

Paper Type: Original Article

Attention-Based Deep Neural Network with Rectified Adam Optimization for Alzheimer's Disease Detection Using Retinal MRI Imaging

Ramin Goudarzi Karim^{1,*} , Sanaz Hami Hassan Kiyadeh²

¹ Department of Computational and Information Sciences, Stillman College, Alabama, USA; rkirim@stillman.edu.

² Department of Mathematics, University of Alabama, Alabama, USA.

Citation:

Received: 20 August 2024

Revised: 08 September 2024

Accepted: 02 March 2025

Goudarzi Karim, R., & Hami Hassan Kiyadeh, S. (2025). Attention-based deep neural network with rectified adam optimization for Alzheimer's disease detection using retinal MRI Imaging. *Computational Engineering and Technology Innovations*, 2(4), 109-121.

Abstract

A chronic, and degenerative brain disease is termed as Alzheimer's Disease (AD). Now, there is no effective treatment for AD. The treatment plans and patient care have been greatly improved by the Early Detection (ED) of the AD. For ED and monitoring, Magnetic Resonance Imaging (MRI) is considered to be an effective tool. Because, MRI is non-invasive and produces high quality images. The Retinal imaging is the most popular non-invasive retinal imaging technique can be used by Deep Learning (DL) models for diagnosing AD dementia. Here, the conventional methods for AD detection experience difficulties in generalizability, interpretability, and computational efficiency. These conventional methods restrict its implications in clinical settings. But, these conventional methods for AD detection models exhibit high Accuracy (Acc). An Attention based Deep Neural Network (ADNN) with rectified Adam Optimizer (AO) using the National Institute of Mental Health and Neurosciences (NIMHANS) dataset was suggested in this study for overcoming those limitations. Image normalization is the initial method for ensuring the data consistency. Then, the Histogram of Oriented Gradients (HOG) with Wavelet Transform (WT) was utilized for robust Feature Extraction (FE). For classification, the ADNN processed the extracted features. Thus, focusing on critical areas related to AD improves the interpretability. Rectified AO optimizes the model, this will result in improving the convergence speed, and reducing overfitting. Significant robustness, interpretability, and efficiency improvements was demonstrated by the outcomes. For Real-Time (RT) application in clinical AD detection, this framework is beneficial.

Keywords: Alzheimer, Magnetic resonance imaging, Image normalization, Rectified adam optimizer, Histogram of oriented gradients.

1 | Introduction

Alzheimer's disease (AD) is a progressive neurological disease. It has the following symptoms; Memory loss, cognitive decline, and behavioural abnormalities. Dementia was commonly caused by AD (*Fig. 1*). It affects millions of people worldwide, especially elder people. Thus, the brain cells are gradually damaged by the AD,

 Corresponding Author: rkirim@stillman.edu

 <https://doi.org/10.48314/ceti.vi.74>



Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).

and losing functions that are needed for daily life. In several countries, one of the major causes of death was found to be AD [1]. To slow down the disease growth, the Early Detection (ED) and accurate diagnosis is vital. This will improve the outcomes of the patients. In *Fig. 1*, the standard clinical tests fail to detect the early symptoms, because it is subtle. For the purpose of improving the diagnostic abilities, the advancements in Medical Imaging (MI), computational techniques, and intense learning were analysed. By doing this, the ED of AD is attained. Deep Learning (DL) is a subset of Artificial Intelligence (AI). For detection and diagnosis of AD, AI is considered to be an effective tool through the study of brain imaging data like Magnetic Resonance Imaging (MRI) [2], [3]. DL methods like Convolutional Neural Networks (CNN) and Deep Neural Network (DNN) are utilized for extracting the complex patterns from High Dimensional (HD) MRI data. When comparing the DL methods with conventional Machine Learning (ML) methods, the DL method performs better than the conventional ML methods by Accuracy (Acc) [4]. Then, the minor structural and textural variations in the brain that indicates AD was detected by these models. This will help in preventing major symptoms. Patterns related to AD was effectively detected by the DL models, as it trains on large sets of labelled MRI images. Thus, more precise classification among healthy brain images and diseased brain images was attained. Improved diagnostic Acc, efficiency, and automation were the benefits that are provided by DL in AD detection [5]. DL models have the potential to process hundreds of MRI scans promptly and simultaneously, but the manual diagnostic methods are time consuming. These manual diagnostic methods may have the possibility of making human error. When dealing with subtle variations, high Acc in differentiating AD from non-AD is attained by the model, as it has the potential to learn complex features autonomously [6], [7]. Still there are many challenges despite the benefits. More generalizability was often needed for DL model because of data variability agreed treatment demographics. DL models especially DNN are black boxes. Lack of interpretability is the meaning for it. In clinical contexts, DNN is crucial, because the comprehension regarding the Decision Making (DM) process is crucial. Here, the practical implications in Real-Time (RT) clinical applications were limited by these models, because these models are computationally intensive.

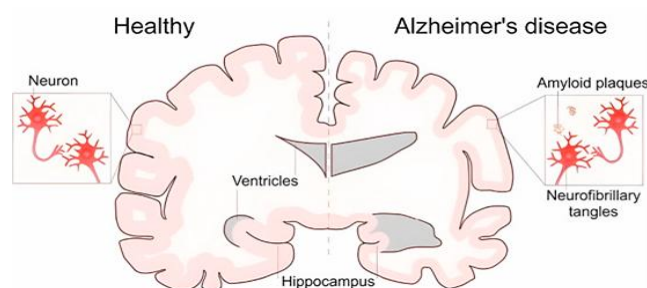


Fig. 1. Healthy brains vs AD.

