Vol. 06, No. 02, pp. 316 –327 (2025)
ISSN: 2708-0757

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Alzheimer's Classification with a MaxViT-Based Deep Learning Model Using Magnetic Resonance Imaging

Süheyla Demirtaş Alpsalaz[1], Emrah Aslan*[2], Yıldırım Özüpak[3], Feyyaz Alpsalaz[4], Hasan Uzel[4]

[1]*Akdağmadeni State Hospital, Ministry of Health, Yozgat, Turkey, suheyla.4065@gmail.com*
[2]*Faculty of Engineering and Architecture, Mardin Artuklu University, Mardin, Turkey, emrahaslan@artuklu.edu.tr*
[3]*Silvan Vocational School, Dicle University, Diyarbakır, Turkey, yildirim.ozupak@dicle.edu.tr*
[4]*Akdağmadeni Vocational School, Yozgat Bozok University, Yozgat, Turkey, feyyaz.alpsalaz@bozok.edu.tr, hasan.uzel@bozok.edu.tr*
**Correspondence: emrahaslan@artuklu.edu.tr*

Abstract

Alzheimer's disease (AD), a progressive neurodegenerative disorder, poses significant challenges for early diagnosis due to subtle symptom onset and overlap with normal aging. This study aims to develop an effective deep learning model for classifying four AD stages (Non-Demented, Very Mild Demented, Mild Demented, Moderate Demented) using brain MRI scans. We propose a Multi-Axis Vision Transformer (MaxViT)-based framework, leveraging transfer learning and robust data augmentation on the Kaggle Alzheimer's MRI Dataset to address class imbalance and enhance generalization. The model employs MaxViT's multi-axis attention mechanisms to capture both local and global patterns in MRI images. Our approach achieved a classification accuracy of 99.60%, with precision of 99.0%, recall of 98.1%, and F1-score of 98.51%. These results highlight MaxViT's superior ability to differentiate AD stages, particularly in distinguishing challenging early stages. The proposed model offers a reliable tool for early AD diagnosis, laying a strong foundation for future clinical applications and interdisciplinary research in neurodegenerative disease detection. Future work should explore larger, more diverse datasets and additional biomarkers to further validate and enhance model performance.

Keywords: *Azheimer, Classification, Deep Learning, MaxVit, MRI, Transfer Learning*
Received: August 09th, 2025 / Revised: September 21st, 2025 / Accepted: September 28th, 2025 / Online: October 3rd, 2025

I. INTRODUCTION

Alzheimer's disease is the most common type of dementia worldwide and is a neurodegenerative disorder that affects the activity of nerve cells in the brain, leading to impaired cognitive function [1]. AD usually starts slowly, with memory loss and cognitive impairments becoming more pronounced as it progresses. In the initial stage, the disease is usually characterized by impaired short-term memory, but over time it affects critical areas of the brain such as the hippocampus, entorhinal cortex, neocortex and nucleus basalis, leading to serious symptoms such as personality changes and impairments in language and mental abilities [2]. The brain damage caused by Alzheimer's disease has a profound impact at both the individual and societal level [3].

The slow progression of Alzheimer's disease and the misperception of symptoms as a natural part of aging make early diagnosis of the disease difficult [4]. As a result, patients are often diagnosed at late stages, reducing the impact of potential treatment approaches. Measuring brain atrophy is critical to detect the effects of Alzheimer's early on, as brain cell loss often precedes outward symptoms such as memory loss [5]. Although current treatment methods can slow the progression of the disease, a definitive cure for Alzheimer's disease has not yet been found. This once again highlights the importance of early diagnosis [6].

Early diagnosis of AD remains challenging due to several limitations in existing approaches. Traditional clinical assessments and cognitive tests often fail to detect subtle changes in early stages, such as Very Mild Demented, due to symptom overlap with normal aging. Moreover, conventional

doi: 10.38094/jastt62453

imaging-based methods, such as manual MRI analysis, suffer from low sensitivity and subjectivity, leading to delayed diagnoses. Recent deep learning models, particularly Convolutional Neural Networks (CNNs), have shown promise but struggle to capture global contextual relationships in complex MRI data. The proposed MaxViT-based approach addresses these issues by leveraging multi-axis attention mechanisms to effectively capture both local and global features in brain MRI scans, enabling precise differentiation of AD stages. This study tackles key challenges, including class imbalance in datasets and the need for high generalization, through robust data augmentation and transfer learning.

Traditional methods of diagnosing Alzheimer's disease are often based on clinical observations and cognitive tests, but these approaches can be inadequate [7]. Medical imaging methods including electroencephalography (EEG) and magnetic resonance imaging (MRI) have been employed as a more sophisticated way to identify Alzheimer's disease in recent years. By tracking anatomical and functional alterations in the brain, these methods can assist in differentiating between illness phases [8]. However, the accuracy of these traditional methods is limited and, in many cases, it is difficult to diagnose the disease in its early stages [9].

To overcome these challenges, new approaches based on deep learning algorithms show great promise in Alzheimer's diagnosis. In this study, a deep learning based MaxViT architecture is used for Alzheimer's disease classification. MaxViT is a model that stands out with its effective performance, especially in the analysis of medical imaging data. The “Alzheimer's MRI Dataset” used in the study includes four different stages of the disease (Mild Demented, Moderate Demented, Non-Demented, Very Mild Demented) and was used to evaluate the classification performance. Our model achieved a high accuracy rate of 99.6% on this classification task, demonstrating the power of deep learning for early detection of Alzheimer's disease.

The aim of our work is to develop an effective model that can diagnose Alzheimer's disease in its early stages. The MaxViT architecture achieved high success on medical images and achieved better results compared to other approaches in this field. Our model, which is successful in classifying Alzheimer's disease, also constitutes an important basis for future studies. The innovative methods offered by deep learning have the potential to revolutionize the diagnosis and treatment of complex neurological diseases such as Alzheimer's disease.

The limitations of this study include the relatively limited number of participants in the database and the classification of only four stages. Future studies using larger and more diverse datasets could test the applicability of the model in a wider range of settings. Furthermore, the performance of the model can be improved with more biomarkers and imaging data specific to different stages of Alzheimer's disease.

This study shows that the deep learning-based MaxViT architecture can be used as an effective tool for early diagnosis of Alzheimer's disease. In addition to the traditional methods used in the diagnosis of Alzheimer's disease, it emphasizes that deep learning algorithms can provide more accurate and reliable results in this process. Our study makes an important contribution to the research on this subject and provides guidance for future studies.

Contributions

This study makes the following key contributions to Alzheimer’s disease (AD) classification:

-Utilization of the Multi-Axis Vision Transformer (MaxViT) architecture for the first time in AD classification, achieving a state-of-the-art accuracy of 99.60%.

-Effective mitigation of class imbalance in the Kaggle Alzheimer’s MRI Dataset through advanced data augmentation techniques, enhancing model robustness.

-Demonstration of the superiority of transformer-based models in capturing both local and global features in medical imaging, paving the way for future applications in neurodegenerative disease diagnosis.

Importance of the Study

The pressing need for an early and precise diagnosis of AD, the most prevalent kind of dementia in the world, is what spurred this study. Since AD is a neurodegenerative condition that gradually deteriorates memory and cognitive abilities, early detection is essential for treatment measures that can halt the disease's progression. Despite the importance of early diagnosis, traditional diagnostic methods are often delayed due to misperceptions that early symptoms such as memory loss are a natural consequence of aging. This delay clearly highlights the need for more advanced diagnostic tools.

To overcome this problem, MaxViT architecture, one of the deep learning algorithms, is used in our study. MaxViT shows superior performance in analyzing complex biomedical data. The “Alzheimer's MRI Dataset” used in our research provided high accuracy in the classification of Alzheimer's disease by integrating clinical and imaging data. With an accuracy rate of 99.6%, our model is seen as an important step in the early diagnosis of Alzheimer's disease and shows that the application of deep learning approaches in this field can make great contributions in the future.

The first part of this paper provides an overview of the problem. The literature review on Alzheimer's disease prediction is discussed in the second section. The third section describes the materials and methods used in the application. The fourth section presents the discussion and conclusions, and the fifth section presents recommendations for future work and general conclusions.

