Received: 18 April 2023 | Revised: 1 July 2023 | Accepted: 6 August 2023

IET Image Processing

DOI: 10.1049/ipr2.12910

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

A robust and clinically applicable deep learning model for early detection of Alzheimer's

Md Masud Rana[1] | **Md Manowarul Islam**[1] | **Md. Alamin Talukder**[1] | **Md Ashraf Uddin**[2] | **Sunil Aryal**[2] | **Naif Alotaibi**[3] | **Salem A. Alyami**[3] | **Khondokar Fida Hasan**[4] | **Mohammad Ali Moni**[5]

[1]Department of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh

[2]School of Information Technology, Deakin University, Geelong Waurn Ponds Campus, Australia

[3]Department of Mathematics and Statistics, Faculty of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

[4]School of Professional Studies, University of New South Wales, Canberra, ACT, Australia

[5]AI and Cyber Futures Institute, Charles Stuart University, Bathurst, NSW, Australia

Correspondence
Mohammad Ali Moni, AI and Cyber Futures Institute, Charles Stuart University, Bathurst, NSW 2795, Australia.
Email: mmoni@csu.edu.au

Funding information
ICT division, Ministry of ICT, Bangladesh, Grant/Award Number: Ref. No. 56.00.0000.053.20.005.22-105

Abstract

Alzheimer's disease, often known as dementia, is a severe neurodegenerative disorder that causes irreversible memory loss by destroying brain cells. People die because there is no specific treatment for this disease. Alzheimer's is most common among seniors 65 years and older. However, the progress of this disease can be reduced if it can be diagnosed earlier. Recently, artificial intelligence has instilled hope in the diagnosis of Alzheimer's disease by performing sophisticated analyses on extensive patient datasets, enabling the identification of subtle patterns that may elude human experts. Researchers have investigated various deep learning and machine learning models to diagnose this disease at an early stage using image datasets. In this paper, a new Deep learning (DL) methodology is proposed, where MRI images are fed into the model after applying various pre-processing techniques. The proposed Alzheimer's disease detection approach adopts transfer learning for multi-class classification using brain MRIs. The MRI Images are classified into four categories: mild dementia (MD), moderate dementia (MOD), very mild dementia (VMD), and non-dementia (ND). The model is implemented and extensive performance analysis is performed. The finding shows that the model obtains 97.31% accuracy. The model outperforms the state-of-the-art models in terms of accuracy, precision, recall, and F-score.

1 | INTRODUCTION

Alzheimer's disease (AD) is a prevalent form of dementia, characterized by the accumulation of amyloid-beta peptide (A β) in the medial temporal lobe and neocortical structures [1]. This leads to the development of neuritic plaques and neurofibrillary tangles [2]. AD encompasses a range of neurological conditions that impact memory, cognition, behavior, and emotions. Early symptoms include memory loss, difficulty with daily tasks, language challenges, and personality changes [3]. Currently, there is a lack of accurate diagnostic methods and approved disease-modifying treatments for AD [4].

Alzheimer's Disease International (ADI) estimates that more than 55 million individuals worldwide are currently afflicted with Alzheimer's. It is projected that the overall global cost of Alzheimer's disease in 2015 was USD 818 billion, or 1.09% of the world's gross domestic product [5]. The annual worldwide cost is currently approximately USD 1.3 trillion and is expected to reach USD 2.8 trillion by 2030. By 2050, 139 (71% of the total) million people will be affected, with the greatest increase in low- and middle-income countries. Every 3 s, a new dementia case is reported somewhere in the world, where up to 75% of people with dementia are undiagnosed. The most alarming statistic is that one in four people feel there is nothing we can do to prevent dementia, and nearly 62% of medical experts worldwide wrongly think dementia is a normal part of becoming older.

WHO [6] estimates that 50 million people worldwide have dementia. A person above 85 eventually has a 50% chance of AD. In the end, AD kills the brain area that controls breathing and heart monitoring, resulting in death. The three stages of AD are extremely mild, mild, and moderate. However, an individual with AD only begins to show symptoms at a moderate stage, affecting neuronal communication.

Alzheimer's disease progresses at varying speeds. After a diagnosis of Alzheimer's, the typical lifetime is 3 to 11 years, but some patients live 20 years or longer. People with Alzheimer's can lead fulfilling lives for many years following an early diagnosis if the disease can be detected early. Early dementia diagnosis enables the patient to comprehend his situation, and the family better comprehend the patient's situation, which facilitates the establishment of acceptable expectations and mutual planning for the upcoming days.

Deep learning (DL) models with multi-level structures are very helpful in extracting reserved information from images [7, 8]. Convolutional Neural Networks (CNNs) can efficiently reduce computation time by using the advantage of the graphics processing unit (GPU) for computation. Several fields, that is, Medical, Agriculture, and Communication, are being developed and managed by utilizing DL techniques [9-12]. Researchers employ imaging techniques like MRI scans using deep learning to classify Alzheimer's disease and aid in the hunt for better treatments. Using magnetic resonance imaging (MRI) to classify AD, multiple articles have recently been published. Numerous works have been done on disease detection and other research areas using machine learning and deep learning algorithms [13-15]. Nagarathna et al. [16] has proposed a deep learning model for early detection of Alzheimer's disease and archives 95.52% accuracy. Sappagh, Shaker [17] has proposed a Hybrid model, and the early detection accuracy is 92.62% . Our work aims to develop an automated deep-learning model for early detection of Alzheimer's disease.

A DL model combining Inception V3 and Custom CNN Model with necessary layers has been proposed to classify Alzheimer's Disease (AD) into four classes: Very Mild Demented (V.M.D), Mild Demented (M.D), Moderate Demented (Mod.D), and Non-Demented (N.D). In addition, the proposed Hybrid model has been trained and tested using a larger image dataset. Before training the model with the Input dataset, The dataset was pre-processed using different image-processing techniques. Even though the dataset is relatively larger, deep learning algorithms require a diverse large amount of balanced datasets to automate the training and testing processes and avoid overfitting. We have applied oversampling techniques to increase the data in the original dataset. The pre-processed and oversampled MRI images have increased the accuracy of the proposed model.

The paper presents the following notable contributions:

- Hybrid Deep Learning Model for AD Diagnosis: The proposed model merges transfer learning and Convolutional Neural Networks (CNN) to create a hybrid approach for Alzheimer's disease (AD) diagnosis. Inception V3 is utilized through transfer learning for feature extraction, while a custom CNN is developed for accurate classification.
- Effective Data Normalization: The paper introduces a robust data normalization strategy to enhance the model's performance. By applying a Bilateral filter, image quality is improved, and relevant features are extracted. To address the class imbalance, the model incorporates SMOTE-ENN (Synthetic Minority Oversampling Technique followed by Edited Nearest Neighbours) to mitigate overfitting.
- Comprehensive Experimental Evaluation: In-depth experiments evaluate the proposed hybrid model using various performance metrics, including accuracy, precision, recall, and others. The results demonstrate that the hybrid model surpasses state-of-the-art approaches, affirming its effectiveness in AD diagnosis.

These contributions highlight the development of a hybrid deep learning model for AD diagnosis, implementing effective data normalization techniques, and comprehensively evaluating the model's performance.

Existing research works have been described in Section 2 highlighting the Machine Learning model, CNN model, as well as some Hybrid Models. In Section 3, the methodology of the model is described. This section discusses the structure of the model and the pre-processing techniques of the datasets. In Section 4, the result of the proposed Hybrid method is analyzed and compared with other existing models. Different types of efficiency measurements are shown graphically. Section 5 concludes this work by indicating limitations and a plan for future work.

2 | LITERATURE REVIEW

The primary difficulty facing AD researchers now is making a premortem diagnosis that is certain. Aside from the visible alterations in the morphology of the cerebral cortex [18], and the symptoms of mood swings, which are unrelated to ageing [19], there are still many unknowns concerning this condition. Although the disease's processes are still unknown, most medical professionals agree that a combination of hereditary and environmental variables appears to be the cause of AD. For example, several investigations have linked it to the type of herpes simplex virus [20] or infections that cause periodontists [21]. The disease is believed to exhibit sexual dimorphism and is age-related but not entirely an age-dependent condition [22]. Due to its effects on society and the lack of consensus on its cause, the disease remains the focus of in-depth research today. Some of the research on Alzheimer's Disease was implemented on the binary classification, some are 4 class classifications of AD, and some use a smaller dataset. Many of these authors have trained their models using a larger dataset. They had performed classification in different manners.

2.1 | Deep learning/ machine learning approach

Ortiz-Garcia et al. [23] used multimodal data that integrate MRI and PET images in Alzheimer's detection using a deep belief network (DBN). They selected 68 NC, only 70 AD, 111 MCI, and 26 late MCI subjects from the dataset. The dataset contains both PET Image and MRI images. Their preprocessing includes registering and resizing images in 1.5 mm for MRI and 3 mm for PET images in each coronal, sagittal, and axial view. They also performed MRI segmentation to segment the images into white matter (WM) and grey matter (GM). For PET images, they normalized the intensity of the images implementing the mean cerebellum activation level as a normalization value. Then they selected the significant voxels from each modality using Welch's t-test and split the brain into 116 brain regions and discarded cerebellum regions. The classification was performed using an ensemble of deep belief networks with four voting systems: majority voting, weighted voting, SVM-based data fusion, and deep belief network-based data fusion. According to their paper, it is about 90% accuracy rate for DBN and SVM-based voting for the classification of NC, and AD subjects was found.

In a study by Xu et al. [24], AD is predicted by developing an SVM-based technique using gene-coding protein sequence information. They briefly described sequence information as the frequency of two consecutive amino acids. The SVM input is the protein related to AD, and the output is the peptide with labels. Bi et al. [25] proposed random SVM using fMRI dataset to classify subjects by binary grouping into AD and NC groups. The proposed technique by the authors randomly selects data samples and features to create multiple SVMs. They used the accuracy of the SVM as the criterion to evaluate the quality of the features and the accuracy of the random SVM as the criterion to select from previously evaluated qualified features, and their proposed model produces a maximum of 94% of overall accuracy.

Bloch and Friedrich [26] combined two datasets (ADNI and AIBL) to evaluate the diagnosis of AD using baseline and follow-up MRI scans. They used RF with 25-fold bootstrapping for classification and used the synthetic minority oversampling technique (SMOTE) to handle class imbalances in their proposed model. Their results show that the classification accuracy resulting from the combined dataset using RF provides higher accuracy compared to the sole dataset using the same RF classifier. Spasov et al. has developed [27] the flexible model which can learn different types of 3D datasets. According to their paper, the accuracy of 86% may give the flexibility to design a computer-aided diagnosis system to predict several medical neuropsychiatric disorders via imaging and tabular clinical data. Feng et al. [28] have implemented three different models 2D CNN+Softmax, 3D CNN+Softmax, and 3DCNN+SVM, for binary AD classification. After comparing the result among the implemented models, they proposed the 3D CNN and SVM provide a wonderful accuracy of 95.74%.