An Attention based Deep Neural Network (ADNN) with the rectified Adam Optimizer (AO) is suggested for addressing these limitations [8], [9]. In AD detection, for the purpose of improving the interpretability, efficiency, and generalizability, the model is designed. To enhance the interpretability, the most relevant areas of MRI has been emphasized by the model via integrating an Attention Mechanism (AM) in the DNN. The AM was employed by clinicians for the better comprehension of DM process. Detecting the relevant brain areas related to the disease was also facilitated by this AM. The training was optimized, the computational load was reduced, and the convergence speed was improved by the adaptive learning rate of the rectified AO [10]. RT analysis is crucial in the clinical settings, and this model is well-suited and enhancing the Acc. When comparing the conventional DL models in AD detection, the Attention based DNN with Rectified AO offers several benefits. The interpretability was enhanced by this AM, by doing this, the common issue black box in DL is addressed, and it facilitates the clinicians in identifying which brain areas are most involved in the model classification. The disease specific areas were emphasized by the model for improving Acc, ED of the AD was also facilitated. The computational efficiency was enhanced by the rectified AO, and it is crucial for RT

clinical applications. This will make the model faster and less resource intensive. An accurate, scalable DL based Alzheimer's detection model was attained by this methodology.

2| Literature Review

An improved Lightweight (LW) DL model for precisely identifying AD from MRI images was put forth by El-Latif et al. [11]. It eliminates the requirement for conventional techniques like Feature Extraction (FE) and feature classification by combining them into a single step, and it provides strong detection performance without the need for deeper layers. With just seven layers, the suggested approach simplifies the system compared to earlier deep models and reduces analysis time. It evaluates the suggested model using a 36-megabyte dataset that is openly accessible on Kaggle and has a large number of records. With scores of 95.93% for Multi Classification (MC) tasks and 99.22% for Binary Classification (BC), the model outperformed previous models. The research is the first to cover every technique used in the openly accessible Kaggle dataset for AD detection, enabling researchers to address novel problems. The LW DL system's ability to classify AD with high Acc is demonstrated by the results. Kim et al. [12] introduced a VUNO Med-Deep Brain Alzheimer's Disease (DBAD) decision support service, which employs a DL algorithm to diagnose AD. It used a conventional diagnostic evaluation to diagnose AD. DBAD and three medical expert panels ME used T1weighted MRI to determine whether subjects had Normal Cognition (NC) or AD. The Korea Veterans Health Service and Seoul Asan Medical Centre received 98 elder adults (60 years of age and over) who were recruited. Comparable Acc (87.1% for DBAD and 84.3% for ME), sensitivity (93.3% for DBAD and 80.0% for ME), and specificity (85.5% for DBAD and 85.5% for ME) were demonstrated by both DBAD and ME in the diagnosis of AD. But in every investigation, DBAD showed a stronger tendency than the diagnosis of ME. The accessibility of AD diagnosis and treatment may be enhanced by DBAD's potential to assist doctors, who are not specialists in the disease in making therapeutic decisions. DenseNet201, Efficient Net B7, and Extremely Randomised Trees (ERT) are three robust models that Al-Rawashdeh et al. [13] suggested as a novel method for detecting AD. It increases DenseNet201's FE capability by incorporating a self-attention technique. In addition, early halting strategies on Efficient Net B7 should be applied to address overfitting. To differentiate between normal and abnormal MRI images, the Bayesian and Hyperband optimisation techniques modify extra tree Hyperparameters (HP). In order to better understand the model's choice, it made advantage of Shapley's Additive explanations. With minimal computational resources, the suggested approach showed 98.9% Acc in identifying AD. For accurately detecting AD using MRI images, the results show the effectiveness of the suggested FE and ERT models and optimisation techniques. To determine if they could identify pertinent AD Biomarkers (BM) from structural MRI and categorise brain images into AD, Mild Cognitive Impairment (MCI), and Cognitively Normal (CN) groups, Folego et al. [14] created a DL technique. Using brain sMRI datasets from internet databases, it tailored and trained CNN [14]. When tested on the Computer Aided Diagnostic (CAD) Dementia risk, ADNet performed better than a number of earlier methods. The Acc of ADNet-DA, the method's configuration with ML domain adaption, was 52.3%. One of its contributions is the creation of a completely independent, comparatively quick DL system that generates competitive results without depending on the domain-specific disease information of each patient. According on the identification of distinctive characteristics in brain scans, an end-to-end CNN system divides participants into groups with AD, MCI, or CN. The method is an effective method for identifying BM to aid in diagnosing AD and, potentially, many other diseases. An AD detection Deep Convolutional Neural Networks (DCNN) was suggested by Naganandini and Shanmugavadivu [15]. CNN exhibits promise in the automated FE and characterisation of AD due to similar brain patterns and pixel intensities. Automating FE and classification from input image sets is the goal of DL approaches. Each image is assigned to either AD or Healthy Control (HC) by a two-way classifier. The Acc of classifying diseases into binary categories was evaluated through experiments utilising the MIRIAD dataset. The effectiveness of the recommended approach for AD detection is confirmed by the recorded outcomes of CNNs with 4-and 5-layer architectures. It is a new brain imaging technology, and Feng et al. [16] suggested an advanced DL method to evaluate its contribution to improving the Acc of detecting AD. Binary and ternary disease classification models are conducted using Three Dimensional Convolutional Neural Networks (3D-CNNs)

