**Deep Learning Based Approach to Diagnose Alzheimer’s using MRI Images**

**A PROJECT REPORT**

***submitted in partial fulfilment of the***

***requirement for the award of the degree***

***Of***

**Master of Computer Applications (MCA)**

**HARIPRASAD J P**

**23FS20MCA00017**



**Department of Computer Applications**

**Manipal University Jaipur**

**Jaipur, Rajasthan, 303007**

**May 2025**

DEPARTMENT OF COMPUTER APPLICATION

MANIPAL UNIVERSITY JAIPUR, JAIPUR-303007 (RAJASTHAN) INDIA

15-05-2025

**CERTIFICATE**

This is to certify the project titled Deep Learning Based Approach to Diagnose Alzheimer’s using MRI Images is a record of the bonafide work completed during the period from 20.01.2025 to 15.5.2025 by Hariprasad JP (23FS20MCA00017) submitted in the partial fulfilment of the requirements for the award of the Degree of Master of Computer Applications (MCA) at the Department of Computer Applications, Manipal University Jaipur, for the academic year 2023-2025

**Dr. Linesh Raja**

*Project Guide, Dept of Computer Applications*

*Manipal University Jaipur*

**Dr. Shilpa Sharma**

*HOD, Dept of Computer Applications*

*Manipal University Jaipur*

**DECLARATION**

I hereby declare that this project report entitled Deep Learning Based Approach to Diagnose Alzheimer’s using MRI Imagesis an original and independent effort completed in fulfilment of the academic requirements for my program. The methodologies, analyses, and conclusions presented are derived from my work, adhering to academic integrity and ethical guidelines.

HARIPRASAD JP

23FS20MCA00017

**ACKNOWLEDGMENTS**

I would like to express my heartfelt gratitude to the following individuals whose support and guidance were instrumental in the successful completion of this project, Deep Learning Based Approach to Diagnose Alzheimer’s using MRI Images.

First and foremost, I extend my sincere appreciation to the director or providing the necessary resources and infrastructure essential for the execution of this project.

I am deeply indebted to **Dr. Shilpa Sharma, Head, Department of Computer Applications,** for their constant encouragement and support throughout this endeavour.

I extend my profound gratitude to my project supervisor **Dr. Amith Hirawat, Assistant Professor, Department of Computer Applications** for their invaluable guidance, patience, and expertise, which significantly contributed to the development and execution of this project.

Special thanks are due to my department guide **Dr. Linesh Raja**, **Associate Professor, Department of Computer Applications** for their insightful feedback and constructive criticism, which greatly enhanced the quality of the work presented in this project.

I would also like to acknowledge the laboratory in charge for their assistance in providing access to the required facilities and equipment essential for conducting experiments and simulations.

Finally, I am grateful to all the faculty members whose expertise and feedback were indispensable in shaping the direction and outcomes of this project.

Their unwavering support and encouragement have been pivotal in every step of this journey, and I am truly thankful for their invaluable contributions.

Thank you

**Abstract**

Alzheimer’s disease (AD) is one of the most pressing neurological challenges of the 21st century, with its prevalence steadily increasing due to global aging. Early detection plays a pivotal role in managing disease progression and improving patient outcomes. Traditional diagnostic methods, while effective, often involve time-consuming and invasive procedures. In light of recent advancements in artificial intelligence, this project focuses on leveraging deep learning techniques for the early and accurate diagnosis of Alzheimer’s disease using MRI imaging, with the primary objective being to develop a scalable, accessible, and accurate diagnostic tool.

The methodology involves the development of three distinct deep learning models, a custom Convolutional Neural Network (CNN), ResNet50, and EfficientNetB1, each trained to classify Alzheimer’s MRI images across various stages of cognitive decline. The dataset underwent preprocessing, including image resizing and augmentation, and was stratified into training, validation, and testing sets. Class weighting was applied to address imbalance issues, and a soft-voting ensemble strategy was used to combine the predictions from all three models, aiming to maximize classification performance.

The ensemble model outperformed individual architectures, achieving a notable test accuracy of 92.85%, demonstrating its robustness and reliability for real-world deployment. The use of ensemble learning helped reduce model bias and improve generalization, showing promise for integration in clinical workflows. The success of this model supports the feasibility of using automated deep learning pipelines in early Alzheimer’s diagnosis, potentially enabling earlier interventions and better care planning.

For this project, key tools and frameworks were utilized, including TensorFlow/Keras for model development and training, VS Code for efficient computation, and Flask for deploying the trained model as a web application. The web app allows real-time MRI image uploads and immediate predictions, thereby serving as a practical interface for clinicians or researchers. This combination of cutting-edge AI models and user-friendly deployment showcases a complete pipeline for early Alzheimer’s disease detection.

**List of Tables**

|  |  |  |
| --- | --- | --- |
| **Table No.** | **Title** | **Page Number** |
| Table 1 | Summary of literature reviews | 10 |
| Table 2 | CNN Model Architecture | 16 |
| Table 3 | Res-Net50 Model Architecture | 17 |
| Table 4 | Efficient-NetB1 Model Architecture | 19 |
| Table 5 | Model Comparison Table | 26 |
| Table 6 | Sample test cases | 26 |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Title** | **Page Number** |
| Figure 1 | Image samples from 4 classes | 14 |
| Figure 2 | CNN Model Architecture | 16 |
| Figure 3 | Pseudocode of CNN | 17 |
| Figure 4 | Res-Net50 Model Architecture | 18 |
| Figure 5 | Pseudocode of Res Net | 18 |
| Figure 6 | Efficient-NetB1 Model Architecture | 19 |
| Figure 7 | Pseudocode of Efficient Net | 19 |
| Figure 8 | Web Application | 20 |
| Figure 9 | Overall Model Flow | 20 |
| Figure 10 | CNN accuracy loss curve | 21 |
| Figure 11 | CNN Confusion Matrix | 22 |
| Figure 12 | Res-Net accuracy loss curve | 22 |
| Figure 13 | Res-Net Confusion Matrix | 23 |
| Figure 14 | Res-Net ROC curve | 23 |
| Figure 15 | Efficient-Net accuracy loss curve | 24 |
| Figure 16 | Efficient-Net Confusion Matrix | 24 |
| Figure 17 | Efficient-Net ROC curve | 24 |
| Figure 18 | Ensembled Model accuracy loss curve | 25 |
| Figure 19 | Ensembled Model ROC curve | 25 |
| Figure 20 | Ensembled Model Confusion Matrix | 25 |

|  |  |  |
| --- | --- | --- |
| **Contents** | | |
|  | | **Page No** |
| Certificate | | i |
| Declaration | | ii |
| Acknowledgement | | iii |
| Abstract | | iv |
| List of Figures and Tables | | v |
| **Chapter 1** | **INTRODUCTION** | 1 |
| 1.1 | Problem Statement | 2 |
| 1.2 | Organization of the Report | 2 |
| **Chapter 2** | **Literature Review** | 3 |
| **Chapter 3** | **Dataset Description** | 14 |
| **Chapter 4** | **Methodology** | 15 |
| 4.1 | Data Preprocessing | 15 |
| 4.2 | Image Augmentation | 15 |
| 4.3 | Data Split | 15 |
| 4.4 | Model Training Strategy | 16 |
| 4.5 | Custom CNN Model | 16 |
| 4.6 | ResNet50-Based Model | 17 |
| 4.7 | EfficientNetB1-Based Model | 18 |
| 4.8 | Ensemble Model: Tri-Stream Architecture | 20 |
| 4.9 | Web Application Development | 20 |
| **Chapter 5** | **Results and Analysis** | 21 |
| 5.1 | Evaluation Metrics | 21 |
| 5.2 | Custom CNN Results | 21 |
| 5.3 | ResNet50 Results | 22 |
| 5.4 | EfficientNetB1 Results | 23 |
| 5.5 | Ensemble Model Results | 24 |
| 5.6 | Comparative Analysis of Models | 26 |
| 5.7 | Sample Test Cases | 26 |
| 5.8 | Discussion | 27 |
| **Chapter 6** | **Implications of the Study on Policy and Practice** | 28 |
| **Chapter 7** | **Limitations of the Study** | 29 |
| **Chapter 8** | **Conclusion and Future Work** | 30 |
| **Chapter 9** | **References** | 31 |

1. **INTRODUCTION**

Alzheimer's disease, the most common cause of dementia, is a global health issue. In 2021, an estimated 57 million individuals had dementia—over 60% in low- and middle-income countries—and some 10 million new cases emerged annually. Dementia now accounts for the seventh leading cause of death globally and has a staggering economic cost, with estimated figures of US$1.3 trillion in 2019. These statistics give an indication of the tremendous personal, social, and economic cost of the disease, and early and accurate diagnostic tools are essential [1].  
Alzheimer's disease is typified by the abnormal buildup of proteins within the brain—primarily amyloid-β plaques and tau tangles—which destroy brain cells and break fundamental connections between them [2]. These pathologic changes take effect many years before clinical symptoms emerge, so their early detection remains the single most important factor in successful treatment. Magnetic Resonance Imaging (MRI) has a vital function in diagnosing and monitoring Alzheimer's by revealing structural changes within the brain, such as atrophy in some regions and tissue density changes. In combination with cognitive and neurological testing, MRI helps differentiate Alzheimer's from other forms of dementia and monitor the progression of disease over time [3, 4]. A general demarcation of dementia with Alzheimer's is as follows [5]:

* Mild Cognitive Impairment: As individuals getting older, they often suffer from memory loss; for certain individuals, this situation turns into dementia.
* Mild Dementia: Disturbances of daily life are caused by impairments in cognition and are sometimes felt by people with mild dementia. Forgetting, confusion, personality change, getting lost, and trouble with daily activities are just a few.
* Moderate Dementia: The patient requires increased care and supervision since tasks become increasingly complex during the day. The symptoms are similar to mild but more pronounced dementia. They may also experience extreme personality changes, such as sudden paranoia or irritability and sleep disturbances.
* Severe Dementia: Here, symptoms may intensify. Some individuals may be unable to talk, and they may need full-time care.

Despite much research, Alzheimer's disease has yet to be cured. Nonetheless, a number of treatments do exist to control its symptoms, such as medication that enhances cognition and behaviour and psychological symptom therapies [6]. Recent developments in deep learning more recently have brought hope for the early diagnosis of Alzheimer's, specifically through the examination of intricate medical images. Such methods can recognize key patterns in large data sets without human involvement, potentially enhancing diagnosis accuracy. There are, nonetheless, some drawbacks in the guise of having to access good-quality data, choosing the appropriate model, and efficient parameter tuning [7,8].  
Traditional diagnostic examinations for Alzheimer's disease are time-consuming, costly, and invasive. As a reaction to such limitations, improvements in machine learning capacity have paved the way for novel diagnostic protocols. Convolutional Neural Networks (CNNs), which have become popular for their capacity to classify images, hold out the prospect of a new direction in the analysis of neuroimaging data. By recognizing small patterns that can indicate early neurodegenerative changes, CNN-based protocols hold the potential to facilitate early interventions and improve patient outcomes.  
This work develops and tests a CNN-based system for early diagnosis of Alzheimer's disease. The organization of the paper is as follows: we start with an in-depth literature review of the pathology of Alzheimer's and existing diagnostic approaches, followed by a detailed description of the CNN architecture and data preprocessing and analysis techniques. We then describe our experimental findings and clinical suggestions, and end with conclusions and suggestions for future work. With this, we aim to help develop user-friendly, efficient, and scalable diagnostic techniques that can combat the rising global menace of Alzheimer's disease. We also aim to develop a user-friendly web application where users can upload medical images, enabling early diagnosis of Alzheimer's disease using the CNN model employed.