From structural MRI to diffusion tensor imaging, Karim Derghal [29] suggested a cross-modal transfer learning model. The model was initially initialized with domain-dependent data augmentation and pre-trained on a structural MRI dataset before being trained on mean diffusivity data. The technique reduces the phenomenon of overfitting, improves the learning efficiency, and therefore increases the prediction accuracy to 83.57%.

Kang, Wenjie suggests an ensemble learning (EL) architecture based on 2D CNNs, using a multimodel and multislice ensemble [30], to overcome the lack of neuroimaging data. First, the top 11 coronal slices for the AD versus CN classifications of the grey matter density maps were chosen. Second, the chosen slices were used to train the discriminator of a generative adversarial network, VGG16, and ResNet50. Multi-model integration decreased the prediction error rate, while multi-slice ensemble learning was created to acquire spatial data. Finally, domain adaptation and transfer learning were utilized to improve those CNNs. When classifying AD versus CN, AD versus MCI, and MCI versus CN, respectively, this approach attained accuracy values of 90.36%, 77.19%, and 72.36%. Zhang, Fan has proposed their hyper-tuned model (Efficient Net-B1) [31] for Alzheimer's classification with 93.32%.

Table 1 provides a portion of an analytical and comparative summary of the works under discussion.

2.2 | Hybrid deep learning/machine learning approach

A Hybrid AI-based model was proposed [32] by testing three different hybrid models (DenseNet201-Gaussian NB, DenseNet201-XG Boost, DenseNet201-SVM). The models were trained and tested using Adam Optimizer and 1000 Epochs. The best model (DenseNet201-Gaussian NB) provides an accuracy of 91.75%. The model produces the expected result only when both the training and testing data are similar.

Nagarathna. has proposed a Hybrid Model [16] by combining VGG19 with CNN. The Hybrid model has been trained after preprocessing the dataset by rescaling, augmentation, and data resampling. They compared the result with a CNN model which produces 82.65% accuracy after pre-processing the data. But the proposed hybrid model produces 95.52%.

The paper [33] by Yildirim, Muhammed has defined a Hybrid model by combining Resnet50 and CNN. The Resnet50 provides an accuracy of 78%, while the proposed hybrid model gives an accuracy of 90%.

This paper [34] proposed by Puente-Castro, uses a DL model to identify Alzheimer's disease in sagittal MRI images. The extracted data are taken from ADNI and OASIS. To address the data imbalance, an initial parameterization was set in the class weights. An evaluation metrics and strategy were defined to determine the output of the proposed model. Therefore, the proposed Hybrid method presents satisfactory results in both sets of sagittal images.

Different types of combinations (CNN and RNN) have been tested by Ebrahimi, Amir, and Luo in their paper [35]. The models use 3D MRI images. Firstly, the 3D images are preprocessed and then sliced into 2D images. The CNN model extracts the features. The RNN models perform binary classification. In every case, the CNN ResNet-18 model has been used. The CNN ResNet-18 has been tested with the same dataset for classification and found 82% accuracy. The tested hybrid models provide the accuracy as (i) ResNet-18 + LSTM, 84%, (ii) ResNet-18 + BiLSTM, 79%, (iii) ResNet-18 + GRU, 82%, (iv)

TABLE 1 ML/DL approach for AD detection.

Ref #	Dataset	Classification technique	Imaging	Accuracy
[23]	ADNI	Ensembled of deep belief networks	MRI	90%
[24]	ADNI	SVM	Protein sequence information	85%
[25]	ADNI	Random SVM	fMRI	94%
[26]	ADNI, AIBL	RF	MRI	75%
[27]	ADNI	Dual learning based on 3D CNN	MRI, demographic, neuropsychological, and APOe4 genetic data	86%
[28]	ADNI	2D CNN, 3D CNN, 3D-CNN-SVM	MRI	82.57%, (2D CNN), 89.76% (3D CNN), 95.74% (3D-CNN-SVM)
[29]	sMRI(416)	Cross-Modal Transfer Learning	MRI	83.57%
[30]	MRI Data(798)	2D-CNN, VGG16	MRI	90.36%
[31]	ADNI	Efficient Net-B1	MRI	93.20%

TABLE 2 Hybrid ML/DL approach for AD detection.

Ref#	Dataset	Classification technique	Imaging	Accuracy
[32]	ADNI	DenseNet201 + Gaussian NB	sagittal MRI	91.75%
[16]	ADNI	VGG19 + CNN	MRI	95.52%
[33]	ADNI	ResetNet50 + CNN	MRI	90%
[34]	ADNI, OASIS	ResetNet + SVM	sagittal MRI	86.81%(OASIS) 78.64%(ADNI)
[35]	MRI Data	CNN + RNN	MRI	91%
[17]	ADNI	CNN-BiLSTM	MRI	92.62%

ResNet-18 + TCN (with 2 residual blocks), 82%, (v) ResNet-18 + TCN (with 3 residual blocks), 88%, (vi) ResNet-18 + TCN (with 4 residual blocks), 91% (vii) ResNet-18 + TCN (with 5 residual blocks), 78%.

The paper by El-Sappagh, Shaker has proposed a hybrid model (ResNet-18 + BiLSTM) and shown an accuracy of 92.62% [17]. The model jointly optimizes two types of tasks, that is, classification of a multiclass dataset and four cognitive scores computation, by simultaneously learning and fusing discriminative features from time-series and BG data. A comparative summary of the works discussed is given in Table 2.

The aforementioned studies' findings show that for both machine learning and deep learning models, ensemble models perform better than single models. Similarly to this, the findings imply that data augmentation improves performance and lowers the likelihood of model overfitting. The application of transfer learning, when a pneumonia-specific pre-trained model is utilized, is said to have good accuracy. The ADNI dataset was mostly used.

3 | METHODOLOGY

We have proposed and built a hybrid deep-learning model using InceptionV3 and a CNN model for Alzheimer's disease detection. Figure 1 represents the overall workflow diagram of the proposal. The main steps of the proposal are given as follows:

- Step-1: The proposal starts with data collection and the images were taken from the ADNI database. This dataset comprises a total of 6,400 images extending four distinct Alzheimer's disease stages, each representing a unique progression of the disease.
- Step-2: Due to the fact that unprocessed real-world data often includes noise and varies in image size, preprocessing is required before the data can be fit into a deep learning model. Various steps, including resizing the image, converting the image into bgr2rgb, feature scaling etc. are applied to the collected dataset.
- Step-3: There is an uneven distribution of Alzheimer's disease progression images across the ADNI dataset. SMOT-ENN is used at this point to fix the problems and normalize the data so that model training can proceed without overfitting.
- Step-4: In this step, we build our proposed hybrid model using the InceptionV3 model and a tune CNN model for Alzheimer's detection. Here, we take the InceptionV3 model for Alzheimer's detection and merge it with a basic CNN model. The next step is to divide the image dataset into a training set and a test set. Finally, the model has been trained and validated with the appropriate images.
- Step-5: Finally, the performance is evaluated based on the accuracy of the proposed hybrid model. In this stage of the proposed model, the performance metrics such as accuracy, precision, recall, f1 score, and confusion matrix have been used to evaluate the model of our experiment and

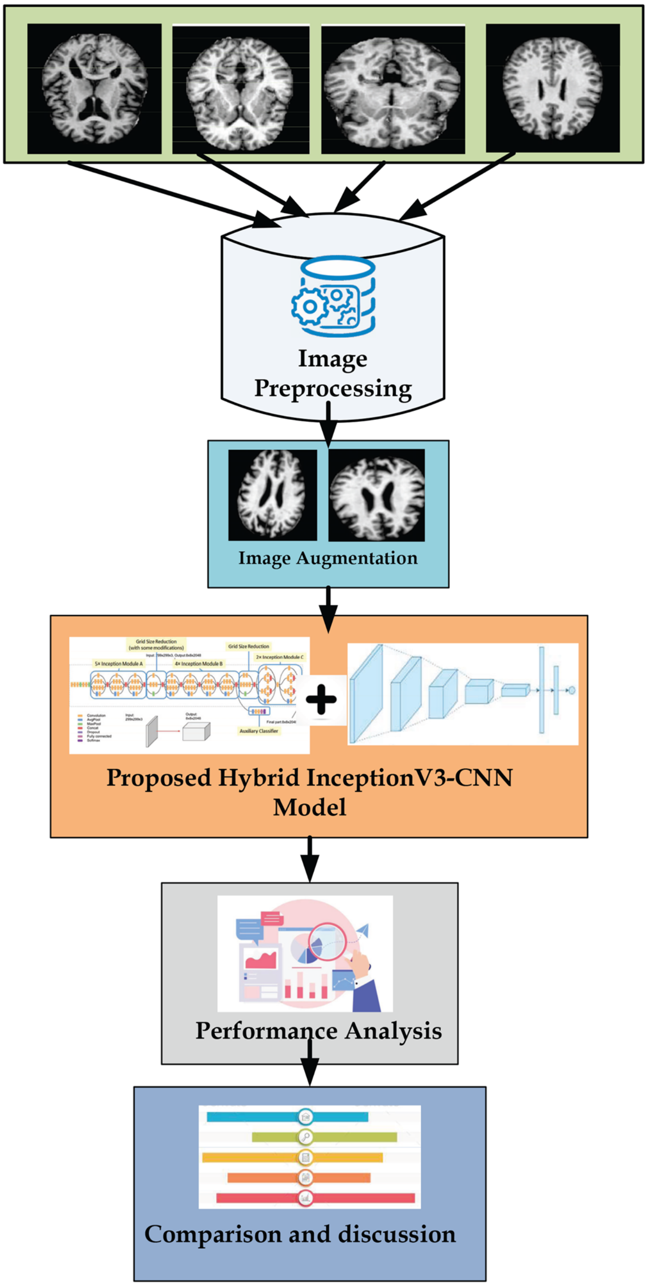


FIGURE 1 The workflow diagram of the proposed InceptionV3-CNN hybrid model.

also showed the comparison analysis with other existing models.

