EARLY STAGE OF ALZHEIMER'S DISEASE PREDICTION

SUDARSAN K , ABHISHIEK B , JEEVA D , NISHA DEVI K

ABSTRACT

Alzheimer's disease is a neurological ailment that progresses over time. Early identification is essential for effective treatments. An overview of current studies in the area of early-stage Alzheimer's disease prediction is provided in this abstract. Our research examines a complex strategy that integrates information from several data sources, such as neuroimaging, cognitive tests, genetic markers, and clinical history. Machine learning algorithms are used to analyse these many data sets in an effort to spot unnoticeable trends and biomarkers indicating the start of the disease. We emphasise the significance of extensive longitudinal studies, which contribute to the creation of prediction models and offer insightful information about the course of Alzheimer's disease. We also talk about the necessity for strict data privacy laws and the ethical issues surrounding the usage of sensitive health data. Early diagnosis of Alzheimer's disease is crucial for prompt interventions, individualised treatment regimens, and raising the standard of living for those who are affected. This abstract provides a look into the exciting lines of inquiry in this crucial domain that have the potential to transform early detection and treatment approaches for Alzheimer's disease.

KEYWORDS: *Early stage of Alzheimer's disease prediciton, Machine Learning, Personalized treatment ,Disease progression , Healthcare advancements.*

INTRODUCTION

Alzheimer's disease (AD) stands as one of the most pressing challenges in modern healthcare, affecting millions of individuals worldwide and placing an ever-increasing burden on both patients and healthcare systems. The importance of an early AD diagnosis may be attributed to a number of factors, including the possibility for more successful treatments, better patient care, and the development of research into disease-modifying medications. Even with significant advancements in our knowledge of the pathogenic processes causing AD, getting a prompt diagnosis in the early stages is still a difficult undertaking.

The goal of this research is to critically examine the state of early-stage Alzheimer's disease prediction. Our study aims to address the urgent need for precise and trustworthy tools to detect AD risk factors in people before overt cognitive decline appears. We seek to shed light on the multiple strategy needed to address this complicated topic by a thorough examination of the most recent developments in neuroimaging, genetics, cognitive evaluation, and upcoming technologies.

Exciting new opportunities for early detection have been made possible by the development of novel biomarkers and advanced machine learning techniques. This research explores the potential of such breakthroughs,

highlighting their contribution to the creation of prediction models that can offer perceptions into illness start years prior to the appearance of clinical symptoms. We also recognize the ethical issues related to data protection and consent, underlining the significance of ethical research procedures in this field.

Alzheimer's disease is a complicated and progressive neurological ailment that causes memory loss, a slow deterioration of cognitive function, and eventually the loss of independence. While older persons are the majority of those affected, a sizeable number of cases start with modest cognitive impairments in midlife or even earlier. As a result, identifying persons who are at risk for AD in its early stages has significant consequences for both clinical practice and public health. Because prospective disease-modifying therapies are most likely to be successful when used at the earliest possible stage of the illness, the need of early identification is further highlighted by this fact. Early diagnosis also gives patients and their families the ability to make future plans, pick the right course of treatment, and take part in clinical trials for novel medications. We hope that this paper will contribute to the ongoing efforts to improve diagnostic accuracy, patient outcomes, and ultimately bring us closer to efficient treatments and preventive strategies for Alzheimer's disease as we set out on this journey through the changing landscape of early AD prediction.

The goal of this paper's structure is to give a thorough review of the approaches and developments in the field of predicting Alzheimer's disease in its early stages. We'll focus on the following significant elements.

- Neuroimaging and Biomarkers: In this section, we'll examine how cutting-edge neuroimaging methods, such structural and functional MRI, PET scans, and cerebrospinal fluid biomarkers, may be used to spot subtle changes in the brain that are related to AD. These technologies provide important new information on the anatomical and metabolic changes that occur before clinical symptoms.

- Genetic Predisposition: Genetic predisposition is a major factor in the risk of AD. We'll look at the most recent developments in genetics and genomics, such as the discovery of genes linked to risk, like APOE, and how genetic data may help develop early prediction models.

- Cognitive Tests and Assessments: For many years, AD has been diagnosed through cognitive tests and evaluations. We'll talk about how cognitive testing has changed over time, its use in spotting early cognitive decline, and new digital tools and technology that improve assessment precision.

