\text{median}(f)}{f_l} + \omega_0 \cdot \mathbb{I}(|\nabla S(x)| > 0) \quad (3)$$

With the vector of each frequency $f = [f_1, \dots, f_N]$.

The first term simulates median frequent balancing and compensates for class disparity issues by increasing the weight of rare classes in an image. The next term put high weights on anatomical boundary area for encouraging accurate segmentation of contour. ω_0 is fixed as $\frac{2 \cdot \text{median}(f)}{f_{\min}}$ for providing high priority to boundaries.

3.3. Radiomics Feature Extraction Process

For the generation of feature vectors, radiomics features are utilized in this work. We have extracted a set of 93 radiomics features. Every single feature class, excluding shape, can be computed either on the derived image or the original image, attained by using multiple filters. The shape descriptor is extracted from the label mask and is independent of gray value. It can be computed separately from the enabled input image varieties and computed on the original image. The radiomics feature is categorized as follows:

3.3.1. First Order Statistics

It describes the distribution of voxel intensity in the image region described by the mask with basic and widely utilized metrics.

3.3.2. Shape-Based (3D)

The attribute was unrelated to the gray-level brightness shipping from the ROI and estimated only on non-derived pictures and masks. The feature was built using an approximation of the shape calculated using a triangle mesh. Initial triangles for mesh construction are found by finding the midpoints of edges connecting voxels between and within the region of interest (ROI). A mesh of linked triangles is acquired by linking the vertices, with every triangle determined by 3 adjacent vertices that share all the sides accurately with other triangles. This mesh was generated with the marching cubes method. The procedure involves stimulating a 2×2 cube with a mask in the region. In every position, the corner of cubes is later marked 'segmentation' (1) or 'not segmentation' (0). By representing each corner with a distinct bit of a binary value, an individual cube index could be calculated (0-255). This index is then utilized for determining whether triangles can be present in the cube that is determined in the lookup table. These triangles can be determined as follows: the normal is continuously oriented in a similar way. It can be essential to attain the accurate signed volume utilized in the computation of MeshVolume.

3.3.3. Shape-Based (2D)

This attribute may only be calculated on masks and images that were not used to create the ROI, as it doesn't rely on the gray-level pattern of intensity within the ROI. The feature was developed in the estimated shape determined by the circumference mesh. First, the vertex edges are found by finding the midpoint of an edge between a pixel that has been added to the ROI and a pixel that has not. A mesh of connected lines is attained by interconnecting the vertices, with all the lines determined by 2 adjacent vertices that share points exactly with other lines.

This mesh was created utilizing a modified version of the marching cubes technique. During this process, a 2×2 square has been stimulated with the mask space (2d). To every position, the corner of squares is next marked 'not segmented' (0) or 'segmented' (1). To give the corners as certain bits in binary numbers, a single square index has been acquired (0-15). This index is then employed for determining whether lines can exist from the square that determines the lookup table. These lines are determined so that the normal triangle described by this point and origin is kept consistent. This outcome has signed value for the surface region of all the triangles; thereby, if summed, the superfluous (positive) region contained by triangles partly inside and outside the ROI has been correctly cancelled out by the (negative) area of triangles completely outside the ROI.

3.3.4. Gray Level Co-Occurrence Matrix

It calculates gray-level zone from the image. A gray-level zone can be determined as a count of interconnected voxels that share similar gray-level intensity. If the distance is 1, a voxel is considered connected based on the infinity norm. It is matrix in that the texture feature is extraction for texture investigation. The GLRLM scheme is a process of extracting higher-order statistical texture features. A gray-level run is determined as a pixel line from a specific direction with a similar intensity value.

3.3.5. Gray Level Run Length Matrix

It calculates gray-level zone from the image. A gray-level zone can be determined as a count of interconnected voxels that sharesimilar gray-level intensity. If the distance is 1, a voxel is considered connected based on the infinity norm. It is matrix in that the texture feature is extraction for texture investigation. The GLRLM scheme is a process of extracting higher-order statistical texture features. A gray-level run is determined as a pixel line from a specific direction with a similar intensity value.

3.3.6. Gray Level Size Zone Matrix

It is the beginning of Thibault matrices. For texture images f with N gray-levels, it can be represented by $GS_f(s, g)$ and provides a statistical demonstration by the estimate of the bivariate conditional probability density function of an image distribution value.

3.3.7. Neighbouring Gray Tone Difference Matrix

It measures the differences between the average gray value and gray value of their neighbors in the distance. The amount of entire differences for gray-level was kept from the matrix.

3.3.8. Gray Level Dependence Matrix

It enumerates gray-level dependency from the image. It can be determined as the quantity of interconnected voxels in the distance is based on center voxels. The main properties of this scheme are as follows: (i) texture feature is simply calculated; (ii) it can be fundamentally invariant in spatial rotation; (iii) it can be invariant in linear gray-level transformation and is insensitive to monotonic gray-level transformations.