in conjunction with MRI. The DL performances of 3D-CNN, 3D-CNN-Support Vector Machine (SVM), and Two-Dimensional (2D) CNN models were compared using the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. A ternary classification Acc for MCI and AD diagnoses was attained by the 3D-CNN-SVM. Additionally, 3D-CNN-SVM gave BC the greatest results. This study demonstrates that 3D-CNN-SVM outperforms the currently employed DL techniques with MRI. Additionally, 3D-CNN-SVM showed effectiveness without the need for manual FE and is completely independent of scanner variability and imaging technique. This suggests that it could be misused and used to virtual patient imaging data by unskilled operators. Furthermore, 3D-CNN-SMV might be a good screening choice because of the MRI modality's non-invasiveness, safety, and nonirradiated features. The community as a whole is impacted by AD. This study enhances value-based patient care in the clinic by distinguishing AD and MCI patients from HC. Early mental retardation, moderate mental impairment, late mild mental impairment, and final Alzheimer's stage are the four stages into which Ni et al. [17] classified AD patients using a Transfer Learning (TL) approach using CNN networks and ResNet50, VGG16, and DenseNet121 on a large dataset. The researchers categorise Alzheimer's patients into different stages using CNN, TL with ResNet50, VGG16, and DenseNet121, and a large dataset. The model classifies MRI images into EMCI, MCI, LMCI, and AD after being trained and evaluated on ADNI data via the Keras Application Programming Interface (API). Compared to the other models, VGG16, DenseNet121, and ResNet50 performed significantly better. The Acc has significantly improved over the previous results. Deep models and topologies, such as 2D and 3D CNNs and Recurrent Neural Networks (RNNs), have been suggested by Ebrahimi et al. [18]. Each MRI scan is separated into 2D slices in order to use a 2D CNN on 3D MRI volumes; the 2D image slices in the MRI volume are not connected [18]. In order for the 2D CNN + RNN model to comprehend the relationship between 2D image slice sequences for an MRI, an RNN may instead be used in conjunction with a CNN model. The issue is that the 2D CNN's FE process operates independently of the RNN's categorization. To solve this problem, voxel-based predictions can be made using 3D CNNs rather than 2D CNNs. The main contribution of the paper is the introduction of TL to 3D CNNs from a batch of 2D images. Sequence based decisions improve the Acc of slice-based decisions in differentiating AD patients from healthy people, as demonstrated by the MRI dataset. The voxel-based approach with TL from ImageNet to MRI datasets using 3D CNNs performed noticeably better than the rest, according to a number of implementations and experiments using CNNs on MRI scans for AD diagnosis. A classification technique based on brain scans was presented by Rao et al. [19]. The main organ impacted by AD is the brain, which presents a viable method for achieving more accurate and consistent outcomes. In order to effectively capture the spatial information represented in 3D MRI scans, it extended convolutional techniques to 3-D. By applying different weights to the contributions of each layer, the classification was carried out by selectively integrating features collected from many layers of the 3D convolutional network. Researchers included TL in this model since they knew it may reduce training time and increase the effectiveness of AD detection. Using brain MRI scans from three distinct classes; AD, MCI, and standard control (NC), the suggested structure included TL and Fine Tuning (FT). ResNet50V2 and InceptionResNetV2 were among the pre-trained DL models for AD classification that were examined. With a better classification Acc than other similar methods, ResNet50V2 became the best method. It achieved a continuously high testing Acc after achieving an exceptional training Acc. These results clearly show how DL techniques, specifically TL with ResNet50V2, can detect AD using 3D MRI brain scans with remarkable Acc. A hybrid DL approach was created by Balaji et al. [20] to detect AD early. A CNN with the long-term memory algorithm and multimodal imaging are used in a technique for early AD detection that combines the results of MRI, PET, and traditional neuropsychological tests [20–23]. To increase Acc, the suggested approach uses AO and updates the learning weights. The method's 98.5% Acc rate in differentiating between EMCI and CN controls is remarkable. These results demonstrate that DNN may be trained to reliably identify AD by automatically detecting imaging BM linked to the disease.

3 | Proposed Methodology

Clinicians can diagnose epilepsy, schizophrenia, depression, brain lesions, Parkinson's disease (9PD), and AD with the help of CAD systems. A neurological disease that affects many people, AD affects memory, behavior, and judgement. In order to effectively stop the course of AD, ED is essential. This paper proposes and implements an adequate model dataset of National Institute of Mental Health and Neurosciences (NIMHANS) extracted features using ADNN with rectified AO, to accurately classify AD cases using MRI analysis. *Fig. 2* depicts the workflow of ADNN with the corrected rectified AO.

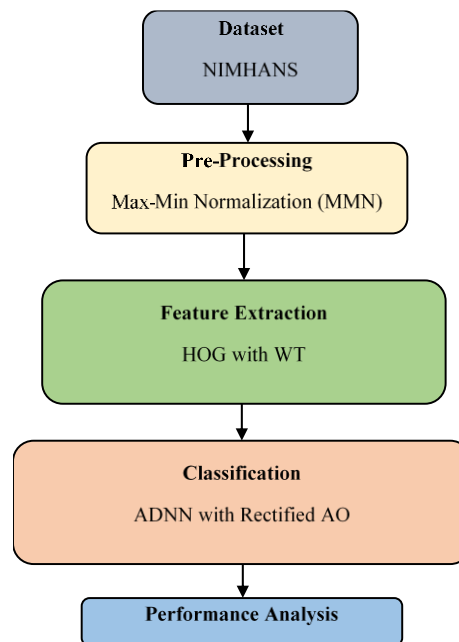


Fig. 2. Suggested workflow model.

3.1 | Data Collection

The experimental analysis in this study makes use of the NIMHANS datasets. The data collection includes 800 brain scans of 99 people, ages ranging from 55 to 87 years, including 60 normal subjects and 39 AD subjects. The subjects in the NIMHANS database are chosen from the outpatient department and geriatric clinic. *Fig. 3* displays a few brain pictures from the NIMHANS dataset.