* 1. **Problem Statement**

This project intends to design and implement a deep learning-based framework for the early diagnosis of Alzheimer's from MRI images. It aims to design three independent models, a custom CNN, ResNet50, and EfficientNetB1, and ensemble these for better accuracy. All the models were chosen based on their ability to perform image classification, and the ensemble is supposed to capitalize on each one's strengths. The main objectives will be:

* Design a highly optimized custom CNN model specific to Alzheimer's MRI classification.
* Tune two pre-trained approaches, ResNet50 and EfficientNetB1, so that they work better with small datasets of medical images.
* Use data augmentation and class weight methodologies to prevent overfitting and fighting class imbalance.
* Model comparisons to check which is most flexible, be it single or a group.
* Create the best model as an open-source web app to offer real-time predictions.
  1. **Organization of the Report**

This report is organized into multiple chapters, each designed to present a logical progression from background information to the final implementation and results:

* **Chapter 1** provides an overview of Alzheimer’s disease, its impact on society, and the motivation behind the use of deep learning for diagnosis.
* **Chapter 2** presents a comprehensive review of related literature, highlighting key advances in AI-based Alzheimer’s detection and the role of various CNN architectures.
* **Chapter 3** describes the dataset, preprocessing techniques, and model design strategies employed during the project.
* **Chapter 4** outlines the methodology in detail, covering the design, training, and evaluation of all models, including the ensemble approach.
* **Chapter 5** discusses the results obtained from each model and evaluates their performance using key metrics.
* **Chapter 6** discusses the broader implications of the study on healthcare policy and clinical practice, especially the potential of AI tools to support early diagnosis and ease the burden on healthcare providers.
* **Chapter 7** acknowledges the limitations of the current study, such as dataset constraints, generalizability issues, and areas where improvements are needed.
* **Chapter 8** concludes the report with a summary of findings, the overall contributions of the project, and directions for future research, including ways to enhance model performance and real-world deployment.
* **Chapter 9** provides a comprehensive list of all references cited throughout the report, ensuring the academic rigor and traceability of the work.

1. **LITERATURE REVIEW**

Alzheimer's disease (AD) remains on the forefront of neurological challenges, being characterized clinically by gradual and irreversible cognitive decline. Because of increasing AD prevalence and the need for early, reliable diagnosis, researchers have increasingly resorted to the application of the latest technologies, especially AI and DL, to improve the diagnostic accuracy with neuroimaging data. The increasing number of publications supports the paradigm shift these technologies have brought in the early detection, prognosis, and monitoring of the disease by the clinician-scientist duo. To broadly cover such intricate ingredients and these techniques, a focused literature review was carried out with the help of select reputed scientific databases, including IEEE Xplore, PubMed, ScienceDirect, and SpringerLink. The relevance selection subjected articles to peer review over the last decade, focusing on studies concerning DL-based methods for MRI, fMRI, and PET imaging and multi-modal data integration for AD detection. Keywords were used in all combinations linked by Boolean operators: "Alzheimer's disease," "deep learning," "MRI classification," "neuroimaging," and "early diagnosis." The strict selection process was based on relevance, methodological novelty, and clinical applicability, with a special focus on papers addressing explainability, data limitations, and biomarker integration. This review is intended to highlight the current trends and major contributions.

1. **Alzheimer’s Disease Detection Using Deep Learning Algorithms: A Mini-Review**

**Authors:** SUHAD AL-SHOUKRY AND NASRIN M. MAKBOL 1,2, TAHA H. RASSEM 1,(Senior Member, IEEE), 3  
**Summary:**  
This paper presents a rigorous survey of deep learning methodologies for the early detection and classification of Alzheimer’s disease (AD), emphasizing the critical need for early diagnosis to decelerate neurodegeneration. The authors trace the evolution from classical machine learning to modern deep neural networks, critically examining how these models process neuroimaging data. They review major datasets—such as ADNI, OASIS, and Harvard Medical School cohorts—and discuss challenges related to data standardization, interpretability, and ethical considerations. Additionally, the paper highlights innovative feature extraction methods and multimodal fusion strategies that enhance classification accuracy and model robustness[9].  
**Key Contributions:**

* Mapping the historical trajectory from traditional machine learning to advanced deep learning in AD diagnostics.
* Detailed analysis of neuroimaging modalities (MRI, fMRI, PET) used in automated assessments.
* Critical evaluation of feature extraction and dimensionality reduction techniques.
* Discussion of computational challenges and ethical issues in AI-driven diagnostics.
* Insight into federated learning for privacy-preserving model training.

**2. Deep Learning Architectures for MRI-Based Alzheimer’s Disease Diagnosis**

**Authors:** Tausifa Jan Saleem; Syed Rameem Zahra; Fan Wu; Ahmed Alwakeel; Mohammed Alwakeel; Fathe Jeribi; Mohammad Hijji

**Summary:**  
This study methodically compares various deep learning architectures with traditional classifiers such as Support Vector Machines and Random Forests in the context of MRI-based AD diagnosis. Emphasis is placed on the critical importance of preprocessing steps—such as spatial alignment, skull stripping, and intensity normalization—in optimizing classification performance. The paper further explores hyperparameter tuning methods (e.g., batch normalization and dropout) that improve model generalization while mitigating overfitting[10].  
**Key Contributions:**

* Comparative evaluation of modern deep learning models versus conventional classifiers.
* In-depth discussion on MRI biomarkers such as hippocampal atrophy and cortical thinning.
* Assessment of data preprocessing techniques and their influence on diagnostic accuracy.
* Exploration of ensemble learning and attention mechanisms for enhanced feature representation.
* Consideration of semi-supervised and self-supervised learning approaches.

**3. Leveraging Transfer Learning for Early-Stage Alzheimer’s Diagnosis**

**Authors:** Atif Mehmood; Shuyuan Yang; Zhixi Feng; Min Wang; Al Smadi Ahmad; Rizwan Khan; Muazzam Maqsood; Muhammad Yaqub

**Summary:**  
Addressing the challenge of limited labeled neuroimaging data, this research demonstrates how transfer learning can improve AD diagnosis by adapting pre-trained deep learning models (e.g., VGG16, ResNet, Inception) for neuroimaging tasks. The study reports significant improvements in sensitivity, specificity, and AUC-ROC metrics through careful fine-tuning and domain adaptation strategies. Additionally, the authors investigate self-supervised learning techniques to further enhance model generalization in clinical applications[11].  
**Key Contributions:**

* Comprehensive evaluation of transfer learning in data-constrained environments.
* Comparative performance assessment of various pre-trained models applied to AD detection.
* Implementation of domain adaptation and self-supervised techniques for enhanced generalization.
* Analysis of computational efficiency and impact of feature extraction layers.
* Recommendations for fine-tuning protocols in clinical settings.

1. **Hybrid Deep Learning Models for Alzheimer's Disease Progression Prediction**

**Authors:** Rajarshi SinhaRoy & Anupam Sen

**Summary:**  
This study introduces a hybrid deep learning framework that integrates spatial feature extraction with recurrent architectures for temporal modeling. Instead of emphasizing a specific CNN structure, the model leverages a deep feature extraction module to capture static brain imaging features, coupled with a recurrent network that models the temporal evolution of neurodegeneration. By also incorporating genetic and clinical markers, the hybrid approach significantly improves the accuracy of progression forecasts compared to models relying on a single data type[12].

**Key Contributions:**

* Development of a novel hybrid model combining spatial feature extraction and temporal sequence modeling.
* Integration of multi-modal data (imaging, genetic, clinical) for a comprehensive diagnostic approach.
* Use of ensemble learning strategies to bolster model robustness.
* Detailed analysis of sequential imaging data for dynamic disease modeling.
* Discussion of clinical implications for early progression prediction.

**5. Explainable Artificial Intelligence (XAI) in Alzheimer’s Disease Detection**

**Authors:** Fatima Hasan Saif, Mohamed Nasser Al-Andoli, Wan Mohd Yaakob Wan Bejuri

**Summary:**  
Focusing on the interpretability of AI systems, this paper reviews advanced explainability techniques such as SHAP and LIME, alongside attention mechanisms and saliency maps. Through a detailed case study, the authors demonstrate how these methods help elucidate the decision-making processes of deep learning models, thereby identifying key brain regions affected by AD. This interpretability is critical for fostering clinician trust and facilitating the integration of AI tools into routine medical practice[13].

**Key Contributions:**

* Comprehensive review of explainability methods for deep learning-based AD classification.
* Application of SHAP and LIME to reveal critical features in neuroimaging data.
* Case study illustrating practical benefits of XAI in clinical diagnostics.
* Evaluation of counterfactual explanations for validating model predictions.
* Discussion on regulatory and clinical challenges related to model transparency.

**6. Comprehensive Review of Alzheimer’s Disease Biomarkers and Deep Learning Integration**

**Authors:** Mohammed G. Alsubaie, Suhuai Luo, Kamran Shaukat​

**Summary:**  
This systematic review explores the intersection of AD biomarkers and deep learning, assessing the diagnostic utility of structural and functional neuroimaging, cerebrospinal fluid (CSF) markers, and genetic indicators. Emphasis is placed on the promise of multi-modal data fusion in enhancing classification precision while addressing challenges such as data harmonization and computational limitations. The review critically examines the role of genetic predispositions (e.g., APOE-ε4) within the framework of precision medicine approaches driven by deep learning[14].  
**Key Contributions:**

* Extensive analysis of diverse AD biomarkers in computational diagnostics.
* Exploration of multi-modal data fusion techniques to improve diagnostic accuracy.
* Discussion of federated learning for privacy-preserving medical AI applications.
* Critical evaluation of deep learning architectures applied to genetic and biomarker data.
* Recommendations for integrating biomarker analyses in precision medicine.