3.4. Classification

Previous research has shown that the CNN architecture can be enhanced by increasing the precision of the methods used for training and enhancing the standard of the input information; furthermore, it is known that the degree to which a model can be optimized depends on the kind of model employed. In the case of multi-layered deep learning models, for instance, it is possible to fine-tune the training parameters. The convolution layer of a convolutional neural network is responsible for obtaining features, and the quality of removing features is proportional to the fourier kernel width. From a compositional standpoint, lines created by adjacent pixels frequently from the borders of a picture. The image's texture is made up of a collection of edge lines, and these lines are used to create various local patterns. The picture is composed mostly of its local pattern. The network model's convolution layer allows for feature extraction and local picture pattern formation in various forms. Although the convolution layer will extract more information with a smaller convolution kernel, overfitting issues may arise. However, if the convolution kernel too large, the convolution layer maybe unable to retrieve as many features, which will reduce the impact of picture classification. As a result, the precision of picture classification may be enhanced by optimizing the convolution kernel. Since the CNN model predominantly gathers the image characteristics layer by layer via multiple convolution layers, the number of convolution layers impacts the degree of accuracy of the model's extraction of characteristics. The features obtained by the model classifier become increasingly fine-grained with increasing numbers of convolution layers, which can result in over fitting, while the features become increasingly coarse-grained with decreasing numbers of convolution layers, which can result in a decline in image classification accuracy. Therefore, convolution layer modification may improve the model's precision for classification.

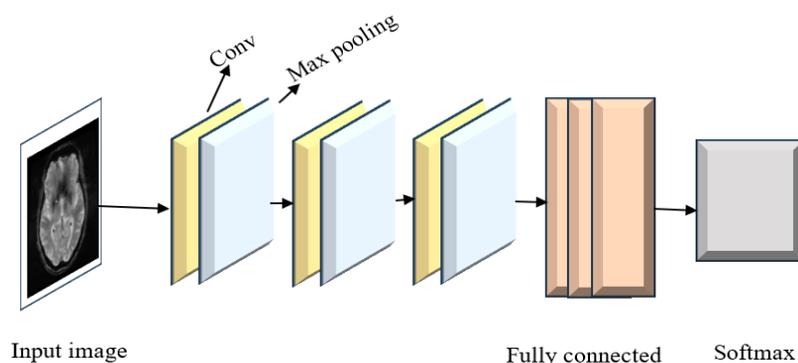


Figure 1. Shows the CNN architecture for the proposed model.

The depth learning model suggested in this research has to be refined to increase the accuracy of picture categorization and identification. A smaller convolutional kernel is selected in the first convolution layer so as to derive further details about the features of the picture. Second, the overfitting issue is dealt with by adopting the maximum pool sampling procedure into the model. Figure 1 shows the updated photo classification model, which has three convolution layers with diminishing convolution kernels in each. Following each convolution layer, the output features are supplied into the optimum pooling layer. After processing the input picture data through three full connection layers, the model produces the categorization result using a Softmax classifier.

The convolution layer's parameters for modeling are the dimension and the number of convolution kernels. Because the first layer of convolution is close to the image layer that contains the input and is mainly used for obtaining the fundamental elements of the image, its parameters significantly affect the feature set. To facilitate the computation of characteristics in the next convolution layer, attribute information such as shadow, boundary, and lighting of the image must be extracted using a smaller convolution kernel.

In the convolution layer, the activation function is used to create a map of the characteristics that were gathered. As a result, the enhanced CNN model's the ReLU activation function looks like the following mathematical function:

$$z(y) = \text{Max}(0, y) \quad (4)$$

When the activation function of ReLU is used for learning characteristics in a typical convolutional neural network model, it may result in the loss of significant characteristic data during picture categorization. Based on the current ReLU activation function, it may be enhanced to prevent the loss of relevant features during picture categorization. An expression for the optimal activation function is,

$$z(y_k) = \begin{cases} \frac{y_k}{d_k}, & y_k < 0 \\ y_k, & y_k \geq 0 \end{cases} \quad (5)$$

The novel activation function calculation formula not only preserves the negative data within the feature map whenever the feature being processed is less than zero, but it also enhances effective distinctive reinforcement learning.

The Softmax function, which uses a supervised learning method to regress the characteristics, is utilized by the optimised CNN to determine the photographs. The picture target category C can take on any one of K possible values throughout the categorization process. The cost function of Softmax regression may be written as, for a given image training set $\{a_1, a_2, \dots, a_n\}$ where a_i is a sample of the training set, b_i is a category for classifying the picture, and $b_i \in \{1, 2, \dots, K\}$ is the cost function.

$$s(\alpha) = \frac{1}{N} \left[\sum_{l=1}^L \sum_{m=1}^M 1\{b_l = n\} \log \left(\frac{\exp(\alpha_m^T a_l)}{\sum_{n=1}^M \exp(\alpha_m^T a_l)} \right) \right] \quad (6)$$

Provided M markers are added to the cost function, we can write the likelihood of classifying training sample as group n as

$$\lambda(b_l = n | a_l; \alpha) = \frac{\exp(\alpha_m^T a_l)}{\sum_{n=1}^M \exp(\alpha_m^T a_l)} \quad (7)$$

