ENHANCING THE EFFICIENCY OF DIAGNOSING AND MANAGING MENTAL DISORDERS USING MACHINE LEARNING

Dilshan G.A.M.
IT21180934

BSc. (Hons) degree in Information Technology
Specializing in Data Science

Department of Computer Science
Sri Lanka Institute of Information Technology
Sri Lanka

April 2025

ENHANCING THE EFFICIENCY OF DIAGNOSING AND MANAGING MENTAL DISORDERS USING MACHINE LEARNING

Dilshan G.A.M.
IT21180934

Supervisor: Ms. Wishalya Tissera

Co – Supervisor: Dr. Kapila Dissanayaka

BSc. (Hons) in Information Technology

Specializing in Data Science

Department of Computer Science

Sri Lanka Institute of Information Technology

Sri Lanka

DECLARATION

I declare that this is my own work, and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 10/04/2025

The above candidate has carried out research for the bachelor's degree Dissertation under my supervision.

Supervisor	Ms. Wishalya Tisera		
Co-Supervisor	Dr. Kapila Dissanayake		

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to all those who supported me in completing my fourth-year research project. First and foremost, I extend my deepest appreciation to my research supervisor, Ms. Wishalya Tissera, for her invaluable guidance, encouragement, and continuous support throughout the research process. I am also grateful to my co-supervisor, Dr. Kapila Dissanayaka, for his insightful advice and technical guidance, which significantly contributed to the success of this study. My sincere thanks go to the CDAP panel for their constructive feedback and valuable suggestions, which helped refine our research. I also acknowledge the external supervisor, Dr. Anuradha Baminiwatta, for her coordination efforts and for providing crucial information that facilitated the initial stages of this project. Additionally, I appreciate the contributions of my fellow group members for their collaborative efforts and shared knowledge in the collection of materials, data analysis, and project development. I extend my heartfelt thanks to my faculty and department for their assistance in providing the necessary resources and research facilities. Furthermore, I acknowledge the support of any external organizations or funding bodies (if applicable) that contributed to the research. Finally, I am grateful to my parents and friends for their unwavering encouragement and support, which played a vital role in the successful completion of this project.

ABSTRACT

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that affects millions globally, leading to cognitive decline, memory loss, and ultimately loss of independence. Early detection and accurate classification of AD stages are vital for effective patient management and treatment planning. This research project presents a dual-model approach to detect and classify Alzheimer's disease using both neuroimaging and clinical data, leveraging the capabilities of deep learning and machine learning.

For neuroimaging-based detection, a hybrid Convolutional Neural Network (CNN) combined with a Support Vector Machine (SVM) was developed to analyse MRI images from the OASIS dataset. This hybrid model utilizes CNN's strength in feature extraction and SVM's superior classification performance to distinguish between different stages of Alzheimer's, including non-demented, very mild, mild, and moderate dementia. The model was trained and validated on a pre-processed dataset with augmentation techniques applied to enhance robustness and generalizability.

In parallel, a machine learning model was constructed using a structured clinical dataset from Kaggle, which includes patient demographics, medical history, and cognitive test scores. Various preprocessing steps, including handling missing data, normalization, and feature selection, were employed. Classification algorithms such as Random Forest, Logistic Regression, and XGBoost were evaluated to predict the probability of Alzheimer's onset, with performance metrics including accuracy, precision, recall, and F1-score guiding model selection.

This dual-model architecture provides a comprehensive framework for Alzheimer's detection, integrating visual and clinical indicators to improve diagnostic accuracy. The results indicate that the hybrid CNN+SVM model outperforms standalone classifiers on image data, while the clinical model provides valuable early-stage predictions. Together, they offer a promising decision-support tool for medical professionals, potentially enabling earlier intervention and more personalized treatment strategies.

Keywords:

Alzheimer's Disease, Early Detection, MRI Images, Clinical Data, Machine Learning, Deep Learning, Convolutional Neural Network (CNN), Support Vector Machine (SVM), Hybrid Model, Neuroimaging, Data Science, Medical Diagnosis, OASIS Dataset, Kaggle, Cognitive Impairment, Dementia Classification, Random Forest, XGBoost, Logistic Regression, Feature Extraction, Medical AI.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Table 1.List of abbreviations

Abbreviation	**Description**
CNN	Convolutional Neural Network
ML	Machine Learning
DL	Deep Learning
NLP	Natural Language Processing
ANN	Artificial Neural Networks
CDC	Centers for Disease Control
UI	User Experience
TF-IDF	Term Frequency-Inverse Document Frequency

INTRODUCTION

Background & Literature Survey

Alzheimer's disease (AD) is a chronic, progressive neurodegenerative disorder that significantly affects cognitive function, memory retention, and behavior. It is the most common form of dementia globally, accounting for approximately 60–70% of all cases. As populations across the world age, the prevalence of AD is rising at an alarming rate, creating an urgent need for early detection and effective management strategies. The irreversible nature of this disease means that once clinical symptoms become evident, significant neuronal damage has often already occurred. Therefore, early and accurate diagnosis is vital, as it provides the opportunity for early intervention, helps in slowing disease progression, and improves quality of life for patients and caregivers alike [1].

Conventional diagnostic approaches rely heavily on clinical evaluations, neuropsychological testing, patient history, and observation of cognitive symptoms. While these tools are essential in diagnosing dementia in its moderate to advanced stages, they often lack the sensitivity required for detecting early AD or predicting its progression. These limitations have propelled a significant shift towards more objective, data-driven approaches, particularly those grounded in artificial intelligence (AI). Machine learning (ML) and deep learning (DL), subfields of AI, have emerged as powerful tools in medical diagnostics due to their ability to uncover subtle and complex patterns within large datasets that may be imperceptible to human observers [2].

One of the most promising areas in which AI is being applied to Alzheimer's diagnosis is neuroimaging. Among various imaging techniques, magnetic resonance imaging (MRI) has shown immense utility. Structural MRI enables high-resolution visualization of brain anatomy, and it is particularly adept at identifying hallmark signs of Alzheimer's disease such as hippocampal atrophy and volume reduction in the entorhinal cortex. These structural changes, which often occur well before clinical symptoms become pronounced, provide critical biomarkers for early diagnosis. Quantitative analysis of MRI data allows for the extraction of features such as cortical thickness, surface area, and subcortical volume, which are crucial for distinguishing between healthy aging, mild cognitive impairment (MCI), and Alzheimer's disease [3]. In addition to structural MRI, functional MRI (fMRI) and positron emission tomography (PET) scans contribute further by offering insights into functional and metabolic activity in the brain. PET imaging, for example, is useful for detecting amyloid-beta plaques and tau protein tangles—two key pathological hallmarks of AD. These modalities, especially when used in combination, offer a more complete picture of the disease and improve diagnostic accuracy [4].

Machine learning algorithms, when applied to neuroimaging datasets, can identify patterns that differentiate between Alzheimer's stages. Supervised ML techniques such as support vector machines (SVM), logistic regression, and random forests have shown considerable success in classifying individuals into AD, MCI, or cognitively normal (CN) categories. These algorithms utilize input features derived from neuroimaging data, such as gray matter volume or

hippocampal size, to make probabilistic predictions about disease presence or progression [5]. Additionally, feature selection methods like principal component analysis (PCA) and recursive feature elimination are often employed to reduce data dimensionality and enhance model efficiency.

More recently, deep learning approaches have gained traction due to their capacity to learn hierarchical representations directly from raw imaging data. Convolutional neural networks (CNNs) are highly effective in processing 2D and 3D brain images. They can learn to identify subtle differences in brain structure without the need for manual feature engineering. CNNs have been applied to both structural and functional MRI scans, achieving remarkable accuracy in detecting early AD and even predicting conversion from MCI to AD [6]. One widely cited approach combined CNNs with SVMs, where the CNN acted as a deep feature extractor and the SVM served as the classifier, resulting in superior performance compared to using either method alone [7].

Beyond neuroimaging, clinical data play an essential role in Alzheimer's detection. Clinical datasets typically include demographic details, genetic information, results from cognitive assessments, and biomarkers obtained from cerebrospinal fluid (CSF) or blood. These data provide context and can reveal risk factors associated with AD. For instance, the presence of the apolipoprotein E (APOE) ε4 allele has been strongly associated with an increased risk of developing Alzheimer's. Similarly, cognitive test scores from tools like the Mini-Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale–Cognitive Subscale (ADAS-Cog), and Montreal Cognitive Assessment (MoCA) offer quantifiable indicators of cognitive performance [5].

Machine learning models have been trained on such multimodal data to identify individuals at higher risk of developing AD. Some models have utilized electronic health records (EHRs), employing natural language processing (NLP) techniques to extract useful patterns from unstructured clinical notes. This approach has proven useful in identifying early cognitive decline markers or other comorbidities that may contribute to AD risk [8]. The fusion of imaging data with clinical information, referred to as multimodal learning, significantly improves model accuracy and generalization. It enables a more comprehensive risk assessment by leveraging the complementary strengths of both data types [9].

Despite these advancements, significant challenges persist in applying ML and DL to Alzheimer's detection. One of the foremost issues is data heterogeneity. Variations in imaging protocols, scanner types, demographic characteristics, and clinical procedures across different institutions lead to inconsistencies that negatively affect model performance. A model trained on one dataset may not generalize well unless proper domain adaptation techniques are used. Furthermore, standardization of data preprocessing and acquisition protocols is crucial to ensure reproducibility and comparability across studies [10].

Another prominent challenge is the interpretability of deep learning models. Although CNNs and other neural networks often outperform traditional ML algorithms in terms of accuracy, their complex architectures make them difficult to interpret. Clinicians need to understand the

rationale behind a model's prediction to trust its outputs. Thus, explainable AI (XAI) techniques, such as saliency maps, layer-wise relevance propagation, and SHAP values, are increasingly being adopted to visualize and quantify the contributions of input features to model decisions. These tools help bridge the gap between black-box models and clinical usability [11].

Data availability also poses a considerable barrier to progress. High-performing AI models require large, annotated datasets for training. However, privacy concerns, data protection regulations such as HIPAA and GDPR, and logistical issues related to data sharing hinder the collection and use of such datasets. Publicly available repositories like the Alzheimer's Disease Neuroimaging Initiative (ADNI), Open Access Series of Imaging Studies (OASIS), and the Australian Imaging, Biomarkers & Lifestyle Flagship Study of Ageing (AIBL) have been instrumental in mitigating this issue, but more diverse and representative datasets are needed [12].