- Machine Learning and Predictive Models: Powerful machine learning algorithms have the capacity to sift through enormous information and spot patterns pointing to the early AD. We will examine these models' potential and drawbacks, as well as how they integrate with different data sources.

- Ethical Considerations: It is crucial to use sensitive health data in a responsible manner. We will discuss the necessity for open and ethical research practices in the pursuit of early AD prediction, as well as ethical issues relating to data protection, informed consent, and these needs.

RELATED WORK

- The paper "Detection of Alzheimer's disease at Early Stage using Machine Learning" S. Pavalarajan, B. A. Kumar, S. S. Hammed, K. Haripriya, C. Preethi and T. Mohanraj was published in the journal Sensors in "2022 International Conference on

  Advanced Computing Technologies and Applications (ICACTA)". The paper discusses the use of Detection of

  Alzheimer's disease at Early Stage using Machine Learning

  Identification of dementia is an important concern in medical image processing. Alzheimer is a common kind of dementia. Four machine learning models were designed for identifying this disease. This is classified as a classification problem, and the classification algorithms tested include logistic regression,

  support vector classifier

- The paper "Machine Learning Based Optimal Feature Selection Technique for Early Stage Prediction of Alzheimer's Disease" by K. Bhatt, N. Jayanthi and M. Kumar on 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC). The paper discusses the Dementia is characterized by degeneration or dysfunction of the brain's nerve cells. Neuronal degeneration causes alteration in memory, behavior, and cognitive capacity of a person. Alzheimer's Disease (AD) is the most conventional type of dementia. AD is currently not curable. Its early detection would be helpful for patients to receive the necessary care at an early stage. It would also prevent the disease from becoming

  deadly. In order to select the most effective features for AD prediction, a novel Extra Tree Classifier (ETC) based feature selection method is proposed. Also, different ML classifiers with features obtained using the proposed feature selection approach are analyzed to identify

  AD in its early stage. Classification accuracy, precision, recall, and F1-score have all been used to evaluate the method's effectiveness and robustness. The Naive Bayes classifier with ETCbased feature selection achieves the highest performance accuracy of 91.96% on OASIS dataset. The proposed work outperformed prior approaches and accurately predicted AD.

- The paper "Early Stage Alzheimer's Disease Diagnosis Method" by M. H. Memon, J. Li, A. U. Haq and M. Hunain Memon was presented at the 2019 16th

  International Computer Conference on Wavelet Active

  Media Technology and Information Processing. The paper discusses The Alzheimer's disease accurate early-stage detection is critically necessary for effective treatment and recovery. Therefore, accurate detection of Alzheimer's disease is a great research problem.

  Different researchers used various techniques to detect Alzheimer's disease effectively however; these methods still have lack of prediction accuracy. In this study we proposed machine learningbased method to diagnosis Alzheimer's disease accurately.

  We used machine learning classifiers for accurate prediction of Alzheimer's disease. Alzheimer's Disease

Neuroimaging Initiative data set has been used to check the proposed method performance. The experimental results demonstrate that logistic regression performance was excellent in terms of accuracy and achieved optimal accuracy 98.12%. Thus, it is recommended for effective early detection of Alzheimer's Disease.

- The paper "Machine Learningbased Alzheimer's Disease Prediction using Personalized Methods" by M. Lavanya, R. R. Chandan, P. Rajasekar, P. R. Rham, M. Deivakani and A. S. Mahesh Kumar was published in 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC).

  Alzheimer's disease (AD) is a neurological illness which occurs over time. Early detection reduces the likelihood of disease worsening condition.

  Unfortunately, the present research focuses on defining disease processes in their existing stages rather than forecasting disease progression. The Generative Personalized embedding model (GPE) is a multi-layer perceptron that is proficient of relating prior knowledge to the prediction process by realizing the patient's behavioural data and helps in forecasting the disease progression. This work suggest a prediction process based on GPE. As a result, a GPA network is built with the connectivity layers and activated layers to observe the possible correlation among information and next generation of Alzheimer's disease. Experiments reveal that the outcomes significantly outperformed the majority of previous designs.