3.1 | Dataset collection

The Alzheimer's Dataset used in this research has been taken from [36], which is a slice and 2D image collected from Alzheimer's Disease Neuroimaging Initiative (ADNI) database [37]. It has been used in most studies alone or in combination with other data sets. The main objective of the ADNI study is to collect more and more related data and test the effectiveness of MRI and other biomarkers in measuring the progress of AD. The data set consists of 6400 MRI images, and all of the Images are utilized in the investigation. In the dataset used, there are 4 classes of data: Very Mild Demented, Mild Demented, Moderate Demented, and Non-Demented. In the experiment, the data set was first organized and pre-processed. The network was first trained using part of pre-processed data, and then testing the model was carried out with the remaining data. Table 3 displays the volume of data and image samples used in the investigation. Figure 2 shows sample MRI images of different stages of Alzheimer's.

TABLE 3 Number of samples in the ADNI dataset.

Class	No of samples
NonDemented	3200
VeryMildDemented	2240
MildDemented	896
ModerateDemented	64

3.2 | Data pre-processing

Data pre-processing is a series of computerized techniques for assembling input data and making them utilizable for target learning models. Data pre-processing is required to clean the noise, identify and rectify missing values, and make the data for the objective of obtaining a final image that is as clean as possible and that contains only the information relevant to the task at hand [38].

There is no skull stripping processing performed on the dataset used in this research. The images in the dataset were already skull-stripped and therefore do not contain any skull. However, for the purposes of this study, registration was not applied as a pre-processing technique to map the images into a standard space.

Further, this pre-processed dataset enhances the efficacy of the training and testing of the whole proposed research model. In this proposed hybrid model, three types of data pre-processing are utilized, namely Image Filtering, data normalization, and data augmentation. Figure 3 shows the In the same way, there are a number of image pre-processing techniques that are common to a wide variety of diagnostic investigations of Alzheimer's. Figure 4 shows the preprocessed images and the original image related.

3.2.1 | Image filtering

An image filter is a tool that can be used to change the size, colors, shading, and other features of an image. The image is changed using several graphical editing methods and image filters. An image filter often modifies the image at the pixel level, which means that each pixel is changed independently. Both 2-D and 3-D images can use it. Typically, the image

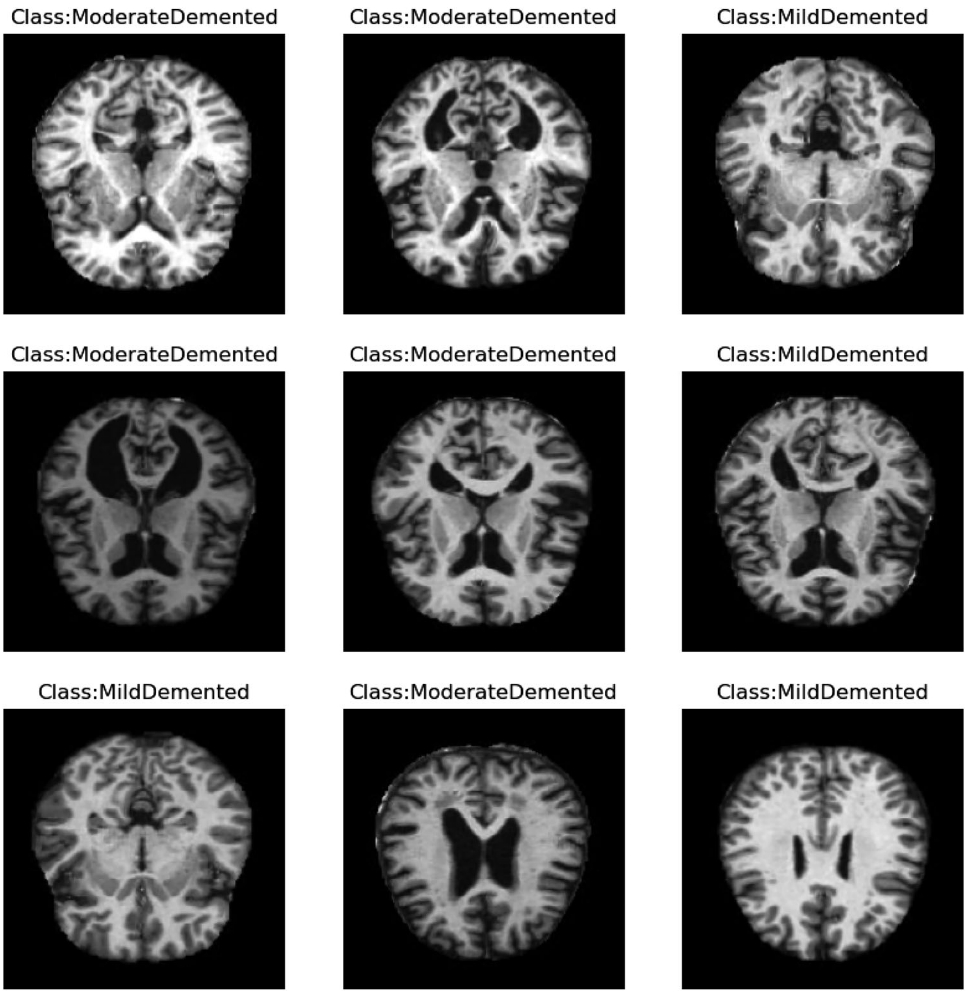


FIGURE 2 Sample MRI images in ADNI dataset.

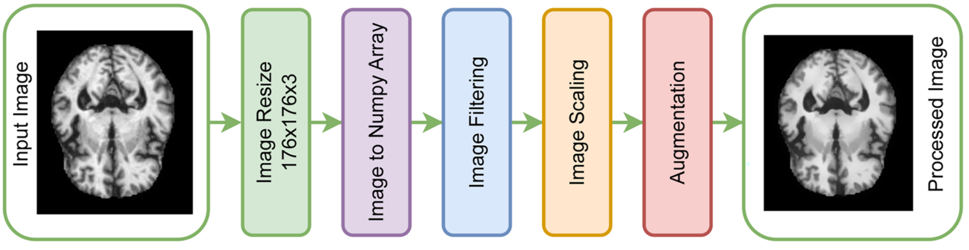


FIGURE 3 Image pre-processing steps.

filtering process offers choices like Editing the image's colour scheme, theme, and contrast, Changing the texture, enhancing the image's effects, and adjusting image brightness.

In this paper, the bilateral filter [39] has been used. A non-linear, edge-preserving, and noise-reduction smoothing filter for images is known as a bilateral filter. It swaps out each pixel's intensity for a weighted average of intensity values from adjacent pixels. Importantly, the weights depend not only on the radiometric differences but also on the Euclidean distance of the pixels (e.g. range differences, such as color intensity, depth distance etc.). Sharp edges of the processed image are preserved as a result.

3.2.2 | Data normalization

The ANDI image dataset for the current study has different attributes [40]. The original images in the dataset are collected and recorded in different dimensions, especially images collected from real life. Therefore, the MRI images are resized to

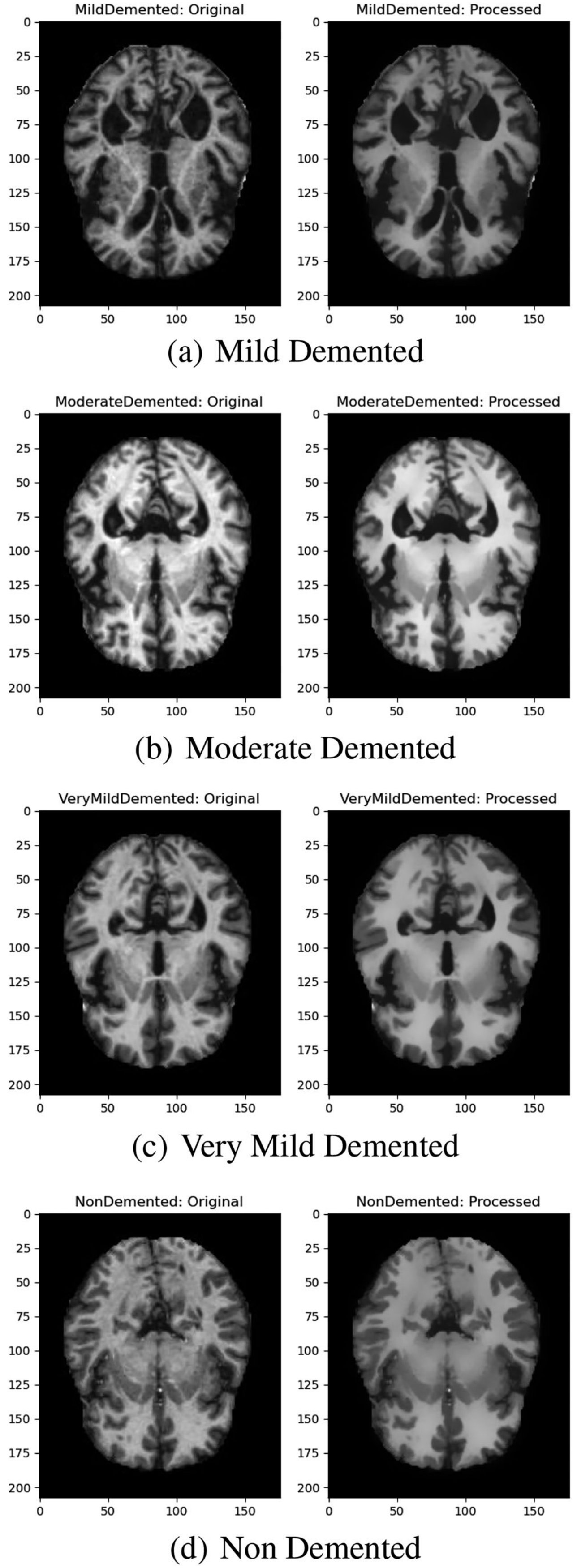


FIGURE 4 Original vs. pre-processed images.

TABLE 4 Number of training, validation, and testing dataset.

Class	Before SMOTE-ENN	After SMOTE-ENN
Training dataset	5120	10208
Validation dataset	640	1135
Testing dataset	640	1261

a target size from a model implementation and computational perspective. All images are converted into sizes of 176 x 176 x 3, indicating height, width, and number of channels, respectively.

