IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 7 March 2024, accepted 12 April 2024, date of publication 17 April 2024, date of current version 1 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3390186

RESEARCH ARTICLE

Dual Attention Convolutional AutoEncoder for Diagnosis of Alzheimer's Disorder in Patients Using Neuroimaging and MRI Features

SHAHA AL-OTAIBI[1], MUHAMMAD MUJAHID[2], AMJAD REHMAN KHAN[2], (Senior Member, IEEE), HAITHAM NOBANEE[3,4,5], JABER ALYAMI[6,7], AND TANZILA SABA[2], (Senior Member, IEEE)

[1]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P. O. Box 84428, Riyadh 11671, Saudi Arabia
[2]Artificial Intelligence and Data Analytics (AIDA) Laboratory, CCIS, Prince Sultan University, Riyadh 11586, Saudi Arabia
[3]College of Business, Abu Dhabi University, Abu Dhabi, United Arab Emirates
[4]Oxford Center for Islamic Studies, University of Oxford, OX1 2JD Oxford, U.K.
[5]Faculty of Humanities & Social Sciences, University of Liverpool, L69 3BX Liverpool, U.K.
[6]Radiological Sciences Department, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[7]Smart Medical Imaging Research Group, King Fahd Medical Research Center, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: Jaber Alyami (jhalyami@kau.edu.sa)

This work was supported by Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, through the Researchers Supporting Project PNURSP2024R136.

ABSTRACT Alzheimer's disease is a neurodegenerative disease causing memory loss and brain protein accumulation. Early diagnosis is crucial for clinical trials and patient care. Magnetic resonance imaging (MRI) methods have improved diagnosis and prognosis, but doctors need to interpret images proficiently. Deep learning technology has shown potential in detecting Alzheimer's disease, but the disease progresses slower in early phases. A new dual-attention convolutional autoencoder model is presented, offering improved detection abilities and potential for real-time use in Alzheimer's disease diagnosis. The study utilized two datasets: the first ADNI dataset, which includes three classes (MCI, CN, and AD), and the second Alzheimer's Disease Neuroimaging Dataset, which includes two distinct classes (AD and MCI). We analyze the effectiveness of our proposed model by evaluating key performance metrics such as accuracy, precision, sensitivity, specificity, F1 score, and AUC score. In addition, we utilize cross-validation and mean absolute error to validate our model while also fine-tuning the parameters. Based on experimental data, the proposed model accurately detected Alzheimer's disease with an accuracy of 0.9902 ± 0.0139. Based on the results, the proposed model demonstrates excellent performance compared to the existing methods described in the literature. The proposed mode achieves precision, sensitivity, and specificity of 0.9882 ± 0.0587, 0.9898 ± 0.0865, 0.9912 ± 0.0872 respectively. The model achieved an AUC score of 0.9992 for MCI and 0.9919 for AD class. Furthermore, the proposed method can enhance the affordability of Alzheimer's disease diagnostics and increase the rate of early AD detection by facilitating remote healthcare.

INDEX TERMS Alzheimer's disease, ADNI dataset, MRI features, dual attention CNN, healthcare.

I. INTRODUCTION

Alzheimer's disease (AD) is a progressive neurological disorder that gradually affects cognitive performance up to death and overall functioning. Alzheimer's disease, is the sixth most common cause of death in the United States [1] and the most widespread neurodegenerative disease. Alzheimer's disease has a profound and severe impact on millions of people worldwide, leading to significant consequences for both the affected individuals and their close ones. This condition

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy.

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

demonstrates a prolonged progression, increasingly affecting cognitive processes such as speech, thinking, brain function and praxis. The experience is psychologically stressful for both the affected patient and their close friends [2]. Initially, those suffering with AD may experience impairments with memory, lack of interest, and the performance of daily tasks. The risk of acquiring Alzheimer's disease increases twofold every five years for persons aged 65 and older [3].

Currently, there is no available cure for Alzheimer's disease that can cure its effects. Nevertheless, certain drugs, such as acetylcholinesterase inhibitors, have the potential to slow down the cognitive and functional decline linked to the condition and improve the overall quality of life for those affected [4]. Therefore, an immediate and precise diagnosis in the early stages of Alzheimer's disease is crucial to minimise the length of therapy. Furthermore, obtaining an accurate diagnosis is crucial in order to avoid unnecessary testing that might provide inaccurate positive or negative conclusions [5]. Alzheimer's is not a common feature of the ageing process. Memory challenges are typically one of the initial signs of Alzheimer's disease and related dementia's. Alzheimer's disease accounts for around 60-80% of cases of dementia [6].

In 2020, it was projected that 5.8 million individuals in the United States who were 65 years or older will be diagnosed with Alzheimer's disease. By the year 2060, it is projected that its population would increase nearly substantially to reach 14 million. The anticipated cost of treating Alzheimer's disease in 2010 ranged from $159 to $215 billion. It is expected that these costs would rise from $379 billion to over $500 billion annually by 2040 [7]. While mortality rates for heart disease and cancer are decreasing, Alzheimer's disease is witnessing an upward trend in death ratios. Alzheimer's disease and other dementia's have been shown to be inaccurately documented on death certificates, leading to a potential underestimation of the prevalence of Alzheimer's disease among the elderly [8], [9]. Symptoms and causes are shown in Figure 1. Magnetic resonance imaging (MRI) is a highly popular imaging technique utilised to examine the inner parts of the body as well as the structure of the brain.

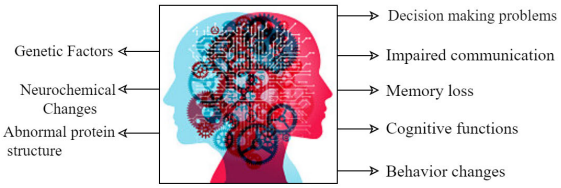


FIGURE 1. Causes and Symptoms of alzheimer's disease. Alzheimer's disease is a complicated neurodegenerative brain disease that gradually impairs thinking. Neuroimaging can reveal brain structure and function changes, making it robust for understanding Alzheimer's disease.

Magnetic resonance imaging is a technique that gathers information from atomic nuclei located inside the body by using strong magnets and low-energy radiofrequency waves [10]. MRI is currently the most commonly employed form of in real life cerebral imaging for Alzheimer's disease analysis because it provides detailed information about the structure of the brain, allows for the identification of factors such as tumors, strokes, and white matter that may contribute to a relapse in cognitive function, and can be used to characterize regional brain atrophy. This is because the most popular method for in real life brain imaging is now magnetic resonance imaging. The ADNI utilizes advanced neuroimaging methods that can accurately identify the early phases of Alzheimer's disease. The development of new neuroimaging techniques, such as MRI, has made it possible to detect and share previously concealed anatomical and molecular indicators associated with Alzheimer's disease [11].

A. RESEARCH MOTIVATIONS

Alzheimer's disease is a long-term, irreversible brain problem that causes patients to lose their ability to think and remember things. The authors suggest a CNN-based model that uses brain MRI data to identify people with Alzheimer's disease and put them into two distinct classes. The suggested model stacks up against other CNN models in terms of accuracy, precision, recall, and F1 score. The main thing this study adds is a 12-layer CNN model that is more accurate, which is better than all the other CNN models [12]. Another study proposes constructing a large number of deep 2D convolutional neural networks in order to extract specific features from localised brain images. By integrating these features, the ultimate classification for the detection of Alzheimer's disease can be made. Transfer learning architectures, Inception-V3 and Xception, were used to analyse the entire brain image. Furthermore, the construction of a CNN required the use of separable convolutional layers, which permitted the automated extraction of generic properties from image input for classification purposes [13].

More research is required to properly grasp the limitations of deep learning in diagnosing Alzheimer's disease, despite its significant potential. A detailed experimental analysis is conducted on recent improvements in deep learning for the detection of Alzheimer's disease, as well as a discussion of the limitations of this technique. Furthermore, we provide a dual-attention convolutional autoencoder mode for precise identification of AD, with the intention of overcoming these constraints. In order to enhance the strength and efficiency of our model, we incorporated other deep transfer learning techniques and sampling methods during its training process. We utilized numerous metrics to evaluate the effectiveness of our model.

Deep transfer learning and sampling methods advance our understanding and solution to this difficult challenge. These technologies illuminate complex medical research. Deep learning analyzes large datasets and finds insights that conventional methods. Furthermore, it not only enables diagnosis but also facilitates in-depth understanding, rendering it a crucial component in unraveling the mechanisms of AD. Sampling methods improve narrative accuracy by creating a plot focused on quick recognition and individual

attention. Every data point adds to a large dataset that could be used to stop disease development. They hope information can eliminate Alzheimer's disease's negative effects, allowing for recovery and better health. Deep learning and sampling represent a new level in Alzheimer's disease research. Understanding becomes influential, while advancement becomes despairing.

B. RESEARCH CONTRIBUTIONS

In light of the absence of a proven treatment for AD at this stage, prompt detection and intervention are essential for dealing with symptoms and impeding the disease's progression. The individual may encounter challenges with verbal expression, cognitive abilities, motor skills, digestion, movement, and interpersonal interaction as the condition worsens. The important contributions of our study are as follows:

- We present a new dual-attention convolutional autoencoder model that will make it easier to diagnose Alzheimer's disease using brain imaging. However, unlike earlier models that needed extra layers to get accurate identification, the proposed model has amazing detection abilities and shows promise for being easily used in real-time situation.
- The proposed model is all-encompassing, covering all subsequent phases; therefore, traditional methods, such as feature extraction, are superfluous. Moreover, our model outperforms its predecessors in several contexts, such as binary and multi-class classification.
- Adaptive Synthetic Sampling (ADASYN) is employed to generate supplementary training data for the minority class, which comprises a comparatively small proportion of the dataset's samples. As a consequence, the procedure for identifying problems related to AD detection is streamlined, a task that is frequently achieved via k-fold cross-validation.
- To demonstrate that the deep transfer learning approach is effective by increasing the classification accuracy of the model and making statistical T-tests and comparisons with other techniques.
- This paper presents a potentially effective method for the automated detection of Alzheimer's disease that has the capacity to enhance the patient experience and facilitate early detection. Training and testing losses, as well as learning challenges, are assessed in this study.

II. LITERATURE REVIEW

Odusami et al. [14] present a multimodal fusion approach that utilises the discrete wavelet transform, a mathematical tool, to examine the data. Transfer learning was employed to enhance the performance of this approach by utilising a pre-trained VGG16 neural network. The IDWT was employed to recreate the merged image. A pre-trained vision transformer was employed to categorise the fused images. According to the benchmark Alzheimer's disease neuroimaging dataset, the MRI test data shows an accuracy of 81.25% for distinguishing between Alzheimer's disease and early mild cognitive impairment. The experiments utilize MRI and PET datasets, and features are retrieved using the ViT model. The sole metric employed to assess accuracy in this study. The primary aim of the study [15] was to provide a fast and efficient diagnostic approach for identifying people who are in a state of good health before the occurrence of moderate cognitive impairment. The experiments were carried out using the MRI multiclass dataset, and deep learning was employed to extract features. The test data was evaluated using four important metrics. The proposed model underwent a comprehensive evaluation by comparing it to the most recent convolutional neural network models published in the literature. It demonstrated impressive F1 score, sensitivity, specificity, and accuracy, with an average of 95.41%, 95.21%, 95.01%, and 97.92%, respectively.

Balaji et al. [16] introduce a mixed-methods deep learning approach for the detection of Alzheimer's disease. The utilisation of multimodal imaging techniques, including positron emission tomography (PET), MRI, the Long Short-term Memory algorithm, and the outcomes of widely administered cognitive assessments, can facilitate the early detection of Alzheimer's disease. To augment the accuracy of the system, the proposed methodology incorporates Adam's optimisation and makes adjustments to the learning weights. The method achieves a good precision score in differentiating between patients who are mentally competent and those who have EMCI. Drawing from these discoveries, it is feasible to instruct deep neural networks to independently identify distinctive imaging characteristics linked to Alzheimer's disease and utilise them in order to produce an imprecise prognosis for the ailment. The evaluations were transmitted using the MRI multiclass dataset, in which CNN was used to extract features and LSTM was used for classification. Four fundamental metrics were applied in order to assess the test data.