- The paper "Alzheimer Disease Prediction using Machine Learning Algorithms," by J. Neelaveni and M. S. G. Devasana in 2020 in 6th International Conference on Advanced Computing and Communication Systems (ICACCS). Alzheimer disease is the one amongst neurodegenerative disorders. Though the symptoms are benign initially, they become more severe over time. Alzheimer's disease is a prevalent sort of dementia. This disease is challenging one because there is no treatment for the disease. Diagnosis of the disease is done but that too at the later stage only. Thus if the disease is predicted earlier, the progression or the symptoms of the disease can be slow down. This paper uses machine learning algorithms to predict the Alzheimer disease using psychological parameters like age, number of visit, MMSE and education.

Dataset and Methodology

This section discusses in detail the used dataset and provides visual examples of the data. In addition, all stages of the proposed model with all hyperparameters are also discussed in this section.
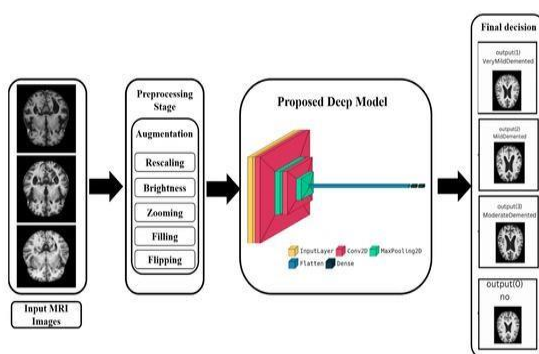
Dataset

In this research, the Alzheimer's dataset was utilized, which is a dataset that has been collected by hand and consists of MRI images that have been verified and labeled by experts. The data is divided into four different classes: Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented. These images can be used to train deep learning models to accurately predict the stage of Alzheimer's disease. The dataset provides an opportunity for researchers to develop algorithms that can accurately diagnose

Alzheimer's disease and aid in the development of effective treatments. The dataset is publicly available on Kaggle and is easily accessible. By making this dataset available to the public, the creators hope to encourage more research in the field and support the development of better algorithms for the diagnosis and treatment of Alzheimer's disease. This dataset was chosen for its availability, its different classes, and its small size on a hard disk compared to other common datasets in this field.
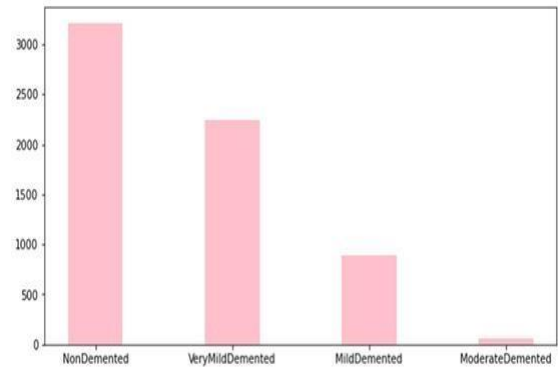
Figure 1. sample image from the database
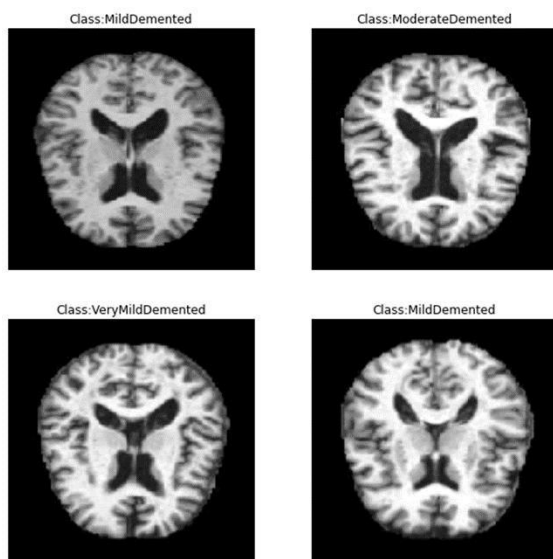
issue. The main block diagram of our approach is shown in Figure 3, and it is covered in more detail in the next section.

Figure 3. Block diagram of all stages of our method.