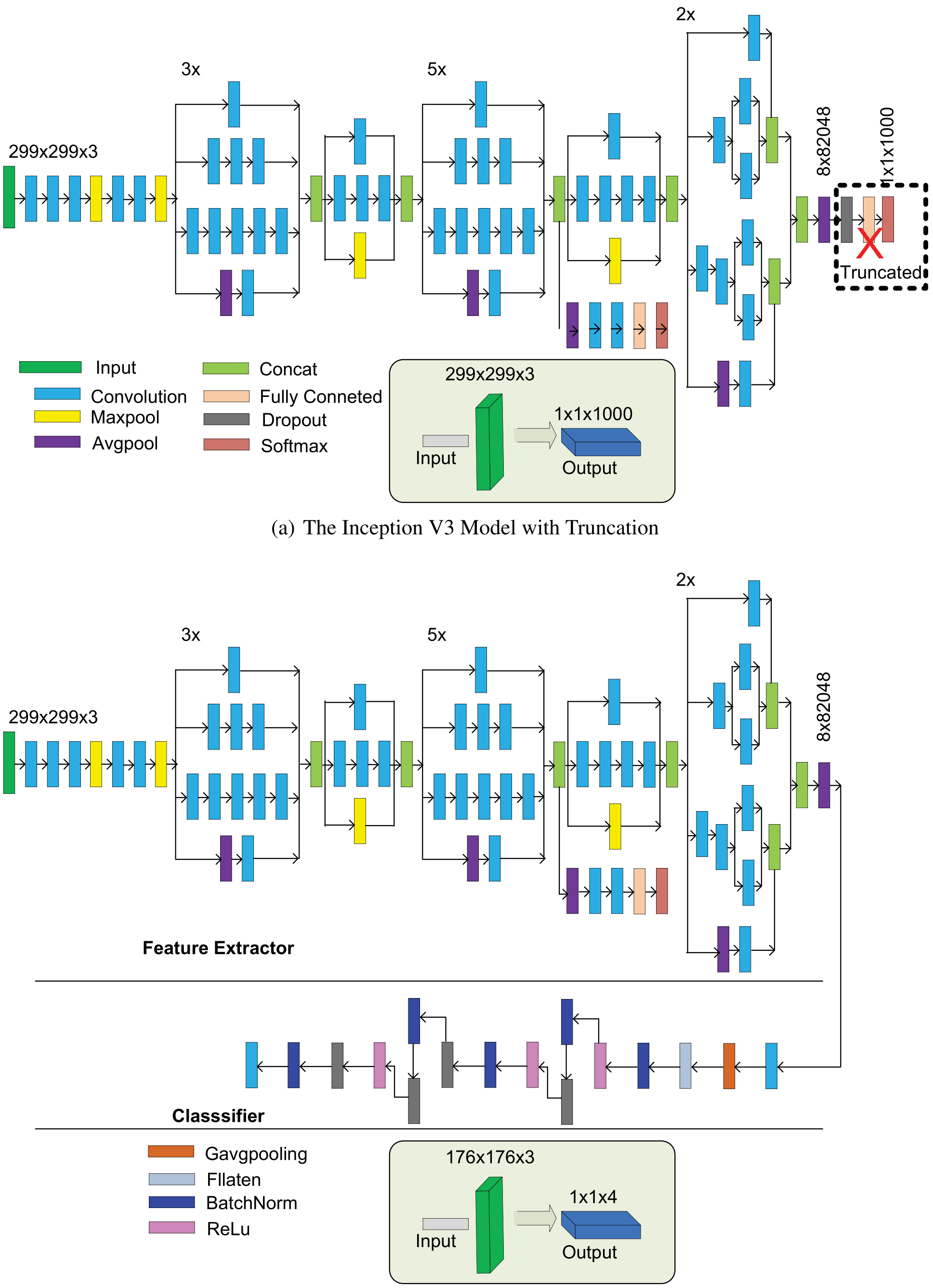
3.2.3 | Data augmentation

We may uniformly increase the number of examples in our dataset using this statistical technique. Based on minority conditions that are supplied as input and are currently present, the module generates new instances. The number of majority cases remains the same. The Synthetic Minority Oversampling Technique (SMOTE), is arguably the most used technique for creating new samples. Nitesh Chawla wrote about this method in his 2002 work [41]. Instead of just producing duplicates of existing minority cases, the method generates new examples by sampling the feature space for each target class and its closest neighbours, integrating the characteristics of the target case with those of its neighbours. But occasionally, oversampling can alter the objective of the endeavour. To manage this in this paper, SMOTE-ENN [42] has been used. This is a hybrid technique that aims to clean overlapping data points for each of the classes distributed in the sample space. After SMOTE oversampling is performed, the class clusters may invade each other's space. As a result, the classifier model will be overfitted. Now, to get better class clusters, Edited Nearest Neighbor (ENN) is applied to under-sample minority class samples done by SMOTE.

3.2.4 | Data splitting

The dataset was divided into a training set and a testing set to train the models and assess their performance. The training set was used to fit the models, while the testing set was used to generate predictions, which were evaluated against the original labels to measure the effectiveness of the approach. During the process, the dataset underwent oversampling and undersampling using SMOTE-ENN.

The data splitting was performed at the subject level to define the training, validation, and test sets. To train and test the proposed model, the entire dataset was split into 80% for training data and 20% for testing and validation data. The allocation of samples for each set is illustrated in Table 4, showcasing the number of samples before and after the application of SMOTE-ENN. After preprocessing the image dataset, a data split was conducted, where 80% of the data was assigned to training, 10% to testing, and 10% to validation.



(a) The Inception V3 Model with Truncation

(b) The Proposed Hybrid Model

FIGURE 5 Proposed model vs. original model.

3.3 | Proposed deep learning model

In the Proposed Hybrid Model, the pre-trained networks Inception V3 [43] learned from the ImageNet data set have been used to extract information from MRI images. Figure 5a shows the truncated model of the original pre-trained Inception V3 model.

Inception-V3 [43] has proven to be a popular and effective architecture for various computer vision tasks. It has demonstrated superior performance in image classification, object detection, and image recognition competitions. The Inception-V3 architecture is renowned for its ability to capture multi-scale features by utilizing inception modules, which consist of parallel convolutional layers with different filter sizes. One of the key reasons for selecting Inception-V3 is its exceptional accuracy on diverse datasets and tasks. Its deep and complex architecture allows it to learn intricate patterns and features, enabling it to achieve state-of-the-art performance on challenging image recognition tasks. This high accuracy is particularly crucial in our research focused on early detection of Alzheimer's disease. Another one is its computational efficiency. Despite its depth and complexity, Inception-V3 has been optimized to reduce computational overhead. This makes it computationally feasible for real-time applications, where efficient processing is essential. Furthermore, it offers the flexibility of learning where pre-trained weights on large-scale datasets, such as ImageNet, provide a valuable starting point for our task. By leveraging these pre-trained weights, we can benefit from the learned features and reduce the amount of training required on our specific dataset. Considering these factors, including Inception-V3 as our feature extractor was a well-founded choice. Its impressive accuracy, computational efficiency, and transfer learning capabilities make it an ideal candidate for our research on the early detection of Alzheimer's disease.

The Keras module applications allow downloading such a network with the weights of ImageNet, with target image size, and by ignoring the final output layers. Finally, we have added 16 necessary CNN layers including a fully connected output layer to the downloaded truncated Inception V3 model. The final output layer is added for sorting with four output units (AD/NAD/MAD/VMAD). In the first training phase, all layers of the base model are frozen and only the CNN layers added manually are trained. This is done for 100 epochs with the Adam optimizer and a learning rate of 0.001.

These models have been prepared by the fundamental building elements, that is, CNN layers [44]. When a filter is applied to an input, the result is an activation. This is how convolution works at its core. A feature map is created after applying the same filter to the same input numerous times. This feature map shows the locations and intensities of any patterns that were found in the input as well as an image of the patterns. Another part of CNN architecture is a pooling layer. The proportions of the feature sets are reduced with the help of the pooling layer. As a result, there are fewer parameters to learn and fewer network processing requirements. The convolution layer forms the feature map and the pooling layer adds the features in the map in a specific area.

The proposed CNN model contains 16 layers: 3 Dropout Layer with a dropout rate of 0.5, 1 GlobalAveragePooling2D Layer with a 4X4 pool size, 1 Flatten Layer, 5 Batch Normalization layers, 5 Dense Layers with a ReLU activation function, 1 final dense with softmax prediction layer.

The CNN input Dropout layer with a 0.5 dropout rate takes the Input from the output layer of the Inception V3 and deactivates half of the input and hidden neurons of the network by setting the input to zero value. The Dropout layer chooses the subset of the feature. This will remove the overfitting of the Network. In backpropagation, the weight of only activated neurons will be updated. The GlobalAveragePooling2D layer will average the full image. This down-sample of the network. The following Flatten layer will convert the Input image into data of a 1-dimensional array. The distribution of layers changes during taring. This may allow occurring Gradient Vanishing problem. The BatchNormalization layer normalizes the value between 0 and 1 and avoids/minimizes the problem. The following dropout layer again selects the subset of the feature with a dropout rate of 0.5. The following Dense layer with ReLU works by matrix-vector multiplication. The layer gets the 2048 dimension of input and produces 512 as output. The ReLU outputs the value if it is positive; otherwise, it returns 0. The following combination of BatchNormalization, Dropout, and Dense layers functions in the same manner and reduces the dimensions to extract the feature map. Finally, the fully connected output layer (dense with Softmax) classifies the output into the 4 desired classes. Figure 6 shows the proposed hybrid model's summary view with layer details.

The proposed model classifies Alzheimer's disease by MRI images very efficiently. The efficiency of the classification of the proposed model is better than the existing models. But the environment we have with Microsoft Windows takes 10–12 h to train the Hybrid Model.

TABLE 5 Confusion matrix.

Predicted results	Actual positive	Actual negative
Yes	TP	FP
No	FN	TN

4 | EXPERIMENTAL RESULTS ANALYSIS

4.1 | Experimental environment

The experiments were carried out on a machine running Microsoft Windows 11 Pro, with an Intel(R) Core (TM) Intel(R) Core(TM) i5-8350U CPU @ 1.70GHz 1.90 GHz CPU running at 5 cores, 8 logical processors, and 256 GB SSD, 16GB of RAM and 8 GB Shared GPU.

4.2 | Programming language and tools

The implementation of this thesis was done in Python programming language. The SMOTE-ENN, data augmentation, and data filtering libraries were made exclusively in Python, using the core Python library along with some of the popular imported libraries, including Sklearn, NumPy, Seaborn, and TensorFlow. Data analysis and the testing of the model were done in Python via Anaconda Navigator's Jupyter Notebooks. The CNN was built using Keras with a TensorFlow backend. All visualizations were created using matplotlib of Python.

4.3 | Performance metrics

Understanding how well the system works is important. Different evaluation metrics are used for this. The model divides the data into four categories: true positive (TP), true negative (TN), false positive (FP), and false negative (FN) as shown in table 5.