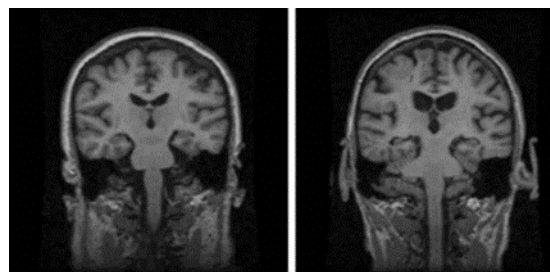


Fig. 3. Sample image of NIMHANS dataset normal and Alzheimer control.

3.2 | Data Pre-Processing using Max-Min Normalization (MMN)

After acquisition of the brain pictures from NIMHANS, normalizing procedure is used to improve the image's quality. MRI images vary in brightness, contrast, and intensity due to different imaging conditions, patient anatomy, and machine settings (*Fig. 4*). Normalization addresses this variability by adjusting the intensity levels across all images to a consistent scale. This step ensures that the model focuses on actual

disease related patterns rather than differences caused by imaging conditions. One common approach in normalization is linear intensity scaling, where the pixel intensity values are rescaled to a fixed range, such as $[0, 1]$ or $[0, 255]$. This standardizes the images and makes the features more comparable across different scans.

$$I_{\text{norm}} = \frac{I - I_{\min}}{I_{\max} - I_{\min}} \times (\text{new_max} - \text{new_min}) + \text{new_min}, \quad (1)$$

where *Eq. (1)*, I is the original intensity of each pixel in the MRI image. $I_{\min}$ and $I_{\max}$ are the minimum and maximum intensity values in the original image. new_max and new_min are the desired minimum and maximum intensity values after normalization (e.g., 0 and 1 for $[0, 1]$ range). This scaling adjusts each pixel intensity in the MRI image to fall within a standardized range, helping reduce inconsistencies. For more complex adjustments, a sigmoid based normalization function can be used for NIMHANS. This method scales intensities based on a non-linear transformation, which can be particularly useful when some brain regions need to be emphasized over others.

$$I_{\text{norm}} = \frac{\text{new_max} - \text{new_min}}{1 + e^{-\beta(I - \alpha)}} + \text{new_min}, \quad (2)$$

where, β controls the steepness of the sigmoid curve (higher values make the transition between low and high intensities sharper) can be expressed as *Eq. (2)*. α shifts the curve's center along the intensity axis to align with desired features in the MRI image. This equation emphasizes certain intensity ranges, making subtle features more visible and helping the model better recognize Alzheimer's related patterns. In many ways, normalization is applied not only on individual images but across the entire dataset to ensure that all images are comparable. In this case, global values $I_{\min}$, $I_{\max}$, β and σ are computed across the dataset, and each image is normalized using these global values. This ensures consistent contrast and brightness levels for images used in training the Alzheimer's detection model. Image normalization for AD detection standardizes MRI images, making pathological patterns clearer and reducing variability across scans. By applying these normalization techniques, including linear scaling, sigmoid-based transformations, and z-score normalization, the model receives images with uniform intensity levels, enabling it to identify subtle yet crucial features associated with Alzheimer's pathology more accurately.

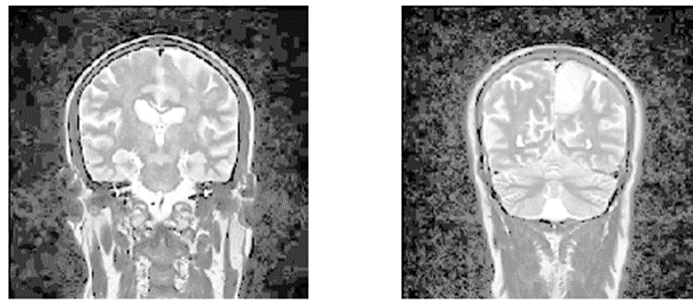


Fig. 4. Pre-processed images.

3.3| Feature Extraction

Histogram of Oriented Gradients (HOG) processes the output of normalization, ensuring that any detected gradients and features are more reflective of anatomical structures rather than imaging inconsistencies from NIMHANS. It analyses the gradients in an image, focusing on edge directions and intensities. AD often alters brain structures in ways that affect edges and gradients in MRI images, making HOG particularly useful for capturing these changes. Wavelet Transform (WT) is applied first to each MRI image to capture multi scale details. This technique decomposes the image into various frequency components, which enables the identification of fine and coarse patterns and is useful for detecting subtle changes in brain tissue structure associated with Alzheimer's. Approximation (LL), horizontal (LH), vertical (HL), and diagonal (HH) are the

four sub-bands that make up the widely used Discrete Wavelet Transform (DWT). The decomposition equation for WT in a 2D MRI image can be expressed as *Eq. (3)*.

$$I(x, y) = \sum_{j,k} \left(C_{LL} \phi_{j,k}(x, y) + C_{LH} \psi_{j,k}^{(H)}(x, y) + C_{HL} \psi_{j,k}^{(V)}(x, y) + C_{HH} \psi_{j,k}^{(D)}(x, y) \right), \quad (3)$$

where, C_{LL} , C_{LH} , C_{HL} , and C_{HH} are the coefficients for each sub-band, $\phi_{j,k}$ is the scaling function for the approximation component, $\psi_{j,k}^{(H)}$, $\psi_{j,k}^{(V)}$, $\psi_{j,k}^{(D)}$ are the wavelet functions for the horizontal, vertical, and diagonal components. The LL sub band retains most of the image's structural information, while LH, HL, and HH capture finer texture details. These sub bands are essential for highlighting Alzheimer's affected areas, as they often contain unique textural changes. HOG is then applied to each decomposed sub-band, particularly focusing on the LL sub band since it contains most structural information relevant to Alzheimer's. HOG extracts feature by computing the gradient magnitude and orientation within small regions (cells) across the MRI image. The horizontal and vertical gradients of each pixel are calculated using convolution with filter kernels in *Eqs. (4) and (5)*.