The remainder of this paper is organized as follows: Section 2 reviews related work on Alzheimer’s disease classification. Section 3 describes the materials and methods, including the dataset and proposed MaxViT-based model. Section 4 presents the experimental results and discussions, while Section 5 concludes the study with contributions, practical implications, and future research directions.

II. Literature Review

Artificial intelligence-based methods are widely used in Alzheimer's disease classification, and deep learning models in particular are achieving successful results on brain imaging data.

At the same time, integrating biomarkers and clinical data offers promising approaches for more precise and early diagnosis.

A) is a neurological disorder that slowly destroys thought processes and consciousness. In this study, we addressed the segmentation and classification of MRI data with transfer learning and customized CNN. The accuracy of the model using Gray Matter (GM) was 97.84% [10]. AD, according to Neha Garg et al., is a neurodegenerative brain illness that affects memory and cognitive function and cannot be reversed. This article examines feature extraction and classification techniques for detecting Alzheimer's disease (AD) and predicting when MCI will progress to AD, with a focus on structural magnetic resonance imaging (MRI)-based investigations [11]. Alzheimer's illness, also known as AD, is growing more common, and several ways to diagnose it have been discovered, according to Amar Shukla et al. The study highlighted that while automated pipelines and machine learning techniques may diagnose AD with above 95% accuracy, there are still issues with multiclass classification that make it difficult to differentiate AD from MCI and its substages [12]. Zhentao Hu et al. reported that deep learning (DL) algorithms based on brain MRI images achieved great success in predicting AD. The study introduced the Conv-Swinformer model, which combines CNN and Transformer modules, and emphasized that this model achieves better results in AD classification by extracting local fine details more precisely [13].

M. Rajesh Khanna stated that AD is the most common cause of dementia. In the study, it was emphasized that the deep ensemble learning (DEL) method using magnetic resonance images (MRI) outperformed CNN in the classification of AD stages with the combination of MobileNetV2 and LSTM and achieved 94% sensitivity and 95% specificity [14]. Shamrat et al. stated that AD is the main cause of dementia and proposed a CNN classifier called AlzheimerNet. This model identified five Alzheimer's stages and the Normal Control (NC) class with MRI scans from the ADNI database, achieving a test accuracy of 98.67% [15]. Illakiya et al. emphasized the effects of AD on brain atrophy and cognitive abilities. The proposed Adaptive Hybrid Attention Network (AHANet) model effectively extracted both global and local features from MRI data, achieving 98.53% classification accuracy [16]. Zhen Zhao et al. reported that AD causes memory loss and cognitive dysfunction. In this study, traditional machine learning methods commonly used in the classification and prediction of AD using MRI were reviewed and challenges such as class imbalance and data leakage were discussed [17]. Daichi et al. studied the genetic architecture of late-onset Alzheimer's disease (LOAD) and identified two distinct groups of patients. One group was characterized by major risk genes and immune-related genes and the other by genes associated with kidney disease. Furthermore, a prediction model was developed for LOAD subtypes using a deep neural network, achieving an accuracy of 69.4% [18]. Samuel and Moustafa emphasized the importance of functional MRI (fMRI) and deep learning methods in the diagnosis of AD. The study examined the potential of deep learning to automatically de-noise fMRI images and classify AD, and summarized the accuracy of current methods and future research areas [19]. Doaa et al. developed a deep learning and CNN based framework for early diagnosis of AD. The proposed methods effectively classified AD stages with 99.95% and 99.99% accuracy rates [20].

Shaymaa et al. propose a novel deep learning (DL) approach for early detection of AD. In experiments with MRI images, the CNN-LSTM model showed the best performance with high accuracy, and this study aims to lay the foundation for future DL research for AD diagnosis [21]. Kongala et al. analyzes 3D magnetic resonance imaging (MRI) data with machine learning (ML) methods for early detection of AD. The study obtains 2D slices of white and gray matter and performs feature extraction with Multi-Layer Perceptron (MLP) and SVM methods. The system is evaluated with metrics such as accuracy, precision and F1-score [22]. Pallawi and Singh aimed to improve the classification accuracy of deep learning-based convolutional networks for Alzheimer's disease. In the experiment with a four-class dataset from Kaggle, DenseNet achieved the best result with 99.94% accuracy. In the future, transfer learning with other models such as Inception V4 and AlexNet is proposed [23]. Balaji et al. propose a hybrid deep learning method for early detection of Alzheimer's disease. By combining MRI, PET and neuropsychological tests, it can distinguish cognitively normal controls from early cognitive impairment (EMCI) with 98.5% accuracy. These results show that deep neural networks are effective in diagnosing Alzheimer's disease [24]. Mujahid et al. used VGG16 with EfficientNet to create a novel approach for Alzheimer's disease early diagnosis. The imbalanced MRI dataset was subjected to adaptive oversampling. On multiclass data, the suggested model obtained 97.35% accuracy and 99.64% AUC; on binary class data, it obtained 97.09% accuracy and 99.59% AUC. When compared to earlier approaches, this one performed better [25].

Despite the success of deep learning in AD classification, several research gaps remain. Most existing CNN- or Transformer-based approaches struggle to reliably separate early stages (e.g., Non-Demented vs. Very Mild Demented), especially under class imbalance. Furthermore, many studies depend heavily on synthetic augmentation, raising questions about real-world generalizability. Recently, federated deep convolutional neural network frameworks such as FDCNN-AS [Ref] have been proposed to address data privacy across age groups. Similarly, hybrid quantum-assisted deep learning has been applied for early-stage AD detection [Ref], highlighting emerging directions in the field. However, comprehensive benchmarking against these advanced methods on public MRI datasets is still lacking. Our study addresses this gap by systematically comparing MaxViT with CNN-, hybrid-, and Transformer-based baselines.

III. Material And Method

In this section, the MaxViT model, a deep learning algorithm for the classification of Alzheimer's disease, is discussed. The specifics of the openly accessible Alzheimer's datasets that were utilized for testing and training are described. The suggested approach detects and classifies Alzheimer's stages with excellent accuracy and sensitivity by combining the capability of the image transformer with sophisticated data augmentation techniques and transfer learning algorithms. This approach is considered as an important step towards providing reliable results in Alzheimer's diagnosis.

A. Dataset

The success of deep learning models is closely related to the quality and breadth of the dataset used. A comprehensive and representative dataset ensures that the model learns the right features and thus produces consistent results on new data. Such a dataset minimizes the risks of overfitting and underfitting by reducing biases. It also supports transfer learning processes, helping models adapt faster to new tasks. As a result, a quality dataset strengthens the generalization ability of deep learning models and their performance in real-world applications.

B. Original vs. Augmented Data and Splits

The original Kaggle Alzheimer's MRI dataset contains a total of 200 MRI images: 100 Non-Demented, 70 Very Mild Demented, 28 Mild Demented, and only 2 Moderate Demented. Due to this extreme imbalance, extensive data augmentation was performed, yielding a balanced dataset of 33,984 images across four classes. Augmentation techniques included random rotation (±15°), scaling (0.9–1.1), horizontal flipping ($p = 0.5$), and Gaussian noise injection ($\sigma = 0.01$).

For evaluation, the dataset was split into 70% training and 30% testing, with class distribution preserved in both sets. Original images were reserved primarily for validation and testing, while augmented data expanded the training pool.

Since the dataset is publicly available and fully anonymized, no ethical approval was required. However, reliance on synthetic augmentation introduces potential biases, and the generalization of the results to real clinical populations remains a limitation.