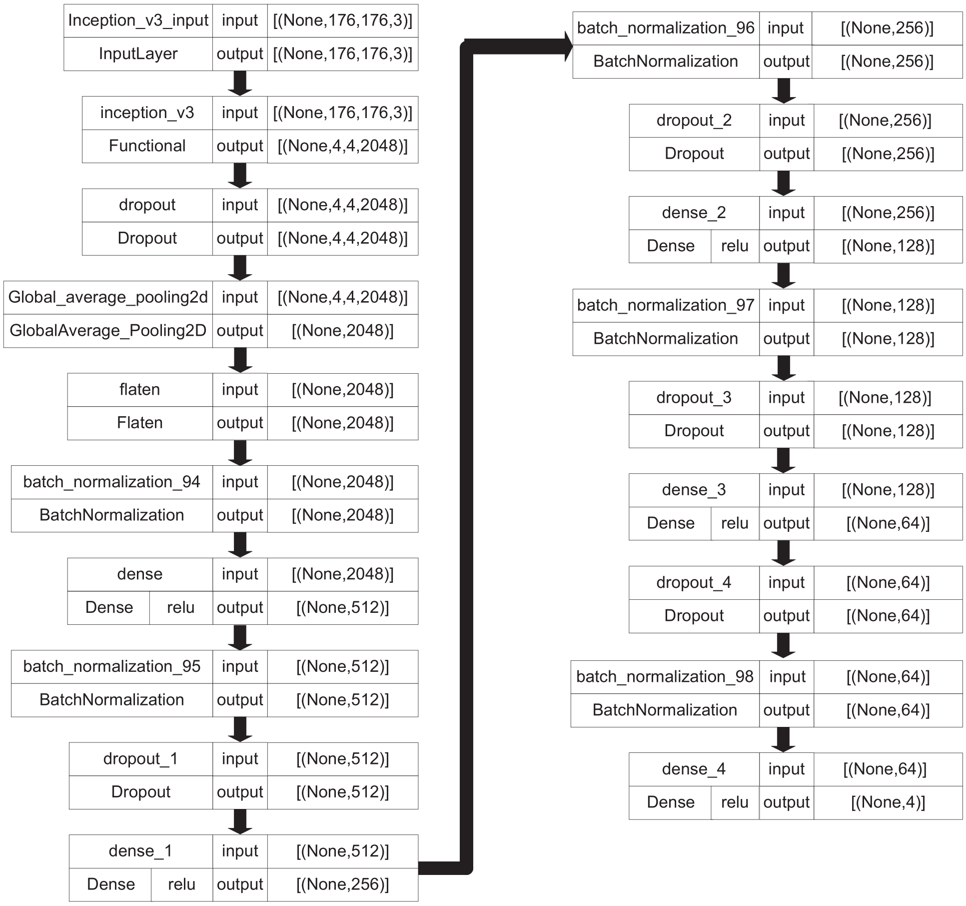


FIGURE 6 The proposed hybrid model summary with layer details.

TP shows accurately classified positive instances, TNs correctly detected negative instances, while FPs and FNs are incorrectly predicted as positive and negative, respectively.

Accuracy is the number of correctly classified predictions by a model over the entire number of classified instances and can be defined as below:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}. \tag{1}$$

Precision is the number of correct positive classifications from the total number of actual classified predictions by the model as positive and can be statistically defined as

$$Precision = \frac{TP}{TP + FP}. \tag{2}$$

Recall is the score of true positive predictions to the instances that actually belong to the positive class.

$$Recall = \frac{TP}{TP + FN}. \tag{3}$$

F1-score is an evaluation measure to estimate the performance of the model based on the average precision and recall and is represented as:

$$F1 - score = 2 * \frac{precision * recall}{precision + recall}. \tag{4}$$

The network parameters' specific values were chosen after carefully considering their effects on the model's training dynamics, convergence speed, and generalization capabilities. For example, the learning rate of 0.001 was selected as it has been widely used in similar deep-learning tasks and has shown good performance in balancing convergence and avoiding overfitting. Regarding the optimizer's choice, we employed both Adam and RMSprop optimizers to compare their performance in our experiments. These optimizers have been widely adopted in deep learning models and have effectively optimized the model's weights and biases. Other parameters, such as data augmentation techniques (e.g. horizontal flipping, rotation, and zooming) were incorporated to enhance the model's ability to learn diverse image patterns and improve its robustness against variations in the input data. The data splitting ratio of 80% for training and 20% for testing and validation follow common practices to ensure adequate data for model training while preserving a sufficient evaluation dataset for performance assessment.

TABLE 6 Setting of experimental parameters.

Parameters	Values
Learning rate	.001
Optimizer	Adam and RMSprop
Hearing	−0.2 to +.2
Horizontal flipping	True
Rotating	−20 to +20
Zooming	0.8 to 1.5
Splitting	Train-80%, Test and valid-20%

Table 6 shows the different parameters used in the experiments.

4.4 | Experimental setup

We have trained and tested the model in different ways. Firstly the dataset was not augmented, and we trained the model using the RMSprop optimizer with a learning rate of 0.001 and the Adam optimizer with the same learning rate. Then the whole data set was taken by SMOTE-ENN. The model was then trained and tested again by both the RMSprop optimizer and Adam. The input size, preprocessing by Image Filter, and Dataset Splitting ratio were the same for all combinations. We consider 4 different experiments as follows:

4.5 | Result analysis

In our experiment, we have conducted four extensive experiments and found the best one for our proposals. The four experiments are as follows:

- **Experiment 1**: In this case, we perform pre-processing on the data set and apply filtering to improve image quality. The hybrid model was then trained and tested using *RMSprop Optimizer* with a learning rate of 0.001 to achieve the best results.
- **Experiment 2**: In this case, we perform preprocessing on the data set and apply filtering to improve image quality. The hybrid model was then trained and tested using *Adam Optimizer* with a learning rate of 0.001 to achieve the best results.
- **Experiment 3**: In this case, we use the SMOTE-ENN algorithm to balance the initial dataset in an effort to reduce the impact of the overfitting problem. The *RMSprop Optimizer* is then applied with the same rate of 0.001.
- **Experiment 4**: In this particular experiment, we employ the SMOTE-ENN algorithm to normalize the primary data set. After that, the *Adam Optimizer* is utilized with the same rate of 0.001 as before.

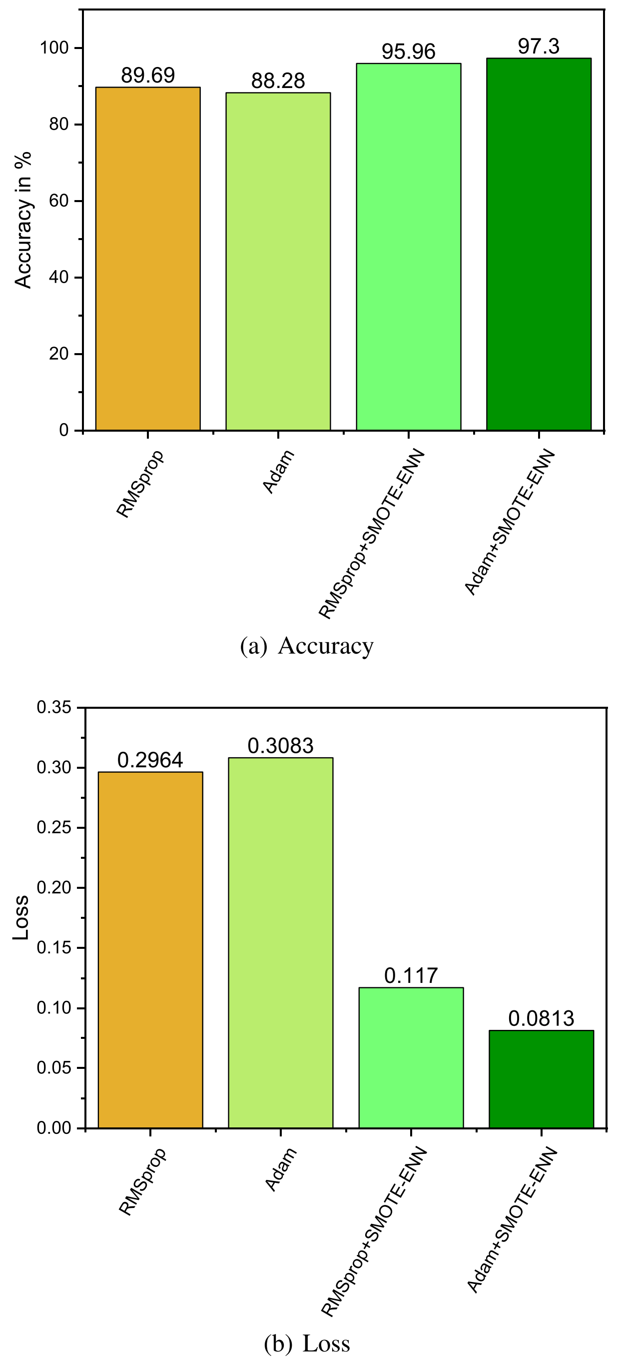


FIGURE 7 Performance results.

The accuracy and loss graph is illustrated in Figure 7 of RMSprop, Adam, RMSprop+SMOTE-ENN and Adam+SMOTE-ENN for AD classification.

Figure 7a reveals that the highest accuracy of 97.3% is achieved when employing the Adam optimizer in conjunction with SMOTE-ENN. This combination outperforms the other configurations, including RMSprop alone (89.69%), Adam alone (88.28%), and RMSprop with SMOTE-ENN (95.96%). These findings suggest that integrating SMOTE-ENN into the training process improves the model's ability to classify AD

TABLE 7 Testing result comparison in different experiments.

Experiment #	Optimizer	Augmentation	Accuracy	Loss	F1 score
1	RMSprop	NA	89.69	0.2964	0.8753
2	Adam	NA	88.28	0.3083	0.8663
3	RMSprop	SMOTE-ENN	95.96	0.1170	0.9589
4	Adam	SMOTE-ENN	97.31	0.0813	0.9705

accurately. Furthermore, the superior performance of the Adam optimizer, particularly when coupled with SMOTE-ENN, highlights its effectiveness in this AD diagnosis task.

Figure 7b focuses on the loss values obtained from the different optimization techniques and data augmentation methods. Lower loss values indicate better convergence and potential for improved generalization of the model. The results show that using SMOTE-ENN in combination with either RMSprop or Adam leads to lower loss values compared to using the optimizers alone. Specifically, the lowest loss value of 0.0813 is achieved when employing the Adam optimizer with SMOTE-ENN, followed by 0.117 for RMSprop with SMOTE-ENN. In contrast, using RMSprop alone results in a loss value of 0.2964, while Adam alone yields a loss value of 0.3083.