The study [17] demonstrates that deep learning models may be utilised for Alzheimer's disease classification. XGBoost was employed for performing classification tasks, whereas DenseNet-201 was utilized for extracting features. The training dataset was obtained from Kaggle, and the final classifier results were assessed using several permutations of the voting mechanism. DenseNet201 surpasses DenseNet121 in some aspects. The DenseNet201 model applied to Gaussian NB produced an accuracy of 91.75%, a specificity of 96.5%, and an F1 score of 90.25%. These results would enhance the capabilities of radiologists to pursue a second opinion or utilise a simulator. The work advanced three two-system techniques to accurately diagnose and predict Alzheimer's disease. The Feed Forward Neural Network design independently implements DenseNet-121 and GoogLeNet features. FFNN uses principal component analysis. The GoogLeNet and Dense-121 models were merged to extract features, and an FFNN was employed for the categorization of AD disease. Multiple assessments

were employed to evaluate the model. Identifying AD and predicting its progression were both highly successful [18].

For direct voice-based detection and early-stage AD severity prediction, they created, tried, and evaluated a new AI-driven model from start to finish. Their proof-of-concept study shows a community-based screening method that was easy to use, open to everyone, and very sensitive. This test can be used to figure out the risk and find problems early, before they get worse. Notably, AI experts and doctors need to work together in well-designed, representative, large-scale studies to make sure that AI-enabled solutions are truly useful before they are widely used to help diagnose dementia. The ADReSSo binary dataset, along with held-out and external test sets, was utilized to identify AD patients using acoustic features. In the end, this will allow AI to get better at early diagnosis, which is very important for improving the quality of life for people with dementia [19].

Based on their findings, they were able to apply a DL-based LSTM model to properly predict who will get Alzheimer's disease based on MRI brain images. With an impressive AUC of 0.97, the model was able to distinguish between positive and negative groups with remarkable accuracy. This demonstrates how well it worked and how useful it may be as a diagnostic tool for early disease detection. Patients and healthcare workers can better plan appropriate treatments and interventions if they have the ability to correctly predict AD from the start. Many medical fields have demonstrated success with deep learning methods. The body of research demonstrating that these techniques work in medical care is growing, and this study adds to it. The Kaggle dataset of MRI brain images was utilized for a four-class classification task, and the data was split using shuffle shift cross-validation. It was also planned to make access to the DL model easier using an Internet-based tool. The deep learning model predicts that they will get Alzheimer's disease. Users can share images of magnetic resonance [20].

Illakiya et al. [21] present a novel attention-based adaptive feature fusion framework for the classification of moderate cognitive impairment and Alzheimer's disease in this study. In contrast, the overwhelming majority of prior studies have neglected the criticality of distinguishing between global and local attributes as distinct entities. Due to this, they have developed AHANet, a technique for extracting significant attributes and subsequently integrating them adaptively so as to take advantage of the additional attributes of the attention modules. The ENLA module was designed to gather data at a global level, as opposed to the Coordinate Attention module, which exclusively acquires spatial information to enable precise localization. The ADNI dataset was utilized to apply enhanced non-local attention and coordinated attention to image features. The evaluation was conducted using the DenseNet-169 model, employing various metrics.

Enrichment analysis was utilised to elucidate the underlying mechanisms of AD, while six AI methods were applied for the detection and identification of AD. A notable disparity in the methylation of 3684 CpGs was observed between the AD and control groups, with a p-value of less than 0.05. All six AI methods in a distincT-test group exhibited high prediction accuracy, with an AUC range of 0.949. Deep learning achieved a sensitivity and specificity of 94.5% and an AUC (95% CI), respectively [22]. The dataset of MRI scans acquired from the ADNI database was employed to train and assess the proposed model. The ADNI six-class dataset was partitioned using random splitting and utilized for deep feature extraction for classification purposes. The performance of the classification model was evaluated on a separate test set. As a consequence, the final dataset consisted of 60,000 image data points, which were evenly distributed among the six classes. Five existing models were trained and tested, and their initial test accuracies were 78.84%, 86.85%, 78.87%, 80.98%, and 96.31%, in that order: VGG16, MobileNetV2, AlexNet, ResNet50, and InceptionV3. By implementing an RMSprop optimizer and modifying the InceptionV3 model, the AlzheimerNet is generated with a learning rate of 0.00001 [23].

One of the best methods now available for identifying Alzheimer's disease is magnetic resonance imaging. It is a crucial and challenging effort, however, to detect early changes in the brain caused by Alzheimer's disease. The dataset utilized for this study was 6219 MRI images. The dataset was divided into three subsets: training, testing, and validation. The hold-out validation technique was employed to evaluate the performance of the model. Performance metrics and losses were used to assess the model's performance. They employ four different deep learning models: the CNNN, the DenseNet121, the ResNet101, and the VGG-16. According to the results, CNN outperformed the other models by a wide margin, with an AUC of 99.26%, an accuracy of 97.60%, a recall of 97%, and a nominal loss of 0.091 [24]. The research [25] presents an innovative approach that combines many modes of data fusion. The experiments conducted on the AD neuroimaging initiative dataset demonstrate that the accuracy of distinguishing between cognitive normal and Alzheimer's disease is 97%, between CN and moderate cognitive impairment is 94%, and between AD and MCI is 97.5%.

The study [16] presents a hybrid deep learning approach for Alzheimer's disease early diagnosis. The method integrates MRI, PET, and neuropsychological test results using multimodal imaging and a CNN with the Long Short-term Memory algorithm. The methodology makes use of Adam's optimization algorithm and updates learning weights in order to improve accuracy. When it comes to telling cognitively normal controls. The study [26] developed a method for the identification of Alzheimer's disease using a fully convolutional network architecture and MRI images from the Kaggle dataset. The pipeline started by using a modified U2-net to segment the hippocampal regions. Following the extraction of 851 radiomic features, 37 relevant features associated with Alzheimer's disease were selected.

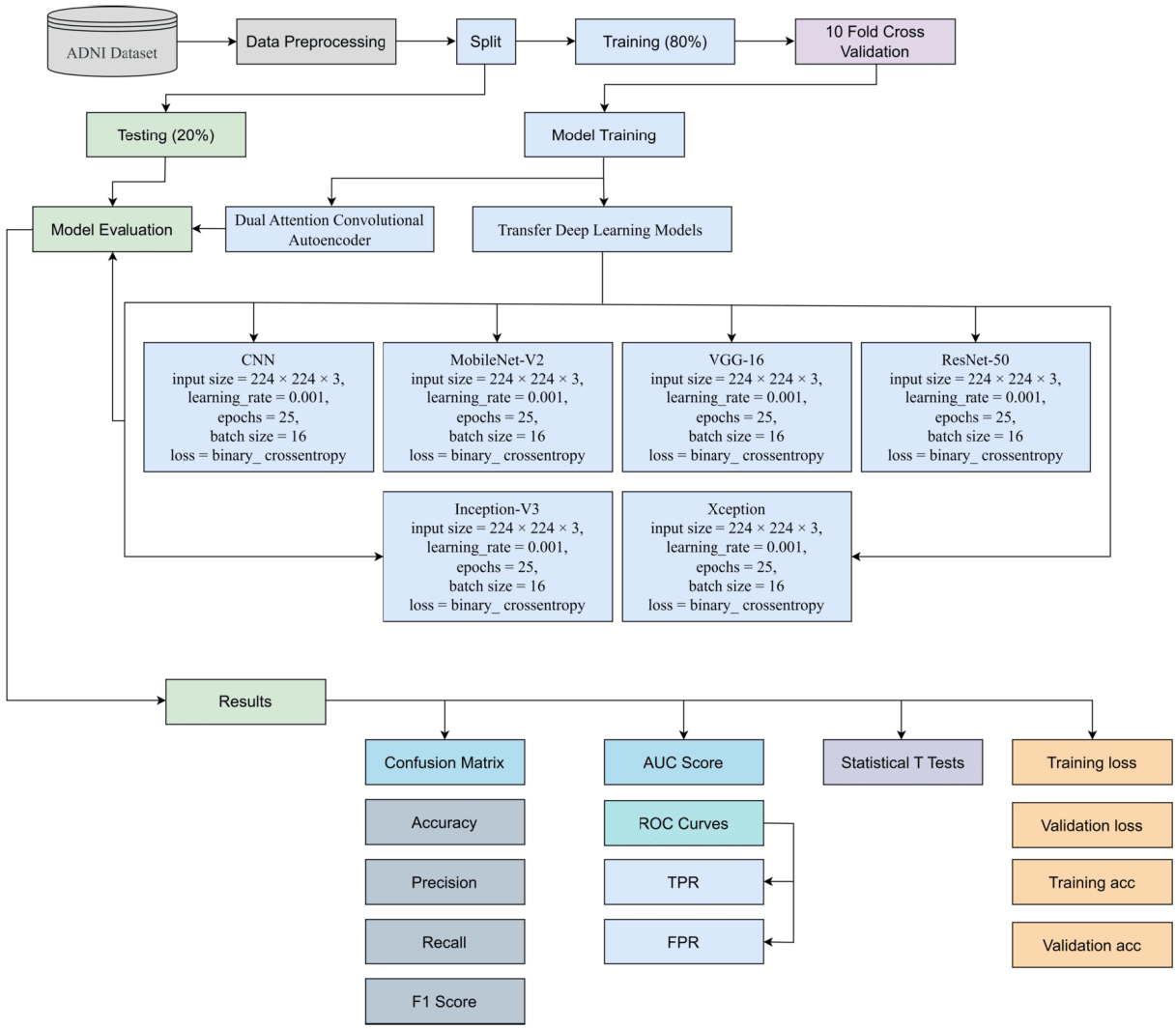


FIGURE 2. Flow Diagram for Automatic Detection of Patients with Neuroimaging Alzheimer's Disease.

To distinguish between people with Alzheimer's disease and normal controls, four classifiers were used [27].

III. PROPOSED METHODOLOGY

Figure 2 shows the work flow for Alzheimer's disease detection. The Alzheimer's disease neuroimaging dataset was collected from the public Kaggle data repository. Then we resized the images according to the transfer deep learning model requirements. The dataset is split into training and testing sets with an 80:20 ratio. The dataset will follow a 10-fold cross-validation procedure, where each observation will be trained and tested k-1 times and once, respectively. Despite resampling at each iteration, it is feasible for Split to repeatedly select the same components for the test set. There is a possibility that certain data points were removed from the ultimate sample. When it comes to preventing model overfitting, K-fold cross-validation is better than splitting the dataset into separate training and test sets. After this, we applied several deep transfer learning models. All deep transfer models are trained using 224 × 224 input size and 3 channels, 0.001 learning rate, 25 epochs, 16 batch sizes, loss categorical crossentropy for multiclass, and binary crossentropy for binary class.

The development of a deep learning architecture [28] is inherently a formidable task. Weights have been assigned and subsequently updated iteratively in deep learning prior to the training period. Time is consumed significantly during the deep learning process when weights are adjusted iteratively; this results in the model becoming overfit. Transfer learning the most efficient approach to addressing the aforementioned challenges. Transfer learning utilises the knowledge acquired from pre-trained models that have undergone rigorous training using large datasets. Additionally, it modifies the hyperparameters and optimises the hidden layers of the pre-trained models. We can enhance the efficiency of deep learning with transfer learning.

Convolutional neural networks [29] are extensively used in the study and practice of mental image processing. They are able to recognise and classify particular features that are depicted in images. They have a wide range of applications,

some of which include computer vision, medical analysis of images, and image classification. The functioning of CNN is carried out through the use of neural networks. The activation, convolution, and pooling layers make up the network's structural support structure. The term "VGG" refers to a specific type of neural network that is capable of categorising and identifying a wide variety of different diseases. In Inception V3, there are both symmetrical and asymmetrical building components, for a grand total of around 20 million different qualities. Each layer is made up of either a dropout layer, a convolutional layer, an average pooling layer, a maximum pooling layer, a totally connected layer, or a fully connected layer, depending on whether it is a fully connected layer or a completely connected layer.