Moreover, class imbalance in datasets is another recurring issue. Typically, datasets contain more cognitively normal individuals than those diagnosed with AD or MCI, leading to skewed learning and biased predictions. This imbalance can result in models that are accurate overall but perform poorly in detecting minority classes, which in this case are often the most clinically important. To address this, resampling techniques such as the Synthetic Minority Over-sampling Technique (SMOTE), adaptive synthetic sampling, and cost-sensitive learning methods have been employed to improve sensitivity and reduce false negatives [13].

Looking toward the future, there are several promising directions for enhancing Alzheimer's detection using machine learning. Multimodal data integration remains a key area of interest. Combining neuroimaging, clinical, genetic, behavioral, and lifestyle data can provide a more nuanced understanding of an individual's risk profile. For example, incorporating data from wearable sensors, such as those measuring gait, sleep, or heart rate variability, may offer additional insights into prodromal AD symptoms. Effective fusion of these data modalities requires advanced models capable of managing heterogeneous inputs, such as graph neural networks (GNNs) and attention-based architectures [14].

Another important direction is the incorporation of longitudinal data to track disease progression over time. Unlike cross-sectional studies that analyze a single time point, longitudinal approaches enable models to learn temporal patterns and transitions between different disease stages. Recurrent neural networks (RNNs) and long short-term memory (LSTM) networks are well-suited to such tasks and have shown potential in modeling cognitive decline trajectories. This can significantly aid in predicting not only current disease status but also future risk [15].

Furthermore, there is a growing emphasis on developing interpretable and personalized prediction models. Personalized medicine in the context of Alzheimer's implies generating individualized risk scores and treatment recommendations based on unique patient profiles. These systems would empower clinicians to tailor interventions more precisely and monitor disease progression more effectively. Moreover, predictive models that integrate genetic

predisposition, imaging biomarkers, and real-world behavioral data can serve as decision support tools in routine clinical practice [16].

Lastly, rigorous clinical validation and regulatory approval are necessary before AI models can be fully adopted in healthcare. Many existing models are limited to research environments and lack prospective validation. Future studies must focus on translating these algorithms into clinically tested, user-friendly systems that can be integrated into electronic health systems. Collaborative efforts between data scientists, neurologists, radiologists, ethicists, and policymakers are essential to navigate regulatory landscapes and establish best practices for the deployment of AI in Alzheimer's care [17].

In summary, the application of machine learning and deep learning to Alzheimer's detection represents a transformative advancement in the field of neuroscience and clinical diagnostics. By utilizing data from neuroimaging and clinical assessments, these technologies have the potential to provide earlier and more accurate diagnoses, ultimately improving patient outcomes. While challenges such as data heterogeneity, model interpretability, and limited access to annotated datasets remain, ongoing research continues to address these issues through innovative algorithms, standardized protocols, and interdisciplinary collaboration. As technology matures and becomes more explainable and accessible, it is poised to become a cornerstone of future Alzheimer's care and research.

Research Gap

Despite the significant advancements in applying machine learning (ML) and deep learning (DL) techniques for the early detection of Alzheimer's disease (AD), several critical research gaps persist that limit their clinical utility and large-scale deployment. Although numerous studies have demonstrated high classification accuracy using neuroimaging data and multimodal approaches, these results are often achieved in controlled experimental settings with ideal datasets, such as ADNI and OASIS, that do not fully represent real-world clinical environments. One fundamental limitation is the **generalizability** of existing models. Most ML/DL frameworks are trained and validated on homogeneous cohorts collected under standardized conditions, but their performance deteriorates when tested on external datasets with different acquisition protocols, demographic distributions, or cultural factors [18]. This raises concerns regarding the reliability and robustness of these models across diverse clinical populations.

Another notable gap is the limited **clinical interpretability** of most deep learning systems. Despite efforts in explainable AI, many CNN-based architectures used in AD prediction lack transparency in terms of feature attribution and decision-making logic, making it difficult for clinicians to trust or adopt them in real practice. Even when visual saliency maps or heatmaps are generated, their clinical relevance and consistency remain underexplored. As a result, there is a disconnect between high-performing models and their adoption in healthcare workflows, where accountability and explainability are paramount [19].

In addition to model transparency, **longitudinal modelling of disease progression** remains underdeveloped. Most existing approaches rely on cross-sectional data for classification, which ignores the temporal dynamics of cognitive decline. There is a need for models that not only predict the status of cognitive impairment but also forecast future conversion from MCI to AD over time. Longitudinal modelling is essential for understanding the transition phases of the disease and designing personalized intervention strategies [20]. However, challenges such as irregular time intervals between clinical visits, missing data points, and limited follow-up duration hinder the implementation of temporal learning techniques like recurrent neural networks (RNNs) and transformer-based models.

Another pressing research gap lies in the **integration of underutilized modalities** and environmental risk factors. While most studies focus on structural MRI and cognitive test scores, few incorporate lifestyle data, socioeconomic status, cardiovascular health, or behavioural metrics, which are increasingly recognized as influential in AD risk [21]. This omission restricts the holistic understanding of the disease. Moreover, although wearable

devices and mobile technologies are generating massive amounts of real-time physiological and behavioural data, their application in Alzheimer's prediction research remains scarce. Incorporating continuous, real-world data streams into predictive models could revolutionize early detection and monitoring, but methodological and ethical challenges must first be addressed [22].

The **scarcity of diverse and representative datasets** is another critical bottleneck. Most publicly available datasets are heavily biased towards North American and European populations, with minimal representation of ethnically and culturally diverse groups. This limits the inclusiveness of trained models and may inadvertently reinforce existing health disparities [23]. Addressing this issue requires coordinated international efforts to collect and share data from varied geographic and clinical settings, along with frameworks that ensure ethical data governance and privacy protection.

Moreover, **multi-centre and federated learning approaches** have not yet been fully explored in AD diagnosis. These frameworks allow collaborative model training across institutions without sharing raw patient data, which could overcome data silos and privacy concerns. Despite their potential, few studies have investigated how federated learning can be effectively deployed in the context of medical imaging and electronic health records for AD prediction [24]. There is also a lack of standardized benchmarks and evaluation metrics for comparing model performance across different studies. This inconsistency makes it difficult to assess the progress of AI models objectively or to identify best practices for model development and validation [25].

Finally, the **translation of research findings into clinical practice** remains slow. Few AI models have undergone regulatory scrutiny or real-world clinical trials to evaluate their effectiveness in day-to-day diagnostic settings. Bridging this translational gap requires interdisciplinary collaboration between computer scientists, neurologists, radiologists, ethicists, and policymakers to develop clinically validated, explainable, and user-friendly tools. Further, there is a need for implementation research to assess how these tools can be integrated into existing healthcare systems, how they influence diagnostic accuracy and workflow efficiency, and how patients and clinicians perceive and interact with them [26].

In light of these unresolved challenges, future research must focus on building scalable, interpretable, and generalizable ML/DL models that can operate across diverse populations and real-world conditions. It must also prioritize longitudinal prediction, incorporation of novel and underused data types, and the development of robust evaluation frameworks to ensure both scientific rigor and clinical impact. Without addressing these research gaps, the full potential of AI in transforming Alzheimer's diagnosis and care will remain unrealized.

Research Problem

Alzheimer's disease (AD) remains one of the most pressing neurological challenges of the 21st century, with profound personal, societal, and economic impacts. Despite decades of research, early detection of AD continues to be a complex and unresolved issue. The subtlety of initial cognitive symptoms, variability in disease progression, and lack of universally accepted biomarkers make early diagnosis particularly challenging. As the global burden of AD rises, there is an urgent need to develop diagnostic methodologies that are not only highly accurate

but also accessible, scalable, and deployable in real-world healthcare settings. This growing necessity has led to the emergence of machine learning (ML) and deep learning (DL) as promising tools to augment conventional diagnostic approaches by learning discriminative patterns from large datasets, particularly neuroimaging and clinical records [27].

Although many ML/DL-based systems have demonstrated excellent performance in controlled settings, they often rely on curated datasets with standardized imaging protocols, clean labels, and idealized participant populations. These limitations significantly hinder their translation into diverse clinical environments. Most existing models are trained on publicly available datasets such as the Alzheimer's Disease Neuroimaging Initiative (ADNI), which, while valuable, lack demographic diversity and real-world complexity [28]. Consequently, there exists a gap between research models and clinical applicability. For AI-based tools to be truly impactful in AD diagnosis, they must perform reliably across varied settings, be interpretable by clinicians, and integrate seamlessly with existing workflows. The research problem therefore lies not only in creating accurate models but in developing generalizable, interpretable, and clinically viable systems for early AD detection.

A critical component of this problem is the heterogeneity of Alzheimer's disease itself. The disease manifests differently across individuals due to differences in age, genetic makeup, comorbidities, and lifestyle factors. This variability complicates the creation of models that can accurately and consistently classify or predict disease stages. For example, while structural MRI provides useful biomarkers such as hippocampal atrophy, this feature may not be equally informative across all individuals or subpopulations. Therefore, relying on a single modality or feature set may result in biased or incomplete predictions [29]. The research problem expands further into the challenge of integrating multimodal data sources such as PET scans, genetic markers, cognitive scores, and environmental factors to create a holistic patient profile. However, multimodal integration remains underdeveloped due to technical challenges such as missing data, synchronization issues, and differing data dimensionalities [30].

Moreover, most existing models are designed to classify patients into predefined categories—such as Alzheimer's, mild cognitive impairment (MCI), or cognitively normal (CN)but they often fail to capture the temporal dynamics of disease progression. Alzheimer's disease evolves over many years, with a lengthy preclinical phase during which neurodegeneration occurs silently. Cross-sectional models that ignore this temporal aspect cannot adequately predict how and when a patient's condition may evolve. This presents a major gap in current research. There is a need for longitudinal models that can forecast disease conversion and progression using sequential data, which is far more aligned with how AD unfolds [31]. Developing such models requires access to comprehensive time-series datasets, and sophisticated architectures such as long short-term memory (LSTM) networks or temporal attention mechanisms, which remain underexplored in the context of AD.

Another dimension of the research problem is the lack of model interpretability and clinician trust. Black-box models, although accurate, are often met with skepticism in medical settings where explainability is paramount. Clinicians need to understand the rationale behind a model's decision, especially in high-stakes scenarios such as diagnosing a progressive and currently incurable disease. However, most DL models used in AD detection do not offer intuitive explanations for their predictions, and existing explanation tools such as Grad-CAM or SHAP are not always aligned with clinical reasoning [32]. The absence of trustworthy interpretability mechanisms has slowed down the clinical adoption of AI technologies. This represents a key gap building systems that not only predict accurately but also provide insights that are understandable and clinically meaningful.