These findings further support the effectiveness of combining SMOTE-ENN with the Adam optimizer for AD classification. The reduced loss values indicate improved convergence and better generalization, underscoring the proposed approach's advantages. The results suggest that the combination of Adam optimizer and SMOTE-ENN facilitates more.

Our study conducted four experiments to examine the effects of different optimizers, both pre and post the SMOTE-ENN augmentation technique shown in Table 7 . The optimizers we used were RMSprop and Adam, which are widely used in neural network training. Additionally, we applied the SMOTE-ENN technique to augment the data and improve the model's performance. The mathematical equations for RMSprop and Adam optimizers, as well as a general overview of the activation function used, are as follows:

RMSprop: RMSprop is an optimization algorithm that adapts the learning rate based on the magnitudes of recent gradients. It utilizes the following equation:

$$v_t = \beta \cdot v_{t-1} + (1-\beta) \cdot g^2,$$

$$\text{weight}(t) = \text{weight}(t-1) - \left(\frac{\alpha}{\sqrt{v_t + \epsilon}}\right) \cdot g.$$

In the equation above, v_t represents the exponentially weighted moving average of the squared gradients, weight_t denotes the updated weight, g represents the gradient of the objective function, α is the learning rate, β is a decay rate parameter, and ϵ is a small constant for numerical stability.

Adam: Adam is another optimization algorithm that computes adaptive learning rates for each parameter. It combines the concepts of momentum and RMSprop. The equations for Adam are as follows:

$$m_t = \beta_1 \cdot m_{t-1} + (1-\beta_1) \cdot g,$$

$$v_t = \beta_2 \cdot v_{t-1} + (1-\beta_2) \cdot g^2,$$

$$\hat{m} = \frac{m_t}{1-\beta_1^t},$$

$$\hat{v} = \frac{v_t}{1-\beta_2^t},$$

$$\text{weight}\, t = \text{weight}\, t - 1 - \left(\frac{\alpha}{\sqrt{\hat{v}} + \epsilon}\right) \cdot \hat{m}.$$

Here, m_t and v_t represent the exponentially weighted moving averages of the gradients and squared gradients, respectively. $\hat{m}$ and $\hat{v}$ are bias-corrected estimates of the moments, β_1 and β_2 are decay rate parameters, and the remaining variables have the same meanings as in RMSprop.

As shown in Table 7, Experiments 1 and 2 were performed using RMSprop and Adam optimizers without any augmentation (NA). Experiments 3 and 4 used the same optimizers but incorporated the SMOTE-ENN augmentation technique. In terms of accuracy, Experiment 4 achieved the highest accuracy of 97.31%, followed by Experiment 3 with an accuracy of 95.96%. This suggests that the combination of the Adam optimizer and the SMOTE-ENN augmentation yielded the best performance in terms of accuracy. Regarding the loss metric, Experiment 4 also demonstrated the lowest loss value of 0.0813, indicating better convergence and smaller prediction errors. Experiment 3 exhibited a slightly higher loss value of 0.1170, which can be attributed to the data augmentation and the optimization process. When considering the F1 score, Experiment 4 outperformed the other experiments with a score of 0.9705, indicating a better balance between precision and recall. Experiment 3 achieved an F1 score of 0.9589, demonstrating the effectiveness of the SMOTE-ENN augmentation in improving the model's ability to handle imbalanced data. The experiments highlight the impact of different optimizers and utilizing the SMOTE-ENN augmentation technique. Experiment 4, which combined the Adam optimizer with SMOTE-ENN, achieved the highest accuracy, lowest loss, and highest F1 score, indicating its effectiveness in improving the model's performance.

It is evident from the table that the error rate was 0.2964 and 0.3083 without SMOTE-ENN in Experiments 1 and 2, respectively. While the SMOTE-ENN was utilized, the error rate decreased to 0.1170 and 0.0813, respectively. For both scenarios, the f1 score improves from 0.8753 to 0.9589 for *RMSprop* and from 0.8663 to 0.9705 for *Adam* Optimizer. Furthermore, the accuracy test results for both cases improved by 6.27% and 9.03%, respectively, demonstrating the impact of data balancing.

Thus, we have chosen an Adam optimizer with a learning rate of 0.001 for the ultimate output. SMOTE-ENN has been incorporated with the data augmentation to handle class imbalance data. Table 8 shows the class-wise results of the Hybrid model with other parameters like precision, recall, and f1 score.

TABLE 8 Testing result of the proposed model.

Classification	Precision	Recall	F1-score	Support
NonDemented	0.99	1.00	1.00	337
VeryMildDemented	1.00	1.00	1.00	351
MildDemented	0.95	0.93	0.94	274
ModerateDemented	0.94	0.95	0.95	301
micro avg	0.97	0.97	0.97	1263
macro avg	0.97	0.97	0.97	1263
weighted avg	0.97	0.97	0.97	1263
samples avg	0.97	0.97	0.97	1263

TABLE 9 Results of the new dataset (AD-5).

Models	Dataset	Accuracy rate (%)
RMSprop	AD-5	96.07
Adam	AD-5	97.09
RMSprop + SMOTE-ENN	AD-5	97.98
Adam + SMOTE-ENN	AD-5	98.57

Furthermore, we have included another dataset to make the model robust for AD detection. The dataset [45] consists of five stages of Alzheimer's Disease (AD) that are split into two directories: training and testing. The stages are as follows:

- EMCI (Early Mild Cognitive Impairment): Early stage of cognitive decline.
- LMCI (Late Mild Cognitive Impairment): Advanced stage of cognitive decline.
- MCI (Mild Cognitive Impairment): Mild cognitive decline not yet classified as AD.
- AD (Alzheimer's Disease): Diagnosed with Alzheimer's disease.
- CN (Cognitively Normal): Reference group without significant cognitive impairments.

This dataset consists of 1296 images with 5 classes that can be used for tasks such as AD classification or prediction. Table 9 shows the new dataset (AD-5) results in tabular format.

Figure 8a shows the learning curve during the training and validation of the model for *RMSprop optimizer*. The validation and loss curve is well-fitted and shows ideal characteristics as they pass smoothly with increasing epochs during the experiments. Similarly, the loss curve shown in Figure 8b depicts ideal results as the training and validation loss remain steady after 40 epochs and has a minimum distance.

The model training and validation learning curve for *RMSprop optimizer* is depicted in Figure 8a. As the number of epochs increases in the experiments, the training and validation curves smoothly pass, demonstrating ideal characteristics. Also, as shown in Figure 8b, ideal results were achieved when the distance between the training and validation loss remained small after 40 iterations.

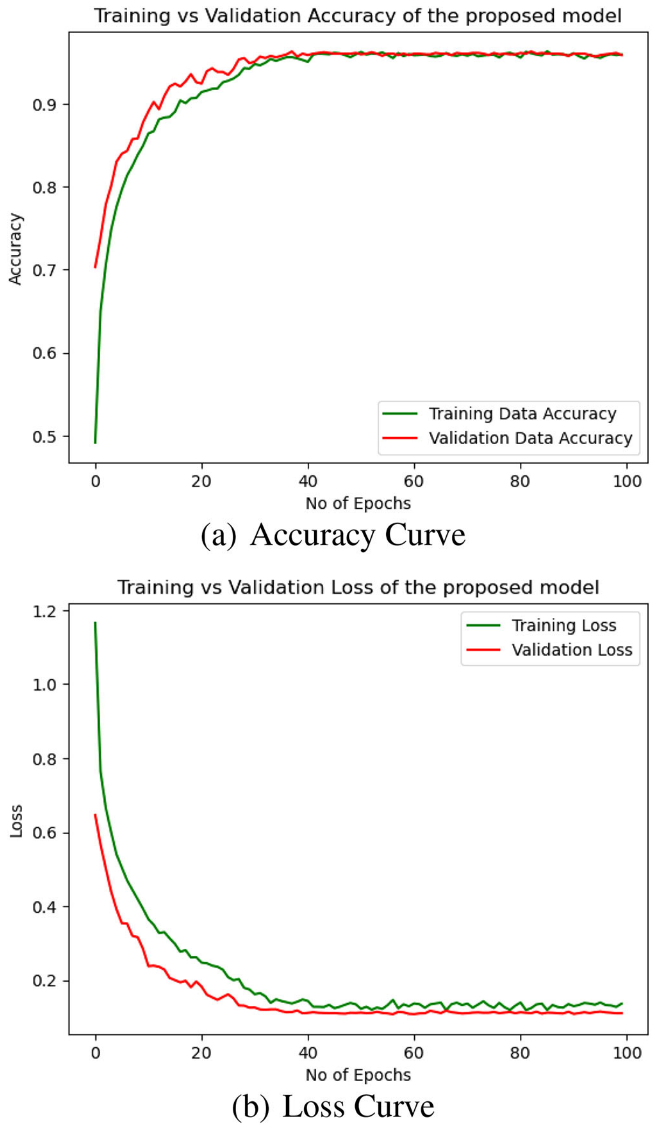


FIGURE 8 Accuracy and loss curve for *RMSprop* optimizer.

Figure 9a depicts Adam Optimizer's progression along various learning curves. The training and validation learning rate accelerates with increasing epoch count, reaching a steady state after 40 iterations. The usefulness of the curves is demonstrated by their output as they pass smoothly during the whole experiment. The loss curve similarly represents the model's loss rate. After 45 epochs, they level off and remain relatively constant with no discernible trend.

The confusion matrices for *RMSprop* and *Adam* optimizer are shown in Figures 10 and 11. When using the *RMSprop optimizer*, the model accurately identifies 328 of 330 healthy patients. Only 29 out of 231 and 18 out of 309 patients, respectively, with mild and moderate symptoms, could not be classified 10. However, the best performance with the proposal is provided by the *adam* optimizer.