Figure 2. Statistics of Kaggle dataset

When working on a multi-classification job with this dataset, we can observe that the class Moderate Demented has a very low amount of photos compared to other classes, which leads to false positives and influences the results. In order to expand the amount of photographs in this class and address the imbalance issue at the same time, we use a data augmentation strategy to resolve this

2. *Preprocessing Stage*

In this step, we utilized the Image Data Generator class from the Keras library, which allows for the application of a number of image augmentation methods to the input data to produce a fresh set of enhanced pictures that can be used for training. The particular augmentation methods used in this research include pixel scaling, brightness modifications, magnification adjustments, constant value filling of newly formed pixels, and random horizontal flipping of pictures. These methods seek to both artificially expand the amount of the training dataset and strengthen the model's resistance to changes in the input data. The input picture data may be fed into our deep model once the Image Data Generator instance has been

established, and the enhanced data can then be utilized to train a deep learning model. This process is crucial to ensure that the model can generalise successfully to novel or unexplored data.

### 3. Proposed Deep Model for Binary-Classification

A high-level neural network API called Keras, which is based on TensorFlow, was used to develop the suggested deep learning model [37]. The model is intended for binary classification, with the objective of determining whether or not an input picture belongs to a certain class. The visualisation of our model for a binary classification challenge is shown in Figure 4. A 150 x 150 x 3 (height, width, and depth) picture that depicts a colour image with three channels (red, green, and blue) is the model's first input layer. In order to extract features from the picture, a succession of convolutional layers (Conv2D) and pooling layers (MaxPooling2D) are applied to the input image. The pooling layers down-sample the feature maps that the convolutional layers output after applying filters to the input picture. The features are then flattened and sent through two dense layers (Dense), which modify the features in non-linear ways by activating them using the 'ReLU' activation function. The final prediction is then generated by the output layer using the sigmoid activation function, which converts the input to a probability-like output between 0 and 1. The model is built using the binary cross-entropy loss function and the 'Adam' optimizer, and it is trained using the fit technique on the training set of data. A loss of 0.061 and an accuracy of 0.993 are shown in the training results, demonstrating the model's capacity to provide precise predictions based on the training data. The executive overview of the suggested

binary classification model
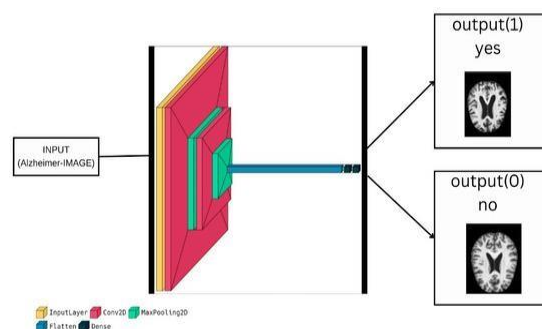
Figure 4. Model visualization to binary classification task

3.4. Proposed Deep Model for Multi-Classification

The model in this study accepts images of dimensions (150, 150, 3), indicating that each image is 150 x 150 pixels with three color channels (red, green, blue). The model applies a sequence of Conv2D and MaxPooling2D layers to decrease the spatial dimensions of the image and extract significant features. These extracted features are then flattened and passed through two dense layers with 'ReLU' and 'SoftMax' activation functions. The 'SoftMax' activation function provides the final probability scores for each class in the classification task. The model is compiled with an 'Adam' optimizer and a categorical cross-entropy loss function. It is trained on the training data for 100 epochs and evaluated on the validation data, achieving an accuracy of 96%. The model's visualization for multi-classification tasks is shown in Figure 6.
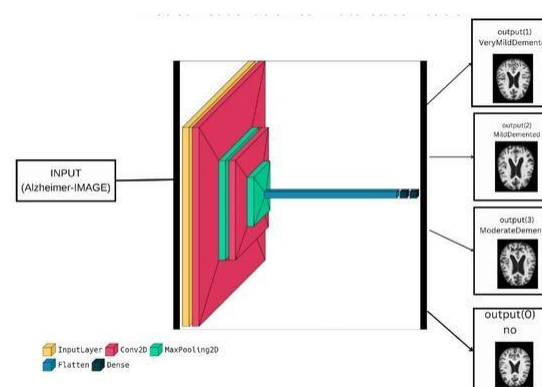
Figure 6. Model visualization to multi-classification task.

Algorithm:

Step 1: BEGIN

Step 2: INPUT: dataset_directory, training_percentage, image_augmentation_parameters, model_parameters, optimizer, loss_function, performance_metrics.