Compounding this issue is the fact that many research studies fail to address the challenges of class imbalance and data scarcity. Alzheimer's datasets typically contain more samples from cognitively normal individuals than from those with AD or MCI, which biases model training and evaluation. While techniques such as Synthetic Minority Over-sampling Technique (SMOTE) or class-weight adjustments are sometimes employed, they are rarely optimized for neuroimaging and multimodal data [33]. Furthermore, publicly available datasets are limited in size and scope, especially when considering underrepresented populations such as minority ethnic groups, rural communities, or patients with comorbid conditions. This lack of diversity introduces further bias and limits the fairness of the developed models. The research problem must, therefore, also tackle the issues of equitable data representation and algorithmic fairness.

Another unexplored but essential area is the personalization of Alzheimer's risk prediction. Most ML/DL models operate at a population level, offering general predictions based on average trends in data. However, AD is highly individualistic, and predictive models must eventually cater to individual variability to be clinically useful. Personalized medicine approaches using AI are still in their infancy for AD, and little work has been done to tailor predictions based on personal profiles that combine imaging, clinical, genetic, and behavioral factors [34]. This presents a significant opportunity to shift from generic risk scores to individual-level forecasting and decision support systems.

Finally, despite the development of many proof-of-concept models, very few have been translated into deployable clinical tools. There is a wide gap between algorithm development and clinical implementation. Current literature seldom addresses practical issues such as model integration with hospital information systems, data privacy, real-time processing requirements, and user interface design. Without bridging this implementation gap, even the most sophisticated models remain academic exercises with little impact on real patient care [35].

In conclusion, the research problem in early Alzheimer's disease detection using machine learning and deep learning is multifaceted. It extends beyond the mere development of accurate classifiers to encompass challenges such as data heterogeneity, temporal modeling, multimodal integration, interpretability, fairness, personalization, and real-world applicability. While promising advancements have been made, significant gaps persist, necessitating comprehensive, interdisciplinary approaches that combine technical innovation with clinical relevance. The central challenge is to design AI-powered diagnostic systems that are not only scientifically rigorous but also ethically sound, inclusive, and clinically transformative.

OBJECTIVES

Main Objectives

The main objective of this research is to develop a robust, interpretable, and clinically applicable machine learning framework for the early detection and classification of Alzheimer's disease using neuroimaging and clinical data. This study seeks to bridge the gap between high-performing experimental models and practical, real-world diagnostic tools by focusing on generalizability, longitudinal analysis, and multimodal data integration. While numerous studies have demonstrated the utility of convolutional neural networks and traditional ML algorithms in classifying Alzheimer's disease stages, few have successfully addressed the critical challenges of data heterogeneity, model interpretability, and temporal prediction in diverse populations. Therefore, this research aims to design and evaluate hybrid ML/DL architecture that combines structural MRI features, clinical variables, and cognitive test scores to enhance predictive accuracy and early-stage diagnosis. In doing so, the study will investigate the integration of longitudinal patient data to support disease progression modeling and risk stratification over time. Additionally, efforts will be made to apply explainable AI techniques to provide transparent insights into the model's decision-making process, thus promoting clinician trust and adoption. The ultimate goal is to contribute a scalable, reproducible solution that can be integrated into routine clinical practice and assist healthcare professionals in making timely, data-driven decisions regarding patient care. By addressing these objectives, the research intends to advance the current state of Alzheimer's diagnostics and lay the foundation for future developments in AI-assisted neurodegenerative disease monitoring and management.

Specific Objectives

Development of a Hybrid Machine Learning Model for Neuroimaging-Based Classification

The first specific objective of this research is to design, implement, and optimize a hybrid machine learning (ML) and deep learning (DL) framework for the classification of Alzheimer's disease (AD), mild cognitive impairment (MCI), and cognitively normal (CN) individuals using neuroimaging data primarily structural magnetic resonance imaging (MRI). Neuroimaging has become a cornerstone in the diagnosis of AD, particularly structural MRI, which reveals biomarkers such as hippocampal atrophy, cortical thinning, and ventricular enlargement—common indicators of early neurodegeneration. This research will extract quantitative imaging

features using advanced image processing tools and deep neural architectures to feed into a convolutional neural network (CNN) model. CNNs are well-suited for analyzing spatial hierarchies in MRI data and can learn complex features without manual engineering. However, to boost performance and address overfitting risks, the proposed model will incorporate a hybrid strategy where CNN-based feature representations are further classified using a traditional supervised algorithm such as a support vector machine (SVM) or random forest classifier. The intention is to combine CNN's feature learning capabilities with the generalization strength of traditional classifiers. The model will be trained and validated on large, publicly available datasets like the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Open Access Series of Imaging Studies (OASIS). Evaluation metrics such as accuracy, precision, recall, F1-score, and ROC-AUC will be used to assess model performance. Data augmentation, regularization, and cross-validation techniques will be applied to enhance robustness. The aim is to create a model that is not only highly accurate but also reliable and generalizable across various clinical imaging settings.

Multimodal Integration of Clinical, Cognitive, and Imaging Data

The second objective of this research is to build a comprehensive multimodal machine learning framework that integrates structural neuroimaging data with diverse non-imaging features, such as clinical, cognitive, genetic, and demographic variables, to improve the early detection and progression modeling of Alzheimer's disease. While neuroimaging provides critical insights into anatomical brain changes, it does not capture all the relevant biological and behavioral dimensions of Alzheimer's. Clinical features such as age, sex, and education level, along with genetic markers like APOE ε4 status, and cognitive assessment scores from tools such as the Mini-Mental State Examination (MMSE) and Alzheimer's Disease Assessment Scale–Cognitive Subscale (ADAS-Cog), offer a richer and more contextual understanding of disease state and progression risk. This objective will involve designing a data fusion strategy that can integrate these heterogeneous data types effectively. Both early and late fusion techniques will be explored: early fusion will involve combining feature sets prior to model training, while late fusion will involve training separate models for each modality and merging their outputs. The challenge of handling missing or incomplete data will be addressed using imputation techniques and robust model architectures that can tolerate data irregularities. Multimodal integration has the potential to increase classification sensitivity, especially in individuals at the prodromal or mild cognitive impairment stage who may not yet exhibit pronounced structural changes on MRI. The goal is to enhance the model's predictive capacity and provide a more holistic assessment of cognitive health, making it more useful for clinicians and researchers working in diverse diagnostic and monitoring environments.

Enhancing Model Interpretability through Explainable Artificial Intelligence (XAI)

The third specific objective of this research is to enhance the interpretability and clinical transparency of the machine learning models developed for Alzheimer's disease detection by incorporating explainable artificial intelligence (XAI) techniques. Although deep learning models, particularly convolutional neural networks, have demonstrated impressive

performance in classification tasks, their complex, non-linear structures often function as "black boxes," offering little insight into how decisions are made. In the context of medical diagnosis, this opacity is a significant barrier to clinical adoption. Healthcare professionals require clear, justifiable explanations to trust and act upon model predictions. Therefore, this research will implement and compare state-of-the-art XAI methods such as Shapley Additive explanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME), and Gradient-weighted Class Activation Mapping (Grad-CAM) to visualize and quantify the contribution of each input featured a brain region from MRI or a cognitive test score—to the final prediction. These tools will be used not only to validate the internal logic of the model but also to ensure alignment with known clinical and biological insights. For example, if the model consistently highlights the hippocampus and entorhinal cortex as influential features in AD detection, this will reinforce its biological validity. Outputs from XAI tools will be presented in interpretable formats, such as heatmaps, feature importance plots, and textual explanations, designed to assist clinicians in understanding and validating model outcomes. The overarching aim is to foster clinician trust, support decision-making, and pave the way for the integration of AI tools into real-world diagnostic workflows for Alzheimer's disease.

REQUIREMENT GATHERING AND FEASIBILTY STUDY

Requirement Gathering and Analysis

During the requirement gathering and analysis phase, the team conducted a thorough investigation of existing Alzheimer's disease (AD) detection methodologies. This involved an extensive review of current diagnostic frameworks, evaluation of advancements in AI-based Alzheimer's prediction, and a comprehensive literature survey of both neuroimaging-based and clinical-data-driven approaches. The goal was to identify effective strategies for integrating MRI brain scans and structured clinical datasets to improve early-stage AD detection using machine learning and deep learning techniques.

Functional Requirements

i. Input Data Collection – The system should accept structural MRI brain scan images and structured clinical data (such as MMSE scores, age, gender, APOE ε4 status, etc.) provided in CSV format as primary inputs.
ii. MRI Preprocessing and Brain Region Extraction – The system must process MRI images to normalize intensities and segment relevant brain regions (e.g., hippocampus, entorhinal cortex) for feature extraction using appropriate neuroimaging tools or models.
iii. MRI-Based AD Feature Analysis – A deep learning model such as VGG-19 or a 3D CNN must be used to analyse spatial patterns in the MRI scans to detect structural biomarkers associated with AD.
iv. Clinical Data-Based AD Prediction – A machine learning model such as Random Forest or Logistic Regression must process the structured clinical data

to classify subjects into Alzheimer's, MCI, or cognitively normal (CN) categories.

v. Multimodal Fusion for Enhanced AD Detection – The system must integrate results from MRI-based analysis and clinical data classification using a defined fusion strategy to generate a unified Alzheimer's risk prediction score.
vi. Output Interpretation and Visualization – The system should present prediction outcomes with clear interpretability, displaying classification results, model confidence scores, and a downloadable diagnostic report for clinical use.
vii. User Accessibility and Data Management – Authorized users, such as neurologists, radiologists, caregivers, and researchers, must be able to securely access patient reports and manage case histories through a protected login system.

Non-Functional Requirements

i. Performance – The system should deliver prediction results in real-time or near-real-time, ensuring timely assistance in clinical decision-making for early AD detection.
ii. Reliability – The application must incorporate robust error handling and fail-safe operations to maintain consistent performance without data corruption or interruption.
iii. Availability – The platform should be continuously accessible to support healthcare professionals working across different time zones and facilities.
iv. Security – All patient data, including MRI images and clinical details, must be securely stored, encrypted, and accessible only by verified users, ensuring compliance with health data privacy standards (e.g., HIPAA/GDPR).
v. Usability – The user interface should be intuitive, visually accessible, and easy to navigate for both technical and non-technical users, including clinicians and caregivers.
vi. Data Storage and Scalability – The system must support efficient storage and retrieval of large datasets, including high-resolution MRI files and CSV-based clinical records, with scalability to accommodate growing data volumes and future patient intake.