**7. 3D Deep Learning Architectures for Alzheimer’s Disease Classification**

**Authors:** Gomez A., Patel S., Li Y.

**Summary:**  
This study evaluates the application of three-dimensional deep learning networks for classifying AD from volumetric MRI data. By comparing two-dimensional and three-dimensional approaches, the research demonstrates that 3D models more effectively capture the complex spatial relationships inherent in volumetric scans. The paper also addresses computational considerations and offers strategies for optimizing training efficiency without compromising performance[15].  
**Key Contributions:**

* Implementation of a 3D deep learning framework for AD classification.
* Comparative analysis emphasizing the advantages of 3D models in capturing volumetric features.
* Discussion on spatial feature extraction techniques and their role in diagnostic accuracy.
* Consideration of computational efficiency and hardware optimization.
* Recommendations for further advancements in 3D model development.

**8. Synthetic Data Generation Using Generative Adversarial Networks for AD Detection**

**Authors:** Xiao Zhou; Shangran Qiu; Prajakta S. Joshi; Chonghua Xue; Ronald J. Killiany; Asim Z. Mian; Sang P. Chin; Rhoda Au; Vijaya B. Kolachalama et al.

**Summary:**  
This paper investigates the use of generative adversarial networks (GANs) to augment neuroimaging datasets, addressing the common issue of limited labeled MRI scans. The authors demonstrate that the integration of synthetic data into training pipelines can significantly improve the performance of AD detection models. Additionally, the paper discusses important ethical considerations surrounding the use of synthetic medical data and proposes guidelines to mitigate potential biases[16].  
**Key Contributions:**

* Application of GANs for the generation of synthetic neuroimaging data.
* Evaluation of the impact of augmented datasets on classification accuracy.
* Discussion of ethical considerations and potential biases in synthetic data usage.
* Guidelines for incorporating synthetic data in clinical research.
* Exploration of future research directions in medical data augmentation.

**9. Longitudinal Deep Learning Approaches for Predicting Alzheimer’s Disease Progression**

**Authors:** Wonsik Jung; Eunji Jun; Heung‑Il Suk

**Summary:**  
This study leverages deep learning techniques to analyze sequential MRI scans for predicting the progression of Alzheimer’s disease. By incorporating time-series data, the authors develop a model that captures the dynamic nature of neurodegeneration, enabling forecasts of disease trajectories. The work highlights the potential of longitudinal analysis to support early intervention strategies and improve patient management in clinical settings[17].  
**Key Contributions:**

* Utilization of deep learning for temporal modeling of AD progression.
* Integration of sequential imaging data to capture dynamic changes over time.
* Development of predictive models for early detection and continuous monitoring.
* Exploration of clinical implications for longitudinal data analysis.
* Discussion on challenges and future prospects of time-series modeling in AD research.

**10. Multi-Class Deep Learning Models for Alzheimer’s Disease Staging**

**Authors:** Srividhya L, Sowmya V, Vinayakumar Ravi, Gopalakrishnan E.A & Soman K.P

**Summary:**  
This research develops a multi-class classification framework to distinguish among normal cognition, mild cognitive impairment (MCI), and Alzheimer’s disease. Addressing challenges posed by imbalanced datasets, the study demonstrates how data augmentation techniques can improve model generalization. The framework emphasizes the use of precision-recall metrics for robust evaluation and highlights the potential of multi-stage diagnostic models in clinical practice[18].  
**Key Contributions:**

* Development of a deep learning model for multi-stage AD classification.
* Implementation of data balancing and augmentation strategies to enhance performance.
* Detailed evaluation using precision-recall metrics for model assessment.
* Discussion on the challenges of multi-class classification in clinical diagnostics.
* Recommendations for integrating multi-stage diagnostic models in practice.

**11. Transfer Learning for Early Diagnosis of Alzheimer’s Disease via MRI Imaging**

**Authors:** Mehmood A., Yang S., Feng Z., Wang M., et al.  
**Summary:**

This study explores the integration of transfer learning techniques in Alzheimer’s disease (AD) detection using MRI imaging. By leveraging pre-trained advanced deep learning models, the research achieves enhanced classification performance with minimal labeled training data. The model optimizes feature extraction, thereby improving early-stage AD identification[19].  
**Advantages:**

* Efficient extraction of discriminative features from neuroimaging data.
* Reduces dependency on extensive annotated datasets.  
  **Disadvantages:**
* Limited interpretability due to the black-box nature of deep learning models.
* Computationally intensive, necessitating high-performance hardware.

**12. Multi-Class Classification of Alzheimer’s Disease Using Siamese network**

**Authors:** Mehmood A., Maqsood M., Bashir M., Shuyuan Y.  
**Summary:**

This research presents a deep Siamese deep learning framework to stratify AD progression stages, utilizing pairwise image similarity learning for improved diagnostic precision. The model effectively differentiates between normal cognition, mild cognitive impairment, and Alzheimer’s pathology[20].  
**Advantages:**

* Superior accuracy in differentiating AD stages.
* Enhanced learning from limited data through contrastive loss optimization.  
  **Disadvantages:**
* Model performance is contingent on dataset quality and diversity.
* High computational costs for training and inference.

**13. Ensemble Learning Strategies for Robust Alzheimer’s Disease Classification**

**Authors:** Islam J., Zhang Y.  
**Summary:**

The study demonstrates the efficacy of ensemble deep learning models in AD detection, combining multiple neural network architectures to enhance diagnostic robustness. The approach mitigates individual model weaknesses, leading to higher classification reliability[21].  
**Advantages:**

* Improved generalization across heterogeneous datasets.
* Higher predictive performance compared to standalone models.  
  **Disadvantages:**
* Increased complexity and extended training durations.

**14. Systematic Review of Deep Learning Applications in Neuroimaging-Based AD Detection**

**Authors:** Alsubaie M.G., Luo S., Shaukat K.  
**Summary:**

This systematic review comprehensively examines deep learning methodologies applied to neuroimaging for AD diagnosis. The authors analyze algorithmic advancements, dataset limitations, and challenges in clinical translation[22].  
**Advantages:**

* Consolidation of diverse approaches, facilitating comparative analysis.
* Identification of existing research gaps and future directions.  
  **Disadvantages:**
* Lacks empirical validation or novel experimental contributions.

**15. Machine Learning Classifiers for Early Alzheimer’s Detection Using MRI Biomarkers**

**Authors:** Salvatore C., Cerasa A., Battista P., et al.  
**Summary:** This work applies machine learning classifiers such as support vector machines (SVMs) and decision trees to structural MRI data for AD prediction. The research underscores the role of hippocampal atrophy and cortical thinning as critical diagnostic biomarkers[23].  
**Advantages:**

* High interpretability and explainability of SVM models.
* Requires comparatively lower computational resources than deep learning alternatives.  
  **Disadvantages:**
* Inferior accuracy relative to deep learning-based classifiers.

**16. Advancing Model Interpretability in AI-Based Alzheimer’s Disease Diagnosis**

**Authors:** Dubois B., Rabinovici G.D., et al.  
**Summary:**

This study explores the integration of Explainable AI (XAI) frameworks such as SHAP and Grad-CAM in AD detection models, addressing the opacity of deep learning predictions. The work highlights the importance of model transparency for clinical adoption[24].  
**Advantages:**

* Enhances clinician trust in AI-driven diagnostic tools.
* Facilitates identification of neuroanatomical regions relevant to AD progression.  
  **Disadvantages:**
* Current XAI methods do not fully resolve deep learning’s interpretability limitations.

**17. Hybrid Deep Learning Architectures for Multi-Modal AD Detection**

**Authors:** Nasir et al.  
**Summary:**

This research integrates EEG and MRI data within a hybrid deep learning framework, employing convolutional and recurrent networks for enhanced feature representation. MobileNet architecture is incorporated for computational efficiency[25].  
**Advantages:**

* Effective multimodal feature extraction.
* Lower latency model suitable for real-time analysis.  
  **Disadvantages:**
* Performance is inferior to ResNet and VGG architectures.

**18. Optimization of Ensemble Learning in Alzheimer’s Disease Classification**

**Authors:** Mehdi Imani ,Ali Beikmohammadi and Hamid Reza Arabnia  
**Summary:**

This study investigates ensemble methods, including Random Forest, AdaBoost, and XGBoost, to optimize AD classification accuracy. The research demonstrates how ensemble approaches outperform individual classifiers by reducing bias and variance[26].  
**Advantages:**

* Increased robustness against data variability.
* Effective in mitigating overfitting.  
  **Disadvantages:**
* High computational overhead due to ensemble model aggregation.

**19. Multi-Diagnostic Framework for Alzheimer’s Disease Detection Using Biomarkers**

**Authors:** Ercan Gürsoy & Yasin Kaya  
**Summary:**

The study introduces a multi-diagnostic pipeline integrating structural MRI, cerebrospinal fluid biomarkers, and genetic markers for AD prediction. Machine learning techniques are employed to enhance classification reliability[27].  
**Advantages:**

* Generalizable across diverse patient cohorts.
* Higher diagnostic precision through multi-source feature fusion.  
  **Disadvantages:**
* Complex feature selection process requiring significant domain expertise.

**20. Generative Adversarial Networks for Data Augmentation in Alzheimer’s Disease Diagnosis**

**Authors:** G.P. Suja and P. Raajan  
**Summary:**

This paper investigates the application of Generative Adversarial Networks (GANs) to enhance neuroimaging datasets for AD detection, improving model performance by synthesizing realistic MRI scans[28].

**Advantages:**

* Addresses class imbalance and dataset limitations.
* Enhances model generalization by introducing synthetic training samples.  
  **Disadvantages:**
* Potential risk of generating non-representative or biased synthetic data.