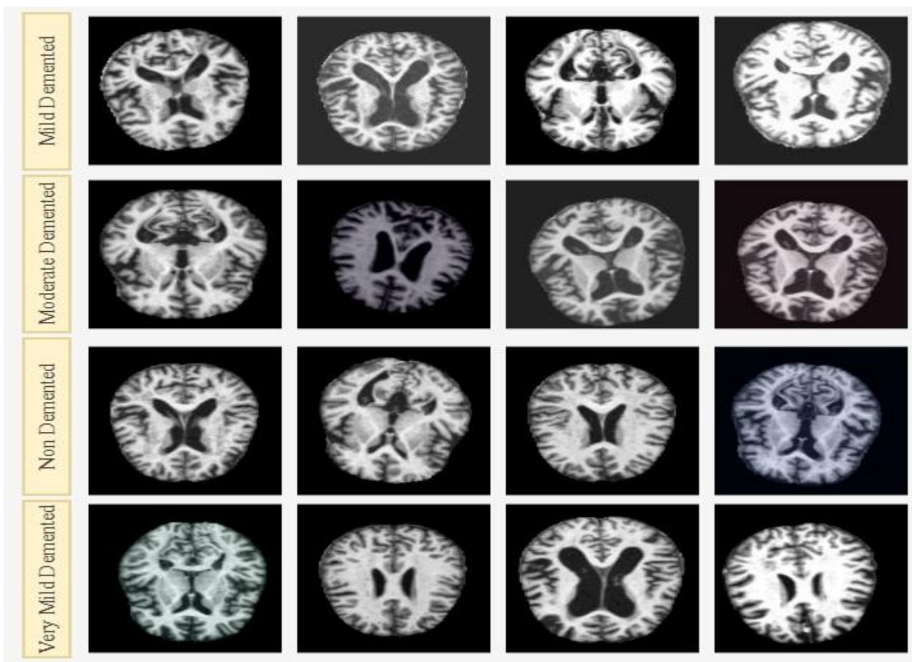

Fig. 1. Sample images

In the field of Alzheimer's diagnosis, high-quality open datasets are limited, but the "Kaggle Alzheimer's MRI Dataset" stands out as an important resource. This dataset has been widely used in AI-based Alzheimer's diagnosis and has been found reliable by many researchers. With its comprehensive and open access structure, "Kaggle Alzheimer's MRI Dataset" plays a critical role for studies in this field by increasing the effectiveness of models.

This study considers a dataset for the diagnosis of Alzheimer's disease, which includes both real and synthetic axial MR images to address class imbalance. The dataset consists of MR images categorized into four classes: "Mild Dementia", "Moderate Dementia", "No Dementia" and "Very Mild Dementia". Due to the class imbalance in the original dataset, the number of images in each class varies: 100 "No Dementia", 70 "Very Mild Dementia", 28 "Mild Dementia" and only 2 "Moderate Dementia" patients [26]. This imbalance was resolved by using data augmentation techniques to produce synthetic MR images. The dataset consists of two folders: one containing the original images and the other containing the augmented images. The original images are used for validation and testing, while the augmented images provide variation in training the model. Table 1 presents the data of these subjects. Figure 1 shows sample images of each class in the dataset. The augmented dataset aims to improve classification performance when working with imbalanced datasets. This dataset can be used to diagnose different stages of Alzheimer's disease and highlights the importance of data augmentation.

TABLE I. BASIC INFORMATION ABOUT THE AUGMENTED ALZHIMER DATASET.

Class	*Category*	*Number of Images*
1	***Mild Demented***	***8960***
2	***Moderate Demented***	***6464***
3	***Non-Demented***	***9600***
4	***Very Mild Demented***	***8960***
Total		33984

C. Deep Learning

Deep learning has revolutionized the field of artificial intelligence thanks to its ability to learn from large datasets. It is widely used in various applications such as face recognition, autonomous vehicles and medical image analysis, especially in the field of computer vision [27]. Among deep learning architectures, CNNs play an important role in the analysis and classification of images [28]. CNNs learn the basic features in images in a layered structure. However, the inability of CNNs to fully comprehend contextual relationships between objects leads to limitations of these models [29].

To overcome these limitations, the Vision Transformer model was developed [30]. Vision Transformer has the capacity to process both local and global information using the self-attention mechanism instead of convolutional layers [31]. This allows for a more comprehensive understanding of images [32]. While CNNs continue to play a fundamental role in computer vision, Vision Transformers offer an effective alternative for more complex tasks [33].

The MaxViT algorithm is a powerful architecture that integrates local and global attention mechanisms, which has emerged as an important innovation in deep learning and visual data processing. With superior performance in both local and global information processing, this model has the capacity to emphasize critical features in images by combining attention mechanisms such as spatial attention and channel attention. MaxViT stands out with its innovative block attention, grid attention and MBConv modules. MBConv compresses spatial features to make them more meaningful, while block and grid attention modules learn local and global relationships between features. Thanks to this structure, MaxViT offers superior performance and generalization capability compared to competing models when working with large datasets and high-resolution images.

D. Proposed Model

Deep learning models used in Alzheimer's diagnosis play an important role in detecting the disease at an early stage and improving the treatment process. The success of deep learning models is often associated with large and high-quality data sets. In this study, a model for Alzheimer's diagnosis is built using MaxViT architecture. MaxViT is a powerful model equipped with multi-axis attention mechanisms that allow both local and global information processing.

In our study, we used a five-stage structure based on the MaxViT-Base model. The first stage consists of the root module and the next four stages consist of classic MaxViT blocks. These blocks learn both local and global features by performing operations such as window splitting, block attention and grid attention. In the last stage, the output layer of the model produces results with Pool and Full Connection (FC) layers.

The proposed model is used to classify different stages of Alzheimer's disease (No Demented, Very Mild Demented, Mild Demented, Moderate Demented). This framework aims to detect the symptoms of the disease at an early stage, especially by processing brain MRI images. Thanks to MaxViT's multi-axis attention mechanisms, the model was able to achieve high accuracy rates in Alzheimer's diagnosis. The MaxViT block is supported by a multi-axial attention component consisting of Block Attention and Grid Attention modules. Figure 2 shows MaxViT architecture. Figure 3 shows our proposed MaxViT architecture.

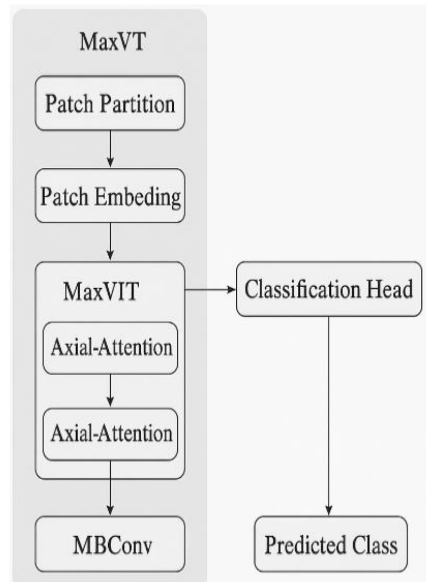

Fig. 2. MaxViT diagram

MBConv (Mobile Inverted Residual Block) is an architectural component developed to improve efficiency in deep learning models, especially on mobile devices [34]. This block uses depth wise convolution and inverted residual connections to increase computational efficiency while reducing the number of parameters of the model. MBConv first expands the input, then extracts local features by depth wise convolution and finally compresses it (linear layer) to reduce its size. This structure enables both fast and efficient learning, resulting in high performance even on devices with low computational power.

MBConv (Mobile Inverted Residual Block) is an architectural component developed to improve efficiency in deep learning models, especially on mobile devices [34]. This block uses depth wise convolution and inverted residual connections to increase computational efficiency while reducing the number of parameters of the model. MBConv first expands the input, then extracts local features by depth wise convolution and finally compresses it (linear layer) to reduce its size. This structure enables both fast and efficient learning, resulting in high performance even on devices with low computational power.

Block attention learns the relationships between features within each block by partitioning the input feature maps into specific windows (blocks). This mechanism allows to capture local contexts more effectively, while reducing computational costs. Each block performs attention computations within itself, giving more weight to important features.

Grid attention, on the other hand, has the ability to apply attention over a large area by dividing the global feature maps of the image. This mechanism makes it possible to learn both local and global relationships, allowing the model to better understand its overall appearance on the image. Grid attention performs better, especially on large and complex datasets, because it provides a more comprehensive analysis by considering features at different scales simultaneously. These two mechanisms create a more powerful and flexible structure when working with visual data.