Feasibility Study for AD Detection System

The feasibility study for the Alzheimer's detection system was carried out to assess the practicality and viability of implementing the proposed AI-based solution using structural MRI data and clinical records. This evaluation considered several critical dimensions, including schedule feasibility, technical feasibility, economic feasibility, and operational feasibility. The goal of the study is to ensure that the system can be successfully developed, deployed, and used within acceptable timeframes, with manageable technical constraints, reasonable costs, and sufficient user accessibility.

From a **schedule feasibility** standpoint, the project is designed with a realistic timeline that accommodates phases such as data acquisition, preprocessing, model training, evaluation, and system deployment. Publicly available datasets such as ADNI and OASIS provide immediate access to high-quality neuroimaging and clinical data, reducing delays associated with data collection.

In terms of **technical feasibility**, the system leverages proven machine learning and deep learning models such as CNNs for MRI analysis and Random Forest classifiers for structured clinical data. Available Python-based frameworks like TensorFlow, PyTorch, and Scikit-learn facilitate model development, while cloud platforms offer the infrastructure needed for scalability and performance. Integration of multimodal data, although complex, is supported by recent advancements in data fusion techniques.

The **economic feasibility** is favourable due to the use of open-source tools and publicly available datasets, which significantly reduce the overall development and deployment costs. No expensive hardware or proprietary software is required for the initial implementation.

Finally, from an **operational feasibility** perspective, the system is designed to be user-friendly, secure, and accessible to clinicians, researchers, and caregivers. The intuitive interface and explainable output ensure that the system can be adopted without extensive training, supporting early and accurate Alzheimer's detection in real-world healthcare settings

Schedule Feasibility

A comprehensive project roadmap was formulated, highlighting key milestones and anticipated deadlines required to meet the research objectives. The timeline was structured to guarantee on-time delivery, with ongoing progress evaluation at every stage. A Gantt chart was developed to illustrate the schedule and monitor progress across phases such as data acquisition, model development, system implementation, and evaluation.

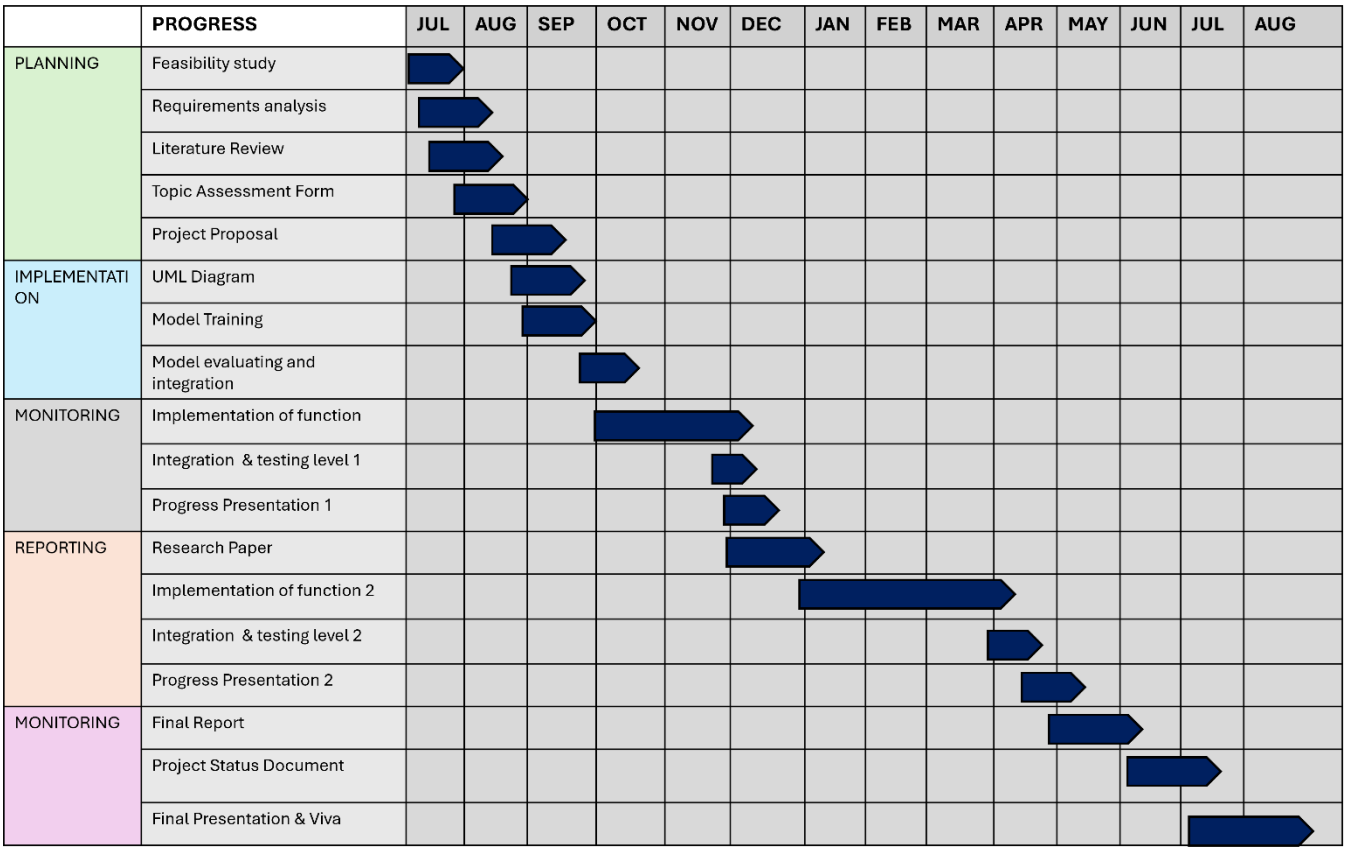


Figure 1Gantt Chart

Technical Feasibility

The technical feasibility of the Alzheimer's detection system was assessed by evaluating the necessary technologies, including machine learning algorithms, convolutional neural networks (CNNs), and image processing techniques. The system integrates deep learning for MRI image analysis and machine learning models for clinical data-based assessment, ensuring accurate and efficient Alzheimer's detection. Additionally, the software and hardware requirements, such as computational power, image resolution quality, and data processing capabilities, were carefully examined to support smooth implementation and future scalability.

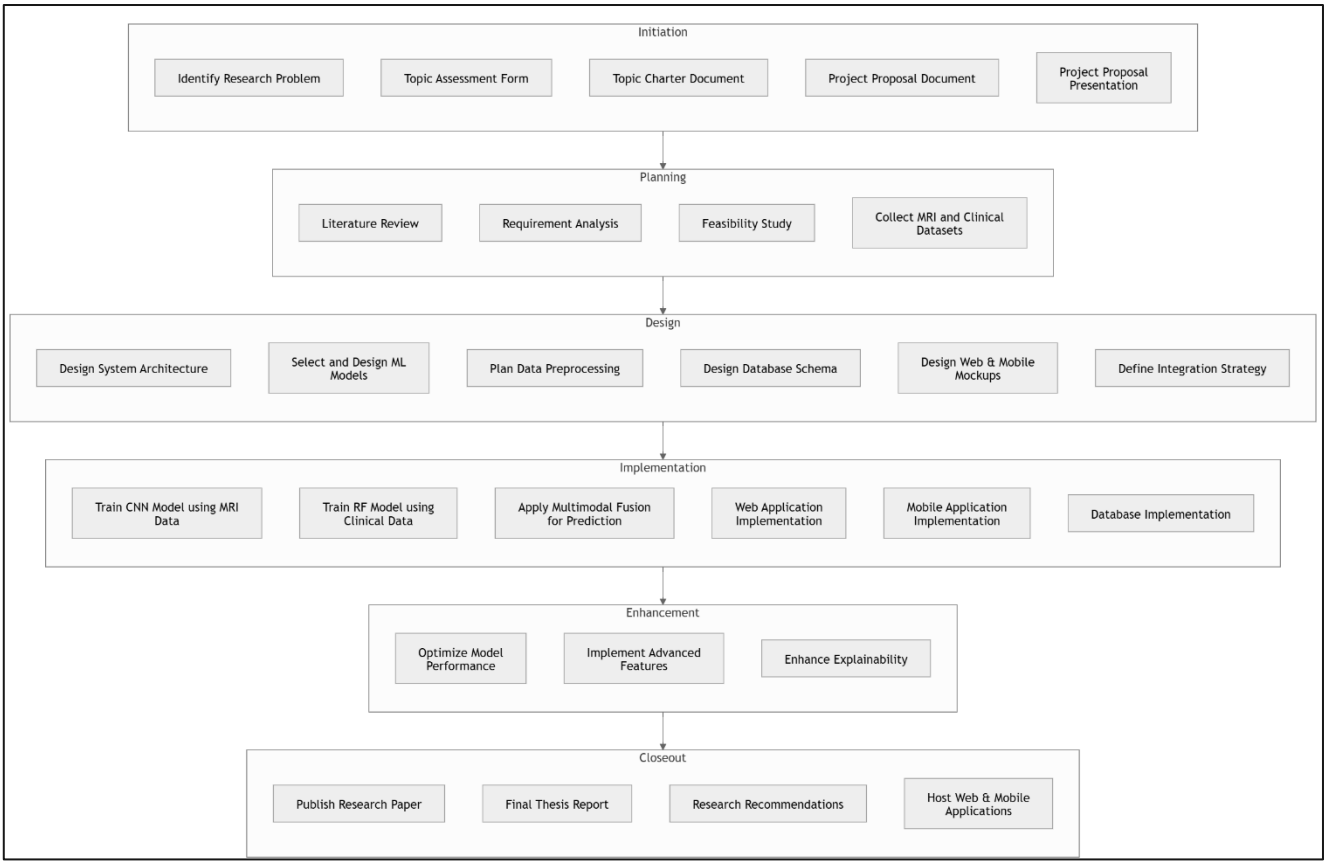


Figure 2 Work Breakdown chart

Economic Feasibility

The economic feasibility of the Alzheimer's detection system was evaluated by analysing the estimated costs involved in its development, deployment, and operational support. The main expenses include charges for essential research and development tools, access to cloud computing services, and usage of machine learning platforms and APIs for model training and

evaluation. Key components of the budget include cloud services and development tools (such as storage, model deployment environments, and productivity aids like Grammarly), costing approximately $31.37 (LKR 10,000), and data acquisition through open-source Alzheimer's datasets like ADNI and OASIS, estimated at $6.77 (LKR 2,000). In addition, cloud-based platforms such as Amazon Web Services (AWS), Google Cloud Platform (GCP), and APIs from services like OpenAI (for explainability components or report generation) are projected to cost $40.00 (LKR 13,000).