$$G_x = I(x+1, y) - I(x-1, y). \quad (4)$$

$$G_y = I(x, y+1) - I(x, y-1). \quad (5)$$

Using the calculated gradients, the magnitude $M(x, y)$ and $\theta(x, y)$ orientation at each pixel is determined as *Eqs. (6) and (7)*.

$$M(x, y) = \sqrt{G_x^2 + G_y^2}. \quad (6)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{G_y}{G_x} \right). \quad (7)$$

Each cell accumulates gradient orientations into a histogram, with bins representing angle ranges (e.g., 0° to 180°). These histograms capture edge and shape details, which are critical for detecting structural abnormalities associated with Alzheimer's to detect from the NIMHANS. HOG histograms from cells are combined into blocks, and these blocks are normalized to account for variations in contrast within the MRI. Features extracted from each wavelet sub-band using HOG are concatenated to form a comprehensive Feature Vector (FV). This vector captures both global and localized patterns texture (from WT) and edges/shapes (from HOG) within the MRI image. The final FV for each MRI image can be represented as *Eq. (8)*;

$$\text{Feature Vector} = [HOG_{LL}, HOG_{LH}, HOG_{HL}, HOG_{HH}]. \quad (8)$$

The ADNN uses this vector as its classification input. Here the HOG-wavelet vector's Alzheimer's-relevant features are the focus of AM. In conclusion, FE employing HOG with WT in Alzheimer's detection entails obtaining gradient based edge features (using HOG) and dividing MRI images from the NIMHANS dataset into frequency components (using WT). This dual strategy offers a comprehensive feature set that captures the structural and textural abnormalities linked to AD, making it a useful input for the classification model.

3.4 | Classification

The input is the FV that has been extracted. In order to distinguish AD features from normal brain architecture, this vector includes edge and texture information. The FV is fed into the ADNN, which highlights important regions in the AD related MRI images from NIMHANS. In *Fig. 5*, the DNN consists of multiple layers (Input Layer (IL), Hidden Layer (HL), and Output Layers (OL)). Each HL contains nodes (neurons) that progressively extract higher level representations from the input features. In a typical DNN, each neuron in a layer is connected to every neuron in the previous layer, allowing comprehensive feature interactions, which helps capture complex patterns in MRI data associated with AD. In order to identify disease relevant brain areas in AD, the AM allows the model to concentrate on the most important aspects

of the input features. Using the input characteristics, the AM calculates the Query (Q), Key (K), and Value (V) vectors. The query vector represents the evaluated current element, while the key and value vectors represent other feature set elements.

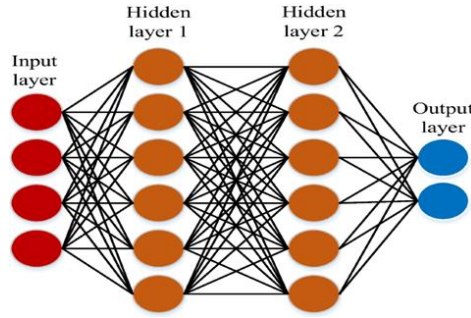


Fig. 5. Structure of DNN.

The model computes attention scores by measuring the similarity between the query and key vectors, which indicates the relevance of each feature element to the current focus. This calculation is typically done as Eq. (9).

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) V, \quad (9)$$

where, $Q \cdot K^T$ represents the dot product between queries and keys, giving a relevance score, d_k is a scaling factor (the dimension of the key vectors), helping stabilize gradients during training. SoftMax normalizes these scores to create a distribution of importance. The resulting scores weight the value vectors so that the most relevant features are given greater emphasis in the network's processing. After applying attention, the features are weighted by importance, enhancing the network's focus on critical regions indicative of Alzheimer's. These attention-enhanced features pass through the HIs of the DNN, where they undergo transformations and nonlinear activations to create more abstract representations. This step captures complex relationships and patterns in the data, which are important for identifying Alzheimer's-specific structural abnormalities in MRI images. In the final layer, the network applies the SoftMax function to convert the output into probabilities across the classes (e.g., Alzheimer's vs. non-Alzheimer's).

$$P(y = \text{Alzheimer's} | \text{Input}) = \frac{e^{z_i}}{\sum_j e^{z_j}}, \quad (10)$$

where Eq. (10), z_i is the output from the last layer for the Alzheimer's class. This gives the probability of the MRI scan belonging to the Alzheimer's class. Rectified Adam (Adaptive Moment Estimation) is used to optimize the network weights during training, enabling faster convergence and more stable training. For each weight, Adam computes the gradient of the loss function to rectify the errors, which indicates the direction in which the weight should be updated to minimize error. Adam maintains two values for each weight. The first moment estimate (mean), m_t updates as Eq. (11).

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad (11)$$

where, g_t is the current gradient and β_1 is the decay rate (typically 0.9). The second moment estimate (uncentered variance), v_t can be expressed as Eq. (12).

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \quad (12)$$

where β_2 is also a decay rate (usually 0.999). To correct the bias in early iterations, the moment estimates are adjusted Eqs. (13) and (14).

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \quad (13)$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t}. \quad (14)$$

Finally, Rectified Adam updates each weight θ_t as Eq. (15).

$$\theta_t = \theta_{t-1} - \frac{\alpha \cdot \hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}, \quad (15)$$

where, α is the learning rate and ϵ is a small constant) prevents division by zero. After training, the model is tested on new MRI images to assess its classification performance (Fig. 6). An MRI scan, the model predicts the probability of it belonging to the Alzheimer's class, focusing on regions identified by the AM. The Attention based DNN with Rectified AO enhances AD detection by focusing on disease relevant features in MRI images and improving training stability and efficiency. The AM directs the network to prioritize critical brain regions, while rectified Adam ensures fast convergence and minimizes overfitting, making this approach highly effective for medical image classification.