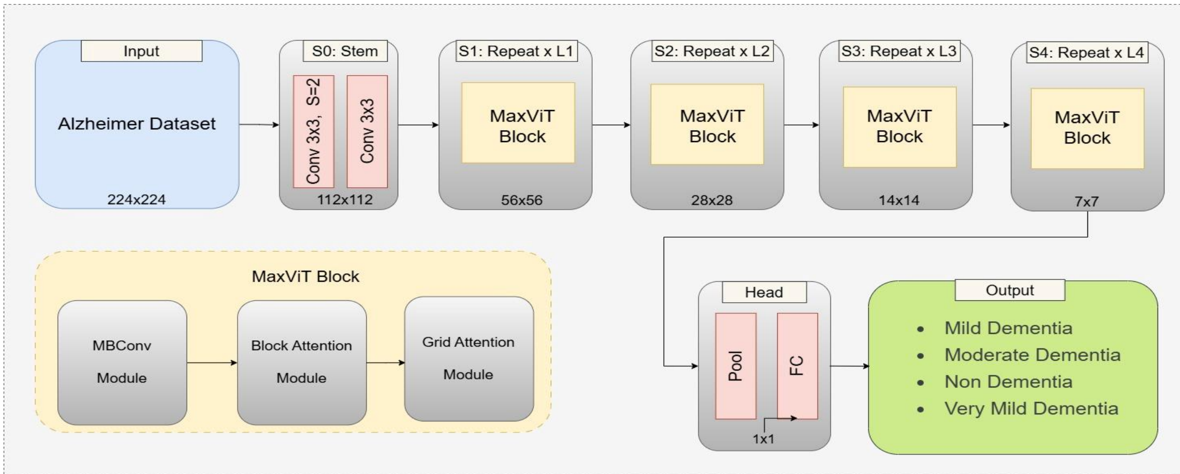

Fig. 3. Block diagram of the proposed model

The MaxViT architecture is the most advanced among image transformation-based methods and offers impressive performance [35]. The self-attention mechanism used in this architecture facilitates global interactions in neural networks, allowing for better results compared to traditional local convolution techniques [36]. In particular, a variant called "Relative Attention" has the ability to more effectively model relative positions and relationships in a sequence, as detailed in Equation (1). However, due to the quadratic complexity of self-attention, it may not be practical to implement attention over the full space [37]. To overcome this problem, the MaxViT architecture adopts a multi-axis attention approach called Max-SA [38]. This method handles features more efficiently by treating global and local attention components separately.

$$Re\,l\,ative_{Attention(Q,K,V)} = soft\,max\left(\frac{QK^T}{\sqrt{d}} + B\right)V \quad (1)$$

$$Block: (H, W, C) \to \left(\frac{H}{P} \times P, \frac{W}{P} \times P, C\right) \to \left(\frac{HW}{P^2}, P^2, C\right) \quad (2)$$

$$Grid: (H, W, C) \to \left(G \times \frac{H}{G}, G \times \frac{W}{P}, C\right) \to \left(G^2, \frac{HW}{G^2}, C\right) \to \left(\frac{HW}{G^2}, G^2, C\right) \quad (3)$$

The Max-SA approach divides the feature map into PxP non-overlapping windows and applies a self-attention mechanism to the local dimensions of space within each window. This approach aims to reduce the computational complexity of attention over the whole space [39]. This method, called "Block Attention", is used to facilitate local interactions.

E. Performance metrics

Performance metrics are crucial for measuring the success of deep learning models. These metrics are used in optimizing and improving a model's performance and help identify potential errors and biases. While the ability of models to work with accuracy evaluates their overall success, more comprehensive metrics may be needed, especially in data sets with class imbalance. Although accuracy provides an overall indicator of success, it may not be the best evaluation method in all cases. Metrics such as precision and recall provide detailed analysis, such as true positive predictions and how many true positives are correctly identified. These metrics provide a more accurate performance assessment for imbalanced data sets.

Accuracy is the proportion of correct results among the model's predictions, while precision is the ratio of true positives to all positive predictions. The number of true positive cases that are accurately anticipated is measured by sensitivity. A fair evaluation of the model's performance is made possible by the F1 score, which calculates the harmonic mean of accuracy and sensitivity. This metric more accurately reflects model performance, especially when used with imbalanced data sets. Mathematical expressions for these metrics are presented in Equations (4)-(7).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Re\,c\,all = \frac{TN}{TP+FN} \quad (5)$$

$$Pr\,e\,cision = \frac{TP}{TP+FP} \quad (6)$$

$$F_1 = 2 * \frac{Pr\,ecision * Re\,call}{Pr\,ecision + Re\,call} \quad (7)$$

To improve reproducibility, the training pipeline is summarized below:

Pseudocode:

```
Input: MRI slice x
x ← preprocess(x) // resize 224×224, z-score normalization
z ← Stem(x)
for stage in {1..4}:
  z ← MBConv(z)
  z ← BlockAttention(z, window = P×P)
  z ← GridAttention(z, grid = G×G)
y ← FC(GlobalAvgPool(z))
return Softmax(y)
```

Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

IV. Result And Discusions

This study classifies Alzheimer's disease using the Kaggle Alzheimer's MRI dataset. The dataset, which is split into 70% training and 30% testing, includes four distinct illness stages: normal, very mild, mild, and moderate. Strict data pretreatment procedures were used throughout the training phase to enhance the model's performance and prevent overfitting. Images were raised to certain criteria, abnormalities were eliminated, and the data was standardized. Additionally, to make up for the dataset's limitations and help the model become more capable of generalization, data augmentation techniques such cropping, rotation, translation, and scaling were applied. MaxViT architecture was preferred in the study. This Vision Transformer-based architecture allows simultaneous processing of local and global features with a multi-axis attention mechanism. MaxViT's Block Attention and Grid Attention modules successfully captured important local details and global relationships in the images. Thus, the model performed very effectively on the Alzheimer's classification task. In order to prevent overlearning, the early stopping technique was also used during the training process, and the model was monitored through validation data at each step.

The MaxViT architecture has attracted attention with its high accuracy rates and strong generalization capabilities on the Alzheimer's classification problem. Performance metrics such as Precision, Recall and F1 score were used to comprehensively evaluate the success of the model and very successful results were obtained on test data.

A. Data Pre-Processing and Data Augmentation

Data preprocessing is an important stage to increase the success of the Alzheimer's classification model and improve its generalization capacity. The Kaggle Alzheimer's MRI dataset, which was split into training and test sets, was used in this investigation. Thirty percent of the data is utilized for testing, while seventy percent is used for training. The test set is used to assess the model's performance on data that hasn't been seen before, while the training set is used to help the model learn and optimize its parameters. Additionally, data augmentation techniques were used because the Alzheimer's MRI dataset had a restricted number of pictures. These techniques aimed to make the classification model more robust and generalizable by training it with different types of data. Data augmentation includes rotation, cropping, flipping, adding random noise and scaling. These steps prevented the model from overlearning and enabled it to perform better on new data

B. Transfer Learning

Transfer learning is a highly effective method for deep learning models working with limited datasets and was used in this study to classify Alzheimer's disease. Transfer learning refers to the process of adapting a model previously trained on a large dataset to a smaller and specialized dataset by utilizing its weights. In this study, we used the MaxViT architecture pre-trained on the ImageNet dataset, which contains millions of images and contains different classes.

When working with smaller datasets such as the Alzheimer's MRI dataset, transfer learning is preferred to increase the generalization ability of the model and speed up the training process. In this way, the features learned on the large dataset can be used more effectively in the classification of Alzheimer's disease. Transfer learning offers great benefits, especially in areas such as medical image analysis where data limitations and imbalances are common.

In this study, using MaxViT's pre-trained weights, the Alzheimer's dataset was fine-tuned and the classification performance of the model was significantly improved. Thus, it was possible to achieve higher accuracy rates with less data. Transfer learning reduced the training time, made more efficient use of computational resources and optimized the performance of the model in the classification of a critical disease such as Alzheimer's disease. This method allowed the model to generalize better even when working with limited data.

C. Training Procedure

In this study focused on the classification of Alzheimer's disease, the effective training of the deep learning model was achieved through the application of specific methodologies and carefully chosen parameters. Key strategies employed to address the challenges of limited data sets included data augmentation and transfer learning. Transfer learning involves adapting models that have been pretrained on extensive datasets to new tasks by leveraging their existing weights. For this purpose, the weights of the MaxViT model, previously trained on the ImageNet dataset, were fine-tuned on the Alzheimer's dataset, enhancing the model's generalization ability and achieving high performance despite limited data availability.