4. EXPERIMENTAL VALIDATION

During this study, the experimental outcomes of the presented method can be tested utilizing the benchmark ADNI database. A primary purpose of ADNI is to examine when the serial MRI, another biological marker, PET, and medicinal and neuropsychological investigation were integrated to measure the progress of MCI and main AD. A particular and sensitive marker of very soon AD progress has been assumed for helping scientists and researchers working to advance innovative remedies, and observing their performance, and minimizing medicinal trial time

and cost. The ADNI uses the findings of numerous co-investigators from large academic institutions and commercial enterprises, and its subjects can be used at more than 50 locations across the US and Canada. One of the main goals of ADNI is to enroll 800 adults between the ages of 55 and 90. Of them, about 200 are older, cognitively regular people who have been involved for three years, 400 are MCI patients who have been involved for three years, and 200 are primary AD patients who have been involved for two years.

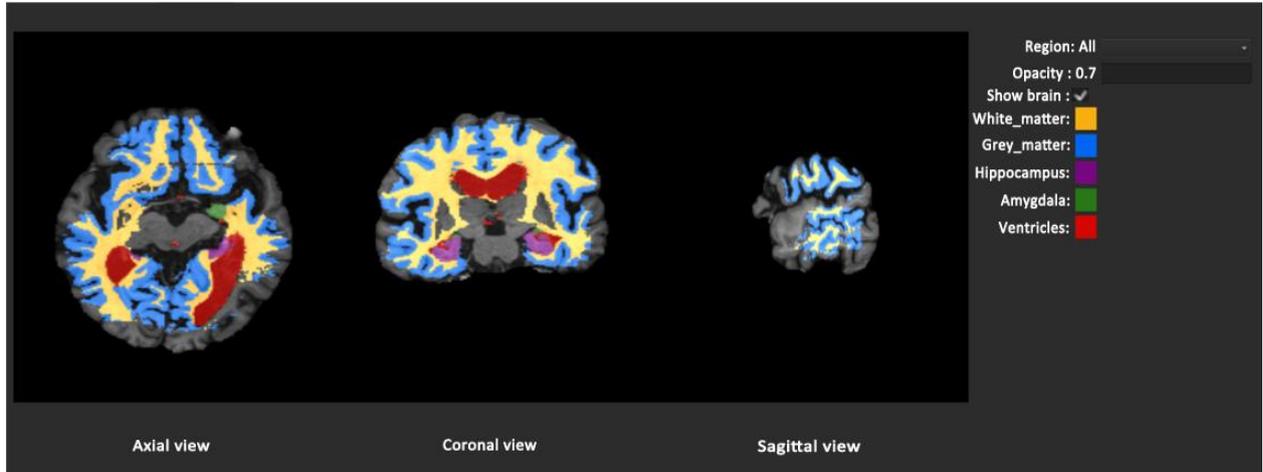


Figure 2. Sample MRI segmentation results.

Figure 2 shows the sample segmentation results obtained by the proposed model. The features extractions by the proposed technique are given below.