Figure 11 demonstrates that it correctly identifies all 337 healthy and 313 Very mildly affected patients. It correctly classified 255 out of 274 patients with a mild case and 286 out of 301

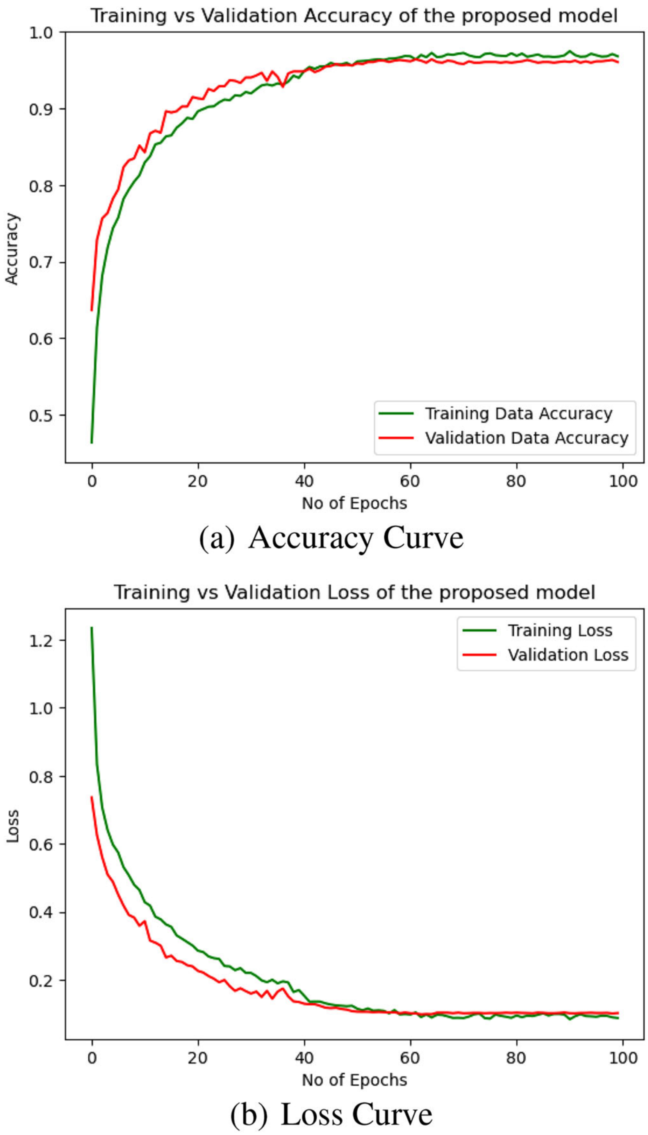


FIGURE 9 Accuracy and loss curve for *Adam* optimizer.

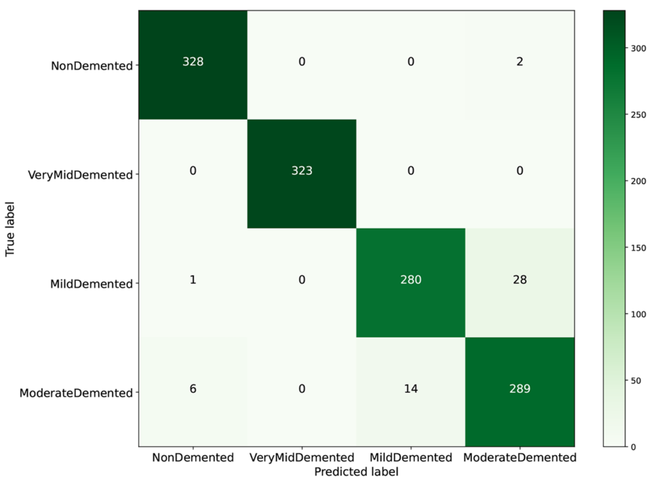


FIGURE 10 Confusion matrix for RMSprop Optimizer.

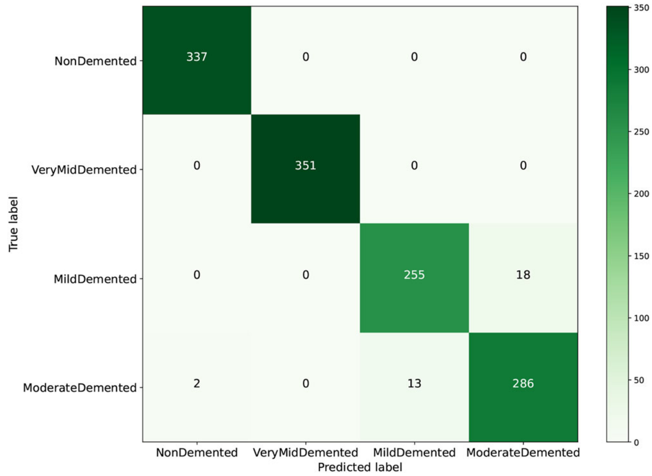


FIGURE 11 Confusion matrix for Adam Optimizer.

with a moderate case. Thus, the effectiveness of the proposal is validated.

Figure 12 shows some test results in the proposed Hybrid model detected. Because the proposal has the potential to achieve an accuracy of 97.31 percent, all of the test samples were identified appropriately.

4.6 | Discussion

The result produced by our Hybrid model has been compared with some other existing similar DL models using different pre-trained networks. Some of them have used very small datasets, some have used similar ADNI data sets, some have used the OASIS dataset, and some have used the collected MRI dataset. The model using OASIS data has produced only 86.81% accuracy, while the accuracy of the model trained by the collected MRI dataset has an accuracy of 91%. The remaining values have an accuracy between 90 and 95%. The proposed model of this article gives the best output with 97.31% accuracy. Table 10 shows the comparison of our performance with the existing works.

4.7 | Complexity analysis

The complexity comparison is shown in Table 11, which provides insights into each experiment's computational complexity based on prediction time.

From Table 11, we observe that Exp4 (Adam + SMOTE-ENN) has a prediction time of 37 s, which is less than the prediction time of Exp3 (RMSprop + SMOTE-ENN) at 38 s. This indicates that Exp4 performs slightly better in terms of prediction time.

- RMSprop and Adam: Both Exp1 and Exp2 do not involve any additional preprocessing techniques. Their prediction times are 35 s and 34 s, respectively.

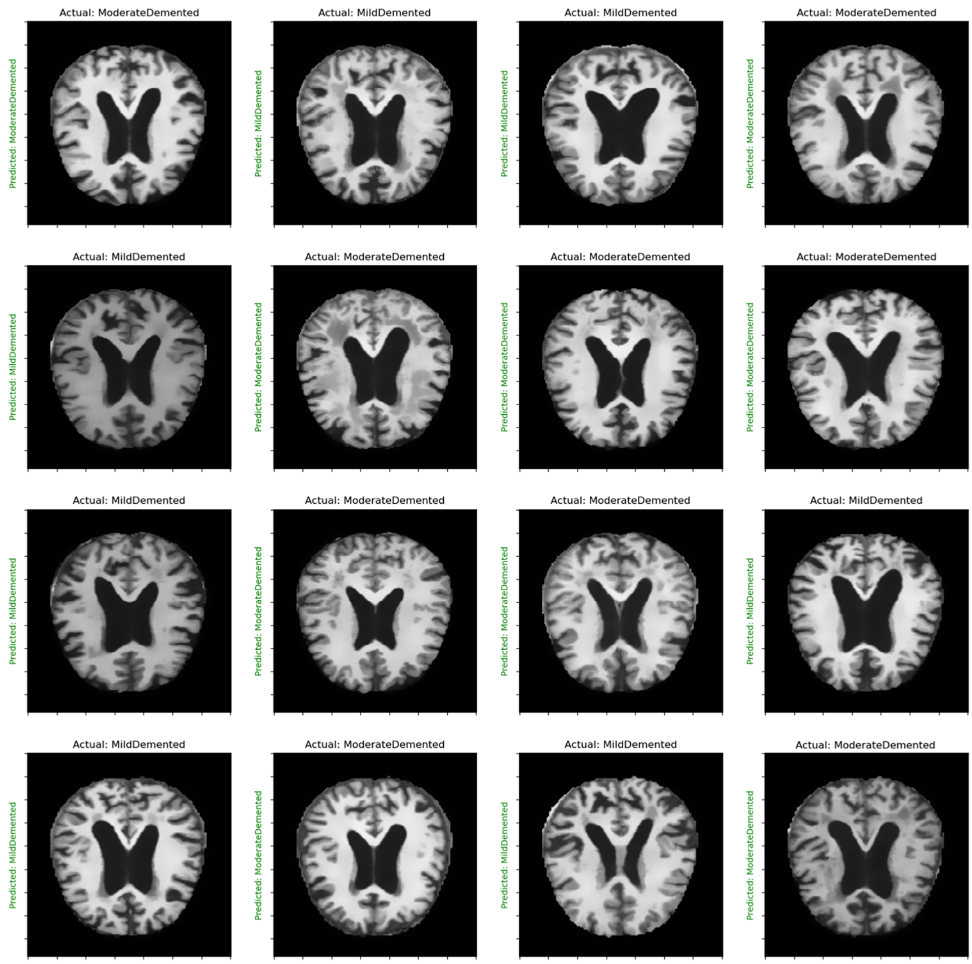


FIGURE 12 Prediction result of the proposed hybrid model.

TABLE 10 Comparison of testing accuracy with existing models.

Ref#	Dataset	Technique	Accuracy
[32]	ADNI	DenseNet201-GaussianNB	91.75%
[16]	ADNI	VGG19 with CNN	95.52%
[33]	ADNI	ResetNet50-CNN	90%
[34]	ADNI and OASIS	ResetNet and SVM	86.81%(OASIS) 78.64%(ADNI)
[34]	MRI (450)	CNN and RNN	91%
[17]	ADNI (1536)	CNN-BiLSTM	92.62%
[3]	ADNI (4094)	VGG-16	83.80%
Proposed Model	**ADNI (6400)**	**InceptionV3-CNN**	**97.31%**

- RMSprop + SMOTE-ENN: Exp3 incorporates the SMOTE-ENN technique as a preprocessing step. SMOTE (Synthetic Minority Over-sampling Technique) and ENN (Edited Nearest Neighbors) are commonly used for addressing class imbalance in machine learning datasets. The addition of SMOTE-ENN may increase the computational complexity of Exp3 compared to Exp1 and Exp2. Exp3 has a prediction time of 38 s, indicating a slightly higher prediction time compared to Exp1 and Exp2.
- Adam + SMOTE-ENN: Exp4 combines the Adam optimization algorithm with the SMOTE-ENN preprocessing technique. Similar to Exp3, Exp4 involves additional steps for handling class imbalance. However, Exp4 demonstrates a lower prediction time of 37 s compared to Exp3. This implies

TABLE 11 Complexity comparison of the four experiments.

Exp. No.	Exp. name	Prediction time (s)
Exp1	RMSprop	35
Exp2	Adam	34
Exp3	RMSprop + SMOTE-ENN	38
Exp4	Adam + SMOTE-ENN	37

that Exp4 provides a slightly better prediction time while incorporating the same preprocessing techniques.

Therefore, we can conclude that Exp4 (Adam + SMOTE-ENN) has a lower complexity compared to the other experiments when considering the additional step of SMOTE-ENN preprocessing.