Fig. 6. Final classification output images.

4 | Results and Discussion

The performance analysis used Python 3.7 software for experimental simulation on a Windows 10 (64-bit) computer with an Intel core i7 processor, 16 GB RAM, and a two-terabyte hard disk. This study compares ADNN with AO performance with the rectified AO to that of an existing landmark-based feature with the AO classifier to evaluate its efficiency using the NIHMANS dataset. The performance analysis includes Acc, f-score, Precision (P), and Recall (R) parameters. Acc refers to how well a calculated or estimated value matches the actual amount. Acc is a straightforward performance statistic that compares total observations to correctly anticipated observations in Table 1. F-scores are weighted; it considers the average recall and precision values, as well as false negatives and false positives, and it can be expressed as Eqs. (16) and (17).

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \times 100. \quad (16)$$

$$\text{F-Score} = \frac{2TP}{2TP+FP+FN} \times 100. \quad (17)$$

Precision measures the fraction of correctly identified negatives. Recall measures the fraction of accurately identified positives and Eqs. (18) and (19) express precision and recall.

$$\text{Precision} = \frac{TP}{TP+FP} \times 100. \quad (18)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100. \quad (19)$$

Here, TP stands for true positives, TN for true negatives, FP for false positives, and FN for false negatives. Fig. 7 illustrates the Acc levels of various ML models and optimization methods. Among the methods compared, the ADNN with rectified AO achieves the highest Acc at nearly 98.60%, followed by DNN with rectified AO with around 98.40% Acc. Traditional DNN and CNN models show moderate performance, achieving roughly 96.26% and 96% Acc, respectively. The RNN model lags, reaching just above 94.26% Acc,

indicating that it is the least effective method among those evaluated. This analysis highlights that models using the rectified AO generally perform better, with the advanced deep neural network (ADNN) architecture achieving the most significant performance boost, suggesting the effectiveness of this optimizer in enhancing model Acc. The precision analysis of *Fig. 8* indicates a similar performance trend to the Acc graph, with ADNN with rectified AO achieving the highest precision at nearly 98.70%. DNN with rectified AO, which has a precision of about 98.50%, comes next. The RNN model has the lowest precision, at about 94.20%, whereas the standard DNN and CNN models have moderate precision levels of about 95.70% and 97.31%, respectively. The outcomes demonstrate that, particularly in the ADNN design, employing the corrected AO greatly increases model precision. The corrected AO performs better than alternative techniques. This implies that when using sophisticated optimization approaches in Neural Network (NN), precision advances are consistent with Acc improvements.

Table 1. Performance analysis of ADNN with AO with the rectified AO on NIMHANS dataset.

Parameters	RNN	DNN	CNN	DNN with Rectified AO	ADNN with Rectified AO
Acc (%)	94.26	96.26	96	98.40	98.60
Precision (%)	94.20	95.70	97.31	98.50	98.70
Recall (%)	92.10	96.00	98	98.40	98.50
F-score (%)	92.10	96.00	97.90	98.40	98.60
Execution time (secs)	50	25	75	19	17

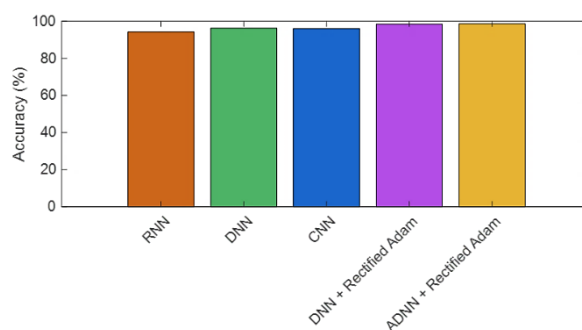


Fig. 7. Comparison of suggested model Acc with current approaches.

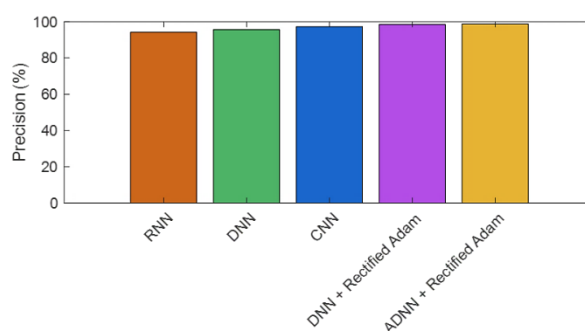


Fig. 8. Comparison of suggested model precision with current approaches.

Fig. 9 illustrates the recall performance comparison, showing that ADNN with rectified AO again leads with the highest recall, reaching just above 98.50%. In contrast, DNN with rectified AO follows closely at around 98.40%. The standard CNN and DNN models exhibit recall rates of approximately 98% and 96%, respectively. However, the RNN model has the lowest recall at about 93%, indicating it may miss more relevant instances compared to other models. This pattern reinforces the effectiveness of the rectified AO in enhancing recall, particularly in ADNN and DNN architectures, where its application provides notable improvements over standard optimizers, ensuring a higher proportion of relevant instances are captured by