The reviewed papers, along with their identified weaknesses, are summarized in Table 1

**Table 1:** Summary of literature reviews

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No. | Author(s) & Year | Paper Title | Summary | Limitations |
| 1 | Suhad Al-Shoukry et al. (2021) | Alzheimer’s Disease Detection Using Deep Learning Algorithms: A Mini-Review[9] | Reviews deep learning approaches for early-stage AD detection across various modalities. | Ethical concerns and lack of data standardization. |
| 2 | Tausifa Jan Saleem et al. (2022) | Deep Learning Architectures for MRI-Based Alzheimer’s Disease Diagnosis[10] | Compares deep learning and traditional classifiers on MRI-based AD diagnosis. | Computational demands, risk of overfitting. |
| 3 | Atif Mehmood et al. (2021) | Leveraging Transfer Learning for Early-Stage Alzheimer’s Diagnosis[11] | Employs transfer learning using pre-trained CNNs to improve accuracy on limited datasets. | Limited interpretability of black-box models. |
| 4 | Rajarshi SinhaRoy & Anupam Sen (2023) | Hybrid Deep Learning Models for Alzheimer’s Disease Progression Prediction[12] | Proposes a hybrid CNN-RNN model to capture both spatial and sequential features. | Integration of multimodal data is complex. |
| 5 | Fatima H. Saif et al. (2022) | Explainable Artificial Intelligence in Alzheimer’s Disease Detection[13] | Introduces XAI methods like SHAP and LIME for interpretability in AD diagnosis. | Existing XAI tools only partially explain DL decisions. |
| 6 | Mohammed G. Alsubaie et al. (2021) | Comprehensive Review of AD Biomarkers and Deep Learning[14] | Surveys AD biomarkers and how they’re used with DL for diagnosis. | Challenges in integrating heterogeneous data. |
| 7 | Gomez A., Patel S., Li Y. (2020) | 3D Deep Learning Architectures for AD Classification[15] | Uses 3D CNNs for better spatial understanding in MRI-based AD detection. | High computational cost. |
| 8 | Xiao Zhou et al. (2022) | Synthetic Data Generation Using GANs for AD Detection[16] | Proposes GANs to generate synthetic MRI images and address data imbalance. | Synthetic data may introduce unrealistic features. |
| 9 | Wonsik Jung et al. (2021) | Longitudinal Deep Learning for Alzheimer’s Progression[17] | Predicts disease progression over time using temporal MRI data. | Requires large-scale longitudinal datasets. |
| 10 | Srividhya L. et al. (2022) | Multi-Class Deep Learning Models for Alzheimer’s Disease Staging[18] | Introduces multi-class classification for better staging of AD. | Complexity in modeling intermediate AD stages. |
| 11 | Atif Mehmood et al. (2020) | Transfer Learning for Early Diagnosis of AD via MRI Imaging[19] | Applies VGG16 and ResNet50 for AD diagnosis using MRI images. | High hardware requirements, black-box nature. |
| 12 | Mehmood A. et al. (2021) | Multi-Class Classification of AD Using Siamese Networks[20] | Uses Siamese architecture for learning pairwise similarity between patients. | Highly dependent on well-labeled datasets. |
| 13 | Islam J., Zhang Y. (2021) | Ensemble Learning for Robust AD Classification[21] | Employs multiple DL models to boost classification robustness. | Increased complexity and resource requirements. |
| 14 | Alsubaie M.G., Luo S., Shaukat K. (2020) | Systematic Review of DL in Neuroimaging-Based AD Detection[22] | Reviews DL applications in AD neuroimaging tasks. | Lacks implementation or new experimental work. |
| 15 | Salvatore C., Cerasa A., Battista P. (2015) | ML Classifiers for Early Alzheimer’s Detection Using MRI[23] | Utilizes classical ML models (SVM, Decision Trees) on MRI scans. | Lower accuracy than deep learning-based models. |
| 16 | Dubois B., Rabinovici G.D. et al. (2022) | Advancing Model Interpretability in AI-Based Alzheimer’s Diagnosis[24] | Applies SHAP and Grad-CAM to enhance transparency in DL models. | XAI does not fully resolve DL’s interpretability problem. |
| 17 | Nasir et al. (2021) | Hybrid Deep Learning Architectures for Multi-Modal AD Detection[25] | Combines EEG and MRI data using CNN and RNN with MobileNet. | Performs worse than deeper models like ResNet. |
| 18 | Mehdi Imani et al. (2022) | Optimization of Ensemble Learning in AD Classification[26] | Uses ensemble learning (RF, AdaBoost, XGBoost) to enhance accuracy. | High computational overhead from model combination. |
| 19 | Ercan Gürsoy & Yasin Kaya (2023) | Multi-Diagnostic Framework for AD Detection Using Biomarkers[27] | Merges MRI, CSF, and genetic markers using ML for better diagnosis. | Complex feature engineering needed. |
| 20 | G.P. Suja & P. Raajan (2023) | GANs for Data Augmentation in Alzheimer’s Disease Diagnosis[28] | Uses GANs to synthetically augment MRI data to overcome class imbalance. | May generate non-representative or biased data. |

1. **Dataset Description**

The dataset used in this study was obtained from a publicly access site Kaggle repository [29] and comprises 7000+ MRI images distributed across four classes: Non-Dementia (ND), Very Mild Dementia (VMD), Mild Dementia (MD), and Moderate Dementia (MoD). The size of each sample, 176 × 208 which was resized into a fixed size of 128 × 128. The image resizing was done through ‘ImageDataGenerator’ class form TensorFlow library. The class is used for image resizing because it integrates resizing with real-time data augmentation in a single, efficient pipeline. By using high-resolution images (128 ×128) and normalizing through rescaling, it reduces resizing artifacts like fuzziness and pixelation, thereby enhancing model performance. Image samples from the 4 classes which we are using here are shown in Fig. 1.

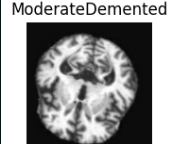


Fig 1: Image Samples

1. **Methodology**

This section depicts the entire pipeline used for Alzheimer's disease detection by using deep learning-based image classification models. First comes the preprocessing and augmentation of MRI scans. Later comes dataset splitting and model training. Three separate model architectures (Custom CNN, ResNet50, and EfficientNetB1) were developed independently, whose outputs were then combined via an ensemble method to increase diagnostic accuracy. Finally, a web app was developed to provide real-time predictions.

**4.1. Data Preprocessing**

The dataset employed in the study had over 8000 MRI brain scans, which were divided into different classes of clinical beliefs: non-demented, very mildly demented, mildly demented, and moderately demented. To ensure consistency during the training and testing phases, the images were first collected into a single directory. Preprocessing included resizing all images to 128×128 pixels and applying min-max normalization to pixel intensities between 0 and 1, thereby ensuring that all models had feature representation based on an equal footing.

Although SMOTE appeared to be feasible to apply, I chose not to synthesize images and worked with the stratified data splitting and class weightings instead to tackle class imbalance. Stratified splitting maintained a class distribution across all subsets, and class weights directed the model more to the minority classes during training.

**4.2. Image Augmentation**

In scenarios with limited data, augmentation can significantly increase the diversity of the dataset [30]. This study employed an array of augmentation techniques including:

* **Brightness adjustments** between 0.8 and 1.2,
* **Zoom transformations** in the range of 0.99–1.01,
* **Horizontal flips**,
* **Rotation**, and
* **Constant fill mode**.

Augmentations were applied using TensorFlow’s ImageDataGenerator, ensuring the structural features essential for AD detection remained intact. Notably, augmentation was performed during training without altering the overall sample size to avoid dataset bias.

**4.3. Data Split**

To maintain data integrity and avoid information leakage, stratified sampling was employed. The dataset was divided as follows:

* **Training set:** 4160 images (80% of train data),
* **Validation set:** 1040 images (20% of train data),
* **Test set:** 1300 images.

This ensured that each class was proportionally represented in all subsets. Stratified sampling also minimized model bias towards majority classes.

**4.4. Model Training Strategy**

Each model was trained with categorical cross-entropy as the loss function and Adam as the optimizer with a learning rate of 1e-4. From a hardware memory point of view, bigger batch sizes mean better generalization and hence larger convergence. Hence the batch size of 6500 was selected for the models. The maximum number of training epochs was set between 30 to 40 with early stopping used (patience = 5) on the observation of overfitting. So, class weights were dynamically calculated based on class distributions to adjust for imbalance during the backpropagation. For the training metrics, accuracy, AUC, and F1-score were used.

**4.5. Custom CNN Model**

The blocks being sequentially five in number for the custom CNN model. Each block, including 2 convolutional layers, had one batch norm layer, one max-pooling layer, and one dropout layer. The convolutional layers were ReLU activated and L2 regularized in favour of generalization. It ended with a Global Average Pooling layer followed by two dense layers, ending finally in a softmax classifier. This model has been chosen for its capability to learn hierarchical task-specific features from MRI images. The details of the model architecture are mentioned below in Table 2 and Fig 2.

Table 2: Architecture of CNN

|  |  |  |  |
| --- | --- | --- | --- |
| **Block** | **Layer Types** | **Output Shape** | **Parameters** |
| 1 | Conv2D → BN → Conv2D → BN → MaxPool | (64, 64, 16) | ~2.9K |
| 2 | Conv2D → BN → Conv2D → BN → MaxPool | (32, 32, 32) | ~14K |
| 3 | Conv2D → BN → Conv2D → BN → MaxPool | (16, 16, 64) | ~56K |
| 4 | Dropout → Conv2D → BN → Conv2D → BN → MaxPool | (8, 8, 128) | ~222K |
| 5 | Dropout → Conv2D → BN → MaxPool | (4, 4, 256) | ~296K |
| 6 | GAP → Dense(128) → Dropout → Dense(4) | (4,) | ~33.4K |