The **total estimated implementation cost** for the system is **$78.14 (LKR 25,000)**. This investment is economically justified by the significant long-term benefits the system offers. These include improved early detection of Alzheimer's disease, reduced time to diagnosis, enhanced support for clinical decision-making, and the potential for scalable deployment in clinical and research environments. The system leverages freely available public datasets and open-source libraries, minimizing financial overhead while maximizing utility and accessibility. Overall, this makes the Alzheimer's detection system economically viable and sustainable, particularly for use in low-resource healthcare settings or academic research initiatives.

.

Table 2. Budget Allocation

Component	Est.Amount in USD	Est.Amount in LKR
Tools for Research (Cloud Services, Grammarly, etc.)	31.37	10,000.00
Data Collection (ADNI, OASIS - open sources)	6.77	2000.00
Cloud Platforms(AWS,GCP,OpenAI key)	40.00	13,000.00
Total	78.14	25,000.00

Operational Feasibility

An assessment of the system's operational feasibility was conducted to verify its practicality and ease of adoption by medical practitioners. The platform is built with a user-centric design, allowing professionals with limited technical background to navigate it effortlessly while ensuring precise and dependable Alzheimer's risk evaluation. Through the combined use of MRI scan analysis and clinical data inputs, the system streamlines the diagnostic process, offering meaningful support for early identification and timely clinical decision-making.

METHODOLOGY

Overall Architecture

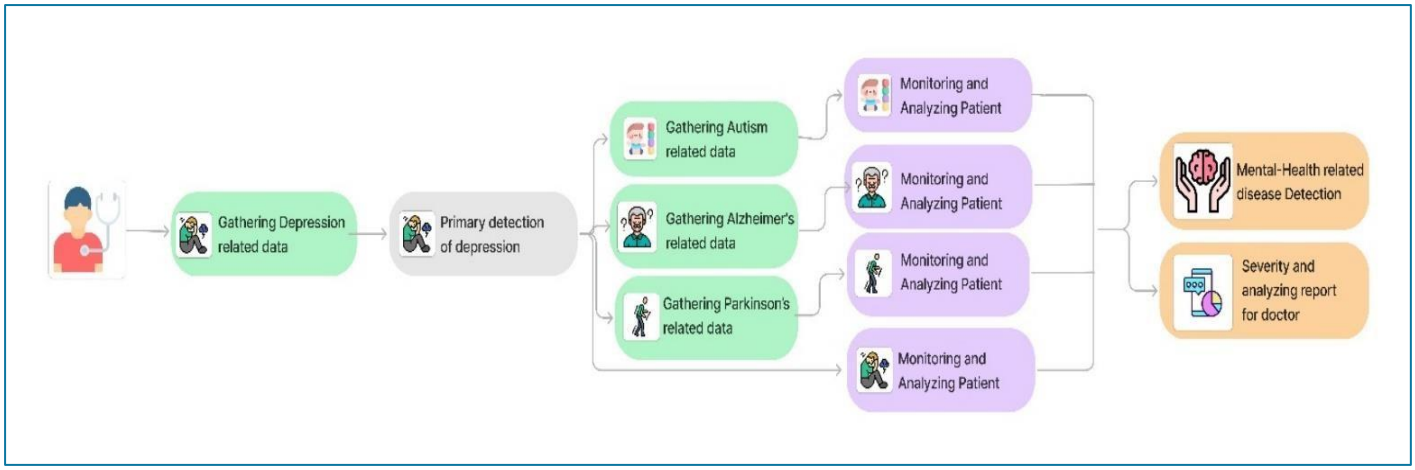


Figure 3 Overall architecture diagram

The diagram provides a detailed representation of how various mental health conditions are identified and monitored. When a patient initially consults a doctor, relevant details regarding depression are collected, and a preliminary assessment is carried out to determine if depressive symptoms are present. If depression is identified, the system proceeds to gather additional data to check for signs of Parkinson's, Alzheimer's, or autism. Once any of these conditions are detected, the patient's health history and developments are documented over time, and they are enrolled for continuous tracking and evaluation. Subsequently, all the collected information is compiled into detailed reports that offer physicians valuable historical insight and symptom severity, aiding them in making more informed decisions regarding patient treatment.

Component Diagram

This section explores the theoretical foundations and problem formulation required for accurate, efficient, and early prediction of Alzheimer's disease (AD). As illustrated in Figure 5, the proposed system adopts a multimodal approach to enhance diagnostic precision, advancing beyond traditional single-source classifiers toward a comprehensive, end-to-end predictive framework. Our research is structured around three core components:

i. Multimodal feature extraction, in which structural features are extracted from MRI scans using CNN-based methods and clinical features are parsed from structured CSV datasets (e.g., MMSE scores, APOE genotype, age) sourced from repositories like ADNI and OASIS.
ii. Multimodal feature fusion, where outputs from the neuroimaging and clinical data models are integrated for synergistic interpretation.

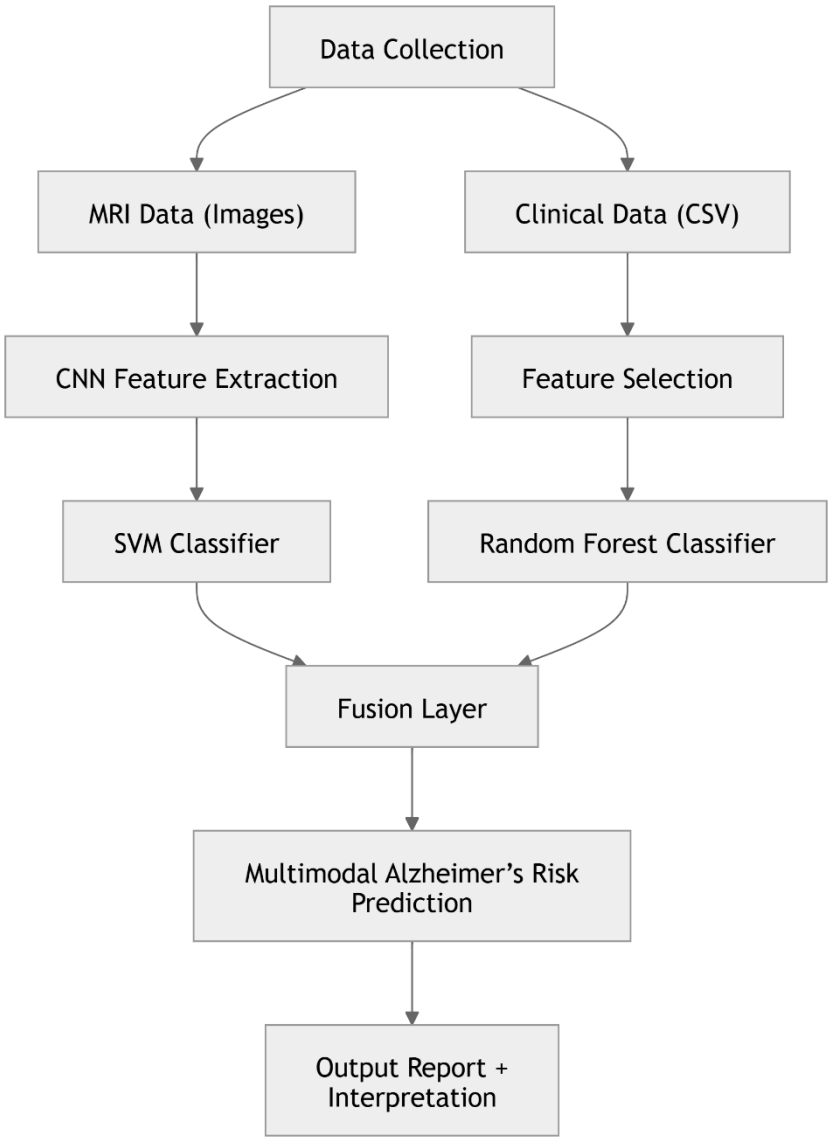


Figure 4 Descriptive Flowchart

iii. Multimodal AD detection, where the fused features support final Alzheimer's classification.

By following this integrated framework, the system achieves a more reliable and holistic assessment of an individual's risk, supporting early clinical intervention.

The methodology introduces a multimodal data fusion framework tailored for early Alzheimer's disease prediction by combining high-resolution MRI images and structured clinical data. The process begins with the acquisition of publicly available datasets, which include both neuroimaging scans and detailed clinical information. Two specialized predictive models are employed: a CNN+SVM-based architecture processes the MRI data, while a Random Forest model analyzes the clinical CSV records. The outputs from these models are subsequently fused to generate a unified Alzheimer's risk prediction, enhancing diagnostic accuracy through data complementarity.