A. DATASET INFORMATION

The dataset employed in this study was acquired from Kaggle. Kaggle is a platform for experts in data science and machine learning, sponsored by Google LLC. These datasets can be used to build a model on an online data science platform. Users can engage in data science projects and cooperate with others. The ADNI dataset contains a total of 5154 images, which were utilised to develop the 2D axial images in this data collection. The classification consists of three classes: MCI (mild cognitive impairment), CN (common normal), and AD (Alzheimer's disease). The AD consists of 1124, the MCI is 2590, and the CN is 1440. As per the dataset source, the images were manually gathered from different sources, and each label was thoroughly validated. The second dataset, named Alzheimer's Disease Neuroimaging_ADNI_Dataset, consists of two distinct classes: AD and MCI. The number of images in AD is 965, while MCI includes 689 images. The samples for neuroimaging Alzheimer's disease is illustrated in Figure 3. The study involved 177 subjects, 69 females and 108 males, aged 48.7 to 90.9 years, with an average age of 74.6. The extensive data collected allowed for rigorous analysis and a significant understanding of the phenomena under investigation. Figure 4 represents the average age group of all subjects.

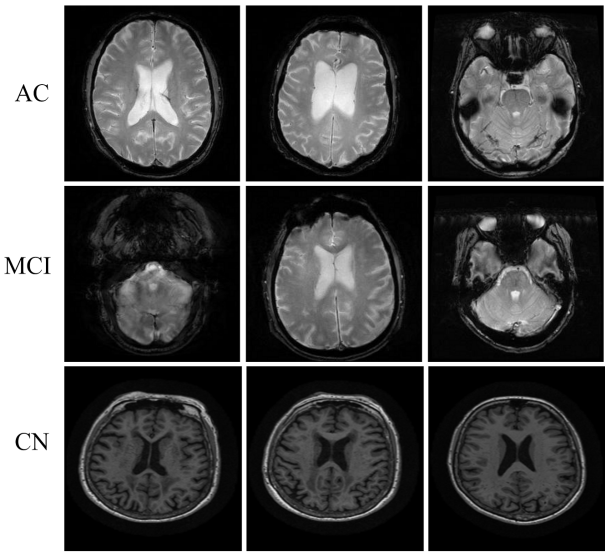


FIGURE 3. Neuroimaging Alzheimer's Disease Sample.

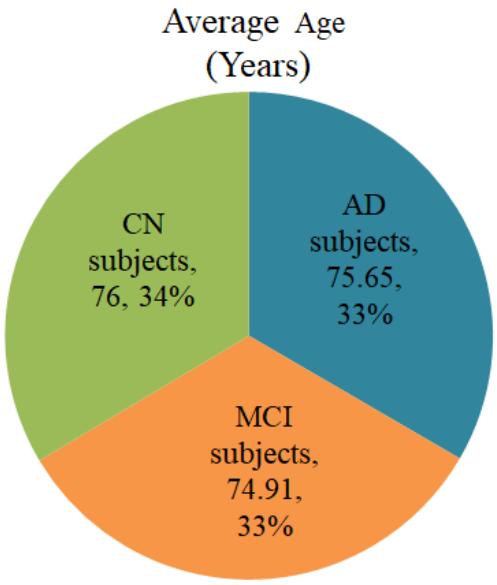


FIGURE 4. Neuroimaging Alzheimer's Disease age group for all subjects.

B. PREPROCESSING

Preprocessing the image enhances its quality, enabling more efficient analysis. To effectively train and make predictions using new data, it is essential that the images align with the input dimensions of the network [30]. The images in the ADNI dataset exhibit a variety of pixel counts. Preprocessing medical images is crucial in order to get superior results and improve image quality in the context of deep learning. Due to the inconsistent width and height of the images, we needed inputs of a consistent size to train the deep learning models. As a result, the specifications of each image were specified as $224 \times 224 \times 3$.

C. ADASYN TECHNIQUE

To tackle the issue of class imbalance in the dataset, it was imperative to identify the minority class, which had a smaller number of instances compared to the dominant class. The distribution of cases between the two classes may be determined by employing computations that utilise the imbalance ratio. The ADADYN algorithm was utilised in an attempt to tackle this problem. ADASYN is an adaptive synthetic oversampling technique that is based on the ideas of SMOTE. The primary objective of ADASYN is to generate synthetic samples for cases belonging to the minority class, which are notoriously difficult to handle. This is achieved by modifying synthetic sample distributions in real-time to align with the density distribution of the instances. This facilitates the attainment of the intended outcome. ADASYN estimates the density distribution of instances from minority classes, prioritising the development of synthetic samples from classes with lower densities. Consequently, when generating synthetic samples, special focus is placed on challenging

scenarios, highlighting the areas that demand the utmost caution. This can be perceived as allocating additional focus to the components that necessitate greater caution. ADASYN addresses the limitations of SMOTE in situations where there is a higher level of class imbalance or when instances of the minority class are spread in complex patterns. This is accomplished through the use of adaptive sample synthesis modification. When the distribution of individuals from the minority class occurs, it is executed [31]. The objective of the ADASYN endeavour is to strike a balance between under-sampling the minority population and overfitting.

We provide Alzheimer's disease data for both binary and multi-class classification. Table 1 displays the distribution of multi-class using both imbalanced and balanced data employed in the current study. After implementing the ADASYN oversampling technique to balance the dataset, AD has 2610 samples, MCI has 2589 samples, and CN has 2592 samples. Table 2 displays the binary class distributions of AD and MCI. For the balance distribution, AD has 976 samples while MCI has 966.

TABLE 1. Class distribution for multi-class Classifiers.

Class	Imbalance Distribution	Balance Distribution
AD	1124	2610
MCI	2590	2589
CN	1440	2592
Total	5154	7791

TABLE 2. Class distribution for binary class Classifiers.

Class	Imbalance Distribution	Balance Distribution
AD	965	976
MCI	689	966
Total	1654	1942

D. PROPOSED DUAL ATTENTION CONVOLUTIONAL AUTOENCODER

The 224 × 224 with 3 colour channels is employed as an input layer for the DACNA architecture for automatic detection of patients with neuroimaging Alzheimer's disease, as shown in Figure 5. The encoder, decoder, and attention mechanism blocks are used. The 3 layers of Con2D and 2 maxpooling in the encoder block are then added to the attention mechanism. After adding the attention mechanism, three conv2D and two-up sampling layers are implemented, followed by an activation layer. After these, a convolutional neural network model was developed and combined with the dual attention mechanism layers. A convolutional network consists of 11 layers. Each layer has its own features.

The CNN model consists of three 2D convolutional layers (2DConv), two max-pooling layers with pool sizes of 2 × 2, two dropout layers with dropout rates of 0.2 and 0.3, and three dense layers. In the first layer of the CNN model, there are 256 filters and a 3 × 3 kernel. The first generation of 2DConv layers had this name. In the activation layer that follows, the ReLU is used as the activation function. Currently, there is a max-pooling layer with a 2 × 2 pixel pool size. The CNN models were developed using the Adam algorithm and the binary_crossentropy loss function for binary class and categorical for multiclass. The model selection procedure involved a total of 25 epochs. The pseudocode of the proposed method is illustrated in algorithm 1.

Encoder-decoder architectures in the medical field lack the capacity to efficiently leverage the overall relationships between objects. Structures based on neural networks, on the other hand, rely heavily on the output of long-term memorization. The dual attention network, a unique framework specifically designed for medical images of natural landscapes, addresses the multiple tasks. This approach generates the spatial attention map via an attention mechanism rather than convolutional layers. This enables the network to directly obtain full global information.

To capture feature interactions in both the spatial and channel domains, DANet incorporates a location attention module and a channel attention module. The channel attention module is intended to reduce noise while improving the quality of critical channels. The DANet position attention module, on the other hand, lets you record important contextual information and dynamically combine relevant data from a global perspective, no matter what scale it is. The addition of spatial and channel interactions results in a more accurate feature representation for scene segmentation.

1) MAXPOOLING LAYER

Feature maps can be reduced in size by employing the pooling layer. These modifications reduce the number of variables that must be collected and the computations that the network must perform. The pooling layer gathers the data that the convolutional layer produces in a specific area of the feature map [32].

$$Y_{a,b} = \text{Max}(X_a \times \text{poolsize}[0], b \times \text{poolsize}[1]) \quad (1)$$

2) FLATTEN LAYER

The data must be flattened, or made into a one-dimensional collection, in order to be forwarded to the next stage. The convolutional layer's output was flattened to create a single, lengthy feature vector. It is related to the final classification model.

3) DROPOUT LAYER

The dropout layer is a mechanism that blocks certain neurons' impact on the following layer while maintaining the activity of all other neurons. A dropout layer can be used to eliminate specific hidden neurons from a hidden layer or to take specific characteristics out of the input vector. Dropout layers prevent the training set from becoming overly well-fitted, which makes them a crucial component of training. The first set of training samples influences learning far more than it ought to

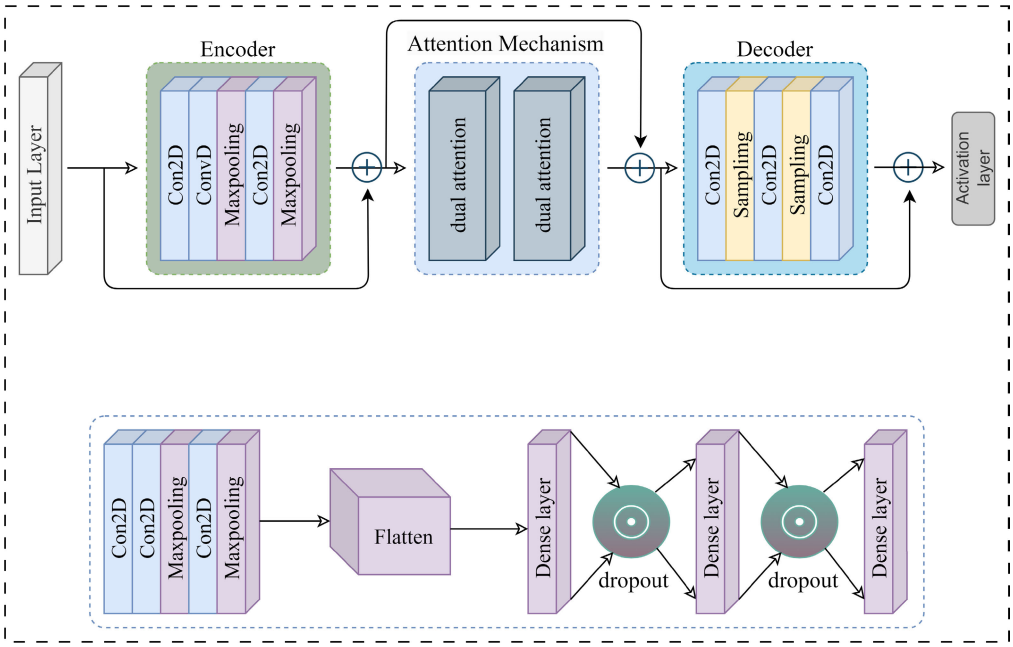


FIGURE 5. Dual attention convolutional AutoEncoder architecture.

in the absence of these data [33].

$$Y = X \odot M \tag{2}$$

4) RECTIFIED LINEAR UNIT

The activation function known as Rectified Linear Units is extensively implemented in deep learning models. The ReLU function is a sequentially linear function that returns the value of the input for positive inputs and zero for negative inputs. In many neural network designs, the default activation function has been chosen because it consistently works better than other options and makes training more efficient. The affordability of ReLU function computations is a common attribute of their utility, resulting from their straightforwardness.