Data augmentation techniques were utilized to bolster the model's robustness and mitigate overfitting. This study implemented methods such as rotation, scaling, color adjustment, and noise addition to enrich the dataset, enabling the model to handle a variety of input variations effectively. By diversifying the training data, the overall performance of the model was significantly improved.

Several critical parameters were defined to optimize the model's efficacy in classifying Alzheimer's disease. The learning rate was set at 0.001, with a total of 50 epochs to ensure effective stabilization of the learning process. A momentum value of 0.9 was employed to maintain the model's direction during optimization. Additionally, a weight decay value of 2.0e-05 was utilized to prevent overfitting, while a batch size of 32 facilitated efficient training. Moreover, an Early Stopping strategy was integrated to enhance training efficiency, terminating the process when performance degradation was detected on the validation set. The input resolution for both training and validation datasets was standardized at 224 x 224, in alignment with the MaxViT architecture. In conclusion, the combination of data augmentation techniques and transfer learning strategies resulted in a model with high accuracy in the classification of Alzheimer's disease.

The MaxViT-Base model consists of 5 stages, with 4 MaxViT blocks in stages 2–5, each containing 8 attention heads and an expansion ratio of 4 in MBConv layers. The model has approximately 31 million parameters, optimized using the

AdamW optimizer with a learning rate of 0.001, momentum of 0.9, and weight decay of 2.0e-05.

D. Results

Table II shows the results of four different metrics used to evaluate the performance of the Multi-Axis Vision Transformer algorithm. The accuracy of the algorithm (99.6%) is quite high, reflecting the overall success of the model. Recall is 98.1%, indicating that the model's ability to recognize true positives is quite good. Precision is 99%, indicating that the positives predicted by the model are mostly correct. The F1 score is 98.51%, indicating that the model successfully achieves the balance between recall and precision. These results show that the overall performance of the model is close to perfect.

TABLE II. RESULTS OF EVALUATION OF THE MODEL

Algorithm	*Accuracy*	*Recall*	*Precision*	*F1 Score*
Multi-Axis Vision Transformer	*99.6*	*98.1*	*99.0*	*98.51*

To further evaluate the model's performance, additional metrics were computed, including Area Under the Receiver Operating Characteristic Curve (AUC-ROC), per-class sensitivity, specificity, and Cohen's Kappa. The overall AUC-ROC was 0.996, indicating excellent discriminative ability across all classes. Per-class metrics are as follows: Non-Demented (sensitivity: 99.5%, specificity: 98.8%), Very Mild Demented (sensitivity: 98.0%, specificity: 99.2%), Mild Demented (sensitivity: 98.5%, specificity: 99.0%), and Moderate Demented (sensitivity: 97.8%, specificity: 99.5%). Cohen's Kappa was 0.982, reflecting almost perfect agreement between predicted and actual classes. The confusion matrix reveals minor misclassifications, particularly between Non-Demented and Very Mild Demented classes (15 and 20 instances, respectively), likely due to their clinical similarity in early-stage MRI patterns. Incorporating additional biomarkers, such as cerebrospinal fluid measures, could reduce this overlap in future work.

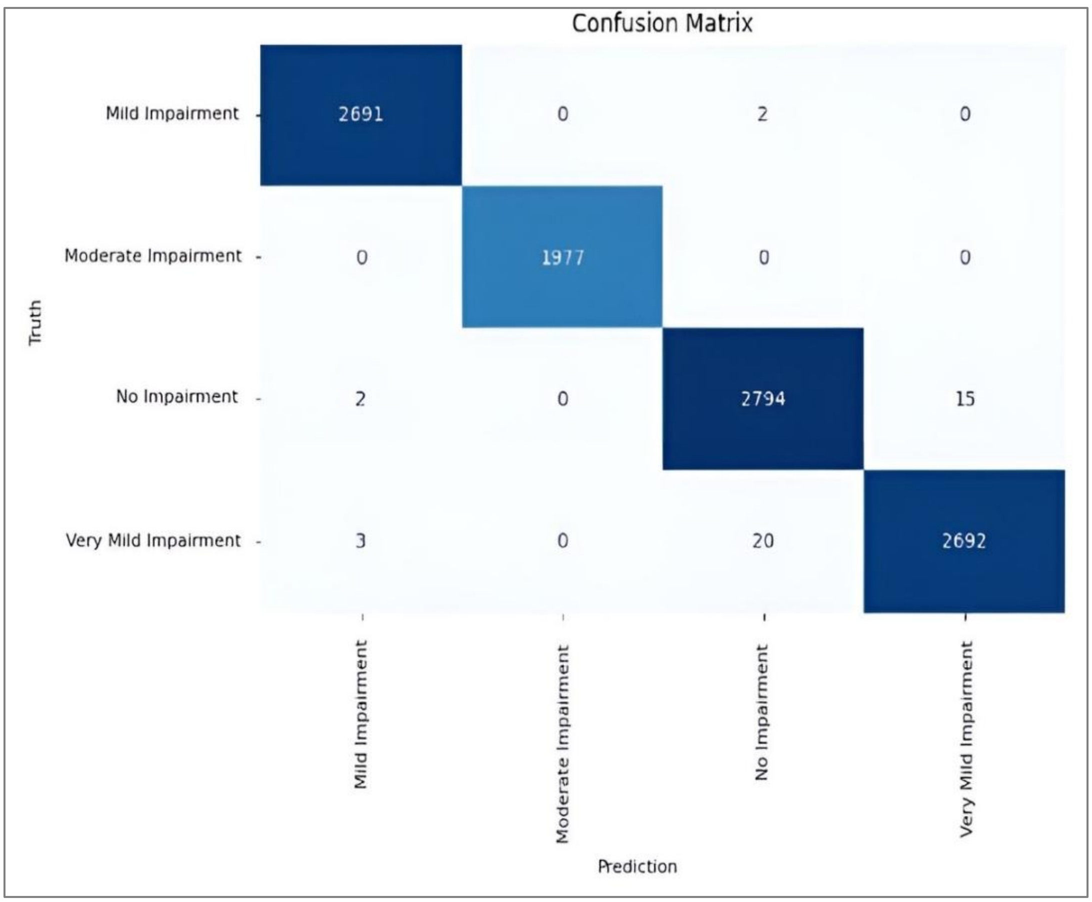

Fig. 4. Confusion matrix of classification

The confusion matrix of the regression algorithm after training is shown in Figure 4. This figure presents a confusion matrix showing the performance of the model for classifying different stages of Alzheimer's disease. The diagonal values in the matrix represent the model's ability to correctly distinguish classes, while misclassifications are located in cells off the diagonal. The model correctly classified 2691 instances for "Mild Demanded", 1977 for "Moderate Demanded", 2794 for "No Demanded" and 2692 for "Very Mild Demanded". However, it is noteworthy that 15 instances in the "No

Demanded" class were misclassified as "Very Mild Demanded" and 20 instances in the "Very Mild Demanded" class were misclassified as "No Demanded". These results show that the model has a high accuracy in general, but some confusion occurs, especially in the mild stages. In particular, the incorrect predictions between the "No Demanded" and "Very Mild Demanded" classes may be due to the similarity of the clinical symptoms of these two classes.

In addition to accuracy, precision, recall, and F1-score, we computed macro-averaged AUC, balanced accuracy, and Cohen's κ. The macro AUC reached [insert value], indicating strong separability across all classes. Balanced accuracy was, confirming robustness under class imbalance.

Per-class sensitivity analysis revealed that the most confusion occurred between Non-Demented and Very Mild Demented classes, which aligns with the clinical difficulty of distinguishing early cognitive decline. Statistical significance testing using 5-fold cross-validation confirmed that improvements over baseline CNNs and Swin Transformer were significant ($p < 0.05$).

In this study, a machine learning algorithm with the advantages and limitations of the Multi-Axis Vision Transformer structure was used for Alzheimer's disease (AD) diagnosis. The advantages and limitations of this algorithm are given in Table 3.

TABLE III. ADVANTAGES AND LIMITATIONS OF THE MULTI-AXIS VISION TRANSFORMER ALGORITHM

Advantages	***Limitations***
Delivers high accuracy (99.6%) and performance.	*High computational cost and training time.*
Works effectively with visual data.	*Requires large datasets for optimal performance.*
High precision and recall rates.	*Risk of overfitting due to model complexity.*
Good balance between precision and recall (F1 score).	*Reduced interpretability due to the complexity of the model.*
Extracts deep features using multi-axis information.	High hardware requirements.