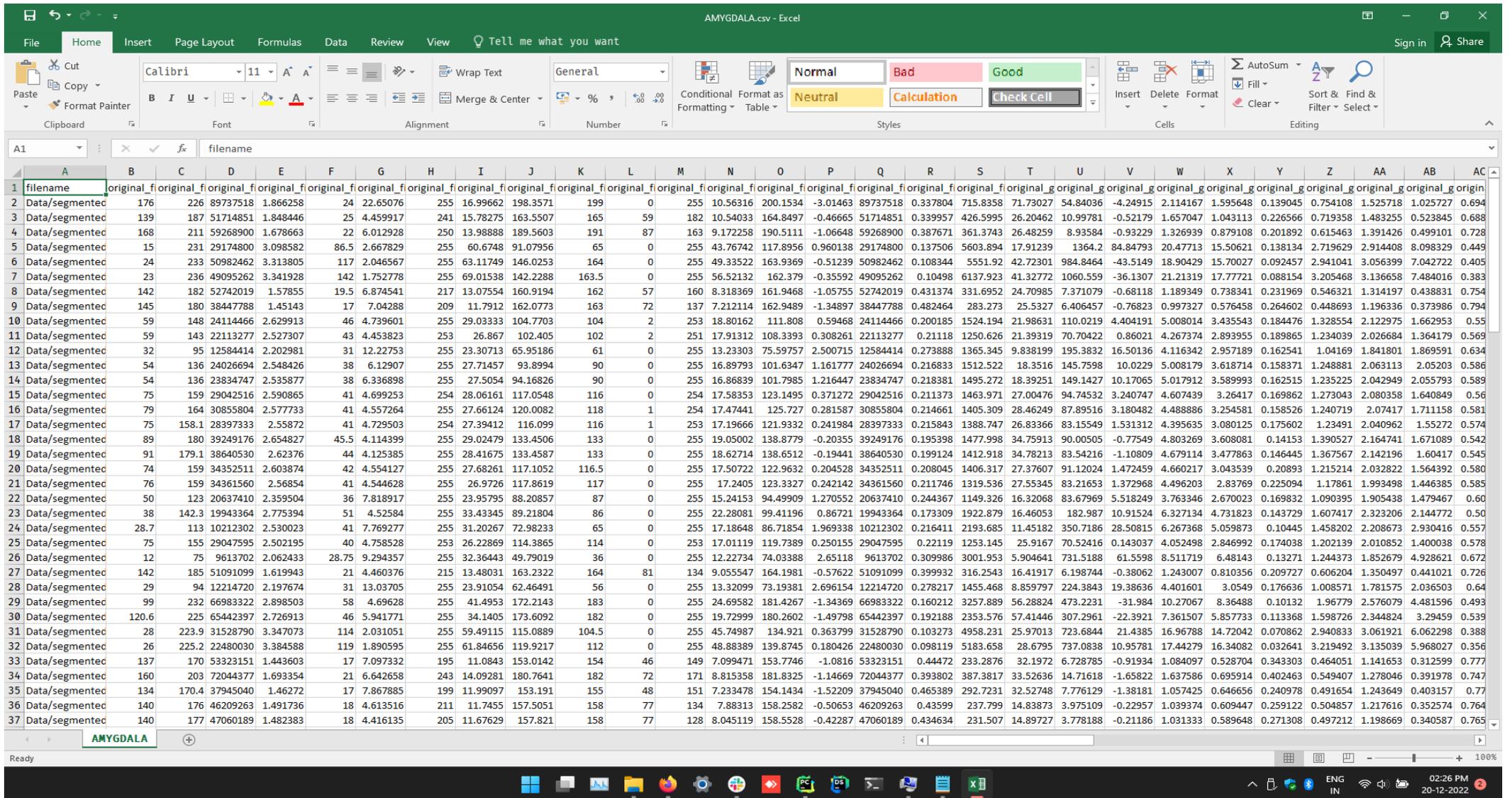


Figure 3 demonstrates the sample results obtained at the time of feature extraction.

original_gldm_GrayLevelNonUniformity, original_gldm_GrayLevelVariance,
 original_gldm_HighGrayLevelEmphasis, original_gldm_LargeDependenceEmphasis,
 original_gldm_LargeDependenceHighGrayLevelEmphasis,
 original_gldm_LargeDependenceLowGrayLevelEmphasis,
 original_gldm_LowGrayLevelEmphasis, original_gldm_SmallDependenceEmphasis,
 original_gldm_SmallDependenceHighGrayLevelEmphasis,
 original_gldm_SmallDependenceLowGrayLevelEmphasis,
 original_glrlm_GrayLevelNonUniformity,
 original_glrlm_GrayLevelNonUniformityNormalized,
 original_glrlm_GrayLevelVariance, original_glrlm_HighGrayLevelRunEmphasis,
 original_glrlm_LongRunEmphasis,
 original_glrlm_LongRunHighGrayLevelEmphasis,
 original_glrlm_LongRunLowGrayLevelEmphasis,
 original_glrlm_LowGrayLevelRunEmphasis, original_glrlm_RunEntropy,
 original_glrlm_RunLengthNonUniformity,
 original_glrlm_RunLengthNonUniformityNormalized,
 original_glrlm_RunPercentage, original_glrlm_RunVariance,
 original_glrlm_ShortRunEmphasis,
 original_glrlm_ShortRunHighGrayLevelEmphasis,
 original_glrlm_ShortRunLowGrayLevelEmphasis,
 original_glszm_GrayLevelNonUniformity,
 original_glszm_GrayLevelNonUniformityNormalized,
 original_glszm_GrayLevelVariance,
 original_glszm_HighGrayLevelZoneEmphasis,
 original_glszm_LargeAreaEmphasis,
 original_glszm_LargeAreaHighGrayLevelEmphasis,
 original_glszm_LargeAreaLowGrayLevelEmphasis,
 original_glszm_LowGrayLevelZoneEmphasis,
 original_glszm_SizeZoneNonUniformity,
 original_glszm_SizeZoneNonUniformityNormalized,
 original_glszm_SmallAreaEmphasis,
 original_glszm_SmallAreaHighGrayLevelEmphasis,
 original_glszm_SmallAreaLowGrayLevelEmphasis, original_glszm_ZoneEntropy,
 original_glszm_ZonePercentage, original_glszm_ZoneVariance,
 original_ngtdm_Busyness, original_ngtdm_Coarseness, original_ngtdm_Complexity, original_ngtdm_Contrast,
 original_ngtdm_Strength

5. RESULT AND DISCUSSION

The results of the classification of the 5 stages of Alzheimer's disease using an optimized CNN algorithm are discussed in this section.