5) LOSS FUNCTION

Binary classification methods employ binary crossentropy loss when there are only two classes, while categorical crossentropy is utilised when there are more than two classes. Transform the provided values to an adjusted range between 0 and 1. Consequently, there exists a minuscule differentiation between the input and output of the two functions. Categorical data is represented as a vector, where each element of the vector represents the probability of a specific class occurring. This vector is used to generate a new vector. It takes information from more than one class and transforms it into a random distribution based on that information [34].

IV. RESULTS

The detection and diagnosis of Alzheimer's disease using neuroimaging were extensively investigated in this paper by employing various deep transfer learning methods. The individual in charge of an organization or institution is usually in charge of making big decisions and keeping an eye on how things are running. A novel Dual Attention Convolutional AutoEncoder (DACNA) model is developed and evaluated in this study to see how it performs in comparison to popular models such as Inception, ResNet, and CNN. The purpose of the research was to find out how well this new method worked. After a comprehensive examination, the subject's potential and limitations are determined. For an accurate assessment of the model's performance, comprehensive research is required. Here, we look at metrics like specificity, accuracy, recall, and F1 score. The study's overarching goal is to advance medical image processing and, more specifically, neuroimaging-based disease diagnosis. The incorporation of deep learning methods has the potential to enhance diagnostic accuracy, which is crucial for the timely and accurate diagnosis of neuroimaging diseases. Hyperparameters Setting for the Proposed method is shown in Table 3.

A. EXPERIMENTAL STRUCTURE

The experiment utilized a Dual Attention Convolutional AutoEncoder model for diagnosing neuroimaging Alzheimer's diseases. The system used an NVIDIA GeForce RTX 3080 graphics card, 32 GB of RAM, and an Intel i7-10750H CPU. TensorFlow and Python 3.10 were used for the DACNA model. The Jupyter Notebook platform was used, and Anaconda was used for the Python 3.10 programming language. The experiment assessed various experiments using five performance metrics, with the initial experiment focusing on binary classification and the subsequent multi-classification task.

Algorithm 1 Proposed DACNA Methodology Framework

Import all necessary libraries
Load Dataset and Split:
Train data path
Test data path
Input Image_size:
Image width, Image height, Channel = (224 × 224 × 3)
ADASYN Technique:
ADA = ADASYN (random state = 42)
Train_data resampled, Train_labels resampled = ADA.fit_resample(Train_data, Train_labels)
Develop DACNA model:
model = input layer
Encoder:
model = Conv2D, Conv2D, Maxpooling, Conv2D, Maxpooling ▷ 5 Different encoder layers
Attention Mechanism:
model = Dual attention, Dual attention ▷ 2 Attention Mechanism layers
Decoder:
model = Conv2D, Sampling, Conv2D, Sampling, Conv2D ▷ 5 Different encoder layers
model = Activation layer
CNN Layers: ▷ Define CNN Different layers
model = Conv2D, Conv2D, Maxpooling, Conv2D, Maxpooling
model = Flatten()
model = Dense(128), Activation = 'Relu'
model = Dropout rate ▷ 20%
model = Dense(128), Activation = 'Relu'
model = Dropout rate ▷ 30%
model = Dense(3), Activation = 'Softmax' ▷ Three class
model = Dense(3), Activation = 'Sigmoid' ▷ Two class
Optimizer = Adam, Loss ← *categorical_crossentropy* ▷ Three class, ← *binary_crossentropy*, ▷ Two class,
Model Training:
Epochs = 30, Learning_rate = 0.001, Batch_size = 16
Evaluation using Test_data:
Accuracy, Precision, Recall, F1 score and Specificity
Predictions (P):
for J, Prediction in Test_data (P) **do**
    if P[0] > P[1] **and** P[0] > P[2] **then**
        print(Case, i=J + 1, : MCI)
    else if P[1] > P[0] **and** P[1] > P[2] **then**
        print(Case, i=J + 1, : AD)
    else
        print(Case, i=J + 1, : CN)
    end if
end for

Evaluation metrics measure the predictive precision of a model in relation to predicted results. The recall measure evaluates a model's ability to correctly identify all positive instances. Precision and the F1 score are important performance indicators that assess the accuracy of positive predictions and overall correctness, respectively. Additionally, the specificity and specific accuracy of negative predictions are also assessed. These indicators are crucial for understanding the model's performance, ensuring its reliability, and driving improvements. The selection of a particular measure is contingent upon the goals of the endeavor. This enables the development of well-informed assessments across a wide range of deep learning tasks.

TABLE 3. Hyper-parameters setting for the proposed method.

Settings	P-Value
Loss	binary and categorical cross entropy
Learning Rate	0.001
Batch Size	16
Epochs	30
Dropout rate	20% and 30%

1) ACCURACY

Accuracy is a crucial metric in classification models, indicating the proportion of correctly predicted events. It is calculated by comparing accurate predictions to incidents. Imbalanced datasets can lead to reduced accuracy and reliability, resulting in incorrect conclusions. Accuracy is most effective when combined with other measures like recall, precision, or F1 score for a comprehensive analysis.

$$\text{Accuracy} = \frac{\text{TrPo} + \text{TrNe}}{\text{TrPo} + \text{FaNe} + \text{FaPo} + \text{TrNe}} \tag{3}$$

2) PRECISION

Precision is a numerical measure that assesses the level of accuracy of a model in predicting positive results by a model. The computation involves dividing the total number of projected positive cases by the number of accurate positive forecasts. There is a negative correlation between increased accuracy and the occurrence of false positives, particularly when the impact of incorrect predictions is substantial.

$$\text{Precision (Prec)} = \frac{\text{TrPo}}{\text{TrPo} + \text{FaPo}} \tag{4}$$

3) RECALL OR SENSITIVITY

Recall, also known as sensitivity is a classification metric that evaluates a model's ability to identify each positive example. It is calculated as the ratio of correctly predicted positive events to actual positive events. A higher recall value means that more positive cases were detected, which is beneficial in situations where correct identification of all positive cases is necessary, like in medical diagnostics.

$$\text{Recall (Rec)} = \frac{\text{TrPo}}{\text{TrPo} + \text{FaNe}} \tag{5}$$

4) F1 SCORE

The F1 score evaluates classification ability by combining recall and precision. It provides a fair assessment of a model's performance by calculating the harmonic mean of these

criteria. Higher scores indicate a balance between precision and recall.

$$\text{F1 score} = 2 * \frac{\text{Prec} * \text{Rec}}{\text{Prec} + \text{Rec}} \quad (6)$$

5) AUC

The Area Under the Curve (AUC) is a metric used to evaluate the effectiveness of a classification model, particularly in multidimensional scenarios. The ROC curve compares the actual positive rate with the false positive rate. Performance is positively associated with greater AUC values, which can vary between 0 and 1.

$$\text{AUC} = (0.5) * (\text{TrPoRa} + \text{TrNeRa}) \quad (7)$$

6) TRPORA

The True Positive Rate (TrPoRa) is a crucial metric that measures the percentage of positive cases accurately identified by a model, indicating improved model performance, especially in scenarios involving substantial costs.

$$\text{TrPoRa} = \frac{(\text{TrPo} + \text{TrPo})}{\text{FaNe}} \quad (8)$$

7) FAPORA

The False Positive Rate (FaPoRa) is a metric used in classification to measure the percentage of negative instances a model incorrectly identifies as positive, with a lower FaPoRa indicating improved model accuracy in avoiding erroneous.

$$\text{FaPoRa} = \frac{(\text{FaPo} + \text{FaPo})}{\text{TrNe}} \quad (9)$$

B. BINARY CLASSIFICATION PROBLEM

Classification results of deep learning and dual-attention Attention Convolutional AutoEncoder model over Binary class is presented in Table 4. For the MCI case, the VGG-16 model achieved a precision of 0.9648, a F1 score of 0.9741, and an accuracy score of 0.9737. For the AD case, the model achieved a precision of 0.9831. CNN achieved an accuracy of 0.9675 and a specificity of 0.9565 for the MCI class, as well as an accuracy of 0.9672 and a specificity of 0.9784 for the AD case. With regard to the MCI case, ResNet-50 achieved an accuracy of 0.9562, a specificity of 0.9378, and a recall of 0.9743. Inception-V3 achieved a score of 0.9710 for its specificity and 0.9802 for its accuracy. The MobileNet-V2 system achieved an accuracy of 0.9634 for the MCI case and a specificity of 0.9682 for the AD case. The Xception algorithm achieved an accuracy of 0.9711 and a specificity of 0.9795. For the case of MCI, the proposed framework achieved an accuracy of 0.9922, a precision of 0.9918, and a specificity of 0.9927, whereas for the case of AD, it achieved a specificity of 0.9918. Although ResNet-50 is an effective model, it has drawbacks such as lower accuracy and constrained representational capacity but also demands more processing power, which can result in overfitting and longer training times. Conversely, the proposed models are more accurate, deliver speed and reduce computing complexity.

TABLE 4. Classification results of deep learning and dual attention convolutional AutoEncoder model over binary class.

Method	Accuracy	Precision	Recall	F1 score	Specificity	Class
VGG-16	0.9737	0.9648	0.9836	0.9741	0.9637	MCI
	0.9737	0.9831	0.9637	0.9733	0.9836	AD
CNN	0.9675	0.9578	0.9784	0.9680	0.9565	MCI
	0.9672	0.9777	0.9565	0.9670	0.9784	AD
ResNet-50	0.9562	0.9406	0.9743	0.9572	0.9378	MCI
	0.9559	0.9731	0.9378	0.9551	0.9743	AD
Inception-V3	0.9809	0.9718	0.9907	0.9812	0.9710	MCI
	0.9809	0.9904	0.9710	0.9806	0.9907	AD
MobileNet-V2	0.9634	0.9593	0.9682	0.9683	0.9585	MCI
	0.9634	0.9676	0.9585	0.9630	0.9682	AD
Xception	0.9711	0.9636	0.9795	0.9715	0.9627	MCI
	0.9711	0.9789	0.9627	0.9707	0.9795	AD
Proposed	0.9922	0.9928	0.9918	0.9923	0.9927	MCI
	0.9922	0.9917	0.9927	0.9922	0.9918	AD

C. MULTICLASS CLASSIFICATION PROBLEM

Table 5 presents the performance of the deep learning and dual-attention Attention Convolutional AutoEncoder models over multi-class. For the MCI scenario, the accuracy, precision, and F1 score of the VGG-16 model were 0.9528, 0.9528, and 0.9298, respectively. The model indicated that the AD example had an accuracy of 0.9232. In determining the MCI class, CNN obtained a specificity of 0.9645 and an accuracy of 0.9560. It obtained an accuracy of 0.9328 and a specificity of 0.9382 for the AD case. ResNet-50 demonstrated an accuracy rate of 0.9469 and a recall rate of 0.9272 in the context of moderate cognitive impairment (MCI). Specificity and accuracy ratings of 0.9625 and 0.9609 were attained using the Inception-V3 model, respectively. In the cases of Alzheimer's disease (AD) and moderate cognitive impairment (MCI), the MobileNet-V2 system demonstrated an accuracy of 0.9060 and a specificity of 0.8888, respectively. The Xception algorithm demonstrated an accuracy of 0.9656 and a specificity of 0.9838. With an accuracy of 0.9830, precision of 0.9832, and specificity of 0.9917, the proposed framework demonstrated exceptional performance in diagnosing mild cognitive impairment (MCI). On the other hand, it obtained a specificity of 0.9792, an accuracy of 0.9760, and a recall score of 0.9756 for AD.

D. VISUAL REPRESENTATION OF ACCURACY AND LOSS

The training and validation loss values provide crucial insights into the variations in learning performance across epochs. This information is essential for recognizing learning issues. Furthermore, the training model weights will provide guidance on the precise epoch to utilize during inference. Figure 6(a) represents the accuracy using binary data at the start of epoch 1, training and validation accuracy above 90, and but after 19th epochs accuracy increases. At the end, training and validation accuracy were above 99. Figure 6(b) represents the accuracy using multi-class data. At the start of epoch 1, training accuracy was above 90 and validation accuracy was below 90, but after 20 epochs, training accuracy increased more than Validation. Figure 6(c) represents the loss using a binary dataset, and Figure 6(d) represents the loss using a multiclass dataset.