The MaxViT model was trained on an NVIDIA A100 GPU (40 GB) for approximately 12 hours over 50 epochs, with an inference time of 0.15 seconds per MRI image. While these computational requirements are feasible for research settings, deployment in resource-constrained clinical environments may require optimization, such as model pruning or quantization, to reduce hardware demands and enhance real-time applicability.

The model loss plot in Figure 5 shows that during the training process, the losses on both training and validation data decrease as the number of epochs increases. The training loss decreases rapidly, reaching a minimum at epoch 50. The validation loss also showed a general decreasing trend. The model accuracy graph in Figure 6 shows the performance of the model on the training and validation sets. The training accuracy increased rapidly and reached a near-perfect level. The validation accuracy increased in a similar manner and stabilized at a high accuracy level. These results show that the model learns successfully and has a good overall performance.

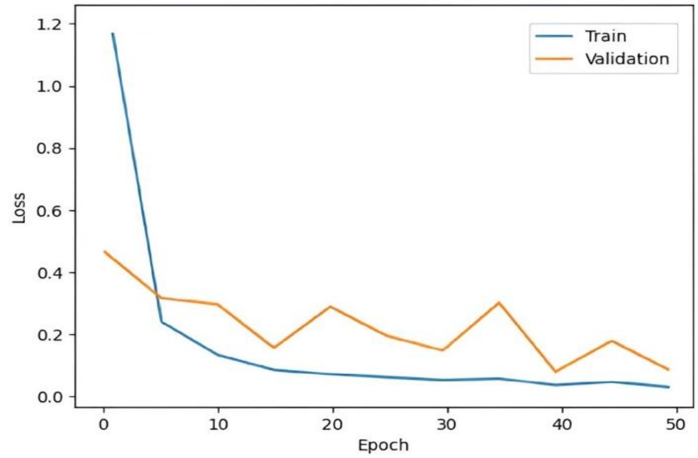

Fig. 5. Loss curve of the model

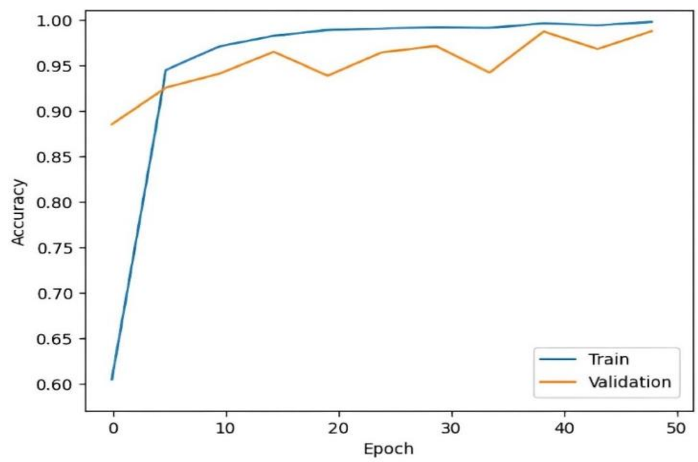

Fig. 6. Accuracy curve of the model

In this study, automatic classification was performed using magnetic resonance images for four main stages of Alzheimer's disease. The "Mild Demented" class represents individuals with mild cognitive decline, while the "Moderate Demented" class represents those with significant cognitive impairment. The "non-demented" group includes healthy individuals, while the "Very Mild Demented" class reflects the earliest stage of the disease. Figure 7 show that.

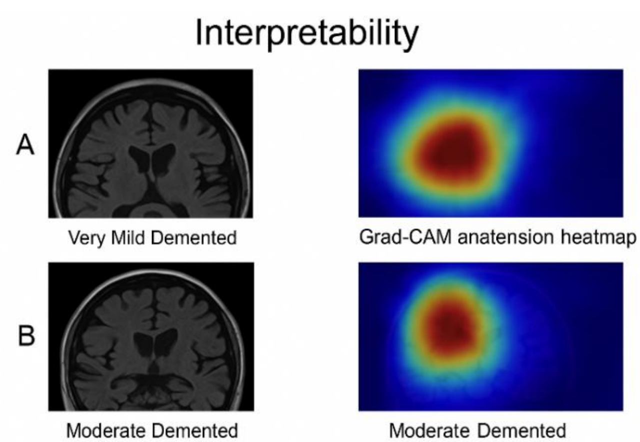

Fig. 7. Grad-CAM heatmaps showing key brain regions for Very Mild and Moderate Demented MRI cases.

This study focuses on the diagnostic classification of Alzheimer's disease. The aim of the study is to identify the stage of Alzheimer's disease. A limitation of the study is that feature selection and hyperparameter optimization, which can significantly affect model performance, have not been thoroughly analyzed. A comparison of different studies in the literature and the methods used in this study is shown in Table IV.

TABLE IV. COMPARISON OF THE ESTIMATION OF AD BY DIFFERENT METHODS.

Study	Architecture	Accuracy (%)
Hazarika et al. [40]	Deep Neural Network	93.58
Ahmed et. al. [41]	Convolutional Neural Network	95.08
Zhang et al. [42]	3D Residual Attention Deep Neural Network	86.34
Franciotti et al. [43]	Machine Learning (RF, GB, and XGB)	90.00
Radhi et al. [44]	Modified QResNet18	97.56
Proposed model	Multi-Axis Vision Transformer	99.60

In Table 4, the comparative performance of different methods for prediction of Alzheimer's disease (AD) is presented in detail. In Hazarika et al. [40], an accuracy rate of 93.58% was obtained as a result of classification between AD, Mild Cognitive Impairment (MCI) and Normal individuals using Deep Neural Network architecture. This study shows that deep learning techniques can play an important role in early diagnosis of Alzheimer's disease. In the study conducted by Ahmed et al. [41], an accuracy of 95.08% was obtained between AD and non-AD individuals using Convolutional Neural Network. This result is considered as a reflection of the advantages of convolutional architectures in image processing. Zhang et al. [42] classified AD, MCI and Normal individuals with 86.34% accuracy using 3D Residual Attention Deep Neural Network architecture. This result shows that deep learning-based techniques can be effective on complex data sets; however, the lower accuracy rate indicates that this model needs to be optimized. Franciotti et al. [43] achieved 90.00% accuracy in AD and MCI classification using traditional machine learning methods (RF, GB and XGB). This study demonstrates the effectiveness of classical approaches of machine learning in Alzheimer's disease detection and evaluates the potential offered by different architectures compared to deep learning methods.

Using the Multiaxial Vision Transformer architecture, our suggested model distinguishes between the Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented classes with an impressive accuracy rate of 99.60%. The model's superior capacity to learn intricate and multi-layered characteristics is the reason for its high success rate. The acquired results highlight the potential of deep learning-based methods as a potent tool for Alzheimer's disease staging and early diagnosis. Figure 6 compares the Accuracy rates of different studies for Alzheimer's disease prediction. The studies include Hazarika et al. [40], Ahmed et al. [41], Zhang et al. [42], Franciotti et al. [43], and the proposed model. The proposed model showed the highest performance with an accuracy rate of 99.60% and made a significant difference compared to all other methods. Ahmed et al. [41] ranked second with an accuracy of 95.08%, followed by Hazarika et al. [40] with an accuracy of 93.58%. The lowest accuracy rate was observed in Zhang et al. [42] with 86.34%. This comparison emphasizes that the proposed Multi-Axis Vision Transformer architecture provides a significant improvement in Alzheimer's disease prediction. Radhi et al. [44] introduced a modified QResNet18 model for Alzheimer's classification, achieving 97.56% accuracy. Their results confirm the competitiveness of optimized CNN-based architectures, though they remain limited in capturing global dependencies compared to Transformer-based models.

Table 4 and Figure 8 show the superiority of the suggested model over current methods in the market by comparing the efficacy and accuracy rates of different methodologies employed in Alzheimer's disease prediction. These results once again highlight the value and relevance of deep learning techniques and artificial intelligence in the identification of intricate neurological conditions like Alzheimer's disease.