the model. *Fig. 10* compares the performance of various ML model mRNN, DNN, CNN, DNN with rectified AO, and ADNN with rectified Aom based on their F-score percentages. Among these methods, the ADNN with the rectified AO achieves the highest F-score, close to 98.60%, indicating superior performance by P and R. The DNN with the rectified AO follows closely, reaching a high F-score below 98.40%. The CNN model ranks next with an F-score below 97.90%, while the standard DNN achieves around 96%, demonstrating moderate performance improvement without the optimizer enhancement. RNN records the lowest F-score, around 92%, highlighting its comparatively lower effectiveness. Including the rectified AO boosts the performance of both DNN and ADNN models, suggesting its significance in enhancing model Acc and robustness in predictive tasks. This analysis underscores the rectified AOs positive impact on NN performance, especially within DNN and ADNN frameworks. *Fig. 11* illustrates the execution time percentages for various ML models RNN, DNN, CNN, DNN with rectified AO, and ADNN with rectified AO. The standard DNN has the highest execution time at around 25%, indicating that it is the most computationally expensive method in this comparison. Following the DNN, the RNN model exhibits a significant execution time of around 55%, while CNN shows a more moderate execution time of roughly 30%. Models optimized with the rectified AO, specifically DNN and ADNN, demonstrate much lower execution times, with ADNN showing the lowest, around 10%, and the optimized DNN close to 20%. This suggests that the rectified AO enhances performance (as seen in the F-score comparison) and significantly reduces computational demands, especially for ADNN, making it highly efficient in Acc and execution time. This optimization could make DNN and ADNN with rectified AO more attractive choices in applications where both high performance and efficiency are critical.

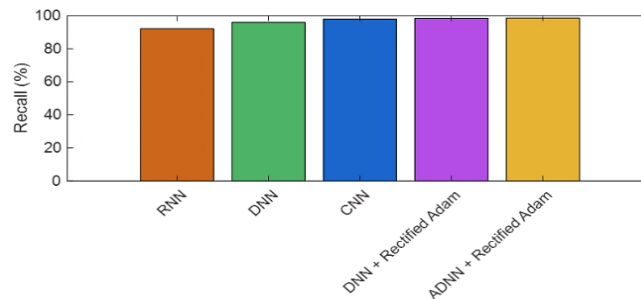


Fig. 9. Comparison of suggested model recall with current approaches.

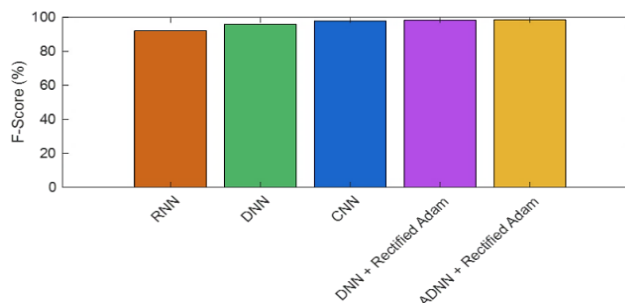


Fig. 10. Comparison of suggested model f-score with current approaches.

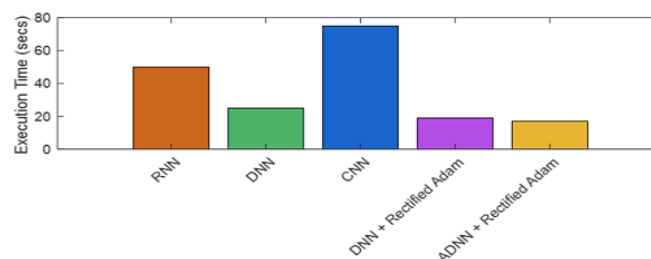


Fig. 11. Comparison of proposed model execution time with existing methods.

5 | Conclusion

In this study, an ADNN model with the rectified AO was developed for AD detection using MRI images from the NIMHANS dataset. Initially, image normalization was applied for data consistency, followed by the HOG combined with WT for FE. These extracted features were then processed by the ADNN, where AM were employed to focus on critical regions associated with Alzheimer's pathology, enhancing the interpretability and relevance of the model. Experimental results demonstrated improved Acc, robustness, and computational efficiency compared to existing models, showcasing the potential of this method for clinical applications. In future work, integrating advanced optimization techniques with the HOG ADNN model and exploring multi modal data sources, such as combining MRI with retinal imaging, could further enhance classification Acc and facilitate RT AD detection in clinical environments. Future work will validate this model across larger, diverse datasets and enhance FE methods to boost diagnostic Acc and generalizability further.

Data Availability

All data supporting the reported findings in this research paper are provided within the manuscript.

Funding

No funding was received to conduct this study.

Conflict of Interest

The authors declares that in this paper, there is no conflict of interests.

References

- [1] Ghazal, T. M., Sagheer, A., Munir, S., Khan, M. A., Ahmad, M., Issa, G. F., ... , & Hasan, M. K. (2022). Alzheimer disease detection empowered with transfer learning. *Computers, materials & continua*, 70(3), 5005-5019. <http://dx.doi.org/10.32604/cmc.2022.020866>
- [2] Afzal, S., Maqsood, M., Khan, U., Mehmood, I., Nawaz, H., Aadil, F., ... , & Nam, Y. (2021). Alzheimer disease detection techniques and methods: A review. *International journal of interactive multimedia and artificial intelligence*, 6(7), 26-38. <https://doi.org/10.9781/ijimai.2021.04.005>
- [3] Kundaram, S. S., & Pathak, K. C. (2020). Deep learning-based alzheimer disease detection. *Proceedings of the fourth international conference on microelectronics, computing and communication systems: Mccs 2019* (pp. 587-597). Springer. https://doi.org/10.1007/978-981-15-5546-6_50
- [4] Pradhan, A., Gige, J., & Eliazar, M. (2021). Detection of alzheimer's disease (AD) in MRI images using deep learning. *International journal of engineering research & technology*, 10(3), 580-585. https://d1wqtxts1xzle7.cloudfront.net/66199526/detection_of_alzheimers_disease_ad_in_mri_images_using_IJERTV10IS030310
- [5] Salehi, A. W., Baglat, P., & Gupta, G. (2020). Alzheimer's disease diagnosis using deep learning techniques. *International journal of engineering and advanced technology*, 9(3), 874-880. <https://doi.org/10.35940/ijeat.C5345.029320>
- [6] Ji, H., Liu, Z., Yan, W. Q., & Klette, R. (2019). Early diagnosis of alzheimer's disease using deep learning. *Proceedings of the 2nd international conference on control and computer vision* (pp. 87-91). Association for Computing Machinery. <https://doi.org/10.1145/3341016.3341024>
- [7] Balne, S., & Elumalai, A. (2021). Machine learning and deep learning algorithms used to diagnosis of alzheimer's. *Materials today: Proceedings*, 47, 5151-5156. <https://doi.org/10.1016/j.matpr.2021.05.499>
- [8] Saleem, T. J., Zahra, S. R., Wu, F., Alwakeel, A., Alwakeel, M., Jeribi, F., & Hijji, M. (2022). Deep learning-based diagnosis of Alzheimer's disease. *Journal of personalized medicine*, 12(5), 1-29. <https://doi.org/10.3390/jpm12050815>