TABLE 5. Classification results of deep learning and Dual Attention Convolutional AutoEncoder model over multi-class.

Method	Accuracy	Precision	Recall	F1 score	Specificity	Class
	0.9528	0.9272	0.9325	0.9298	0.9631	MCI
VGG-16	0.9564	0.9326	0.9366	0.9346	0.9663	CN
	0.9460	0.9232	0.9139	0.9185	0.9621	AD
	0.9560	0.9280	0.9386	0.9333	0.9645	MCI
CNN	0.9353	0.9313	0.8782	0.9041	0.9657	CN
	0.9328	0.8789	0.9216	0.8993	0.9382	AD
	0.9469	0.9156	0.9272	0.9213	0.9569	MCI
ResNet-50	0.9573	0.9269	0.9463	0.9365	0.9628	CN
	0.9541	0.9461	0.9143	0.9299	0.9740	AD
	0.9609	0.9279	0.9578	0.9426	0.9625	MCI
Inception-V3	0.9632	0.9366	0.9540	0.9452	0.9678	CN
	0.9704	0.9708	0.9301	0.9544	0.9678	AD
	0.9060	0.8238	0.8942	0.8576	0.9114	MCI
MobileNet-V2	0.9223	0.9333	0.8306	0.8790	0.9695	CN
	0.8645	0.7942	0.8180	0.8059	0.8888	AD
	0.9656	0.9650	0.9298	0.9471	0.9838	MCI
Xception	0.9511	0.8972	0.9613	0.9282	0.9461	CN
	0.9543	0.9479	0.9161	0.9317	0.9740	AD
	0.9830	0.9832	0.9669	0.9744	0.9917	MCI
Proposed	0.9880	0.9818	0.9822	0.9820	0.9909	CN
	0.9780	0.9590	0.9756	0.9672	0.9792	AD

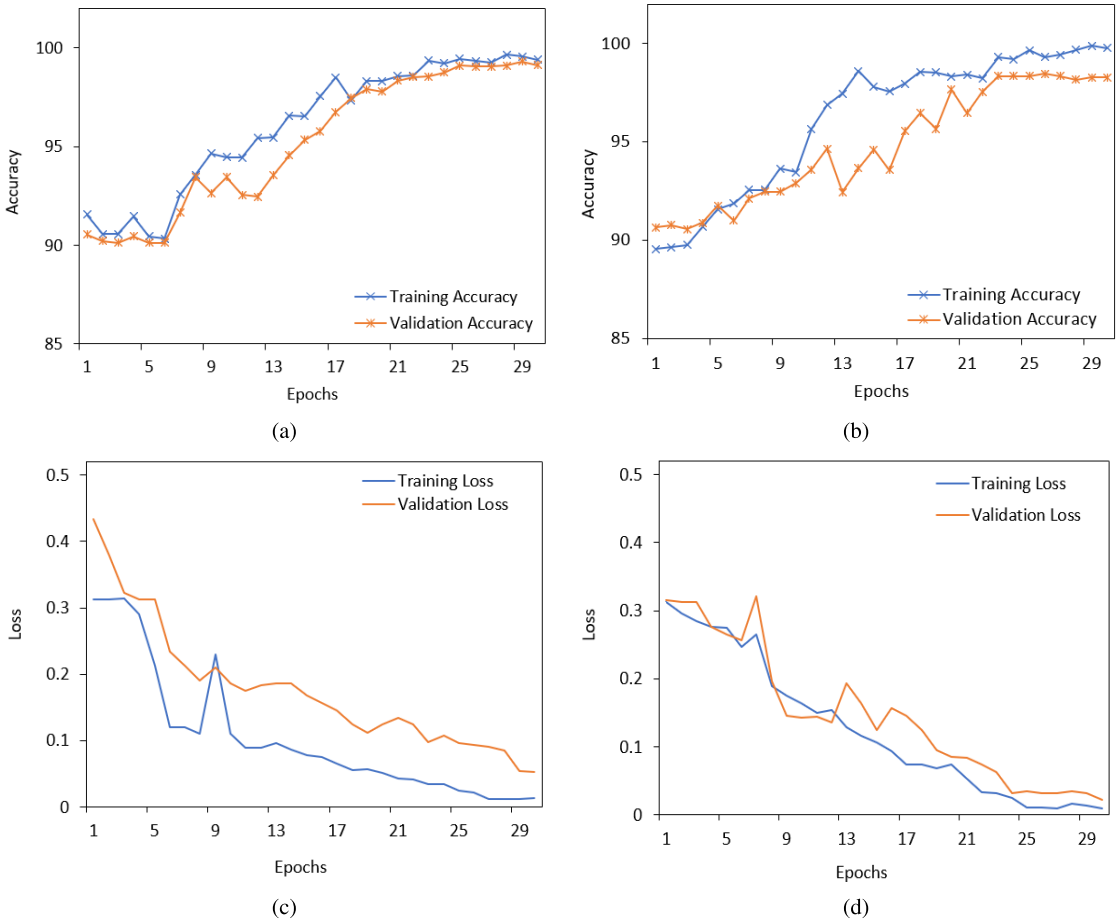


FIGURE 6. Learning performance of DACNA (a) Accuracy using binary dataset, (b) Accuracy using multiclass dataset, (c) Loss using binary dataset, and (d) Loss using multiclass dataset.

E. ROC CURVES

The correlation between TPR and FPR at various levels of categorization is marked on a receiver operating characteristic (ROC) curve. A reduction in the classification threshold yields a rise in the count of positively identified items, consequently contributing to an escalation in both true positives and false positives. ROC curves' specificity and sensitivity can be modified to differing degrees, contingent

upon the circumstances. The graphic presents the true positive rate (TPR) to false positive rate (FPR) ratio for every level of categorization score. The area under the curve, denoted as AUC, is equivalent to the integral of the ROC curve, where the TPR values are utilized, when the FPR values drop between 0 and 1. By utilizing the AUC, one can assess their performance across all conceivable dimensions. An inverse correlation can be observed between the performance of a classifier and the area under the curve (AUC), a metric with a range of 0 to 1. By employing the rocmetrics function, it is possible to divide a problem consisting of multiple classes into a sequence of binary classification problems, with each problem focusing on a single class. To determine the location of the ROC curve, the matching binary problem is implemented for each class in question. One class is regarded as positive in every binary circumstance, whereas the remaining classes are regarded as negative.

Figure 7(a) displays the AUC of the DACNA model. A remarkable value of 0.9991 is achieved with an AUC value, DACNA proves to be a very successful classification model that effectively minimizes false positives and accurately detects true positives. This demonstrates the model's predictive aptitude and its capacity to classify Alzheimer's disease with a balanced combination of sensitivity and specificity. This paper presents six deep learning-based methodologies for identifying and classifying Alzheimer's disease. The DACNA model utilizes transfer learning to extract features, and the average AUC score from the test set is employed as an evaluation measure during model training. We conducted experiments using Neuroimagaing data, which is often employed for evaluating DL-based approaches. However, selecting the appropriate features and parameters, together with having diverse and high-quality training data, might potentially modify the algorithms' effectiveness. Despite these challenges, the proposed deep learning models demonstrated excellent performance in identifying and classifying the disease. Figure 7(b) presents the results of an evaluation that compares six state-of-the-art models using ROC scores for multi-class. The AUC scores for multiclass classification were as follows: VGG-16 achieved a score of 0.9712, CNN achieved 0.9667, ResNet-50 achieved 0.9829, MobileNet-V2 achieved 0.9580, and the proposed model achieved 0.9949.

By evaluating the ROC curve of each method using the AUC metric, we can readily differentiate between them. An optimal discrimination outcome would yield an AUC value of 1.0. If the discrimination is 0.5, it is said to be no more effective than a random chance. The graph offers a quantitative assessment of the algorithms' precision in categorizing and differentiating between binary states. Figure 7(c) shows the AUC score for binary class and Figure 7(d) shows the AUC score for multiclass. For model achieved binary class, DACNA model achieved 0.9992 AUC on MCI case and 0.9919 on AD case.

F. CONFUSION MATRIX

The essential principles of the binary-class matrix and the multi-class confusion matrix are similar. The columns exhibit the initial or anticipated distribution of classes, whereas the rows indicate the predicted or output distribution of the classifier. Figure 8(a) shows the binary class confusion matrix of the proposed approach, and Figure 8(b) shows the multi-class classification matrix. Figure 8(a) shows that the proposed approach has a 99.18% correct prediction for AD and a 99.28% correct prediction for MCI. Only 0.82% and 0.72% of the incorrect predictions were made by the model. Figure 8(b) shows that the proposed approach has a 96.59% correct prediction for MCI, a 98.22% correct prediction for AD, and a 97.57% correct prediction for CN.

In Figure 9(a), the accuracy, precision, recall, and F1 score are presented in relation to Radar. In Figure 9(b), the mean absolute error of various models is contrasted with the proposed framework. VGG-16 achieved an MAE of over 0.06, CNN achieved 0.08, ResNet-50 achieved 0.04 MAE, MobileNet-V2 achieved 0.14 MAE, and the Proposed framework achieved 0.026 MAE. The proposed framework has the lowest MAE score, while MobileNet-V2 has the maximum MAE score.

G. STATISTICAL TEST

Understanding the findings of neural network experiments is enhanced by using statistical tests to assess the statistical importance of the experiments. Additionally, we utilize the information the tools give to clearly communicate the outcomes of our experiment and choose the best settings and algorithms for the challenging predictive modeling task. It is expected that the algorithm with the best average performance will perform better than the others. To ascertain whether there is a statistically significant difference in the average performance of any two algorithms, a statistical hypothesis test is necessary. T-Test ans P-Value of deep learning over the proposed model using accuracy are shown in Table 6 and using ROC in Table 7. Also the visual representation is shown in Figure 10. First we calculate the mean and standard deviation of two models that want to compare, then calculate T-test using formula.

$$\text{T-test} = \frac{(\bar{M}_1 - \bar{M}_2)}{\sqrt{\frac{Std_1^2}{n_1} + \frac{Std_2^2}{n_2}}} \tag{10}$$

In equation 10:

- $(\bar{M}_1 - \bar{M}_2)$ are the mean accuracy of the models being compared
- n_1 and n_2 are the sample size of population

H. TIME COMPLEXITY

The Xception model takes 82 seconds to execute each epoch; in contrast, the proposed model completes epochs more quickly and with optimal accuracy using fewer