Several metrics are frequently used to assess the optimized CNN algorithm's efficiency: Reliability calculates the overall percentage of correctly classified cases across all classes.

$$Accuracy = \frac{T_{pt} + T_{nt}}{T_{pt} + T_{nt} + F_{pt} + F_{nt}} \quad (8)$$

The precision function determines the percentage of accurately predicted examples in a given class relative to all examples anticipated to belong to that class.

$$Precision = \frac{T_{pt}}{T_{pt} + F_{pt}} \tag{9}$$

The percentage of correctly anticipated occurrences in a given class out of all occurrences of that category in the dataset is known as recall (sensitivity).

$$Recall = \frac{T_{pt}}{T_{pt} + F_{nt}} \tag{10}$$

The F1 Score is a consistent indicator of the model's efficiency that is calculated as the harmonic average of precision and recall.

$$F1\ Score = \frac{T2 \times Precision \times Recall}{Precision + Recall} \tag{11}$$

Confusion Matrix allows for the evaluation of certain misunderstandings by offering a thorough analysis of the model's projections over several classes. The confusion matrix values provided in Figure 3 indicate the performance of a classification model for Alzheimer's disease, specifically for the five stages: EMCI, LMCI, MCI, AD, and CN. The values presented, such as 'EMCI=0.98', 'LMCI=0.98', 'MCI=0.99', 'AD=0.99', and 'CN=0.99', represent the classification accuracy for each stage. The value 'EMCI=0.98' suggests that the model attained an accuracy of 98% in correctly classifying instances belonging to the EMCI stage. Similarly, 'LMCI=0.98' indicates that the model accomplished an accuracy of 98% in accurately classifying instances belonging to the LMCI stage.

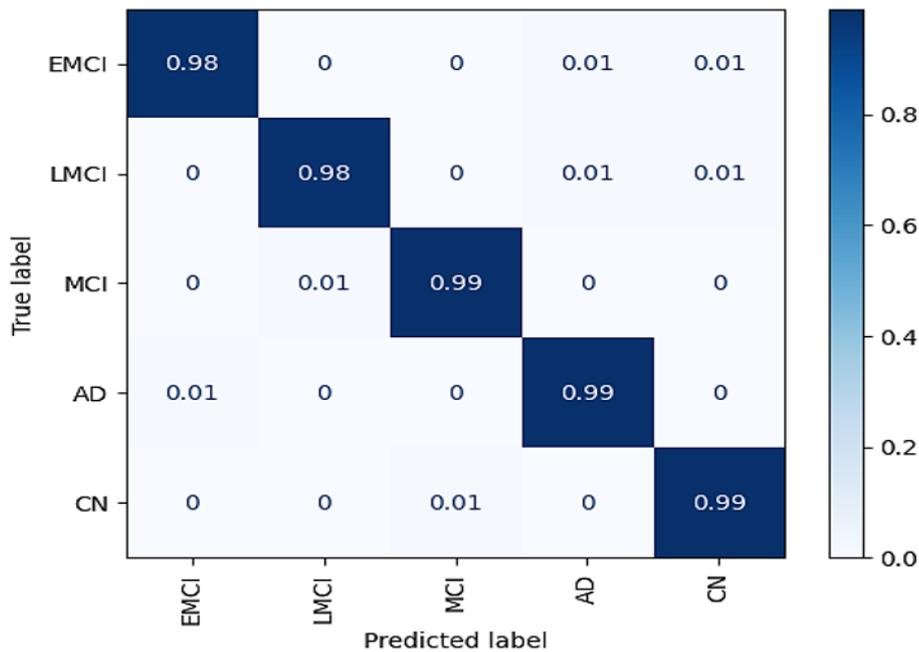


Figure 4. Confusion matrix.

The value 'MCI=0.99' suggests that the model realized an accuracy of 99% in correctly classifying instances belonging to the MCI stage. Similarly, 'AD=0.99' indicates that the model reached an accuracy of 99% in accurately classifying instances belonging to the AD stage. Lastly, 'CN=0.99' suggests that the model completed an accuracy of 99% in correctly classifying instances belonging to the CN stage.

These high accuracy values across all stages (ranging from 98% to 99%) indicate that the optimized CNN model performed very well in classifying Alzheimer's disease stages based on the provided confusion matrix. The

model showed strong discriminatory power in distinguishing between different stages, providing accurate predictions for each stage, as shown in Figure 4.