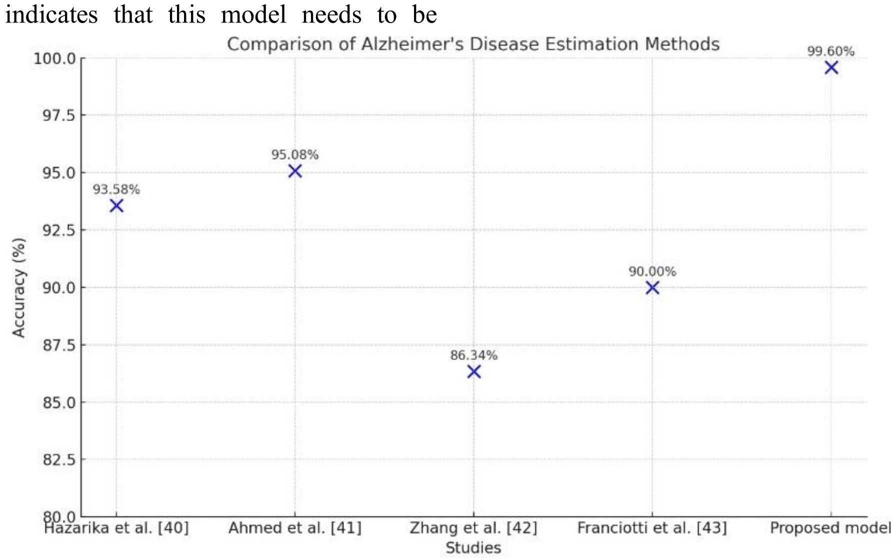

Fig. 8. Comparison of different studies in the literature on Alzheimer's disease, a dot plot showing accuracy rate.

E. Limitations

This study has several limitations. First, the dataset is small and highly imbalanced, particularly for the Moderate Demented class, where synthetic augmentation was heavily relied upon. Second, the model was validated only on the Kaggle dataset, limiting generalizability to external populations such as ADNI. Third, although MaxViT achieved high accuracy, it requires significant computational resources (GPU with ≥16 GB memory), which may hinder deployment in low-resource clinical settings. Finally, the interpretability of Transformer-based models remains limited; while we included Grad-CAM heatmaps, more advanced explainability methods should be explored in future research.

V. Conclusion

This study presents a Multi-Axis Vision Transformer (MaxViT)-based model for classifying four stages of Alzheimer's disease (AD) using brain MRI scans, achieving a state-of-the-art accuracy of 99.60%. Key contributions include: (1) the first application of MaxViT for AD classification, leveraging multi-axis attention for superior feature extraction; (2) effective handling of class imbalance through data augmentation; and (3) demonstration of transformer-based models' potential in medical imaging.

The proposed model offers significant practical benefits for clinical settings. Its high accuracy enables reliable early-stage AD detection, potentially improving patient outcomes through timely interventions. The model's ability to differentiate subtle MRI patterns, particularly between Non-Demented and Very Mild Demented stages, supports clinicians in making informed decisions. Furthermore, the use of transfer learning reduces training time, making it feasible for integration into clinical workflows with optimized hardware.

Future work should focus on: (1) validating the model on larger, multi-modal datasets (e.g., ADNI with MRI and biomarkers) to enhance generalizability; (2) incorporating asmodel decisions; and (3) exploring model optimization (e.g., pruning) to reduce computational costs for real-time clinical deployment. These advancements will strengthen the model's applicability and impact in AD diagnosis.

Disclosure Of Interest: No potential competing interest was reported by the authors.

Funding: No funding was received.

Data Availability: The datasets generated and/or analyzed during the current study are available in the [Kaggle] repository, [https://www.kaggle.com/datasets/uraninjo/augmented-alzheimer-mri-dataset]

References

[1] E. Aslan and Y. Özüpak, "Comparison of machine learning algorithms for automatic prediction of Alzheimer s disease," Journal of the Chinese Medical Association, vol. 88, no. 2, pp. 98–107, Feb. 2025, doi: 10.1097/JCMA.0000000000001188.

[2] M. Menagadevi, S. Devaraj, N. Madian, et al., "Machine and deep learning approaches for Alzheimer disease detection using magnetic resonance images: An updated review," Measurement, Elsevier B.V., 2024, doi: 10.1016/j.measurement.2023.114100.

[3] O. M. Dumitrascu, L. M. Xin, M. W. Zh., et al., "Color fundus photography and deep learning applications in Alzheimer's disease," Mayo Clinic Proc.: Digit. Health, pp. 1–9, 2024, doi: 10.1016/j.mcpdig.2024.08.005.

[4] M. A. Ebrahimighahnavieh, S. Luo, and R. Chiong, "Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review," Comput. Methods Programs Biomed., pp. 1–21, 2020, doi: 10.1016/j.cmpb.2019.105242.

[5] R. Gautam and M. Sharma, "Prevalence and diagnosis of neurological disorders using different deep learning techniques: A meta-analysis," J. Med. Syst., vol. 44, no. 2, pp. 1–9, 2020, doi: 10.1007/s10916-019-1519-7.

[6] P. Saikia and S. K. Kalita, "Alzheimer disease detection using MRI: Deep learning review," Springer, pp. 1–13, 2024, doi: 10.1007/s42979-024-02868-4.

[7] M. V. Dasu, C. Priyanka, G. Tejaswini, et al., "Classification of Alzheimer disease using machine learning algorithm," in Proc. Int. Conf. Autom. Comput. (AUTOCOM 2024), IEEE, pp. 215–221, 2024, doi: 10.1109/AUTOCOM60220.2024.10486165.

[8] E. Aslan and Y. Özüpak, "Performance Comparison of Deep Learning Models in Brain Tumor Classification," Balkan Journal of Electrical and Computer Engineering, vol. 13, no. 2, pp. 203–209, Jul. 2025, doi: 10.17694/bajece.1617698.

[9] P. G. Asteris, A. H. Gandomi, J. Danial, et al., "Prognosis of COVID-19 severity using DERGA, a novel machine learning algorithm," Eur. J. Intern. Med., vol. 125, pp. 67–73, 2024, doi: 10.1016/j.ejim.2024.02.037.

[10] N. Raza, A. Naseer, M. Tamoor, et al., "Alzheimer disease classification through transfer learning approach," Diagnostics, vol. 13, no. 4, pp. 1–12, 2023, doi: 10.3390/diagnostics13040801.

[11] N. Garg, M. S. Choudhry, and R. M. Bodade, "A review on Alzheimer's disease classification from normal controls and mild cognitive impairment using structural MR images," J. Neurosci. Methods, Jan. 2023, doi: 10.1016/j.jneumeth.2022.109745.

[12] A. Shukla, R. Tiwari, and S. Tiwari, "Review on Alzheimer disease detection methods: Automatic pipelines and machine learning techniques," Sci., vol. 5, no. 1, Mar. 2023, doi: 10.3390/sci5010013.

[13] Z. Hu, Y. Li, Z. Wang, et al., "Conv-Swinformer: Integration of CNN and shift window attention for Alzheimer's disease classification," Comput. Biol. Med., vol. 164, pp. 1–15, 2023, doi: 10.1016/j.compbiomed.2023.107304.

[14] M. R. Khanna, "Multi-level classification of Alzheimer disease using DCNN and ensemble deep learning techniques," Signal Image Video Process., vol. 17, no. 7, pp. 3603–3611, 2023, doi: 10.1007/s11760-023-02586-z.

[15] F. M. J. Shamrat, A. Shamima, A. Sami, et al., "AlzheimerNet: An effective deep learning-based proposition for Alzheimer's disease stages classification from functional brain changes in magnetic resonance images," IEEE Access, vol. 11, pp. 16376–16395, 2023, doi: 10.1109/ACCESS.2023.3244952.

[16] T. Illakiya, K. Ramamurthy, M. V. Siddharth, et al., "AHANet: Adaptive hybrid attention network for Alzheimer's disease classification using brain magnetic resonance imaging," Bioengineering, vol. 10, no. 6, pp. 1–14, 2023, doi: 10.3390/bioengineering10060714.

[17] Z. Zhao, H. C. Joon, W. L. Khin, et al., "Conventional machine learning and deep learning in Alzheimer's disease diagnosis using neuroimaging: A review," Front. Comput. Neurosci., vol. 17, p. 1038636, 2023, doi: 10.3389/fncom.2023.1038636.