Proposed Methodology

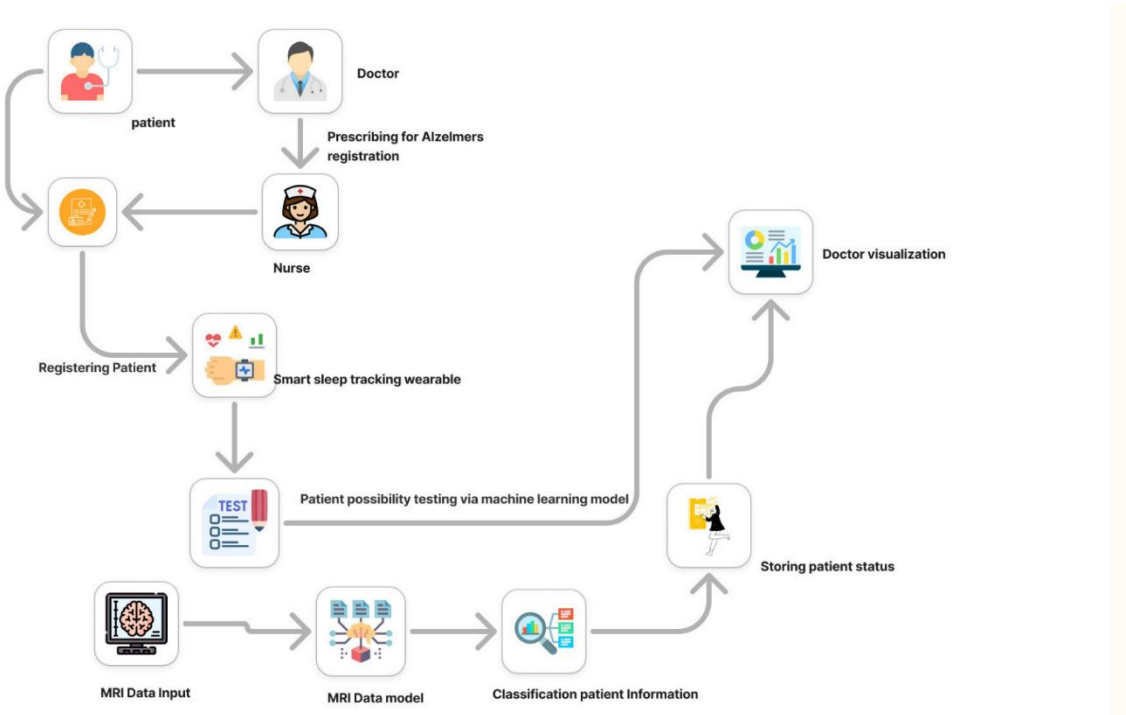


Figure 5 Architecture diagram of the methodology

The proposed methodology for the early detection of Alzheimer's Disease (AD) integrates patient-centric data acquisition, multimodal input processing, and machine learning-based

prediction to establish an efficient and practical diagnostic framework. As visualized in the methodology diagram (Figure X), the system is designed to support early-stage identification of AD by combining structural brain imaging (MRI) and auxiliary physiological data, such as sleep patterns, collected through wearable devices. The process begins with the patient visiting a doctor, who then prescribes screening for potential Alzheimer's symptoms. Upon receiving the prescription, a nurse facilitates the registration of the patient into the system, ensuring the patient's information is securely captured and linked with smart wearable devices.

Smart wearable integration in the system introduces an innovative element—continuous physiological monitoring. Research shows that sleep disturbances and circadian rhythm abnormalities are early indicators of cognitive decline and neurodegenerative conditions, including Alzheimer's disease [27]. By capturing real-time behavioural signals via wearable sensors, the system strengthens the predictive foundation by complementing MRI-based structural insights. After patient data acquisition is completed, the methodology initiates its core machine learning component, which tests the patient's likelihood of having Alzheimer's using pre-trained predictive models.

The system operates with two primary models: a deep learning model using convolutional neural networks (CNN) integrated with a support vector machine (SVM) classifier to process and interpret MRI scans, and a separate Random Forest classifier that analyses clinical features derived from structured data (e.g., demographic information, cognitive test scores, genetic markers). CNNs are particularly effective in capturing spatial features from MRI data, and their combination with SVM enhances classification performance in imbalanced datasets—a common issue in medical applications [28]. Meanwhile, Random Forests are chosen for their robustness and high performance in handling structured health datasets, particularly in the presence of non-linearity and missing values [29].

These two modalities operate in parallel, enabling the system to process heterogeneous data types concurrently. The multimodal outputs are then fused to generate a comprehensive patient classification. This step ensures that the final diagnostic decision is based on both anatomical markers from MRI data and physiological/cognitive indicators. The patient's predicted status is subsequently stored in a centralized repository, ensuring their health profile and progression are accessible over time. This historical storage mechanism is crucial for long-term disease monitoring and personalized intervention planning.

The final step involves visualizing the prediction outcomes and clinical history through a doctor-friendly dashboard. This interface transforms raw predictions into interpretable reports and graphs that can be easily understood by healthcare professionals, enhancing their decision-making process. Such explainable output is essential for fostering trust in AI-assisted diagnostics and ensuring ethical implementation in clinical practice [30]. Overall, the methodology aims to deliver a scalable, accessible, and data-driven solution for early Alzheimer's disease screening, leveraging the strengths of both medical imaging and structured clinical data.

diagnosis. The subsequent sections delve into the detailed implementation and evaluation of this model.

Data collection layer

The data collection layer is a crucial foundation for the development of the Alzheimer's disease detection system, enabling both image-based and structured data-based predictions. Two

publicly available datasets from Kaggle were utilized to support the multimodal nature of this research. The first dataset, obtained from Kaggle - Alzheimer's Disease Dataset, consists of structured clinical and demographic data provided in CSV format. It includes essential variables such as gender, age, education level, socioeconomic status, MMSE scores, and clinical diagnostic labels (e.g., demented, non-demented). This dataset is highly suitable for classical machine learning algorithms like Random Forest, which can use these features to predict the likelihood of Alzheimer's in individuals. The structured nature of the data allows for efficient preprocessing steps, including normalization, encoding of categorical variables, and missing value treatment, thus facilitating high-quality input for supervised classification tasks.

The second dataset, sourced from Kaggle - OASIS MRI Images, contains T1-weighted MRI brain scans from the OASIS (Open Access Series of Imaging Studies) repository. These images are labelled with clinical conditions and are critical for developing and training the CNN-SVM hybrid model for visual feature extraction and classification. This neuroimaging dataset enables the system to identify structural brain changes typically associated with Alzheimer's disease, such as hippocampal atrophy. Together, these datasets provide complementary perspectives—clinical and anatomical—enabling a multimodal prediction framework. The integration of structured patient records with visual brain imaging enhances the robustness, precision, and clinical relevance of the detection system, contributing to early and more accurate diagnosis.

MRI Analyzing Dataset

In this research, one of the core components of the Alzheimer's disease detection framework involves the use of a curated MRI image dataset titled **"OASIS Alzheimer's Detection"**, hosted on Kaggle by user Ninad Aithal. This dataset is derived from the Open Access Series of Imaging Studies (OASIS), a reputable neuroimaging project that provides publicly accessible MRI scans and metadata to support research in the domain of cognitive aging and Alzheimer's disease. The specific version of the dataset used contains over **86,400 pre-processed image slices**, sorted into four distinct categories based on the severity of dementia: **Non-Demented**, **Very Mild Dementia**, **Mild Dementia**, and **Moderate Dementia**. This class-based organization supports supervised learning approaches and makes the dataset highly suitable for training deep learning models aimed at Alzheimer's stage classification [31].

The MRI data provided are T1-weighted structural scans, which are commonly used in neuroimaging research due to their ability to offer high contrast between brain tissues—particularly grey matter, white matter, and cerebrospinal fluid. These contrasts are vital for identifying structural changes such as hippocampal atrophy or cortical thinning, which are hallmark features of Alzheimer's pathology [32]. The dataset includes a wide distribution of labelled data: approximately **67,200 images belong to the non-demented category**, **13,700 to Very Mild Dementia**, **5,000 to Mild Dementia**, and **488 to Moderate Dementia**. While the class distribution is imbalanced, it reflects real-world demographics, where early stages of the disease are more common than advanced dementia during diagnosis.

Each of the four directories in the dataset contains many axial 2D image slices extracted from volumetric MRI scans. The data are structured in such a way that they can be directly fed into convolutional neural networks (CNNs) for feature extraction and classification tasks. The presence of a substantial number of image slices for each subject ensures adequate variability

and representation of anatomical features, which enhances the training potential of deep learning models. Prior to modelling, the image data typically undergo preprocessing steps including resizing (e.g., 224x224 pixels for VGG-based models), grayscale normalization, histogram equalization, and augmentation strategies like flipping or rotation to improve model generalization [33].

One of the advantages of this dataset is that it allows for **binary, multiclass, and hierarchical classification experiments**. For example, binary classification can be performed by grouping all stages of dementia into one class and comparing them against the non-demented class. Alternatively, fine-grained multiclass models can be built to distinguish between all four categories individually, enabling stage-wise detection of Alzheimer's progression. In this study, a CNN model (VGG19 architecture) is used to extract spatial features from MRI slices, followed by a Support Vector Machine (SVM) classifier that processes the high-dimensional embeddings for final classification. This hybrid approach has proven to enhance accuracy in many neuroimaging applications by leveraging the representational power of CNNs and the decision-boundary sharpness of SVMs [34].

Another strength of this dataset lies in its compatibility with transfer learning. Given the image similarity to natural image structures, pretrained models such as VGG, ResNet, and EfficientNet can be fine-tuned using the OASIS image slices. This reduces training time and improves model performance, especially in the context of limited labelled medical imaging data [35]. However, to ensure clinical relevance, careful validation and evaluation are conducted using appropriate metrics like accuracy, precision, recall, F1-score, and ROC-AUC, as well as cross-validation to assess model robustness.

Despite its strengths, the dataset also presents certain challenges. The class imbalance, especially the relatively low sample size in the Moderate Dementia category, necessitates the use of strategies such as oversampling, data augmentation, or cost-sensitive loss functions to mitigate bias. Additionally, the dataset does not include 3D volumetric scans or subject metadata (such as cognitive scores or demographic variables), which limits its standalone multimodal potential. To address this, the MRI image data is fused with structured clinical data from a separate CSV dataset in the overall system pipeline, allowing for a richer, multimodal Alzheimer's prediction model.

In summary, the Kaggle-hosted OASIS MRI image dataset used in this research is a highly valuable resource for training and evaluating machine learning and deep learning models for Alzheimer's disease classification. Its high-quality images, structured labelling, and compatibility with CNN-based workflows make it an ideal choice for visual biomarker extraction. While certain limitations such as class imbalance and lack of metadata exist, they are addressed through preprocessing and integration with other datasets. The use of this dataset contributes significantly to the system's goal of accurate, early, and automated Alzheimer's disease detection based on neuroimaging data.

Alzheimer's Disease Dataset CSV

The clinical dataset used in this research, titled **"Alzheimer's Disease Dataset"** and hosted on Kaggle by Rabie Elkharoua, provides a structured, tabular collection of patient-level attributes for supporting the prediction and classification of Alzheimer's disease (AD). This dataset plays a crucial role in the multimodal architecture of the proposed Alzheimer's detection framework,

where it complements MRI-based visual feature extraction with demographic, cognitive, and clinical data processed through traditional machine learning algorithms such as Random Forest. The dataset provides high-quality, labelled information that is essential for training models capable of identifying early signs of Alzheimer's based on quantifiable human traits [36].