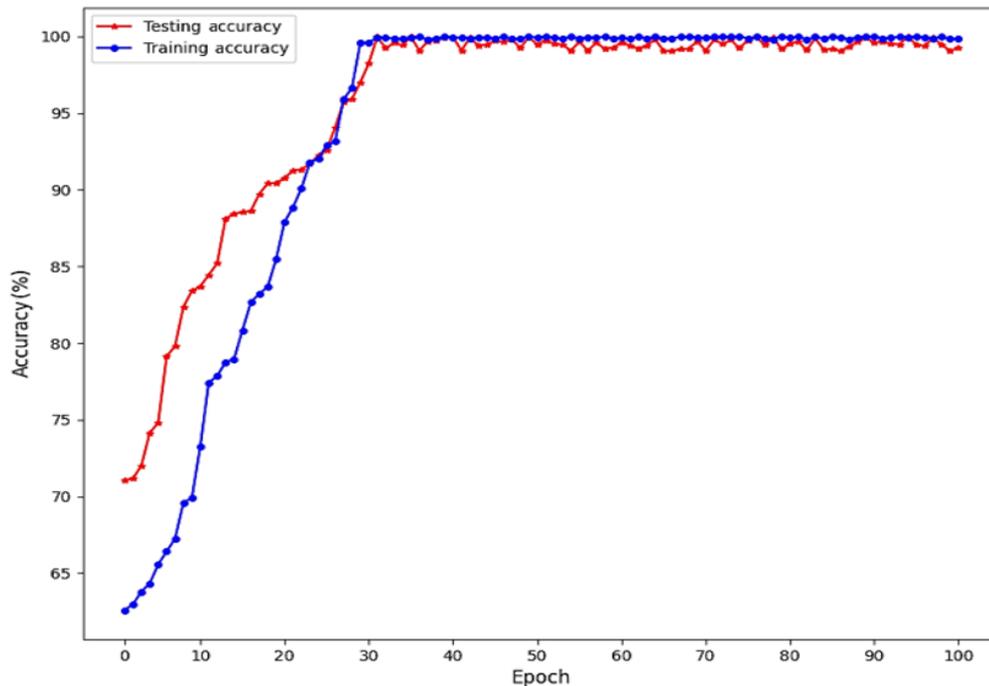


Figure 5. Accuracy of the classification of Alzheimer's disease.

Figure 5 illustrates the categorization of Alzheimer's disease. If the accuracy attained on training data is 100% and on testing data is 99.5%, this suggests that the model has done a very good job of differentiating between the illness's stages. A training accuracy of 100% indicates that the model has learned the training data patterns extremely well. It can perfectly classify the Alzheimer's disease stages in the training set. This high accuracy indicates that the model has captured the intricacies and features specific to each stage during the training process. The testing accuracy of 99.5% suggests that the model's presentation is excellent on unnoticed data, which is a positive outcome. It implies that the model has generalized well and can accurately classify Alzheimer's disease stages in new, unseen instances. The slight decrease in accuracy compared to the training set could be due to the presence of more challenging or diverse examples in the testing set.

Figure 6 illustrates the loss of classification for Alzheimer's disease.

For the categorization of Alzheimer's disease, the loss values attained on the training and testing data are 0.08 and 0.10, respectively. A popular artificial intelligence metric called loss measures the discrepancy between a model's expected outputs and its actual labels. It measures how well the model is fitting the training data and can provide insights into its performance. In this instance, training losses of 0.08 means that the mathematical model's average estimates for the initial data set are extremely close to the actual labels. A lower training loss suggests that the model has successfully learned the patterns and features necessary to classify Alzheimer's disease accurately within the training dataset.

Conversely, the testing loss of 0.10 represents the mean variance among the true labels on the hidden testing results and the predictions made by the model. It is slightly higher than the training loss, which is expected as the model is evaluated on data that it has not been directly trained on. However, a testing loss of 0.10 is still relatively low, indicating that the model is generalizing well and performing reasonably accurately on new, unseen data. It's worth noting that the loss values should be interpreted in the context of the specific problem and dataset. Even though these loss figures indicate that the model is operating effectively, in order to fully comprehend how it performs, one must also take into account additional assessment metrics.

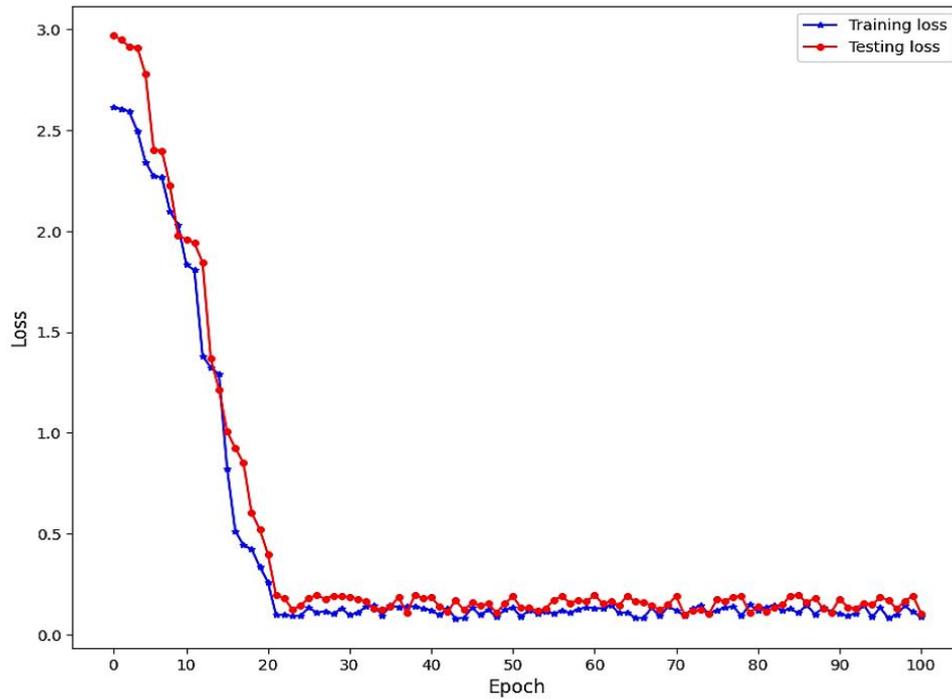


Figure 6. Loss of the classification of Alzheimer's disease.