Step 3: Load input dataset from dataset_directory

Step 4: Split the dataset into training set and validation set with training_percentage

Step 5: Instantiate an ImageDataGenerator object with image_augmentation_parameters

Step 6:
  6.1 IF model_parameters is a pretrained model THEN
  6.2 Load pre-trained model
  6.3 ELSE
  6.4 Define a deep learning model using Keras with model_parameters
  6.5 ENDIF

Step 7: Compile the model using optimizer and loss_function

Step 8: Train the model on the training set for a number of epochs with the compiled model and Image Data Generator object.

Step 9:
  9.1 FOR each epoch in the training process DO.
  9.2 Evaluate the model on the validation set using performance_metrics.
  9.3 IF the validation accuracy is not improving THEN.
  9.4 Reduce learning rate.
  9.5 ENDIF.
  9.6 ENDFOR

Step 10: Test the final model on a separate test set to evaluate its generalization performance using performance_metrics.

Step 11: OUTPUT is the performance_metrics of the proposed method and existing methods for comparison.

Step 12: END

RESULTS AND DISCUSSION:

Results

The proposed deep model is trained on the Kaggle dataset through multiple experiments. A standard approach of cross-validation (10-CV) is used for training and testing to ensure a fair and reliable evaluation of the proposed AD detection model. The approach is implemented on a computer equipped with an NVIDIA Tesla T4 GPU and 14 GB DDR4 RAM, using Keras, a Pythonbased library. The 'Adam' optimizer is applied for training the neural network, with binary cross-entropy as the loss function for model 1 and Categorical Crossentropy as the loss function for model 2. Four evaluation measures are used in this study: Accuracy, Precision, Recall, and F1-score.

$$Accuracy = TP+TN/TP+TN+FP+FN$$
(1)
$$Precision = TP/TP+FP$$
(2)
$$Recall = TP/TP+FN$$
(3)
$$F1\text{-}score = 2 \times Precision \times Recall / Precision + Recall$$
(4)

where TP denotes true positives, FP denotes false positives, TN denotes true negatives and FN denotes false negatives.

1. Experimental Analysis

In this paper, two experiments are evaluated using four metrics. The first experiment is based on the first model, which is used for a binary classification task. The second experiment is based on the second model, which is used for multi-classification tasks. The paper provides details and analysis of each
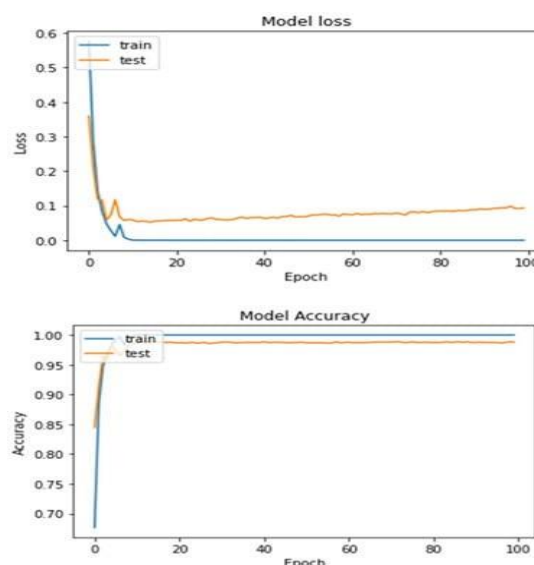
experiment.

A. The first experiment

The proposed deep model was used to classify input MRI images into two groups for a binary challenge (AD or Normal). The confusion matrix of the proposed approach for detecting AD is shown in Figure 9, where class 0 represents normal instances and class 1 represents AD patients.
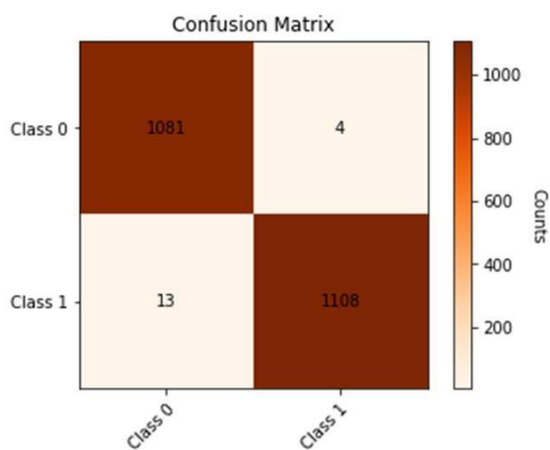
Figure 9. Confusion matrix of the proposed method to detect AD for binary classification tasks.