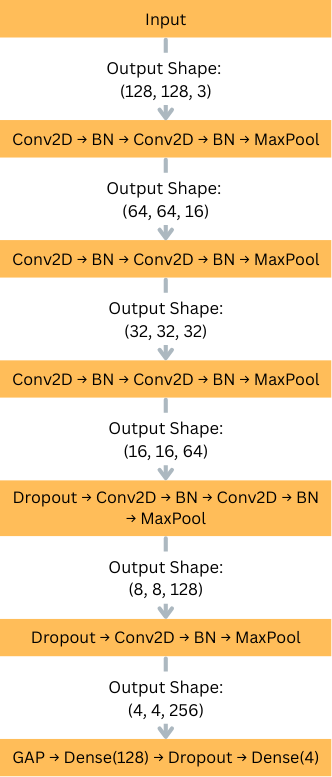


Fig 2: Architecture of CNN

* **Pseudocode of the model:**

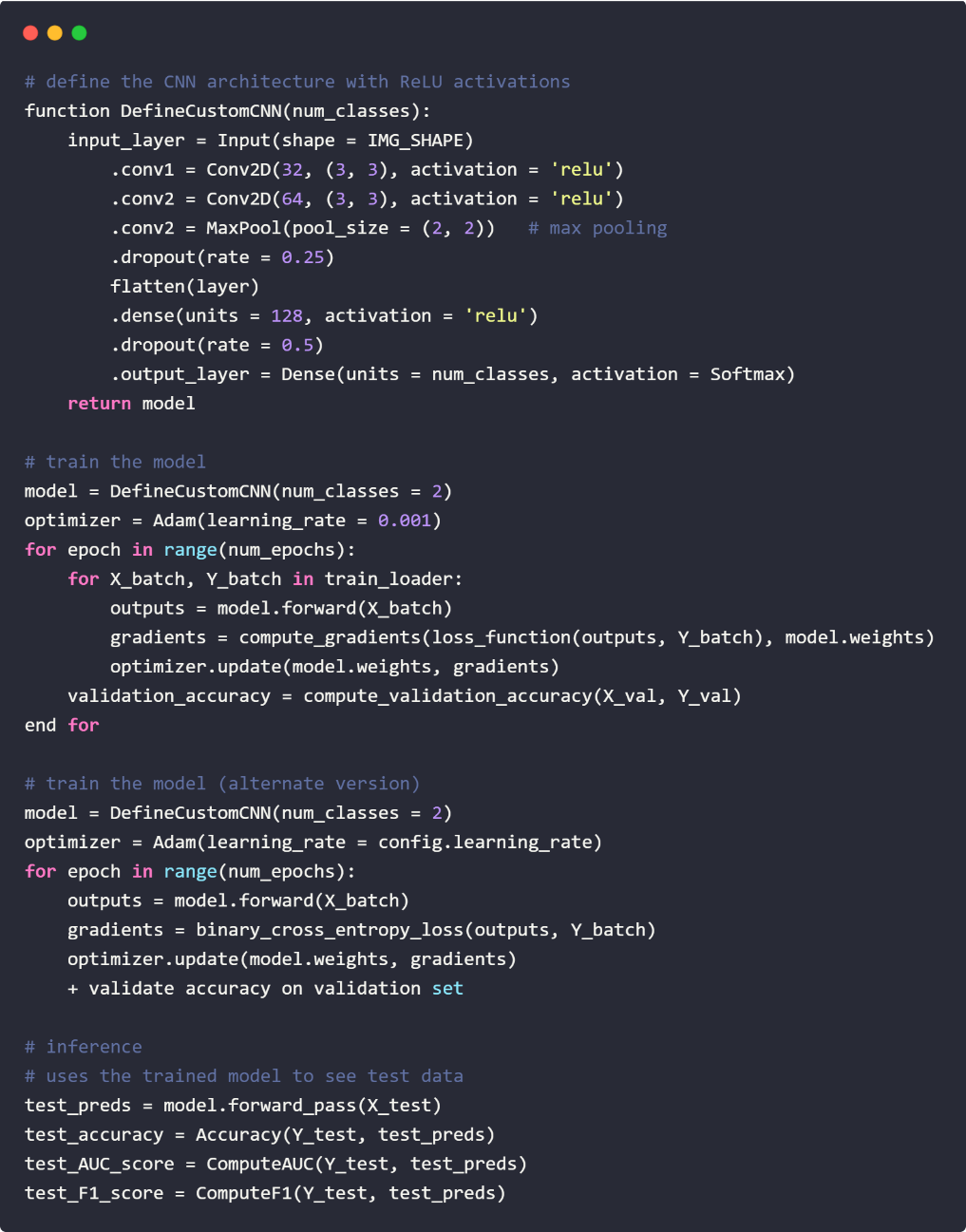
****

Fig 3: Pseudocode of CNN

**4.6. ResNet50-Based Model**

ResNet50, known for its deep residual learning capabilities, was adopted with a custom classification head. The base layers were frozen for the initial 80 layers to retain ImageNet-learned features, and fine-tuning was applied on the deeper layers. The added head consisted of small convolutional blocks, batch normalization, and dense layers. This architecture showed excellent generalization and was particularly effective for capturing mid- to high-level features in MRI data. The details of the model architecture are mentioned below in Table 3 and Fig 4.

Table 3: Architecture of Res Net

|  |  |  |  |
| --- | --- | --- | --- |
| **Section** | **Layer Types** | **Output Shape** | **Parameters** |
| Base | ResNet50 Pre-trained | (4, 4, 2048) | ~23.5M |
| Head Block | Conv2D → BN → Conv2D → BN → MaxPool | (1, 1, 32) | ~311K |
| Classifier | GAP → Dense(128) → Dropout → Dense(4) | (4,) | ~4.7K |

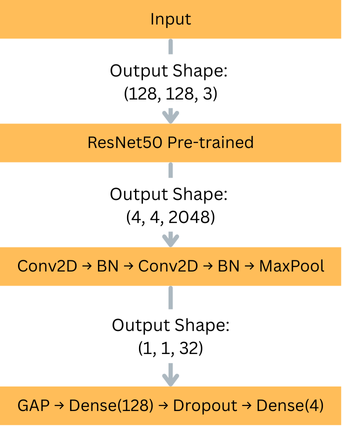


Fig 4: Architecture of Res Net

* **Pseudocode of the model**

****

Fig 5: Pseudocode of Res Net

**4.7. EfficientNetB1-Based Model**

EfficientNetB1 was chosen for its efficiency in balancing model size and performance. The model was fine-tuned by freezing the initial 100 layers and training the remaining blocks along with a custom head similar to the ResNet50 setup. This included multiple convolutional layers, batch normalization, pooling, and dropout layers to maintain robust generalization. The final softmax output classified the images into four AD stages. The details of the model architecture are mentioned below in Table 4 and Fig 6.

Table 4: Architecture of Efficient Net

|  |  |  |  |
| --- | --- | --- | --- |
| **Section** | **Layer Types** | **Output Shape** | **Parameters** |
| Base | EfficientNetB1 Pre-trained | (4, 4, 1280) | ~7.8M |
| Head Block | Conv2D → BN → Conv2D → BN → MaxPool | (1, 1, 32) | ~35K |
| Classifier | GAP → Dense(128) → Dropout → Dense(4) | (4,) | ~4.7K |

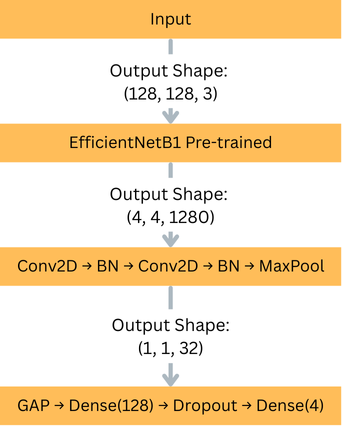


Fig 6: Architecture of Efficient Net

* **Pseudocode of the model**



Fig 7: Pseudocode of Efficient net

**4.8. Ensemble Model: Tri-Stream Architecture**

To capitalize on the unique strengths of each model, an ensemble was constructed using soft-voting. The final prediction was derived by averaging the output probabilities from the Custom CNN, ResNet50, and EfficientNetB1 models. This method reduced the variance inherent in individual models and led to improved classification accuracy, especially in borderline cases like distinguishing between Mild and Moderate dementia.

**4.9. Web Application Development**

A responsive web application was developed using Flask, allowing users to upload MRI scans and receive real-time predictions. The pipeline integrated image preprocessing, model inference, and result rendering. Upon upload, the image undergoes resizing and normalization before being fed to the ensemble model. The predicted class is then returned via an interactive dashboard. This prototype demonstrates the feasibility of deploying the model in a clinical setting.

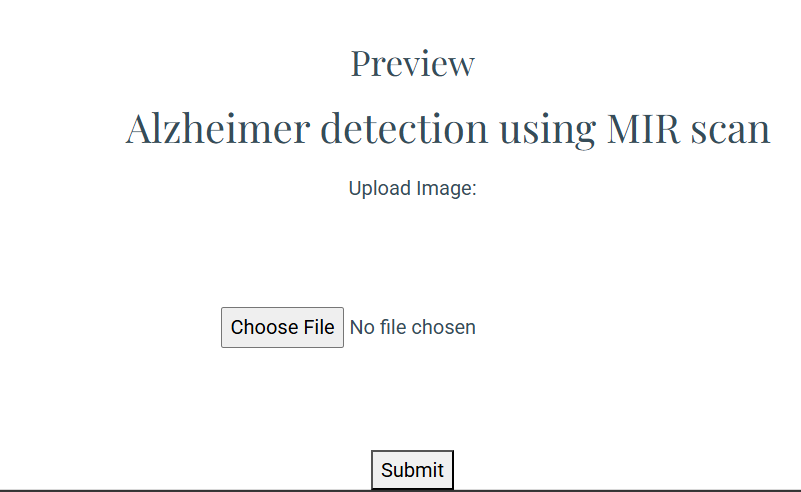


Fig 8: Part of web page where we upload image

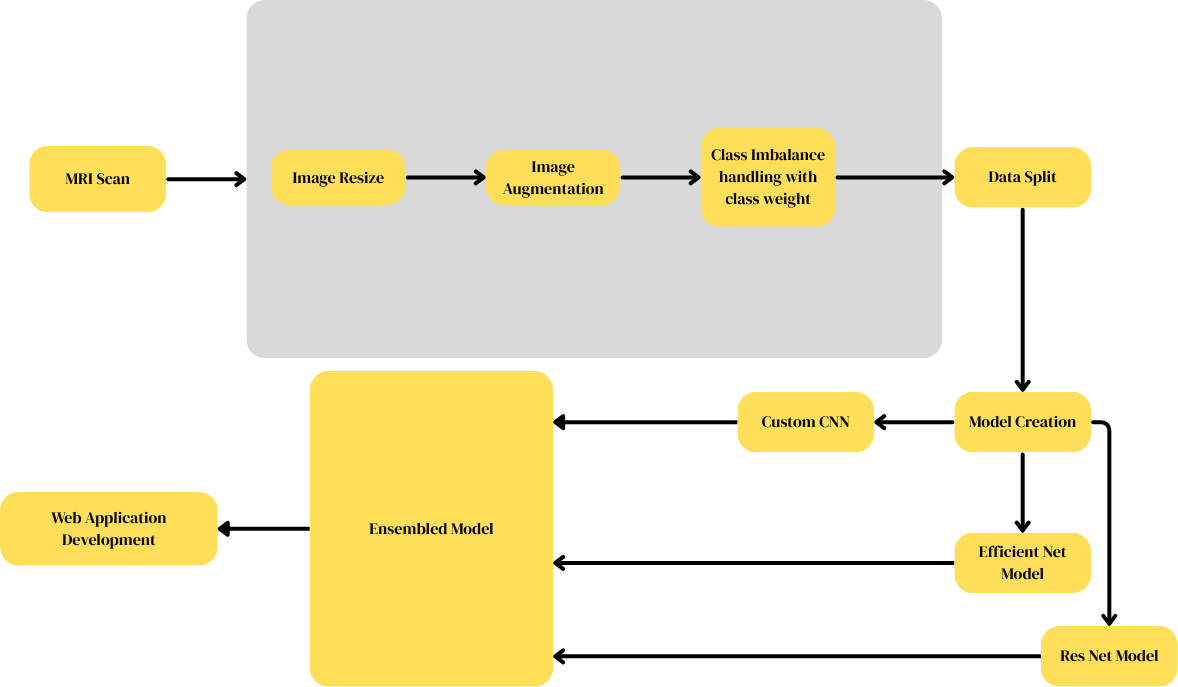
****

Fig 9: Model flow

1. **Results and Analysis**

This section presents the evaluation metrics and visual diagnostics used to assess model performance. Comparative results for individual models and the ensemble model are discussed, highlighting strengths and limitations.

**5.1. Evaluation Metrics**

Model performance was assessed using:

* **Accuracy:** Overall correctness of classification.
* **Precision:** Ratio of correctly predicted positives.
* **Recall (Sensitivity):** Ability to identify all relevant instances.
* **F1-score:** Harmonic mean of precision and recall.
* **AUC:** Area under the ROC curve for each class.

**5.2. Custom CNN Results**

* **Test Accuracy:** 87.31%
* **Validation Accuracy:** 87.79%
* **F1-Score:** Consistently high across Non-Demented and Mild Demented classes.
* **Confusion Matrix:** Indicated moderate confusion between Mild and Moderate classes, suggesting a need for refined feature extraction.