The AUC-ROC achieved by the optimized CNN algorithm, as shown in Figure 7, for EMCI classification is 0.98. This suggests that the model has a good ability to discriminate between cases of EMCI and other phases of Alzheimer's disease. The AUC-ROC obtained for the LMCI classification using the optimized CNN algorithm is 0.98. This suggests that the model performs effectively in identifying cases of LMCI and differentiating them from other stages of Alzheimer's disease.

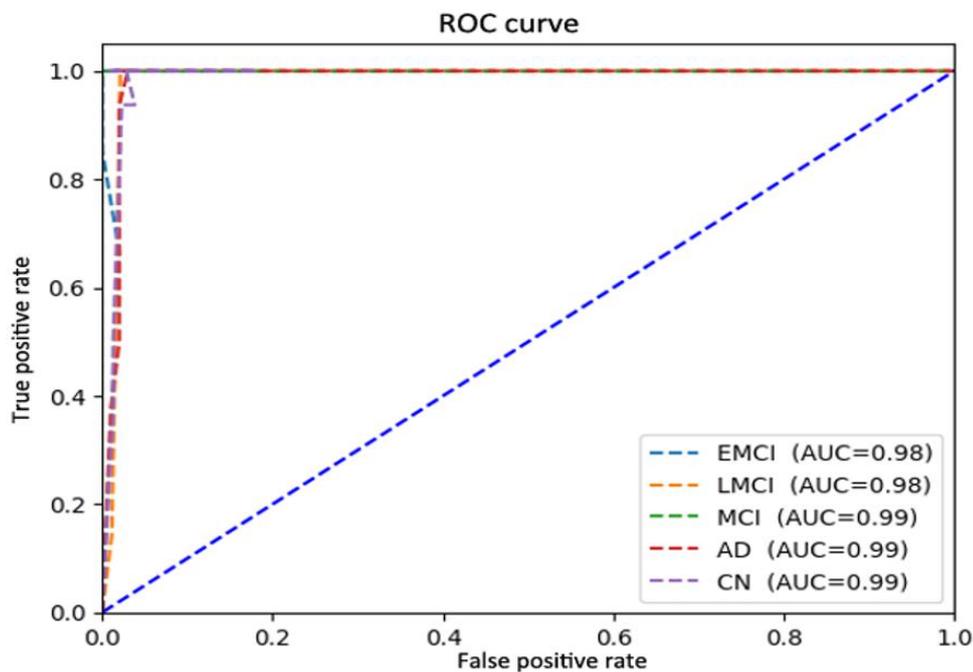


Figure 7. ROC-AUC for classification of Alzheimer's disease.

The optimized CNN algorithm achieves an AUC-ROC of 0.99 for MCI classification. This indicates a high level of accuracy in distinguishing MCI cases from other stages of Alzheimer's disease. The AUC-ROC obtained for the classification of AD using the optimized CNN algorithm is 0.99. This suggests that the model demonstrates

excellent performance in identifying AD cases and distinguishing them from other stages. The optimized CNN algorithm achieves an AUC-ROC of 0.99 for CN classification. This indicates that the model performs exceptionally well in identifying individuals with normal cognitive function and distinguishing them from different stages of Alzheimer's disease.

6. CONCLUSION

In this study, we have presented a novel AADD-DLM system for automated AD diagnosis using brain MRIs. Three main procedures are included in the AADD-DLM approach that is being discussed, namely skull stripping, segmentation, and feature extraction. Initially, the AADD-DLM technique uses the U-Net model for the skull stripping process, which enables the removal of the skull regions in the brain MRI. Next, the QuickNAT model is utilized for an effective brain MRI segmentation process. Moreover, the radiomics feature extraction approach is used to produce a set of feature vectors that are important. For exhibiting the promising performance of the AADD-DLM technique, widespread experimentation analysis is made on the ADNI database. The simulation results demonstrated the AADD-DLM technique's greater effectiveness than other contemporary methods. The optimized CNN classifiers are used for the accurate classification of AD on brain MRIs. The accuracy achieved on the training data of 100% and the testing data of 99.5% for the classification of Alzheimer's disease indicates that the optimized CNN algorithm has performed exceptionally well in accurately predicting the disease's stages.

Funding: This study received no specific financial support.

Institutional Review Board Statement: Not applicable.

Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.

Competing Interests: The authors declare that they have no competing interests.

Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.