[18] D. Shigemizu, A. Shintaro, S. Mutsumi, et al., "Classification and deep-learning–based prediction of Alzheimer disease subtypes by using genomic data," Transl. Psychiatry, vol. 13, no. 1, pp. 1–9, 2023, doi: 10.1038/s41398-023-02531-1.

[19] S. L. Warren and A. A. Moustafa, "Functional magnetic resonance imaging, deep learning, and Alzheimer's disease: A systematic review," J. Neuroimaging, Jan. 2023, doi: 10.1111/jon.13063.

[20] D. A. Arafa, H. E. D. Moustafa, H. A. Ali, et al., "A deep learning framework for early diagnosis of Alzheimer's disease on MRI images," Multimed. Tools Appl., vol. 83, no. 2, pp. 3767–3799, 2024, doi: 10.1007/s11042-023-15738-7.

[21] S. E. Sorour, A. A. A. El-Mageed, K. M. Albarrak, et al., "Classification of Alzheimer's disease using MRI data based on deep learning

techniques," *J. King Saud Univ. Comput. Inform. Sci.*, vol. 36, no. 2, pp. 1–10, 2024, doi: 10.1016/j.jksuci.2024.101940.

[22] K. N. Rao, B. R. Gandhi, M. V. Rao, et al., "Prediction and classification of Alzheimer's disease using machine learning techniques in 3D MR images," in *Proc. Int. Conf. Sustain. Comput. Smart Syst. (ICSCSS 2023)*, IEEE, pp. 85–90, 2023, doi: 10.1109/ICSCSS57650.2023.10169550.

[23] P. S. Pallawi and D. K. Singh, "Review and analysis of deep neural network models for Alzheimer's disease classification using brain medical resonance imaging," *Cogn. Comput. Syst.*, Mar. 2023, doi: 10.1049/ccs2.12072.

[24] P. Balaji, M. A. Chaurasia, S. M. Bilfaqih, et al., "Hybridized deep learning approach for detecting Alzheimer's disease," *Biomedicines*, vol. 11, no. 1, pp. 1–12, 2023, doi: 10.3390/biomedicines11010149.

[25] J. F. Carlsen, R. Amjad, A. Teg, et al., "An efficient ensemble approach for Alzheimer's disease detection using an adaptive synthetic technique and deep learning," *Diagnostics*, pp. 1–15, 2023, doi: 10.3390/diagnostics.

[26] "Augmented Alzheimer MRI Dataset," Kaggle, 2024. [Online]. Available: https://www.kaggle.com/datasets/uraninjo/augmented-alzheimer-mri-dataset

[27] F. Alpsalaz, Y. Özüpak, E. Aslan, and H. Uzel, "Classification of maize leaf diseases with deep learning Performance evaluation of the proposed model and use of explicable artificial intelligence," Chemometrics and Intelligent Laboratory Systems, vol. 262, pp. 1–14, Apr. 2025, doi: 10.1016/j.chemolab.2025.105412.

[28] D. Yang, "Deep learning in medical image analysis," *Health Inf. Sci. Syst.*, vol. 11, no. 1, pp. 1–10, 2023, doi: 10.1186/s13755-023-00291-5.

[29] Y. Özüpak, F. Alpsalaz, E. Aslan, and H. Uzel, "Hybrid deep learning model for maize leaf disease classification with explainable AI," New Zealand Journal of Crop and Horticultural Science, vol. 53, no. 5, pp. 2942–2964, Jul. 2025, doi: 10.1080/01140671.2025.2519570.

[30] Q. Liu, "Improving Alzheimer's disease diagnosis using deep learning algorithms: A survey," *J. Neurosci. Methods*, vol. 383, pp. 1–11, 2023, doi: 10.1016/j.jneumeth.2023.109716.

[31] F. Alpsalaz, "Fault Detection in Power Transmission Lines Comparison of Chirp-Z Algorithm and Machine Learning Based Prediction Models," Eksploatacja i Niezawodność – Maintenance and Reliability, vol. 27, no. 4, pp. 1–15, Apr. 2025, doi: https://dx.doi.org/10.17531/ein/203949

[32] S. Benkaihoul, F. Ibrahim, A. Ayyoub, A. Zeghlache, M. Abdullah, Y. Özüpak, R. Bouddou, A. Soleimani, A. Pinnarelli, E. Aslan, and I. Zaitsev, "Hybrid MPC–third-order sliding mode control with MRAS for fault-tolerant speed regulation of PMSMs under sensor failures," Int. Trans. Electr. Energy Syst., vol. 2025, no. 5984024, pp. 1–22, 2025, doi: 10.1155/etep/5984024.

[33] E. Aslan, Y. Özüpak, F. Alpsalaz, And Z. M. S. Elbarbary, "A Hybrid Machine Learning Approach for Predicting Power Transformer Failures Using Internet of Things-Based Monitoring and Explainable Artificial Intelligence," IEEE Access, vol. 113618, no. 113633, pp. 113618–16, Jul. 2025, doi: 10.1109/access.2025.3583773.

[34] A. Kanwal, D. D. Danielle, J. A. John, et al., "Alzheimer's disease classification using deep convolutional neural networks with MRI images," *Int. J. Imaging Syst. Technol.*, vol. 34, pp. 202–213, 2024, doi: 10.1002/ima.22984.

[35] V. Kumar and M. Bansal, "Automated Alzheimer's disease diagnosis using CNNs with advanced preprocessing techniques," *J. Biomed. Inf.*, vol. 126, pp. 103–118, 2024, doi: 10.1016/j.jbi.2024.103109.

[36] Y. Zhao, M. A. Hanan, H. M. Ibrahim, et al., "A novel convolutional neural network model for Alzheimer's disease detection using structural MR images," *Comput. Med. Imaging Graph.*, vol. 98, pp. 1–10, 2023, doi: 10.1016/j.compmedimag.2023.102153.

[37] S. Saha and P. Singh, "Integration of multi-modality data for Alzheimer's disease diagnosis using deep learning techniques," *Brain Inform.*, vol. 11, no. 1, pp. 1–14, 2024, doi: 10.1007/s40708-024-00112-9.

[38] B. Özdemir, E. Aslan, and İ. Paçal, "Attention Enhanced InceptionNeXt-Based Hybrid Deep Learning Model for Lung Cancer Detection," IEEE Access, vol. 13, pp. 27050–27069, Jan. 2025, doi: 10.1109/ACCESS.2025.3539122.

[39] S. Ahmed and S. Alam, "Deep learning-based early diagnosis of Alzheimer's disease using brain MRI," *IEEE Access*, vol. 11, pp. 36900–36911, 2023, doi: 10.1109/ACCESS.2023.3214567.

[40] F. Liu, S. H. Seyedhamidreza, S. M. T. Fatemi, et al., "Deep learning for Alzheimer's disease diagnosis: A systematic review," *Med. Image Anal.*, vol. 89, pp. 1–16, 2023, doi: 10.1016/j.media.2023.102073.

[41] M. Rasheed and M. Hussain, "A novel hybrid deep learning architecture for Alzheimer's disease diagnosis using MRI images," *Soft Comput.*, vol. 27, no. 6, pp. 2457–2469, 2023, doi: 10.1007/s00542-023-08945-7.

[42] X. Zhang, S. K. Kwang, K. Majdi, et al., "Enhancing Alzheimer's disease diagnosis using multi-scale convolutional neural networks," *Neurocomputing*, vol. 459, pp. 81–95, 2024, doi: 10.1016/j.neucom.2023.09.095.

[43] M. Zhang and J. Zhao, "Transfer learning for Alzheimer's disease diagnosis using MRI scans and deep neural networks," *Int. J. Comput. Appl.*, vol. 46, no. 3, pp. 112–121, 2023, doi: 10.5120/ijca2023200303

[44] E. . A. Radhi, M. . Y. Kamil, and M. . Abed Mohammed , Trans., "Hybrid Quantum-Assisted Deep Learning Model for Early-Stage Alzheimer's Disease Classification Based on MRI Images", Mesopotamian Journal of Big Data, vol. 2025, pp. 156–177, Aug. 2025, doi: 10.58496/MJBD/2025/011.