According to the confusion matrix shown in Figure 10, it can be seen that 1081 normal MRI images were correctly detected as normal, while 0.3% of normal cases were detected as AD cases.

Figure 10. Loss curves (upper) and accuracy curves (lower) for the training and testing data for the proposed model for binary classification task.

B. The second experiment

The proposed deep model was used for multi -classification to categorize input MRI images into four categories: Mild Demented, Moderately Demented, Non -Demented, and Additionally, 98% of AD cases were correctly detected as AD, while 13 MRI images were incorrectly detected as normal cases.

Very Mild Demented. The confusion matrix of the proposed method for detecting demented cases is shown in Figure 11. In this matrix, Class 0 refers to NonDemented cases, Class 1 refers to Very Mild Demented cases, Class 2 refers to Mild Demented

cases, and Class 3 refers to Moderate Demented cases.
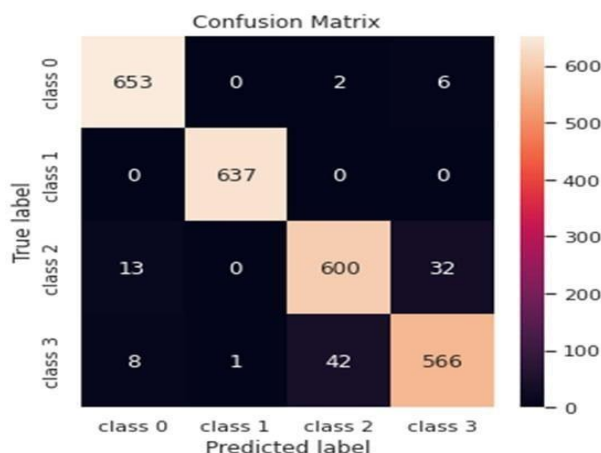
Figure 11. Confusion matrix of the proposed model for multi-classification task.

According to the previous confusion matrix in Figure 11, 653 Non-Demented

cases were correctly detected as NonDemented; 2 MRI images of NonDemented cases were incorrectly detected as Mild Demented cases, and 6 MRI images were correctly detected as Moderately Demented cases. We can also find that 100% of the Very Mild

Demented cases are correctly detected as Very Mild Demented cases. In addition, we can observe that 93% of the Mild Demented cases are correctly detected as Mild Demented, 4.9% of the images are wrongly detected as Moderate Demented, and 2.1% are wrongly detected as Non-Demented cases. Finally, we can also observe from the confusion matrix that 91.7% of the Moderate Demented cases are correctly detected as Moderate Demented, 42 MRI images are wrongly detected as Mild Demented, 1.29% are wrongly detected as Non-Demented cases and 0.16% of the images are wrongly detected as Very Mild

Demented cases.
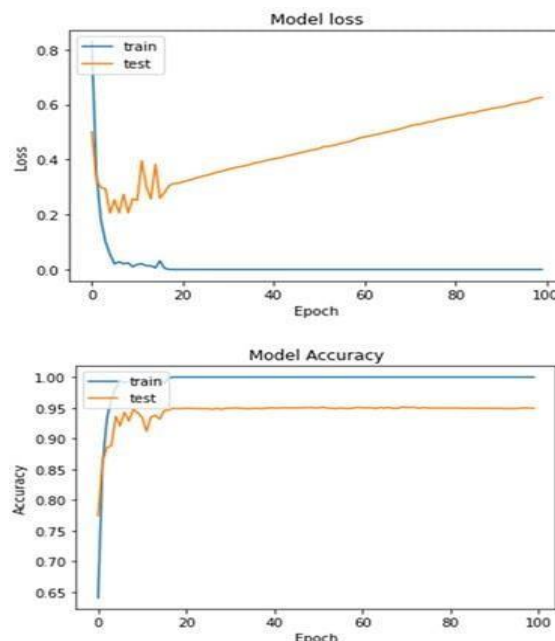
Figure 12. Loss curves (upper) and accuracy curves (lower) for the training and testing data for the proposed model for multi-classification task.