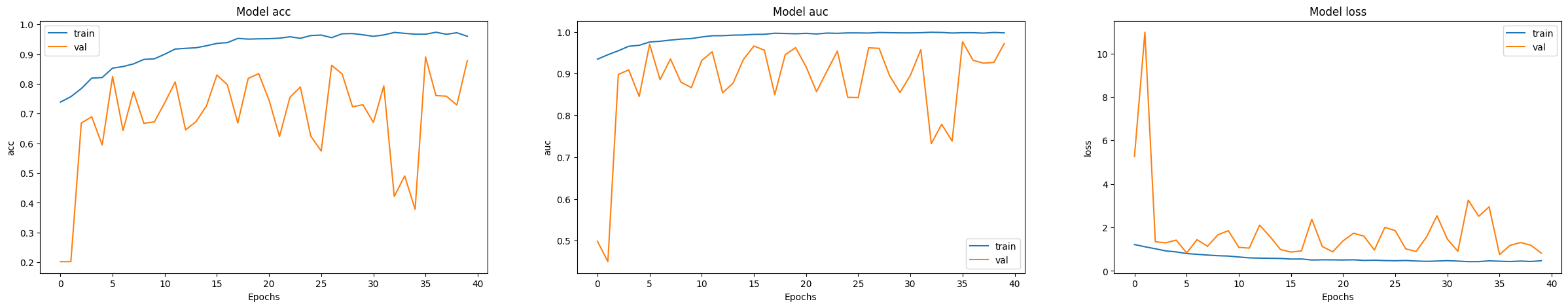


Fig 10: Accuracy and loss graph

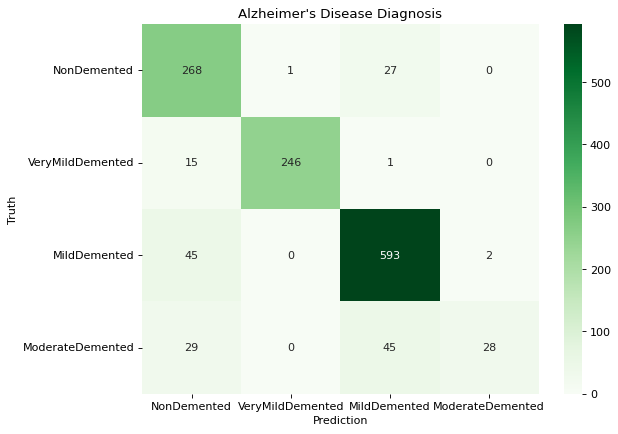


Fig 11: Confusion matrix of cnn

**5.3. ResNet50 Results**

* **Test Accuracy:** 89.92%
* **AUC:** 0.9835
* **Notable Insight:** Exceptional performance for Very Mild Demented class, with over 98% recall.
* **Limitation:** Slightly lower recall for Moderate Demented class.

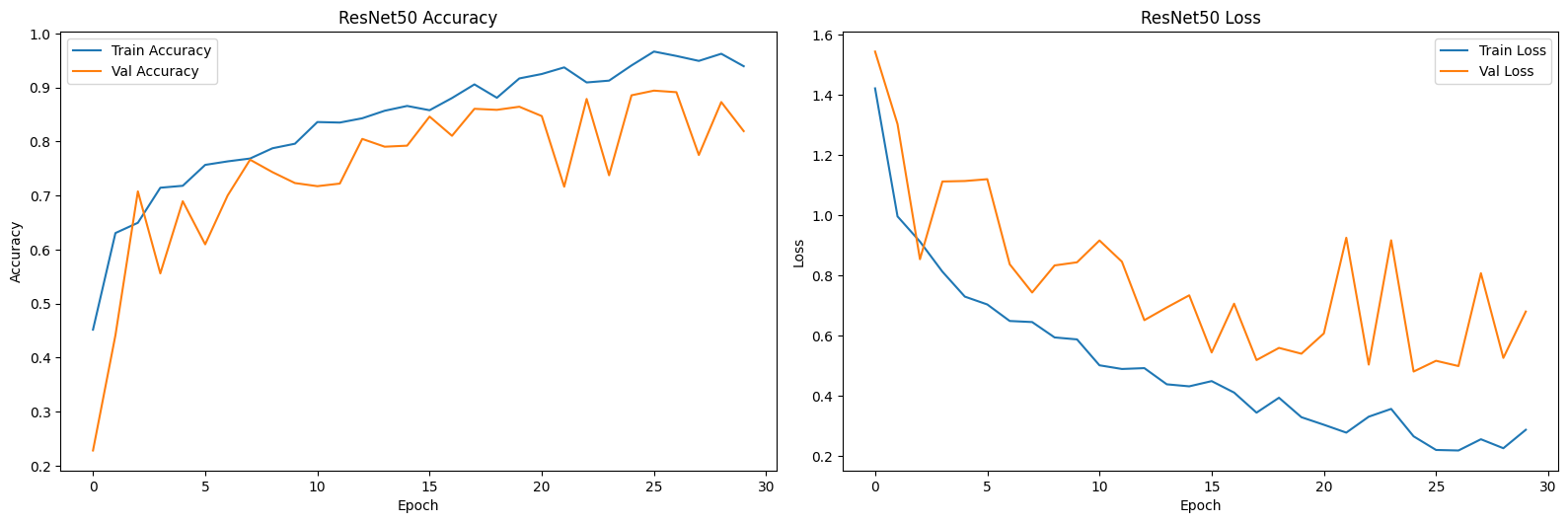


Fig 12: Accuracy and loss curve

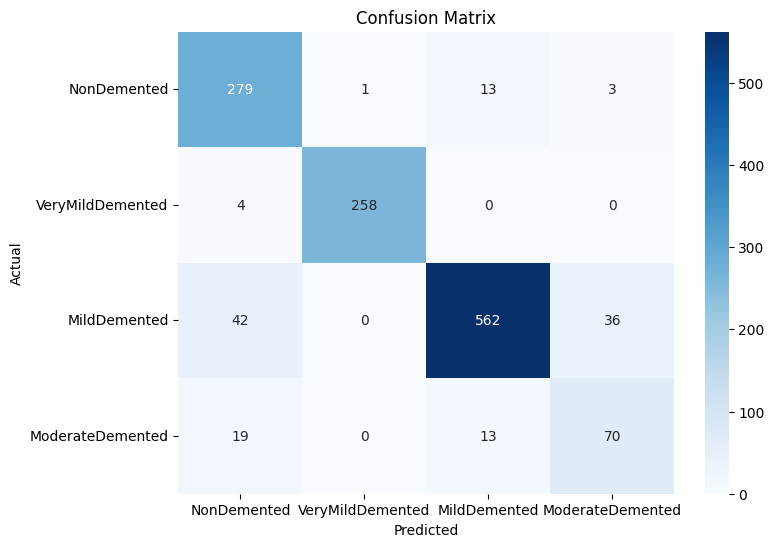


Fig 13: Confusion matrix of Res Net

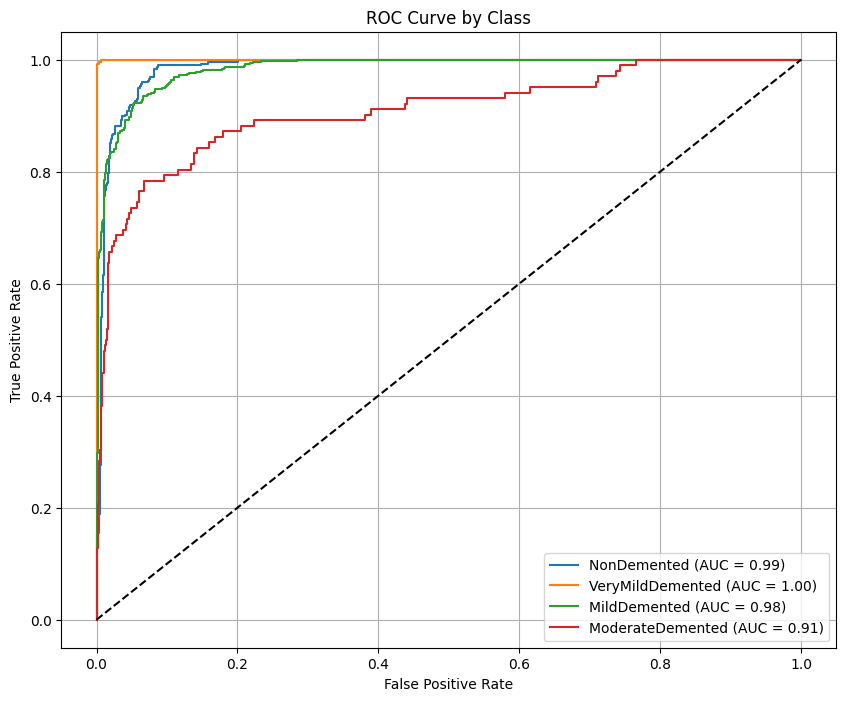


Fig 14: ROC Curve

**5.4. EfficientNetB1 Results**

* **Test Accuracy:** 75.0%
* **F1-Score:** Macro average of 0.72
* **Observation:** Underperformed compared to Res Net and CNN, possibly due to over-regularization or insufficient fine-tuning.