4.8 | Limitations

In our research, we have identified several limitations that deserve attention. First, the dataset used for training and evaluation was relatively limited in size, which may affect the generalizability of our model to larger and more diverse datasets. While efforts were made to ensure its representativeness, further investigation is needed to assess the model's performance on a broader scale. Second, the availability of computational resources posed constraints on the scale and complexity of our experiments. Future studies with greater access to computational resources could explore the model's performance under more extensive scenarios, allowing for a more robust evaluation. Finally, the real-world deployment of the model raises important considerations. Addressing computational efficiency for real-time processing and ensuring ethical considerations, particularly regarding the privacy and security of sensitive data, are vital aspects that require further investigation and discussion.

5 | CONCLUSION

The primary contribution of this research is the design and creation of an effective Hybrid Model for MRI-based early Alzheimer's disease identification. The pre-trained Model Inception V3 was used for extracting the feature of the Input Dataset and the customized CNN layer was used for final classification. The model performs better when using preprocessed data. Additionally, performance improves when the dataset has been oversampled by the SMOTE-ENN technique. Experimental investigations are supported by the theoretical analyses of our suggested model. Using ADNI datasets, performances and behaviours were compared. The proposed model provided 97.31% of accuracy which is 2–10% more compared to other related models. The trials have demonstrated that the CNN model optimizer and input dataset are mostly responsible for the model's performance. The next step would be to challenge these models with more data with less reliable parameters, such as a live stream of data directly from an MRI machine, Data from other sources, and to improve the data preparation process to detect and handle data anomalies. In the future, we will try to collect more real datasets and train the model using only real data.

AUTHOR CONTRIBUTIONS

Md Masud Rana: Data curation, methodology, software, writing - original draft. Md Manowarul Islam: Conceptualization, formal analysis, methodology, supervision, visualization, writing - original draft. Md. Alamin Talukder: Formal analysis, investigation, methodology, visualization, writing - review and editing. Md Ashraf Uddin: Formal analysis, investigation, visualization. Sunil Aryal: Investigation, validation. Naif Alotaibi: Visualization. Salem A. Alyami: Validation. Khondokar Fida Hasan: Validation, visualization. Mohammad Ali Moni: Validation, visualization, writing - review and editing.

ACKNOWLEDGEMENTS

This research was supported by the ICT Innovation Fund, ICT division, Ministry of ICT, Bangladesh; The Research Title "A Deep Learning based hybrid Model for Alzheimer's Disease Detection", Year 2022-2023; Ref. No. 56.00.0000.053.20.005.22-105

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare that they are relevant to the content of this article.

DATA AVAILABILITY STATEMENT

The dataset is collected from a free and open-access source repository such as ADNI Data is available on: https://www.kaggle.com/datasets/sachinkumar413/alzheimer-mri-dataset

ETHICS APPROVAL STATEMENT

Not applicable

PATIENT CONSENT STATEMENT

Not applicable

PERMISSION TO REPRODUCE MATERIAL FROM OTHER SOURCES

Not applicable

CLINICAL TRIAL REGISTRATION

Not applicable

ORCID

Md. Alamin Talukder https://orcid.org/0000-0002-3192-1000
Khondokar Fida Hasan https://orcid.org/0000-0002-8008-8203
Mohammad Ali Moni https://orcid.org/0000-0003-0756-1006

REFERENCES

1. Haque, R.U., Levey, A.I.: Alzheimer's disease: A clinical perspective and future nonhuman primate research opportunities. Proc. Natl. Acad. Sci. 116(52), 26224–26229 (2019)
2. Breijyeh, Z., Karaman, R.: Comprehensive review on alzheimer's disease: causes and treatment. Molecules 25(24), 5789 (2020)
3. Lim, B.Y., et al.: Deep learning model for prediction of progressive mild cognitive impairment to alzheimer's disease using structural mri. Front. Aging Neurosci. 14, 876202 (2022)
4. Zhao, Z., et al.: Conventional machine learning and deep learning in alzheimer's disease diagnosis using neuroimaging: A review. Front. Comput. Neurosci. 17, 1038636 (2023)
5. Alzheimer's Disease International (ADI): Alzheimer's and dementia. (2022). https://www.alzint.org/about/. Accessed 18 October 2021
6. World Health Organization (WHO): Dementia. (2022). https://www.who.int/news-room/fact-sheets/detail/dementia. Accessed 18 October 2021
7. Talukder, M.A., et al.: An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning. Expert Syst. Appl. 230, 120534 (2023)
8. Talukder, M.A., et al.: Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning. Expert Syst. Appl. 205, 117695 (2022)
9. Ahamed, K.U., et al.: A deep learning approach using effective preprocessing techniques to detect covid-19 from chest ct-scan and x-ray images. Comput. Biol. Med. 139, 105014 (2021)
10. Islam, M.M., et al.: A dynamic access-point transmission power minimization method using pi feedback control in elastic wlan system for iot applications. Internet Things 8, 100089 (2019)
11. Sharmin, S., et al.: A hybrid dependable deep feature extraction and ensemble-based machine learning approach for breast cancer detection. In: IEEE Access (2023). https://doi.org/10.1109/ACCESS.2023.3304628
12. Akhter, A., et al.: A robust hybrid machine learning model for Bengali cyber bullying detection in social media. Natural Lang. Process. J. 4, 100027 (2023)
13. Uddin, M.A., et al.: Machine learning based diabetes detection model for false negative reduction. Biomed. Mater. Devices (2023)
14. Ahmed, N., et al.: Machine learning based diabetes prediction and development of smart web application. Int. J. Cogn. Comput. Eng. 2, 229–241 (2021)
15. Talukder, M.A., et al.: A dependable hybrid machine learning model for network intrusion detection. J. Inf. Secur. Appl. 72, 103405 (2023)
16. Nagarathna, C.R., Kusuma, M.I.: Automatic diagnosis of alzheimerā's disease using hybrid model and cnn. Int. J. Innov. Res. Sci. Eng. Technol. 3(1), 1–4 (2022)
17. El-Sappagh, S., et al.: Multimodal multitask deep learning model for alzheimer's disease progression detection based on time series data. Neurocomputing 412, 197–215 (2020)
18. Jarvik, L., Greenson, H.: About a peculiar disease of the cerebral cortex. Alzheimer Dis Assoc Disord 1(1), 7–8 (1987)
19. Förstl, H., Levy, R.: On certain peculiar diseases of old age. History Psychiatry Lond.: Alpha Acad. 2, 71–101 (1991)
20. Itzhaki, R.: Herpes simplex virus type 1, apolipoprotein e and Alzheimer'disease. Herpes: J. IHMF 11, 77A–82A (2004)
21. Dominy, S.S., et al.: Porphyromonas gingivalis in alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. Sci. Adv. 5(1), eaau3333 (2019)
22. Braak, H., Braak, E.: Frequency of stages of alzheimer-related lesions in different age categories. Neurobiol. Aging 18(4), 351–357 (1997)
23. Jain, R., et al.: Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. Cognit. Syst. Res. 57, 147–159 (2019)
24. Xu, L., et al.: An efficient classifier for Alzheimer's disease genes identification. Molecules 23(12), 3140 (2018)
25. Bi, X.a., et al.: Random support vector machine cluster analysis of resting-state fmri in Alzheimer's disease. PLoS One 13(3), e0194479 (2018)
26. Alickovic, E., Subasi, A., Initiative, A.D.N., et al.: Automatic detection of alzheimer disease based on histogram and random forest. In: International Conference on Medical and Biological Engineering, pp. 91–96. Springer, Cham (2019)
27. Spasov, S., et al.: A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. Neuroimage 189, 276–287 (2019)
28. Feng, W., et al.: Automated mri-based deep learning model for detection of alzheimer's disease process. Int. J. Neural Syst. 30(06), 2050032 (2020)
29. Aderghal, K., et al.: Classification of alzheimer disease on imaging modalities with deep cnns using cross-modal transfer learning. In: 2018 IEEE 31st international symposium on computer-based medical systems (CBMS), pp. 345–350. IEEE, Piscataway (2018).
30. Kang, W., et al.: Multi-model and multi-slice ensemble learning architecture based on 2d convolutional neural networks for alzheimer's disease diagnosis. Comput. Biol. Med. 136, 104678 (2021)
31. Zhang, F., et al.: A single model deep learning approach for alzheimer's disease diagnosis. Neuroscience 491, 200–214 (2022)
32. Sharma, S., et al.: HTLML: Hybrid ai based model for detection of alzheimer's disease. Diagnostics 12(8), 1833 (2022)
33. Yildirim, M., Cinar, A.: Classification of alzheimer's disease mri images with cnn based hybrid method. Ingénierie des Systèmes d Inf. 25(4), 413–418 (2020)
34. Puente-Castro, A., et al.: Automatic assessment of alzheimer's disease diagnosis based on deep learning techniques. Comput. Biol. Med. 120, 103764 (2020)
35. Ebrahimi, A., et al.: Deep sequence modelling for alzheimer's disease detection using mri. Comput. Biol. Med. 134, 104537 (2021)
36. Kumar, S.: Alzheimer MRI dataset. https://www.kaggle.com/datasets/sachinkumar413/alzheimer-mri-dataset. Accessed 18 May 2022
37. Alzheimer's Disease Neuroimaging Initiative (ADNI): ADNI data. https://adni.loni.usc.edu/data-samples/access-data/. Accessed 18, May 2021
38. García, S., et al.: Big data preprocessing: methods and prospects. Big Data Anal. 1(1), 1–22 (2016)
39. Elhoseny, M., Shankar, K.: Optimal bilateral filter and convolutional neural network based denoising method of medical image measurements. Measurement 143, 125–135 (2019)
40. Isaksson, L.J., et al.: Effects of mri image normalization techniques in prostate cancer radiomics. Physica Med. 71, 7–13 (2020)
41. Chawla, N.V., et al.: SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357 (2002)
42. Wang, J.: Prediction of postoperative recovery in patients with acoustic neuroma using machine learning and smote-enn techniques. Math. Biosci. Eng.: MBE 19(10), 10407–10423 (2022)
43. Szegedy, C., et al.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826. IEEE, Piscataway (2016)
44. Alzubaidi, L., et al.: Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions. J. Big Data 8(1), 1–74 (2021)
45. Madhucharan, A.: Alzheimer's disease 5 class dataset - ADNI. https://www.kaggle.com/datasets/madhucharan/alzheimersdisease5classdatasetadni. Accessed 29 June 2023

How to cite this article: Rana, M.M., Islam, M.M., Talukder, M.A., Uddin, M.A., Aryal, S., Alotaibi, N., Alyami, S.A., Hasan, K.F., Moni, M.A.: A robust and clinically applicable deep learning model for early detection of Alzheimer's. IET Image Process. 17, 3959–3975 (2023). https://doi.org/10.1049/ipr2.12910