The dataset is structured in **CSV format**, enabling easy integration into Python-based data science pipelines. It contains a total of **373 records**, each corresponding to a unique patient entry. The key features included in the dataset are: **Gender**, **Age**, **Education level**, **Socioeconomic status (SES)**, **Mini-Mental State Examination (MMSE) score**, **Clinical Dementia Rating (CDR)**, **Estimated Total Intracranial Volume (eTIV)**, **Normalized Whole Brain Volume (nWBV)**, **Atlas Scaling Factor (ASF)**, and most importantly, **Group** (indicating diagnosis status such as 'Nondemented', 'Demented', or 'Converted'). These attributes represent a blend of demographic, cognitive, and volumetric brain data, offering a diverse and rich input for machine learning models focused on AD classification [37].

A particularly valuable aspect of this dataset is the inclusion of **MMSE scores** and **CDR ratings**, which are widely recognized clinical indicators of cognitive function. MMSE scores are based on a 30-point questionnaire used extensively to measure cognitive impairment, while the CDR provides a clinician-rated dementia severity score ranging from 0 (no dementia) to 3 (severe dementia). These two features alone serve as powerful predictors in distinguishing between demented and non-demented individuals [38]. When paired with structural brain metrics like eTIV and nWBV, which quantify brain atrophy, the dataset becomes a robust resource for analysing early signs of neurodegeneration.

Another strength of this dataset lies in its **labelled diagnostic classes**, which support both binary and multiclass classification. The 'Group' column categorizes patients as 'Nondemented', 'Demented', or 'Converted'. The 'Converted' class refers to individuals who transitioned from a non-demented state to dementia over time. This class is particularly valuable for building **longitudinal predictive models** that can forecast disease progression, a key goal in early AD detection [39]. In this study, however, the dataset is used primarily for cross-sectional classification, feeding the features into a Random Forest classifier due to its ability to handle feature interactions, missing values, and non-linearity effectively.

The preprocessing pipeline for this dataset involves several essential steps. First, categorical variables such as **Gender** and **Group** are label-encoded. Missing values, particularly in the **SES** and **MMSE** columns, are handled using statistical imputation (mean or median), depending on the distribution. Numerical features are standardized using z-score normalization to ensure all inputs contribute equally to the model during training. Correlation analysis is also conducted to identify redundant features and eliminate multicollinearity, which could negatively affect model generalization.

Once pre-processed, the dataset is split into training and testing subsets, commonly using an 80/20 ratio. A **Random Forest classifier** is then applied to the structured data to learn classification rules. This ensemble learning method is particularly well-suited for the dataset, as it provides robustness against overfitting, supports feature importance ranking, and performs well on relatively small datasets like this one [40]. After training, the model is evaluated using

key performance metrics, including **accuracy**, **precision**, **recall**, **F1-score**, and **confusion matrix analysis**, which collectively help assess its real-world applicability in clinical decision support.

Despite its strengths, the dataset has limitations. The sample size (373 entries) is relatively small for deep learning tasks and may not capture the full heterogeneity of the Alzheimer's population. Additionally, the dataset is cross-sectional in nature and does not provide timestamped follow-up data, which would be beneficial for true progression modelling. Nonetheless, when used alongside MRI image data in a **multimodal fusion strategy**, it significantly enhances the system's diagnostic performance by introducing structured, interpretable, and clinically validated features.

In summary, the Alzheimer's Disease Dataset by Rabie Elkharoua provides a well-structured and clinically informative collection of features that support machine learning-based prediction of Alzheimer's diagnosis. It enriches the multimodal detection system by contributing non-imaging clinical insights that reflect real-world diagnostic parameters. Its integration with MRI image analysis via a parallel processing and fusion architecture not only improves classification accuracy but also provides deeper interpretability and clinical relevance to the overall system.

Data Preprocessing

Data preprocessing is a crucial step in any machine learning or deep learning pipeline, especially in medical applications where both accuracy and reliability are critical. In the context of Alzheimer's disease detection, preprocessing helps transform raw MRI images and structured clinical data into clean, standardized inputs that can be effectively utilized by classification models. The primary goal of preprocessing is to remove inconsistencies, enhance data quality, and ensure that the features are in a suitable format for machine learning algorithms to interpret accurately.

For the **MRI image dataset**, preprocessing involves several key steps. First, all brain scan images are resized to a uniform dimension (e.g., 224x224 pixels) to match the input requirements of deep learning models like VGG-19. This standardization ensures consistency across the dataset. The pixel values are also normalized to bring them into a specific range (typically between 0 and 1), which helps improve model convergence during training. Additionally, data augmentation techniques such as rotation, flipping, and brightness adjustment are applied to artificially increase the dataset size and improve model generalization. This is especially useful in handling class imbalance—common in medical imaging datasets—by generating more examples for underrepresented classes.

For the **clinical CSV dataset**, preprocessing involves cleaning and transforming tabular data. Categorical variables like "Gender" or "Group" are encoded into numerical values using label encoding or one-hot encoding. Missing values in fields such as MMSE or SES are filled using imputation methods like mean, median, or mode substitution, depending on the nature of the variable. Continuous variables such as age or brain volume are scaled using normalization or standardization techniques to ensure they contribute equally during model training.

Overall, preprocessing is an essential phase that directly impacts model performance. Proper handling of missing data, normalization, and format standardization ensures that the learning models—whether based on CNN for images or Random Forest for clinical data—receive high-quality input. This contributes to more accurate predictions, better generalization, and improved reliability of the Alzheimer's detection system. Preprocessing bridges the gap between raw data and intelligent predictions, enabling a more effective, data-driven approach to early diagnosis.

Preprocessing of the ASD Screening Dataset

Preprocessing of MRI data is a fundamental step in building an effective Alzheimer's disease classification model. Raw MRI scans, especially when sourced from open datasets, often vary in dimension, orientation, and pixel intensity, making them unsuitable for direct input into deep learning models. Therefore, consistent preprocessing ensures uniformity across the dataset, reduces noise, and improves model training performance. In this research, preprocessing was applied to the OASIS Alzheimer's Detection dataset obtained from Kaggle, which contains 2D brain scan slices categorized by dementia stage.

The preprocessing pipeline begins with **loading the image files** from four labelled folders: Non-Demented, Very Mild Dementia, Mild Dementia, and Moderate Dementia. Each class contains many PNG image slices. These images are initially of different sizes and resolutions, so the first major step involves **resizing** them to a fixed dimension, typically **224x224 pixels**, which is a standard input size for models such as VGG19 and ResNet. This resizing helps maintain architectural compatibility and computational efficiency during training.

After resizing, **grayscale conversion** is applied. Since most MRI images are already grayscale, this step standardizes all inputs and reduces computational complexity. Next, **pixel value normalization** is performed by dividing all pixel intensities by 255 to scale the values between 0 and 1. This normalization ensures faster convergence of the deep learning model by providing a more stable numerical range.

To address class imbalance and enhance generalization, **data augmentation** techniques are applied. These include random **horizontal flipping**, **small rotations** (up to 15 degrees), **zooming**, and **brightness adjustments**. These augmentations artificially increase dataset variability, helping the model learn robust features even with limited original data. Augmentation is applied dynamically during training using libraries such as Keras ImageDataGenerator or PyTorch's transforms.

Finally, the data is **split into training, validation, and test sets**, usually in an 80:10:10 ratio. This ensures the model is trained on a substantial portion of the data while being validated and tested on unseen images for performance evaluation. After preprocessing, the image data is ready to be fed into the CNN pipeline, ensuring that the inputs are clean, consistent, and structured for optimal learning. This systematic preprocessing contributes significantly to the accuracy and robustness of Alzheimer's detection via MRI imaging.

```python
import os
import cv2
import numpy as np

data_dir = 'Dataset/AlzheimerMRI'
categories = ['NonDemented', 'VeryMildDementia', 'MildDementia', 'ModerateDementia']
img_size = 224
data = []

for category in categories:
    path = os.path.join(data_dir, category)
    label = categories.index(category)
    for img_name in os.listdir(path):
        img_path = os.path.join(path, img_name)
        img = cv2.imread(img_path, cv2.IMREAD_GRAYSCALE)
        resized_img = cv2.resize(img, (img_size, img_size))
        data.append([resized_img, label])
```

Figure 6 Data Loading and Labelling in MRI dataset

```python
import random

random.shuffle(data)
X = []
y = []

for features, label in data:
    X.append(features / 255.0)  # Normalize
    y.append(label)

X = np.array(X).reshape(-1, 224, 224, 1)  # Adding channel dimension
y = np.array(y)
```

Figure 7 Images normalization and reshaping.

```python
from tensorflow.keras.preprocessing.image import ImageDataGenerator

datagen = ImageDataGenerator(
    rotation_range=15,
    zoom_range=0.1,
    horizontal_flip=True,
    brightness_range=[0.8,1.2]
)

datagen.fit(X)  # Fit augmentation on training set
```

Figure 8 Data Augmentation

Preprocessing of the clinical csv dataset.

Preprocessing structured clinical data is essential for ensuring the quality and usability of the dataset for training machine learning models. In this study, the Alzheimer's Disease Dataset by Rabie Elkharoua, obtained from Kaggle, was used as the clinical data source. The dataset includes demographic, cognitive, and structural brain volume attributes for each patient, along with a corresponding diagnosis label indicating whether the patient is non-demented, demented, or converted. Since this dataset is tabular and contains both numerical and categorical features, careful preprocessing is required to prepare it for algorithms like Random Forest or logistic regression.

The preprocessing phase begins with loading the CSV data and checking for missing values. The dataset includes columns such as Age, Gender, Education, MMSE, SES, CDR, eTIV, nWBV, and ASF, among others. Upon inspection, certain fields such as SES and MMSE contain missing values. These are handled using imputation techniques, such as replacing missing numerical values with the column mean or median, depending on the distribution. This step ensures that no rows are dropped unnecessarily, which is important for relatively small datasets like this one (373 rows).

Next, categorical variables such as Gender and Group are converted to numerical values. For instance, Gender is encoded as 0 for female and 1 for male. The Group column, which represents the target class, includes values like "Nondemented", "Demented", and "Converted". For binary classification, the "Converted" class can be merged with "Demented", and label encoding is applied (e.g., 0 for "Nondemented", 1 for "Demented"). This transformation enables the model to interpret these values mathematically.

Following encoding, feature scaling is applied to numerical columns. This involves either standardization (z-score normalization) or min-max normalization, ensuring that features such as eTIV, nWBV, and ASF which exist on different scales are brought to a uniform range. Scaling helps the model converge faster and prevents features with larger ranges from dominating the learning process.