DISCUSSION:

Alzheimer's disease (AD) early-stage prediction is a hotly debated issue in the world of medical research. Alzheimer's disease is difficult to anticipate in its early stages, yet early therapy is more efficient and results in less minor damage than late treatment1. In order to determine the most accurate parameters for Alzheimer's disease prediction, a variety of algorithms including Decision Tree, Random Forest, Support Vector Machine, Gradient Boosting, and Voting classifiers have been used. The Open Access Series of Imaging Studies (OASIS) data is used to generate predictions for Alzheimer's disease, and the performance of ML models is gauged using metrics including Precision, Recall,

Accuracy, and F1-score. Machine learning (ML), a subfield of Artificial Intelligence (AI), uses various probabilistic and optimization techniques to help computers learn from huge and complicated data sets. To diagnose AD in its early stages, researchers generally use machine learning. The survey provides a broad overview of current research in this field and analyses the classification methods

used by researchers working with ADNI data sets. It discusses essential research topics such as the data sets used, the evaluation measures employed, and the machine learning methods used. The proposed classification scheme can be used by clinicians to make diagnoses of these diseases. It is highly beneficial to lower annual mortality rates of Alzheimer's disease in early diagnosis with these ML algorithms. The proposed work shows better results with the best validation average accuracy of 83% on the test data of AD. This test accuracy score is significantly higher in comparison with existing works.

CONCLUSIONS:

The goal of this study is to assess the performance of deep learning models in detecting and classifying Alzheimer's disease (AD) using MRI images. The results obtained in the binary classification task, with an accuracy of 99.30%, and in the four-class classification task, with an accuracy of 95.96%, demonstrate the potential of deep learning models for accurately detecting and differentiating between the different stages of AD. The use of image data with a shape of 150 x 150 x 3, as well as image augmentation techniques and a SoftMax activation function with a dense four-output layer, were found to be critical factors in achieving these results. This study contributes to the growing body of literature on the use of deep learning models for AD detection and classification. Specifically, it demonstrates the potential of using MRI images and deep learning models to accurately detect and classify AD, which has important implications for early diagnosis and treatment. However, there are some limitations to this study that should be considered. The dataset used is relatively small and may not be representative of the entire population. Additionally, only a single modality (MRI images) was considered, and future studies could explore the use of other imaging modalities in combination with deep learning models.

Future work could focus on addressing these limitations and exploring the use of deep learning models in other areas of medical imaging. The development of more explainable deep learning models that can provide insights into the underlying biological mechanisms of AD could further advance our understanding of this disease.

REFERENCES:

- S. Pavalarajan, B. A. Kumar, S. S. Hammed, K. Haripriya, C. Preethi and T. Mohanraj review of Detection of Alzheimer's disease at Early Stage using Machine Learning. 2022 International Conference on Advanced Computing Technologies and Applications (ICACTA).

- K. Bhatt, N. Jayanthi and M. Kumar. Machine Learning Based Optimal Feature Selection Technique for Early Stage Prediction of Alzheimer's Disease ,pp. 715-719.

  M. H. Memon, J. Li, A. U. Haq and M. Hunain Memon, "Early Stage Alzheimer's Disease Diagnosis Method," 2019 16th International Computer Conference on Wavelet Active Media Technology on pp. 222225.

- M. Y. Marusina and A. D. Bukhalov, "Convolutional Neural Networks for Early Prediction of Alzheimer's Diseases", 2021 International Conference on Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS), pp. 394-397.

- M. Lavanya, R. R. Chandan, P. Rajasekar, P. R. Rham, M. Deivakani and A. S. Mahesh Kumar, "Machine Learning-based Alzheimer's Disease Prediction using Personalized Methods," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), pp. 1278-1283.

- J. Neelaveni and M. S. G. Devasana, "Alzheimer Disease Prediction using Machine Learning Algorithms," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 101-104.

- T. Subetha, R. Khilar and S. K. Sahoo, "An Early Prediction and Detection of Alzheimer's Disease: A Comparative Analysis on Various Assistive Technologies," 2020 International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE), pp. 1-5.

- P. Jadhao, P. Palsodkar, R. Raut, K. Chaube, D. Rathod and P. Palsodkar, "Prediction of Early Stage Alzheimer's using Machine Learning Algorithm," 2023 4th International Conference for Emerging Technology (INCET), pp. 1-5.

- R. Sivakani and G. A. Ansari, "Machine Learning Framework for Implementing Alzheimer's Disease," 2020 International Conference on Communication and Signal Processing (ICCSP), doi: 10.1109.