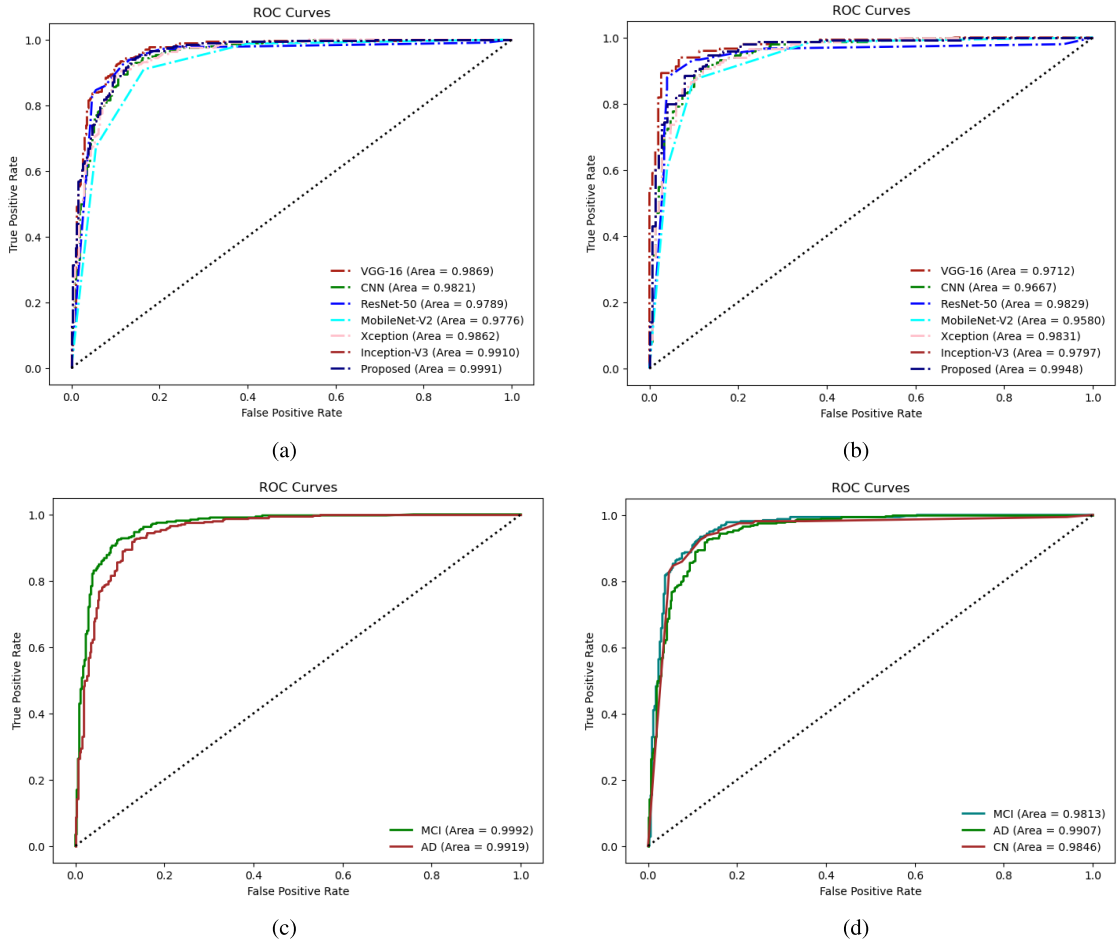


FIGURE 7. Receiver Operating Characteristic (ROC) curves for different deep learning models (a) AUC-ROC using binary dataset, (b) AUC-ROC using multiclass dataset, (c) AUC-ROC of the Proposed DACNA for MCI and AD cases, and (d) AUC-ROC of the Proposed DACNA for MCI, AD and CN cases.

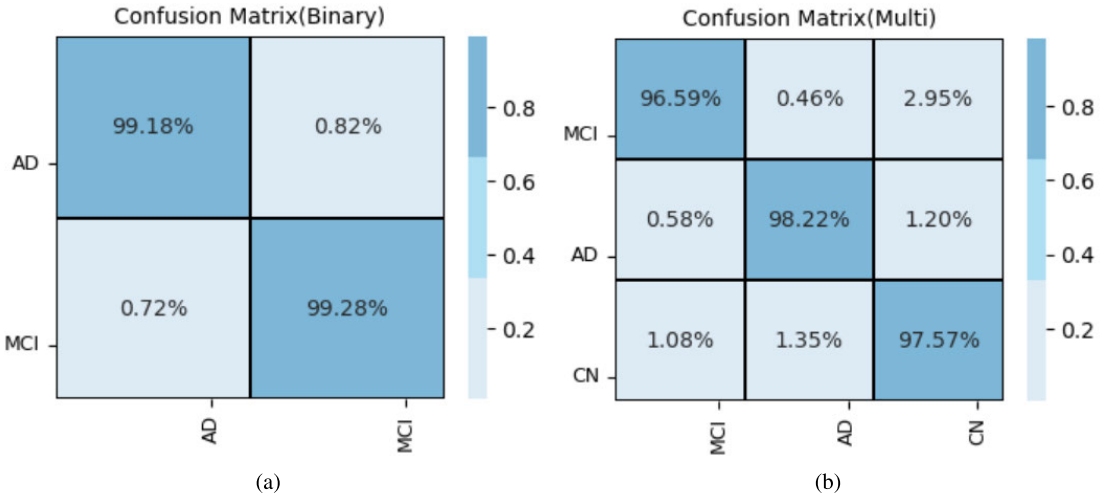
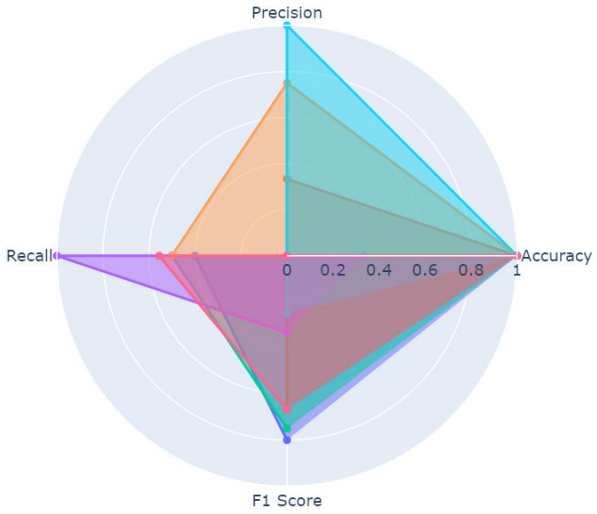
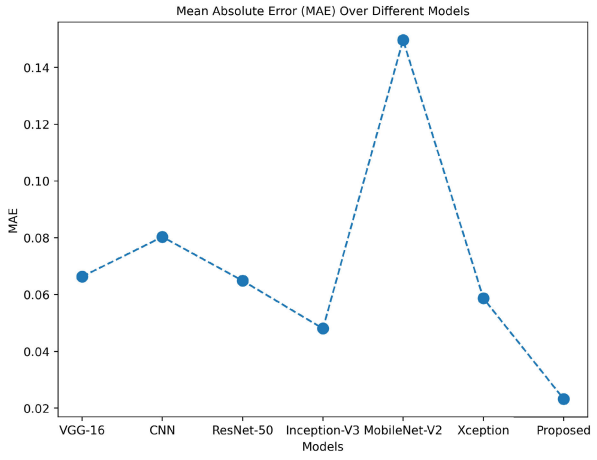


FIGURE 8. Confusion matrix of DACNA (a) using binary dataset, (b)using multiclass dataset.

computing resources as presented in Table 8. Because of its parameterization, which makes scaling simple, it can be used with limited resources. Its precision, cost-effectiveness, and parameter optimization make it a flexible approach to classifying neurological diseases that satisfies real-world needs without compromising accuracy.



(a)



(b)

FIGURE 9. Performance of deep learning (a) using four key metrics for the Proposed DACNA, (b) Mean absolute error over different models.

TABLE 6. Statistical T-Test comparison with the proposed model using Accuracy.

Method	T-Test	P-Value	Hypothesis
VGG-16 versus DACNA	-9.03425	2e-05	Rejected
CNN versus DACNA	-20.24706	0.00	Rejected
ResNet-50 versus DACNA	-22.86084	0.00	Rejected
Inception-V3 versus DACNA	-14.68524	0.00	Rejected
MobileNet-V2 versus DACNA	-14.09898	0.00	Rejected
Xception versus DACNA	-15.16423	0.00	Rejected

TABLE 7. Statistical T-test comparison with the proposed model using ROC.

Method	T-Test	P-Value	Hypothesis
VGG-16 versus DACNA	-470.99991	0.00	Rejected
CNN versus DACNA	-560.99999	0.00	Rejected
ResNet-50 versus DACNA	-166.87720	0.00	Rejected
Inception-V3 versus DACNA	-57.36898	0.00	Rejected
MobileNet-V2 versus DACNA	-164.04877	0.00	Rejected
Xception versus DACNA	-300.99999	0.00	Rejected

I. CROSS VALIDATION

One of the most well-known methods for learning from data is called K-fold cross-validation, and it involves dividing the dataset into K different subgroups. During the training process, the model is trained using K-1 folds in an iterative way, with each iteration focusing on evaluating the remaining fold. K repetitions of this procedure are carried out during each cycle, and performance measurements are generated as a result of these practices. Considering the average performance throughout all of the iterations is what is used to arrive at the end result. Through the utilization of this method, the impact of the unpredictability of the data is diminished, the model is evaluated on a number of subsets, and its capacity to generalize to new cases is enhanced. The accuracy of model evaluations is improved as a result of the combination of these parameters. 10 Fold Results using binary data is presented in Table 9 and using multiclass data in Table 10.

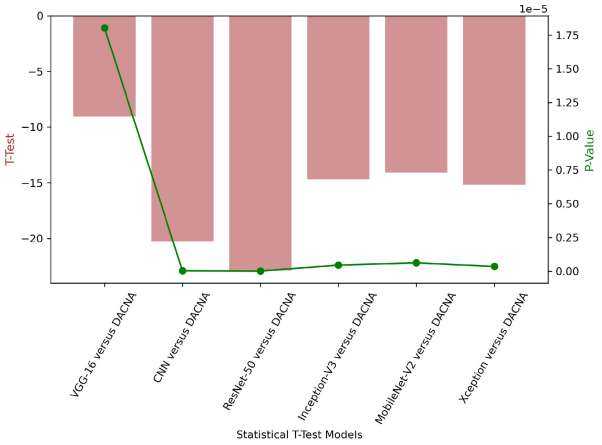


FIGURE 10. T-Test ans P-Value of deep learning over the proposed model. Vertical columns shows the T-Test values and green line curve shows the P-Value.

TABLE 8. Comparison of time complexity.

Model	Time Complexity
VGG-16	62 seconds per epochs
CNN	56 seconds per epochs
ResNet-50	71 seconds per epochs
Inception-V3	79 seconds per epochs
MobileNet-V2	68 seconds per epochs
Xception	82 seconds per epochs
Proposed DACNA	52 seconds per epochs

J. COMPARATIVE ANALYSIS WITH SOTA STUDIES

This section will present comparison between the proposed DACNA model and other existing models intended for the same goal, making use of the most advanced deep learning techniques currently available. In order to highlight the relative performance of each model, the comparison will make use of the same data. An examination of proposed model in comparison to other cutting-edge approaches to AD classification is shown in Table 11. Table 11 that have been presented provide evidence that the model that has been proposed is more reliable than the models that have been used in the previous studies. This is because it achieved the highest level of accuracy among the various methods. In spite of the fact that Shamrat et al. [23] achieved a high level of

TABLE 9. Proposed model evaluation results in 10 Fold cross-validation test option for two-class classification.

Method	Accuracy	Precision	Recall	F1 score	Specificity
VGG-16	0.9594 ± 0.1731	0.9422 ± 0.1823	0.9654 ± 0.1273	0.9512 ± 0.1452	0.9538 ± 0.1765
CNN	0.9521 ± 0.1637	0.9523 ± 0.1926	0.9412 ± 0.2138	0.9443 ± 0.2312	0.9591 ± 0.2135
ResNet-50	0.9410 ± 0.2163	0.9461 ± 0.1832	0.9502 ± 0.1865	0.9532 ± 0.1724	0.9612 ± 0.1423
Inception-V3	0.9712 ± 0.1028	0.9658 ± 0.1932	0.9529 ± 0.1259	0.9741 ± 0.1423	0.9783 ± 0.1646
MobileNet-V2	0.9445 ± 0.2385	0.9503 ± 0.3134	0.9413 ± 0.2413	0.9442 ± 0.1983	0.9578 ± 0.2154
Xception	0.9639 ± 0.1426	0.9617 ± 0.1652	0.9536 ± 0.1736	0.96.47 ± 0.1532	0.9671 ± 0.1192
Proposed DACNA	0.9902 ± 0.0139	0.9882 ± 0.0587	0.9898 ± 0.0865	0.9906 ± 0.0231	0.9912 ± 0.0872

TABLE 10. Proposed model evaluation results in 10 Fold cross-validation test option for three-class classification.