- [9] Chui, K. T., Gupta, B. B., Alhalabi, W., & Alzahrani, F. S. (2022). An MRI scans-based alzheimer's disease detection via convolutional neural network and transfer learning. *Diagnostics*, 12(7), 1-14. <https://doi.org/10.3390/diagnostics12071531>
- [10] Lei, B., Liang, E., Yang, M., Yang, P., Zhou, F., Tan, E. L., ... , & Wang, S. (2022). Predicting clinical scores for Alzheimer's disease based on joint and deep learning. *Expert systems with applications*, 187, 115966. <https://doi.org/10.1016/j.eswa.2021.115966>
- [11] El-Latif, A. A. A., Chelloug, S. A., Alabdulhafith, M., & Hammad, M. (2023). Accurate detection of alzheimer's disease using lightweight deep learning model on MRI data. *Diagnostics*, 13(7), 1216. <https://doi.org/10.3390/diagnostics13071216>
- [12] Kim, J. S., Han, J. W., Bae, J. Bin, Moon, D. G., Shin, J., Kong, J. E., ... , & Kim, K. W. (2022). Deep learning-based diagnosis of Alzheimer's disease using brain magnetic resonance images: An empirical study. *Scientific reports*, 12(1), 18007. <https://doi.org/10.1038/s41598-022-22917-3>
- [13] Al-Rawashdeh, H. S., Usman, A., Dutta, A. K., & Sait, A. R. W. (2024). Hybrid feature extraction technique-based Alzheimer's disease detection model using MRI images. *Journal of disability research*, 3(6), 20240073. <https://doi.org/10.13039/501100019345?sid=sciencopen>
- [14] Folego, G., Weiler, M., Casseb, R. F., Pires, R., & Rocha, A. (2020). Alzheimer's disease detection through whole-brain 3D-CNN MRI. *Frontiers in bioengineering and biotechnology*, 8, 534592. <https://doi.org/10.3389/fbioe.2020.534592>
- [15] Naganandhini, S., & Shanmugavadivu, P. (2024). Alzheimer's disease detection in MRI images using deep convolutional neural network model. *EAI endorsed transactions on pervasive health and technology*, 10, 1-6. <https://doi.org/10.4108/eetpht.10.6435>
- [16] Feng, W., Halm-Lutterodt, N. Van, Tang, H., Mecum, A., Mesregah, M. K., Ma, Y., ... , & Guo, X. (2020). Automated MRI-based deep learning model for detection of Alzheimer's disease process. *International journal of neural systems*, 30(06), 2050032. <https://doi.org/10.1142/S012906572050032X>
- [17] Ni, J., Chen, Y., Gu, Y., Fang, X., & Shi, P. (2022). An improved hybrid transfer learning-based deep learning model for PM2. 5 concentration prediction. *Applied sciences*, 12(7), 1-18. <https://doi.org/10.3390/app12073597>
- [18] Ebrahimi, A., Luo, S., & Disease Neuroimaging Initiative, for the A. (2021). Convolutional neural networks for alzheimer's disease detection on MRI images. *Journal of medical imaging*, 8(2), 024503-024503. <https://doi.org/10.1117/1.JMI.8.2.024503>
- [19] Rao, B. S., Aparna, M., Kolisetty, S. S., Janapana, H., & Koteswararao, Y. V. (2024). Multi-class classification of Alzheimer's disease using deep learning and transfer learning on 3D MRI images. *Traitement du signal*, 41(3), 1397-1404. <https://doi.org/10.18280/ts.410328>
- [20] Balaji, P., Chaurasia, M. A., Bilfaqih, S. M., Muniasamy, A., & Alsid, L. E. G. (2023). Hybridized deep learning approach for detecting alzheimer's disease. *Biomedicines*, 11(1), 1-16. <https://doi.org/10.3390/biomedicines11010149>
- [21] Buvaneswari, P. R., & Gayathri, R. (2021). Deep learning-based segmentation in classification of alzheimer's disease. *Arabian journal for science and engineering*, 46(6), 5373-5383. <https://doi.org/10.1007/s13369-020-05193-z>
- [22] Liu, J., Li, M., Luo, Y., Yang, S., Li, W., & Bi, Y. (2021). Alzheimer's disease detection using depthwise separable convolutional neural networks. *Computer methods and programs in biomedicine*, 203, 106032. <https://doi.org/10.1016/j.cmpb.2021.106032>
- [23] L, S., V, S., Ravi, V., E. A., G., & K. P., S. (2024). Deep learning-based approach for multi-stage diagnosis of alzheimer's disease. *Multimedia tools and applications*, 83(6), 16799-16822. <https://doi.org/10.1007/s11042-023-16026-0>