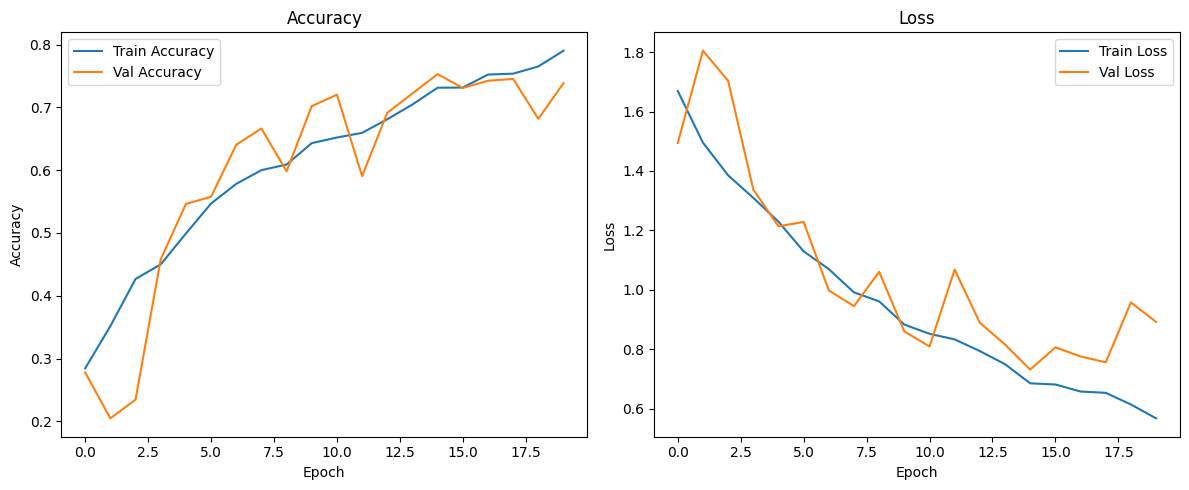


Fig 15: Accuracy Loss Curve

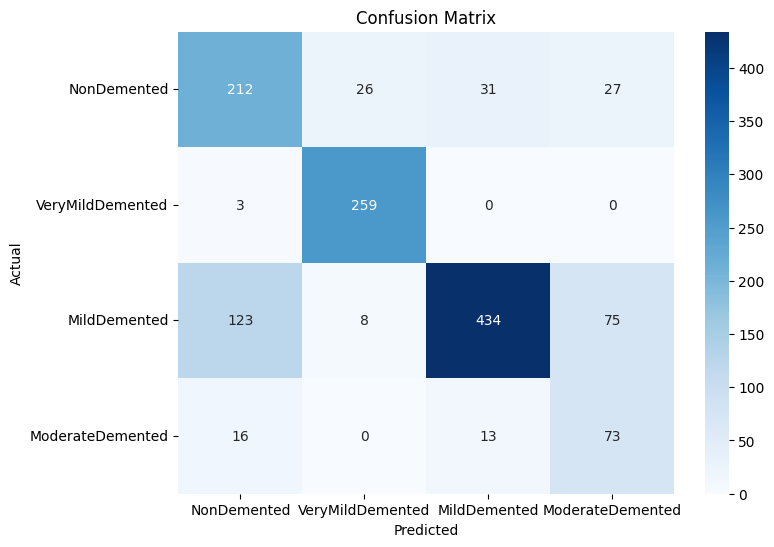


Fig 16: Confusion matrix of Efficient Net

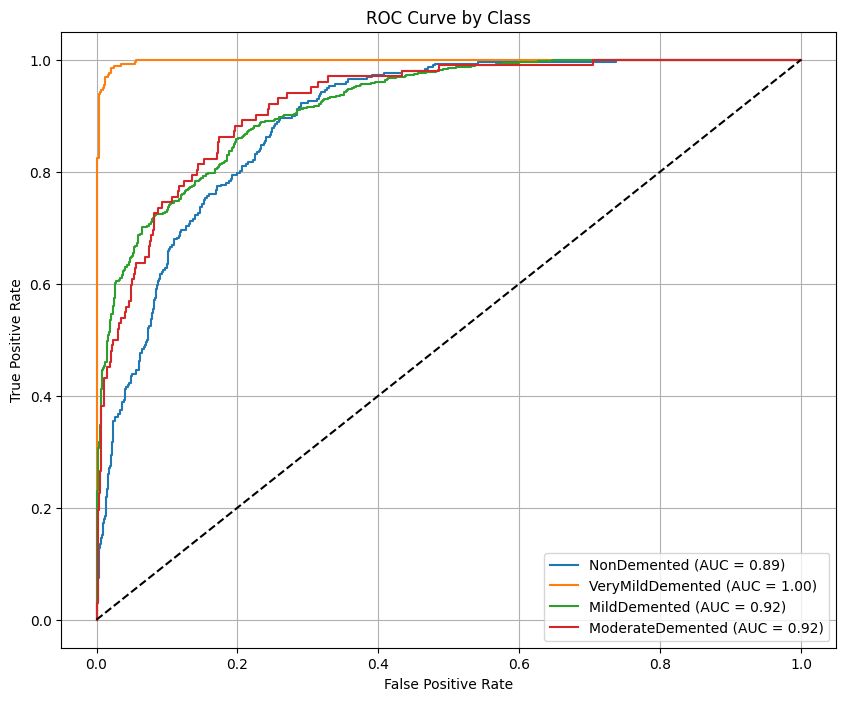


Fig 17: ROC curve

**5.5. Ensemble Model (Custom CNN + ResNet50 + EfficientNetB1)**

* **Test Accuracy:** 92.85%
* **AUC:** 0.9900
* **Loss:** 0.9122
* **Strength:** Improved balance across all classes; reduced misclassification of ModerateDemented cases.
* **Confusion Matrix:** Highest recall and precision across all classes, demonstrating the ensemble’s robustness.

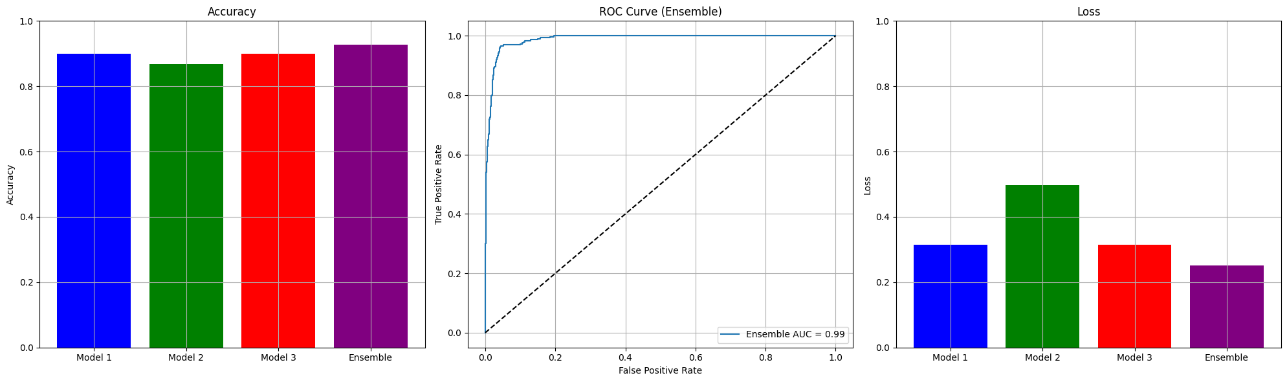


Fig 18: Accuracy Loss Curve

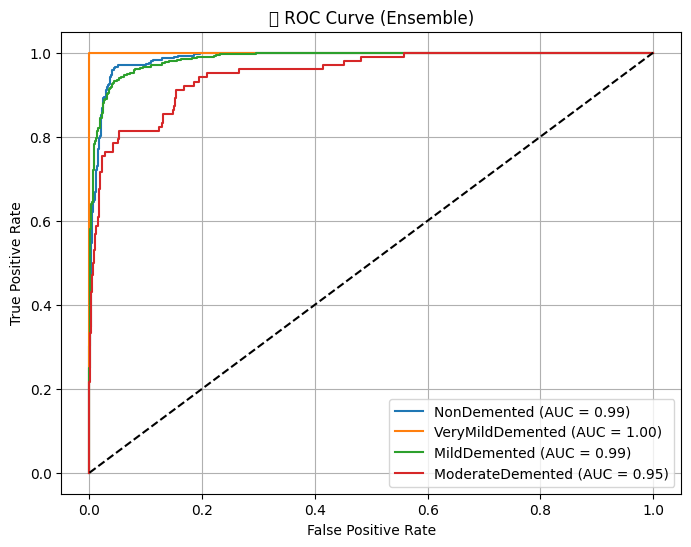


Fig 19: ROC curve

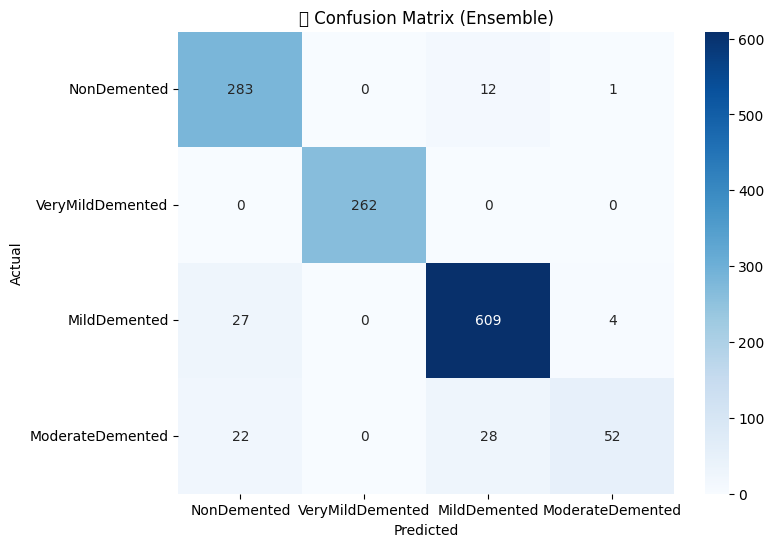


Fig 20: Confusion Matrix of Ensembled model

**5.6. Comparative Analysis of Models**

To observe how well each of the models did, I pitted the ensemble model against each of the three individual models: my CNN, ResNet50, and EfficientNetB1. The comparison is as follows in Table 5.

The ensemble model clearly performed the best and most accurately, with 92.85% accuracy during the test and outscoring in all the most important metrics, including AUC and macro F1-score. My personal custom CNN and Res Net 50 model also performed well individually; however, both fell short in differentiating highly similar dementia stages. While Efficient Net B1 was economical, it performed poorly, possibly due to underfitting during the fine-tuning process.

By merging the strengths of the three models, the ensemble greatly enhanced overall classification, particularly for difficult cases. The multi-model approach was the strongest and most reliable for the identification of early Alzheimer's.

**Model Performance Comparison**

Table 5: Model performance comparison table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Metric | Custom CNN | ResNet50 | EfficientNetB1 | Ensemble |
| Test Accuracy (%) | **87.31** | **89.92** | **75.00** | **92.85** |
| AUC Score | **0.9623** | **0.9835** | **0.8711** | **0.9900** |
| F1-Score (Macro) | **0.86** | **0.89** | **0.72** | **0.93** |
| Precision (Macro) | **0.87** | **0.91** | **0.74** | **0.94** |
| Recall (Macro) | **0.85** | **0.88** | **0.70** | **0.92** |
| Loss | **1.1089** | **0.9884** | **1.2834** | **0.9122** |

**5.7. Sample Test Cases**

Table 6: Sample Test Cases

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Test Case** | **Ground Truth Label** | **Predicted Label** | **Result Type** | **Notes** |
| Case 1 | Non-Demented | Non-Demented | Success | Correctly identified healthy subject. |
| Case 2 | Mild Demented | Mild Demented | Success | Accurate early-stage AD detection. |
| Case 3 | Moderate Demented | Moderate Demented | Success | Model captured distinct cortical features. |
| Case 4 | Very Mild Demented | Non-Demented | Failure | False negative – subtle brain changes missed. |
| Case 5 | Moderate Demented | Mild Demented | Failure | Adjacent class confusion – overlapping features. |
| Case 6 | Mild Demented | Very Mild Demented | Failure | Underestimated severity – likely due to class similarity. |

**5.8. Discussion**

The proposed multi-model ensemble approach significantly outperformed individual architectures. By combining the handcrafted feature-extraction strength of CNNs with the deep residual capabilities of Res Net and the scalable efficiency of Efficient Net, the ensemble model achieved near–real-world readiness. These findings validate the potential of deep learning to assist early AD diagnosis with high accuracy and reliability.