REFERENCES

- [1] N. Yamanakkanavar, J. Y. Choi, and B. Lee, "MRI segmentation and classification of human brain using deep learning for diagnosis of Alzheimer's disease: A survey," *Sensors*, vol. 20, no. 11, p. 3243, 2020. <https://doi.org/10.3390/s20113243>
- [2] A. Loddo, S. Buttau, and C. Di Ruberto, "Deep learning based pipelines for Alzheimer's disease diagnosis: A comparative study and a novel deep-ensemble method," *Computers in Biology and Medicine*, vol. 141, p. 105032, 2022. <https://doi.org/10.1016/j.combiomed.2021.105032>
- [3] L. Nanni *et al.*, "Comparison of transfer learning and conventional machine learning applied to structural brain MRI for the early diagnosis and prognosis of Alzheimer's disease," *Frontiers in Neurology*, vol. 11, p. 576194, 2020. <https://doi.org/10.3389/fneur.2020.576194>
- [4] J. Islam and Y. Zhang, "Early diagnosis of Alzheimer's disease: A neuroimaging study with deep learning architectures," in *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1881-1883.
- [5] M. B. T. Noor, N. Z. Zenia, M. S. Kaiser, S. A. Mamun, and M. Mahmud, "Application of deep learning in detecting neurological disorders from magnetic resonance images: A survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia," *Brain Informatics*, vol. 7, pp. 1-21, 2020. <https://doi.org/10.1186/s40708-020-00112-2>
- [6] S. Qiu *et al.*, "Multimodal deep learning for Alzheimer's disease dementia assessment," *Nature Communications*, vol. 13, no. 1, pp. 1-17, 2022.
- [7] E. Altinkaya, K. Polat, and B. Barakli, "Detection of Alzheimer's disease and dementia states based on deep learning from MRI images: A comprehensive review," *Journal of the Institute of Electronics and Computer*, vol. 1, no. 1, pp. 39-53, 2020.

- [8] E. Lee, J.-S. Choi, M. Kim, H.-I. Suk, and A. s. D. N. Initiative, "Toward an interpretable Alzheimer's disease diagnostic model with regional abnormality representation via deep learning," *NeuroImage*, vol. 202, p. 116113, 2019. <https://doi.org/10.1016/j.neuroimage.2019.116113>
- [9] C. L. Saratxaga *et al.*, "MRI deep learning-based solution for Alzheimer's disease prediction," *Journal of Personalized Medicine*, vol. 11, no. 9, p. 902, 2021. <https://doi.org/10.3390/jpm11090902>
- [10] E. Hussain, M. Hasan, S. Z. Hassan, T. H. Azmi, M. A. Rahman, and M. Z. Parvez, "Deep learning based binary classification for alzheimer's disease detection using brain mri images," presented at the In 2020 15th IEEE Conference on Industrial Electronics and Applications, IEEE, 2020.
- [11] Z. Kong, 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, 2022. <https://doi.org/10.1016/j.bspc.2022.103565>
- [12] A. M. Alhassan, "Alzheimer's disease neuroimaging initiative and Australian imaging biomarkers and lifestyle flagship study of ageing, Enhanced fuzzy elephant herding optimization-based OTSU segmentation and deep learning for Alzheimer's disease diagnosis," *Mathematics*, vol. 10, no. 8, p. 1259, 2022.
- [13] S. A. Ajagbe, K. A. Amuda, M. A. Oladipupo, F. A. Oluwaseyi, and K. I. Okesola, "Multi-classification of Alzheimer disease on magnetic resonance images (MRI) using deep convolutional neural network (DCNN) approaches," *International Journal of Advanced Computer Research*, vol. 11, no. 53, p. 51, 2021. <https://doi.org/10.19101/ijacr.2021.1152001>
- [14] E.-G. Marwa, H. E.-D. Moustafa, F. Khalifa, H. Khater, and E. Abdelhalim, "An MRI-based deep learning approach for accurate detection of Alzheimer's disease," *Alexandria Engineering Journal*, vol. 63, pp. 211-221, 2023. <https://doi.org/10.1016/j.aej.2022.07.062>
- [15] X. Fang, Z. Liu, and M. Xu, "Ensemble of deep convolutional neural networks based multi-modality images for Alzheimer's disease diagnosis," *IET Image Processing*, vol. 14, no. 2, pp. 318-326, 2020. <https://doi.org/10.1049/iet-ipr.2019.0617>
- [16] M. Liu *et al.*, "A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease," *Neuroimage*, vol. 208, p. 116459, 2020. <https://doi.org/10.1016/j.neuroimage.2019.116459>
- [17] J. Zopes, M. Platscher, S. Paganucci, and C. Federau, "Multi-modal segmentation of 3D brain scans using neural networks," *Frontiers in Neurology*, vol. 12, p. 653375, 2021.
- [18] A. G. Roy, S. Conjeti, N. Navab, and C. Wachinger, "Alzheimer's disease neuroimaging initiative, QuickNAT: A fully convolutional network for quick and accurate segmentation of neuroanatomy," *NeuroImage*, vol. 186, pp. 713-727, 2019.

Views and opinions expressed in this article are the views and opinions of the author(s), Review of Computer Engineering Research shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.