Method	Accuracy	Precision	Recall	F1 score	Specificity
VGG-16	0.9434 ± 0.3241	0.9456 ± 0.3187	0.9501 ± 0.29723	0.9424 ± 0.2514	0.9532 ± 0.2173
CNN	0.9551 ± 0.1983	0.9552 ± 0.1898	0.9512 ± 0.1752	0.9491 ± 0.1937	0.9589 ± 0.1635
ResNet-50	0.9381 ± 0.4132	0.9412 ± 0.4213	0.9381 ± 0.3742	0.9482 ± 0.3613	0.9456 ± 0.3034
Inception-V3	0.9545 ± 0.2134	0.9468 ± 0.1983	0.9545 ± 0.2046	0.9543 ± 0.2074	0.9618 ± 0.1982
MobileNet-V2	0.9132 ± 0.4823	0.9132 ± 0.4335	0.9201 ± 0.4752	0.9156 ± 0.4078	0.9383 ± 0.2741
Xception	0.9539 ± 0.2843	0.9539 ± 0.2128	0.9471 ± 0.1897	0.9540 ± 0.1653	0.9617 ± 0.1253
Proposed DACNA	0.9824 ± 0.1283	0.9823 ± 0.1176	0.9820 ± 0.1083	0.9814 ± 0.1472	0.9892 ± 0.1072

accuracy, our technique was ultimately superior. In addition, they employed a model that was fairly complicated and had a number of different classifiers, which is not suited for practical applications in comparison to our technique, which is made more straightforward. In conclusion, it is possible to verify that our DACNA model for the classification of Alzheimer's disease has more durability in comparison to other deep learning models that have been presented in the past. Comparison of different models based on overall performance through accuracy, precision, recall and F1 score to evaluate the efficacy of the framework is shown in Figure 11.

K. DISCUSSION

The study analyzes the deployment of deep learning and a dual-attention convolutional auto-encoder model in binary and multi-class classification tasks. In binary classification, Inception-V3 earned a specificity of 0.9710 and an accuracy of 0.9802, whereas MobileNet-V2 achieved an accuracy of 0.9634 and 0.9682, respectively. The Xception method attained an accuracy of 0.9711 and a specificity of 0.9795. The proposed model attained an accuracy of 0.9922, precision of 0.9918, and specificity of 0.9927 and 0.9918 in the cases of MCI and AD. An impressive AUC value of 0.9991 demonstrates the effectiveness of DACNA as a classification model in reducing false positives and reliably identifying real positives. The proposed approach accurately predicts Alzheimer's disease with a 99.18% success rate and mild cognitive impairment with a 99.28% efficiency rate.

The MobileNet-V2 model achieved an accuracy of 0.9060 and a specificity of 0.8888 for multi-class categorization in AD and MCI. The Xception method showed an accuracy of 0.9656 and a specificity of 0.9838. The proposed model showed outstanding performance in detecting MCI with an accuracy of 0.9830, precision of 0.9832, and specificity of 0.9917. The model achieved a specificity of 0.9792, an accuracy of 0.9760, and a recall score of 0.9756 for AD. The AUC values for multiclass classification

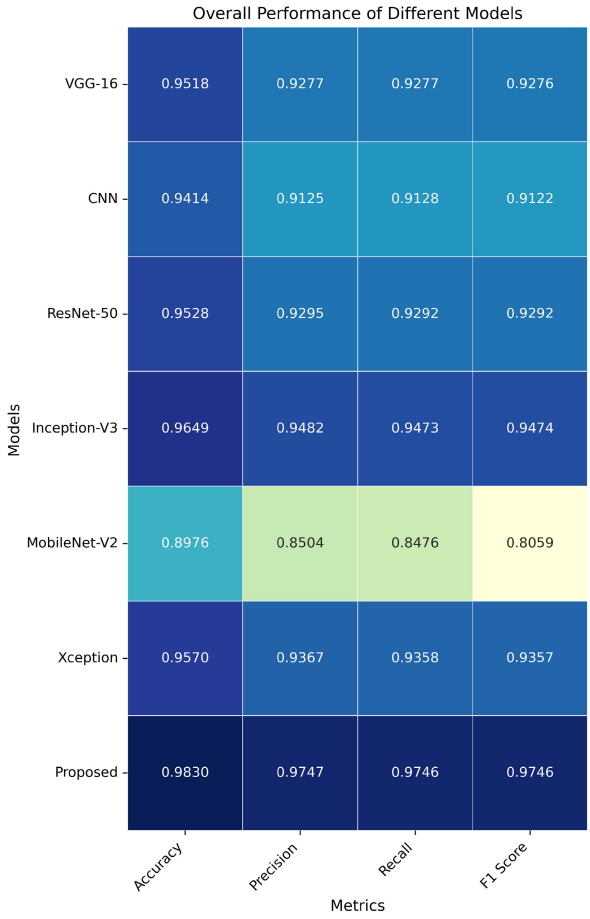


FIGURE 11. Comparison of different models based on overall performance through accuracy, precision, recall and F1 score to evaluate the efficacy of the framework.

were: VGG-16: 0.9712; CNN: 0.9667; ResNet-50: 0.9829; MobileNet-V2: 0.9580; and the suggested model: 0.9949.

The proposed model also evaluated using statistical T-test using accuracy and ROC metrics. Our proposed framework shows a statistically significant difference in

TABLE 11. Comparative Analysis with SOTA studies.

Technique	Objective	Performance
A pre-trained vision trans- former was employed to classify the fused images [14].	The objective is to prevent Alzheimer's disease and early mild cognitive impairment.	Their vision transformer achieved 81.25% accuracy.
Xception architecture was utilized for the classification of Alzheimer's disease [15].	The primary aim of the study was to provide a fast and efficient diagnostic approach for identifying people who are in a state of good health before the occurrence of moderate cognitive impairment.	Authors achieved impressive F1 score, sensitivity, specificity, and accuracy, with an average of 95.41%, 95.21%, 95.01%, and 97.92%, respectively.
The DenseNet201 model applied to Gaussian NB to classify the AD cases [17].	The findings would improve the abilities of radiologists to seek a second opinion or utilize a simulator.	The authors attained a precision rate of 91.75%, a specificity rate of 96.5%, and an F1 score of 90.25%.
The BILSTM deep technique was implemented to classify maladies into multiple classes [19].	Their proof-of-concept study demonstrates a community-based screening tool that is simple to use, accessible to everybody, and highly sensitive. This test may be used to determine the risk and identify problems early.	The authors achieved an AUC of 0.97.
The top six AI techniques for promptly identifying AD instances [22].	Enrichment analysis was employed to clarify the underlying processes of AD.	The AUC value falls within the range of 0.949. Deep learning demonstrated a sensitivity and specificity of 94.5% and an AUC (95% CI).
InceptionV3 technique [23].	The AlzheimerNet is created through modifications to the InceptionV3 model and using an RMSprop optimizer with a learning rate of 0.00001.	They achieved 96.31% accuracy.
ResNet-50 technique [25].	The authors offer a novel strategy that integrates many data fusion techniques.	They achieved 97% accuracy.
Proposed approach.	We present a new dual-attention convolutional auto-encoder model with transfer learning will make diagnosis easier. We enhance the training data using the ADASYN sampling technique and perform T-tests to compare with other methods.	Our model achieved 0.9902 ± 0.0139 and 0.9824 ± 0.1283 accuracy for two and three class classifications, respectively. Also, the proposed approach achieved highest AUC score and lowest validation loss.

average performance compared to previous deep models. The proposed framework executes epochs in 52 seconds, achieving faster completion with optimal accuracy and utilizing less computational resources. Its parameterization simplifies scaling, making it cost-effective and versatile for identifying Alzheimer's disease, meeting real-world requirements while maintaining accuracy. The proposed framework outperformed other models in terms of Mean Absolute Error (MAE) scores, with VGG-16 having the highest MAE, followed by CNN at 0.08, ResNet-50 at 0.04, MobileNet-V2 at 0.14, and the proposed framework at 0.02, indicating its superiority.

L. LIMITATIONS OF THE PROPOSED STUDY

The limitations of the study are significant as they provide possibilities for further research. The following are some limitations of the study:

- Basic image scaling can overlook essential features for Alzheimer's disease classification, leading to a potential overlook of delicate qualities for accurate diagnosis. The process may cause distorted or muddled patterns, reducing the accuracy of classification as it may not accurately represent the progression of the disease.
- We used ADASYN to handle the imbalance problem and enhance the dataset size, but the small dataset used in the study may not accurately reflect the broader Alzheimer's disease population, which might result in biased model results. The small sample size might hinder the ability to provide accurate and thorough diagnoses.
- The research only used magnetic resonance imaging (MRI) for research motives. Further research might use other imaging techniques, such as PET scans, to further understand the disease biology and improve the effectiveness of deep learning models for diagnosing Alzheimer's disease. Utilizing multi-modal approaches might enhance diagnostic accuracy and provide a more thorough comprehension of the issue.

To evaluate the efficacy of our method in healthcare, we intend to develop customized, large datasets for monitoring neurological patients. Participation with healthcare practitioners may enhance the pertinence and practicality of this research. Further research on enhanced deep learning techniques and feature engineering might improve accuracy and efficiency. Our next initiatives focus on improving the implementation methods for several brain diseases and their effects on health. More research is required to enhance precision and effectiveness, despite the positive outcomes.

V. CONCLUSION AND FUTURE DIRECTIONS

This study presents an automated method that may be useful in diagnosing Alzheimer's disease. The method attempts to increase patient satisfaction while also allowing for faster detection of the disease. The proposed model is a dual-attention convolutional autoencoder designed to aid in the diagnosis of Alzheimer's disease through brain imaging. It has exceptional detection capabilities and eliminates the need for feature extraction. Adaptive Synthetic Sampling (ADASYN) is used to provide additional training data for the minority class, making the detection process more efficient. The model's efficacy is validated through statistical tests and comparisons with alternative methods, enhancing classification accuracy.

The study's findings revealed that the proposed model outperformed the current alternatives in terms of performance. The experiments made it clear that the proposed DACNA was robust enough to differentiate between AD and MCI. It got 0.9824 ± 0.1283 accuracy in the three-class classification task and 0.9902 ± 0.0139 in the binary classification task. The purpose of this study is to evaluate the accuracy of deep learning models for diagnosing and classifying AD using MRI images. However, it is critical to recognize our study's limitations. For starters, the dataset used in the study is quite small, limiting its ability to accurately portray the whole population.

Furthermore, future research may use alternative imaging methods, such as PET scans, alongside deep learning models, whereas our study was limited to MRI images. To conclude, the unique deep learning structure and hyperparameters used in this study may impose limits on the findings. Future research will strive to integrate data from more sources into the proposed DACNA system to increase classification accuracy. We will use a variety of samples to evaluate the overall efficacy of the proposed DACNA system. Furthermore, a broader range of scenarios will be examined in order to assess the viability of deploying various deep learning neural networks.