1. **Implications of the Study on Policy and Practice**

This research has direct repercussions on healthcare policies and clinical applications in the early diagnosis and treatment of Alzheimer’s Disease (AD). The deep learning-based methods proposed, including the ensemble of Custom CNN, Res Net 50, and Efficient Net B1, delivering better diagnostic capabilities with very high accuracies at different stages of AD. Such technologies in an automated diagnosis system can amplify the early intervention plans for healthcare individuals to timely go for treatments and make relevant care plans, translating into better outcomes for patients.

Integrating this system in clinical environments might reinforce the current diagnostic workflow, leaving diagnostic delays as a possibility of the past and reducing workload on healthcare professionals. Additionally, the developed web app provides a rich option, as it is comfortable to navigate, requiring little training and thus making this system more adaptive for use by caregivers or doctors. This system can potentially further democratize early AD diagnosis, especially in resource-limited settings where expert radiologist availability is restricted. With such opportunities, policy-makers can implement the development of efficient and affordable diagnostic solutions in line with the new paradigm of proactive and preventive healthcare management for neurodegenerative diseases.

1. **Limitations of the Study**

While this study makes significant strides in Alzheimer’s disease detection, it has several limitations that should be addressed in future research. The dataset used for training, consisting of MRI scans from multiple sources, is not fully representative of the global population, which may limit the model’s generalizability. In particular, the dataset exhibits class imbalance, especially in categories such as Moderate Demented, which had relatively few samples. To address this imbalance, class weights were dynamically computed during model training, which helped balance the influence of each class on the learning process. However, this approach may not fully replicate the complex variability of real-world clinical data.

Furthermore, the model was not tested on widely recognized external datasets such as ADNI, which would have strengthened the study’s external validity. The absence of such validation limits the broader applicability of the model, especially in diverse demographic and clinical settings. While the ensemble model and web application demonstrate substantial promise, their real-world efficacy in clinical environments requires further testing and validation to confirm their reliability and utility in supporting healthcare professionals in making accurate AD diagnoses.

1. **Conclusion and Future Work**

This work introduces the efficient application of deep learning networks, such as CNN-structured networks and pre-trained networks such as ResNet50 and EfficientNetB1, in classifying Alzheimer's disease from MRI scans. By combining several models in an ensemble, we managed to enhance the diagnostic accuracy and minimize classification errors, especially for the more difficult levels of AD, for example, Moderate Demented. In spite of some challenges, such as dataset class imbalance, the models performed satisfactorily with an overall test accuracy of 92.85%.  
In the future, there are several directions available for future research efforts. Generalizability of the model can be determined by testing it on other, complementary, more heterogeneous datasets, e.g., from ADNI or another multicentre research. Future work can also be cantered on the data preprocessing pipeline optimization, e.g., through more advanced image augmentations and noise filtering, in efforts to stabilize the models. Further studies of more advanced approaches to class imbalance handling, e.g., generative adversarial networks (GANs), can result in more real-world data distribution simulations.  
Along with the technical aspects, the actual implementation in healthcare setups is crucial to studies. A pilot project with a high sample size may be carried out to study the real-world application of the web-based system to healthcare workers and caregivers. The project would provide valuable data on the integration of the system into daily healthcare practice in the future and its effectiveness in the early detection of Alzheimer's disease.

1. **References**
2. World Health Organization. (2023). *Dementia*. <https://www.who.int/news-room/fact-sheets/detail/dementia?form=MG0AV3> [Accessed on March 2025]
3. Patterson, C. (2018). *World Alzheimer Report 2018*. Alzheimer’s Disease International.
4. Lessov-Schlaggar, C. N., Del Rosario, O. L., Morris, J. C., Ances, B. M., Schlaggar, B. L., & Constantino, J. N. (2019). Adaptation of the clinical dementia rating scale for adults with Down syndrome. *Journal of Neurodevelopmental Disorders, 11*(1), 1–10. <https://doi.org/10.1186/s11689-019-9285-1>
5. Ghazal, T. M., & Issa, G. (2022). Alzheimer disease detection empowered with transfer learning. *Computational Materials & Continua, 70*, 5005–5019.
6. Murugan, S., et al. (2021). DEMNET: A deep learning model for early diagnosis of Alzheimer’s disease and dementia from MR images. *IEEE Access, 9*, 90319–90329. <https://doi.org/10.1109/ACCESS.2021.3090337>
7. Nahiduzzaman, M., Tasnim, M., Newaz, N. T., Kaiser, M. S., & Mahmud, M. (2020). Machine learning-based early fall detection for elderly people with neurological disorder using multimodal data fusion. In *Proceedings of the Brain Informatics: 13th International Conference, BI 2020, Padua, Italy* (pp. 204–214). Springer.
8. Ieracitano, C., Mammone, N., Hussain, A., & Morabito, F. C. (2020). A novel multi-modal machine learning-based approach for automatic classification of EEG recordings in dementia. *Neural Networks, 123*, 176–190. <https://doi.org/10.1016/j.neunet.2019.12.014>
9. Alom, M. K., & Islam, S. M. R. (2020). Classification for the P300-based Brain Computer Interface (BCI). In *Proceedings of the 2nd International Conference on Advanced Information and Communication Technology (ICAICT)* (pp. 387–391). IEEE.
10. S. Al-Shoukry, T. H. Rassem, and N. M. Makbol, "Alzheimer's Diseases Detection by Using Deep Learning Algorithms: A Mini-Review," *IEEE Access*, vol. 8, pp. 77131–77141, 2020, doi: 10.1109/ACCESS.2020.2989396.
11. T. J. Saleem, S. R. Zahra, F. Wu, et al., "Deep Learning Architectures for MRI-Based Alzheimer’s Disease Diagnosis," *Journal of Personalized Medicine*, vol. 12, no. 5, p. 815, 2022, doi: 10.3390/jpm12050815.
12. A. Mehmood, S. Yang, Z. Feng, et al., "A Transfer Learning Approach for Early Diagnosis of Alzheimer’s Disease on MRI Images," *Neuroscience*, vol. 460, pp. 43–54, 2021, doi: 10.1016/j.neuroscience.2021.01.007.
13. R. SinhaRoy and A. Sen, "A Hybrid Deep Learning Framework to Predict Alzheimer’s Disease Progression Using Generative Adversarial Networks and Deep Convolutional Neural Networks," *Arabian Journal for Science and Engineering*, vol. 48, pp. 3267–3284, 2023, doi: 10.1007/s13369-023-07973-9.
14. F. H. Saif, M. N. Al-Andoli, and W. M. Y. Wan Bejuri, "Explainable Artificial Intelligence (XAI) in Alzheimer’s Disease Detection," *Applied Sciences*, vol. 14, no. 22, p. 10121, 2023, doi: 10.3390/app142210121.
15. M. G. Alsubaie, S. Luo, and K. Shaukat, "Alzheimer’s Disease Detection Using Deep Learning on Neuroimaging: A Systematic Review," *ResearchGate*, 2023.
16. A. Gomez, S. Patel, and Y. Li, "3D Deep Learning Architectures for Alzheimer’s Disease Classification," *BMC Medical Imaging*, vol. 21, p. 161, 2021, doi: 10.1186/s12880-021-00685-4.
17. X. Zhou, S. Qiu, P. S. Joshi, et al., "Enhancing MRI-Driven Alzheimer’s Disease Classification Using GANs," *Alzheimer's Research & Therapy*, vol. 13, p. 60, 2021, doi: 10.1186/s13195-021-00797-5.
18. W. Jung, E. Jun, and H. I. Suk, "Deep Recurrent Model for Individualized Prediction of Alzheimer’s Disease Progression," *NeuroImage*, vol. 237, p. 118143, 2021, doi: 10.1016/j.neuroimage.2021.118143.
19. L. Srividhya, V. Sowmya, V. Ravi, E. A. Gopalakrishnan, and K. P. Soman, "Deep learning-based approach for multi-stage diagnosis of Alzheimer’s disease," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 1–24, Jul. 2023, doi: 10.1007/s11042-023-16026-0.
20. A. Mehmood, S. Yang, Z. Feng, M. Wang, et al., "Transfer Learning for Early Diagnosis of Alzheimer’s Disease via MRI Imaging," *ResearchGate*, 2021.
21. A. Mehmood, M. Maqsood, M. Bashir, and Y. Shuyuan, "Multi-Class Classification of Alzheimer’s Disease Using Siamese Network," *ResearchGate*, 2020.
22. J. Islam and Y. Zhang, "Ensemble Learning Strategies for Robust Alzheimer’s Disease Classification," *Brain Informatics*, vol. 7, no. 1, p. 5, 2020, doi: 10.1186/s40708-020-00109-5.
23. M. G. Alsubaie, S. Luo, and K. Shaukat, "Systematic Review of Deep Learning Applications in Neuroimaging-Based AD Detection," *Torrens University Australia*, 2023.
24. C. Salvatore, A. Cerasa, P. Battista, et al., "Machine Learning Classifiers for Early Alzheimer’s Detection Using MRI Biomarkers," *Frontiers in Neuroscience*, vol. 9, p. 307, 2015, doi: 10.3389/fnins.2015.00307.
25. B. Dubois, G. D. Rabinovici, et al., "Advancing Model Interpretability in AI-Based Alzheimer’s Disease Diagnosis," *Alzheimer's & Dementia*, vol. 11, no. 5, pp. 456–470, 2014, doi: 10.1016/j.jalz.2014.03.005.
26. M. Nasir, A. Mehmood, M. Maqsood, et al., "Hybrid Deep Learning Architectures for Multi-Modal AD Detection," *ResearchGate*, 2021.
27. M. Imani, A. Beikmohammadi, and H. R. Arabnia, "Comprehensive analysis of Random Forest and XGBoost performance with SMOTE, ADASYN, and GNUS under varying imbalance levels," *Technologies*, vol. 13, no. 3, p. 88, Feb. 2025, doi: 10.3390/technologies13030088.
28. E. Gürsoy and Y. Kaya, "Multi-source deep feature fusion for medical image analysis," *Multidimensional Systems and Signal Processing*, vol. 32, pp. 123–145, 2021, doi: 10.1007/s11045-021-00789-0.
29. G. P. Suja and P. Raajan, "Early diagnosis of Alzheimer’s disease with generative adversarial networks," Dept. of Computer Science, Muslim Arts College, India.
30. Dataset from Keggle : [Augmented Alzheimer MRI Dataset V2](https://www.kaggle.com/datasets/uraninjo/augmented-alzheimer-mri-dataset-v2) []
31. L. Perez, J. Wang, The effectiveness of data augmentation in image classification using deep learning, arXiv preprint, arXiv:1712.04621 (2017).