ACKNOWLEDGMENT

This research is supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R136), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors are also thankful to AIDA Lab CCIS Prince Sultan University, Riyadh Saudi Arabia for support.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] D. Caine, "Posterior cortical atrophy: A review of the literature," *Neurocase*, vol. 10, no. 5, pp. 382–385, Oct. 2004.
[2] N. Goenka and S. Tiwari, "AlzVNet: A volumetric convolutional neural network for multiclass classification of Alzheimer's disease through multiple neuroimaging computational approaches," *Biomed. Signal Process. Control*, vol. 74, Apr. 2022, Art. no. 103500.
[3] S. Sheladia and P. H. Reddy, "Age-related chronic diseases and Alzheimer's disease in Texas: A hispanic focused study," *J. Alzheimer's Disease Rep.*, vol. 5, no. 1, pp. 121–133, Feb. 2021.
[4] D. M. Cammisuli, G. Cipriani, and G. Castelnuovo, "Technological solutions for diagnosis, management and treatment of Alzheimer's disease-related symptoms: A structured review of the recent scientific literature," *Int. J. Environ. Res. Public Health*, vol. 19, no. 5, p. 3122, Mar. 2022.
[5] P. M. Rossini, R. Di Iorio, F. Vecchio, M. Anfossi, C. Babiloni, M. Bozzali, A. C. Bruni, S. F. Cappa, J. Escudero, F. J. Fraga, P. Giannakopoulos, B. Guntekin, G. Logroscino, C. Marra, F. Miraglia, F. Panza, F. Tecchio, A. Pascual-Leone, and B. Dubois, "Early diagnosis of Alzheimer's disease: The role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts," *Clin. Neurophysiol.*, vol. 131, no. 6, pp. 1287–1310, Jun. 2020.
[6] K. A. Matthews, W. Xu, A. H. Gaglioti, J. B. Holt, J. B. Croft, D. Mack, and L. C. McGuire, "Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged ≥65 years," *Alzheimer's Dementia*, vol. 15, no. 1, pp. 17–24, Jan. 2019, doi: 10.1016/j.jalz.2018.06.3063.
[7] M. D. Hurd, P. Martorell, A. Delavande, K. J. Mullen, and K. M. Langa, "Monetary costs of dementia in the United States," *New England J. Med.*, vol. 368, no. 14, pp. 1326–1334, 2013.
[8] B. Tejada-Vera, *Mortality From Alzheimer's Disease in the United States: Data for 2000 and 2010*, no. 116. Hyattsville, MD, USA: National Center for Health Statistics, 2013.
[9] I. Campbell-Taylor, "Contribution of Alzheimer disease to mortality in the United States," *Neurology*, vol. 83, no. 14, p. 1302, Sep. 2014.
[10] J. Hsu, Y. Wei, C. H. Toh, I. Hsiao, K. Lin, T. Yen, M. Liao, and L. Ro, "Magnetic resonance images implicate that glymphatic alterations mediate cognitive dysfunction in Alzheimer disease," *Ann. Neurol.*, vol. 93, no. 1, pp. 164–174, Jan. 2023.
[11] C.-L. Wu, T.-J. Lin, G.-L. Chiou, C.-Y. Lee, H. Luan, M.-J. Tsai, P. Potvin, and C.-C. Tsai, "A systematic review of MRI neuroimaging for education research," *Frontiers Psychol.*, vol. 12, May 2021, Art. no. 617599.
[12] E. Hussain, M. Hasan, S. Zafrul Hassan, T. Hassan Azmi, M. Anisur Rahman, and M. Zavid Parvez, "Deep learning based binary classification for Alzheimer's disease detection using brain MRI images," in *Proc. 15th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Nov. 2020, pp. 1115–1120.
[13] A. B. Tufail, Y.-K. Ma, and Q.-N. Zhang, "Binary classification of Alzheimer's disease using sMRI imaging modality and deep learning," *J. Digit. Imag.*, vol. 33, no. 5, pp. 1073–1090, Oct. 2020.
[14] M. Odusami, R. Maskeliūnas, and R. Damaševičius, "Pixel-level fusion approach with vision transformer for early detection of Alzheimer's disease," *Electronics*, vol. 12, no. 5, p. 1218, Mar. 2023.
[15] S. B. Çelebi and B. G. Emiroğlu, "A novel deep dense block-based model for detecting Alzheimer's disease," *Appl. Sci.*, vol. 13, no. 15, p. 8686, Jul. 2023.
[16] P. Balaji, M. A. Chaurasia, S. M. Bilfaqih, A. Muniasamy, and L. E. G. Alsid, "Hybridized deep learning approach for detecting Alzheimer's disease," *Biomedicines*, vol. 11, no. 1, p. 149, Jan. 2023.
[17] S. Sharma, S. Gupta, D. Gupta, A. Altameem, A. K. J. Saudagar, R. C. Poonia, and S. R. Nayak, "HTLML: Hybrid AI based model for detection of Alzheimer's disease," *Diagnostics*, vol. 12, no. 8, p. 1833, Jul. 2022.
[18] A. Khalid, E. M. Senan, K. Al-Wagih, M. M. A. Al-Azzam, and Z. M. Alkhraisha, "Automatic analysis of MRI images for early prediction of Alzheimer's disease stages based on hybrid features of CNN and handcrafted features," *Diagnostics*, vol. 13, no. 9, p. 1654, May 2023.
[19] F. Agbavor and H. Liang, "Artificial intelligence-enabled end-to-end detection and assessment of Alzheimer's disease using voice," *Brain Sci.*, vol. 13, no. 1, p. 28, Dec. 2022.
[20] W. Salehi, P. Baglat, G. Gupta, S. B. Khan, A. Almusharraf, A. Alqahtani, and A. Kumar, "An approach to binary classification of Alzheimer's disease using LSTM," *Bioengineering*, vol. 10, no. 8, p. 950, Aug. 2023.
[21] T. Illakiya, K. Ramamurthy, M. V. Siddharth, R. Mishra, and A. Udainiya, "AHANet: Adaptive hybrid attention network for Alzheimer's disease classification using brain magnetic resonance imaging," *Bioengineering*, vol. 10, no. 6, p. 714, Jun. 2023.
[22] R. O. Bahado-Singh, U. Radhakrishna, J. Gordevič ius, B. Aydas, A. Yilmaz, F. Jafar, K. Imam, M. Maddens, K. Challapalli, R. P. Metpally, W. H. Berrettini, R. C. Crist, S. F. Graham, and S. Vishweswaraiah, "Artificial intelligence and circulating cell-free DNA methylation profiling: Mechanism and detection of Alzheimer's disease," *Cells*, vol. 11, no. 11, p. 1744, May 2022.
[23] F. M. J. M. Shamrat, S. Akter, S. Azam, A. Karim, P. Ghosh, Z. Tasnim, K. Hasib, F. De Boer, and K. Ahmed, "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.
[24] M. Mamun, S. Bin Shawkat, M. S. Ahammed, M. M. Uddin, M. I. Mahmud, and A. M. Islam, "Deep learning based model for Alzheimer's disease detection using brain MRI images," in *Proc. IEEE 13th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Oct. 2022, pp. 0510–0516.
[25] S. Dwivedi, T. Goel, M. Tanveer, R. Murugan, and R. Sharma, "Multimodal fusion-based deep learning network for effective diagnosis of Alzheimer's disease," *IEEE MultiMedia*, vol. 29, no. 2, pp. 45–55, Mar. 2022.
[26] D. Ganesh, M. S. Kumar, M. C. Aparna, M. C. J. Royal, M. D. Vinay, and S. H. Sari, "Implementation of convolutional neural networks for detection of Alzheimer's disease," *BioGecko, J. New Zealand Herpetol.*, vol. 12, no. 1, pp. 71–82, 2023.
[27] I. Abunadi, "Deep and hybrid learning of MRI diagnosis for early detection of the progression stages in Alzheimer's disease," *Connection Sci.*, vol. 34, no. 1, pp. 2395–2430, Dec. 2022.

[28] T. J. Saleem, S. R. Zahra, F. Wu, A. Alwakeel, M. Alwakeel, F. Jeribi, and M. Hijji, "Deep learning-based diagnosis of Alzheimer's disease," *J. Personalized Med.*, vol. 12, no. 5, p. 815, 2022.
[29] V. Patil, M. Madgi, and A. Kiran, "Early prediction of Alzheimer's disease using convolutional neural network: A review," *Egyptian J. Neurol., Psychiatry Neurosurg.*, vol. 58, no. 1, pp. 1–10, Nov. 2022.
[30] G. Battineni, M. A. Hossain, N. Chintalapudi, E. Traini, V. R. Dhulipalla, M. Ramasamy, and F. Amenta, "Improved Alzheimer's disease detection by MRI using multimodal machine learning algorithms," *Diagnostics*, vol. 11, no. 11, p. 2103, Nov. 2021.
[31] A. A. A. El-Latif, S. A. Chelloug, M. Alabdulhafith, and M. Hammad, "Accurate detection of Alzheimer's disease using lightweight deep learning model on MRI data," *Diagnostics*, vol. 13, no. 7, p. 1216, Mar. 2023, doi: 10.3390/diagnostics13071216.
[32] M. Sajjad, F. Ramzan, M. U. G. Khan, A. Rehman, M. Kolivand, S. M. Fati, and S. A. Bahaj, "Deep convolutional generative adversarial network for Alzheimer's disease classification using positron emission tomography (PET) and synthetic data augmentation," *Microsc. Res. Technique*, vol. 84, no. 12, pp. 3023–3034, Dec. 2021, doi: 10.1002/jemt.23861.
[33] H. Wu and X. Gu, "Towards dropout training for convolutional neural networks," *Neural Netw.*, vol. 71, pp. 1–10, Nov. 2015.
[34] Q. Zhu, P. Zhang, Z. Wang, and X. Ye, "A new loss function for CNN classifier based on predefined evenly-distributed class centroids," *IEEE Access*, vol. 8, pp. 10888–10895, 2020.

SHAHA AL-OTAIBI received the M.S. degree in computer science and the Ph.D. degree in artificial intelligence from KSU. She is currently an Associate Professor with the Department of Information Systems, College of Computer and Information Sciences, PNU, Saudi Arabia. She has a Senior Fellow Recognition from U.K. Higher Education Academy (SFHEA). Currently, she is a reviewer in some journals and editorial board member in other journals. Her main research interests include data science, artificial intelligence, machine learning, bio-inspired computing, cybersecurity, and information security.

MUHAMMAD MUJAHID received the master's degree in computer science from the Islamia University of Bahawalpur, Pakistan. He is currently pursuing the Ph.D. degree. He is a Reviewer in the *Journal of Super Computing*, *PLOS One*, IEEE journals, and Springer. His main research interests include text mining, machine learning, data science, artificial intelligence, bio-informatics, the Internet of Things, and sentiment analysis.

AMJAD REHMAN KHAN (Senior Member, IEEE) received the Ph.D. and Postdoctoral degrees (Hons.) from the Faculty of Computing, University Teknologi Malaysia, with a specialization in forensic documents analysis and security, in 2010 and 2011, respectively. He is currently a Senior Researcher with the Artificial Intelligence and Data Analytics Laboratory, College of Computer and Information Sciences (CCIS), Prince Sultan University, Riyadh, Saudi Arabia. He is the author of more than 200 ISI journal articles and conferences. He is also a PI in several funded projects and also completed projects funded from MOHE Malaysia and Saudi Arabia. His research interests include data mining, health informatics, and pattern recognition. He received the Rector Award for the 2010 Best Student from Universiti Teknologi Malaysia.

HAITHAM NOBANEE received the Ph.D. degree from the University of Manchester. He is currently a Professor in finance and financial technology (Fintech) with Abu Dhabi University, United Arab Emirates. An Honorary Professor with the University of Liverpool, U.K., and a Visiting Research Professor with the University of Oxford, U.K. He is a fellow of the Higher Education Academy (HEA) and Learning and Performance Institute (LPI). His work has been published in the *International Review of Economics & Finance*, *The North American Journal of Economics and Finance*, *Renewable and Sustainable Energy Reviews*, *Big Data*, and *Corporate Governance*; and he has also been published in professional journals, such as *Harvard Business Review*. He is also serving the Academic Community as an Editor (*Arab Gulf Journal of Scientific Research-Emerald*), an Associate Editor [*International Review of Financial Analysis* (Elsevier) and *Heliyon* (Elsevier)], a Guest Editor [*International Review of Economic & Finance* (Elsevier)], a member of the Editorial Board [*Financial Innovation* (Springer)], and a reviewer for various international refereed journal.

JABER ALYAMI received the Ph.D. degree from the University of Nottingham, U.K. He is currently a Registered Consultant in radiology and an Assistant Professor with the Radiological Sciences Department, King Abdulaziz University, Jeddah, Saudi Arabia.

TANZILA SABA (Senior Member, IEEE) received the Ph.D. degree in document information security and management from the Faculty of Computing, Universiti Teknologi Malaysia (UTM), Malaysia, in 2012. She is currently an Associate Chair with the Information Systems Department, College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia, where she is also the Artificial Intelligence and Data Analytics Research Laboratory Leader. Her primary research interests include medical imaging, pattern recognition, data mining, MRI analysis, and soft computing. She is an Active Professional Member of ACM, AIS, and IAENG organizations. She is the PSU Women in Data Science (WiDS) Ambassador with Stanford University and the Global Women Tech Conference. She was a recipient of the Best Student Award from the Faculty of Computing, UTM, in 2012.