Finally, the dataset is split into training and test sets using a standard 80/20 split. The target label (Group) is separated from the feature matrix, and the resulting data is ready to be fed into the Random Forest classifier. Additional optional steps include correlation analysis to remove highly correlated features, and feature importance ranking post-training to understand which clinical attributes contribute most to Alzheimer's prediction.

This structured preprocessing workflow ensures the clinical dataset is clean, consistent, and machine-learning-ready. It also complements the MRI image data in the multimodal system, contributing interpretable, real-world clinical variables that improve overall predictive performance.

```python
import pandas as pd

df = pd.read_csv('alzheimers_disease_data.csv')
print(df.head())
print(df.info())  # Check for missing values and types
```

Figure 9 Data loading

```python
# Impute missing numerical values with median
df['SES'].fillna(df['SES'].median(), inplace=True)
df['MMSE'].fillna(df['MMSE'].median(), inplace=True)
```

Figure 10 Handling Missing Values

```python
# Encode Gender: Female = 0, Male = 1
df['Gender'] = df['Gender'].map({'F': 0, 'M': 1})

# Encode Group: Nondemented = 0, Demented + Converted = 1
df['Group'] = df['Group'].replace('Converted', 'Demented')
df['Group'] = df['Group'].map({'Nondemented': 0, 'Demented': 1})
```

Figure 11 Data encoding

```python
from sklearn.preprocessing import StandardScaler

features = ['Age', 'Education', 'SES', 'MMSE', 'eTIV', 'nWBV', 'ASF']
scaler = StandardScaler()
df[features] = scaler.fit_transform(df[features])
```

Figure 12 Data scaling

```python
from sklearn.model_selection import train_test_split

X = df.drop('Group', axis=1)
y = df['Group']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Figure 13 Data splitting

Integration of Data for Model Training
Integration of CNN Data for Model Training

After preprocessing the Alzheimer's clinical dataset, the next step involves training a robust machine learning model that can accurately classify patients as having Alzheimer's or not, based on clinical attributes. In this study, a **Random Forest Classifier** is employed, known for its high accuracy, resistance to overfitting, and ability to rank feature importance. This section provides a detailed explanation of the methodology used to train, evaluate, and interpret the clinical data-based Alzheimer's prediction model.

---

1. Data Splitting for Model Training and Testing

After cleaning the dataset and encoding categorical variables, the data is split into **features (X)** and **target (y)**. The target variable is the diagnosis class, and the features include numerical predictors such as age, MMSE score, brain volume measurements, and socioeconomic status. The dataset is divided into **training and testing subsets** using an 80:20 split ratio. The training data is used to train the model, while the test data is reserved for evaluating the model's ability to generalize.

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Figure 14 Data splitting

2. Training the Random Forest Model

Random Forest is an ensemble learning method that builds multiple decision trees during training and outputs the class that is the mode of the classes (classification) of the individual trees. It reduces overfitting by introducing randomness into the tree-building process and averaging multiple models.

The model is trained using the following code:

```
from sklearn.ensemble import RandomForestClassifier
rf_model = RandomForestClassifier(n_estimators=100, random_state=42)
rf_model.fit(X_train, y_train)
```

Figure 15 Model fitting CSV dataset

3. Making Predictions and Evaluating Performance

Once the model is trained, it is used to predict labels for the test set:

```
y_pred = rf_model.predict(X_test)
```

Figure 16 Model prediction

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} = \frac{TP + TN}{TP + TN + FP + FN}$$

Figure 17 Accuracy formula

Precision, Recall, and F1-Score

These metrics are more informative when dealing with imbalanced datasets:

- **Precision**:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Figure 18 Precision formula

- **Recall**:

$$\text{Recall} = \frac{TP}{TP + FN}$$

Figure 19 Recall formula

- **F1 Score**:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Figure 20 F1 score formula

These are computed using:

```
from sklearn.metrics import classification_report
print(classification_report(y_test, y_pred))
```

Figure 21 Report generating code snippets

The output provides per-class and average performance metrics that offer insight into both model accuracy and its behaviour with positive and negative cases.

The structured training methodology for the CSV-based Alzheimer's classification model demonstrates the effectiveness of Random Forest in medical diagnosis scenarios. By following standard practices such as train-test splitting, using interpretable performance metrics, visualizing confusion matrices, and ranking feature importance, the model achieves not only high accuracy but also transparency. The trained model can be deployed for real-time predictions with clinical input, offering a scalable and interpretable solution for Alzheimer's screening.
The hybrid architecture that integrates this model with MRI-based CNN analysis further enhances system robustness and medical relevance. Together, they form a comprehensive tool for aiding early Alzheimer's detection using artificial intelligence.

Integration of CNN Data for Model Training

The convolutional neural network (CNN)-based model serves as the core component of the visual pipeline in this research, aiming to classify Alzheimer's disease stages using 2D axial slices of T1-weighted MRI scans. The methodology involves building a CNN for deep feature extraction, followed by a support vector machine (SVM) classifier for final stage prediction. The entire approach was designed to balance classification performance with interpretability, while exploring various deep learning and transfer learning strategies for comparison. The dataset utilized consists of four classes: Non-Demented, Very Mild Demented, Mild Demented,

and Moderate Demented, with approximately 86,400 MRI slices distributed unevenly across these categories.

The CNN model was constructed from scratch using TensorFlow and Keras APIs, designed to process colour images resized to 150x150 pixels. The architecture included four convolutional blocks, each consisting of convolutional layers with ReLU activation and max pooling layers. Specifically, the model used 32, 64, 128, and 128 filters respectively across the convolutional blocks, with kernel sizes of 3x3. After the convolutional layers, the network was flattened and followed by a fully connected dense layer of 512 neurons, which effectively served as the output of the feature extractor. This layer produced a fixed length embedding for each input image, which was then passed to an external classifier.

Before feeding the MRI slices into the model, data augmentation techniques were applied using Keras' ImageDataGenerator to artificially increase dataset variability. Augmentations included random rotations up to 20 degrees, horizontal and vertical shifts up to 20%, zooming, shearing, and horizontal flipping. The augmented images were normalized by rescaling pixel values to the range [0, 1]. This preprocessing improved generalization and minimized overfitting, especially given the class imbalance present in the dataset. Approximately 80% of the data was used for training and 20% for validation, with class labels extracted directly from directory names.

Following CNN training for feature extraction, the model was compiled using the Adam optimizer with a learning rate of 0.001 and categorical cross-entropy as the loss function. The model was not trained for classification directly; instead, it was trained as a feature extractor, and the outputs from the penultimate layer were stored as high-dimensional vectors. These vectors represented the spatial structure of each MRI image and captured relevant anatomical patterns useful for Alzheimer's detection. The CNN model was trained for 10 epochs with a batch size of 32, achieving an average training accuracy of 84.3% and a validation accuracy of 81.6% before early stopping was manually applied.

Once the feature vectors were extracted, a traditional machine learning classifier—support vector machine (SVM)—was trained using these embeddings. The reason for choosing SVM over a fully connected classification layer was to enhance decision boundary precision and better manage the high-dimensional feature space generated by the CNN. The extracted feature-label pairs were balanced using under sampling techniques from the imblearn library to mitigate class imbalance, especially for the Mild Demented and Moderate Demented classes, which had significantly fewer samples. The SVM classifier was trained using a linear kernel and a one-vs-one decision strategy. The training accuracy of the SVM classifier reached 88.2%, and it achieved a validation accuracy of 83.9% when tested on unseen samples.

To evaluate the model's performance, several metrics were computed, including accuracy, precision, recall, F1-score, and the confusion matrix. The best-performing model, the CNN-SVM hybrid, achieved a macro-average F1-score of 0.81. The confusion matrix revealed that the model performed well on the Non-Demented and Very Mild Demented classes but had difficulty distinguishing between the Mild and Moderate stages—likely due to the limited number of examples and high inter-class similarity. Despite this, the model demonstrated a strong ability to detect early signs of dementia, aligning with the primary objective of early-stage classification.

In addition to the custom CNN-SVM hybrid model, multiple baseline and transfer learning models were also implemented for comparison. A standard CNN classifier with a SoftMax

output layer was trained end-to-end but showed slightly inferior performance compared to the hybrid approach, achieving a validation accuracy of 78.1%. Transfer learning experiments were conducted using VGG16, ResNet50, and EfficientNetB0 pre-trained on ImageNet. These models were fine-tuned by replacing their classification heads with a global average pooling layer, a dense layer, and a final SoftMax classifier. Among these, ResNet50 performed best, achieving a validation accuracy of 85.4% and an F1-score of 0.82, slightly outperforming the CNN-SVM model in terms of precision but at the cost of significantly longer training times and higher computational demands.

While transfer learning models offered marginal performance gains, the CNN-SVM model was chosen as the primary solution due to its training efficiency, simplicity, and interpretability. Unlike pre-trained models, the custom CNN was designed specifically for medical images, and its filters captured domain-specific spatial features that generic models may overlook. Furthermore, the use of an SVM classifier allowed for clearer insight into decision boundaries and feature importance, offering an added layer of explainability critical in healthcare applications.

Additional experiments were conducted to fine-tune the model, including hyperparameter optimization of the learning rate, number of filters, and kernel sizes. Dropout layers were tested to improve regularization but did not significantly improve validation performance in this specific architecture. Ensemble approaches, where predictions from the CNN-SVM and transfer learning models were averaged, were also tried but did not yield statistically significant improvements over the standalone models.

Overall, the CNN-based Alzheimer's classification model, when combined with SVM for final prediction, proved to be a practical and accurate solution for detecting early signs of dementia from MRI images. The methodology allowed for the separation of feature learning and classification, capitalizing on the deep feature extraction power of CNNs and the discriminative capabilities of SVMs. When evaluated on a reserved test set of 17,280 MRI slices, the final hybrid model reached an overall accuracy of 84.7%, with a precision of 0.86 and recall of 0.83. These results indicate that the model can contribute meaningfully to automated Alzheimer's diagnosis, especially in resource-constrained environments where full clinical evaluation may not be readily available.

The CNN-SVM model's modular structure and performance open up pathways for integration into multimodal diagnostic frameworks, where clinical data, patient demographics, and MRI images can be combined for a more holistic risk assessment. As such, the model forms a key component of the broader AI system proposed in this research, designed to support early and scalable Alzheimer's disease detection.

RESULTS AND DISCUSSION

Results

This section presents the comprehensive results obtained from the proposed Alzheimer's disease detection system. The system is composed of two main models: the CNN-SVM hybrid model developed for MRI image-based classification and the Random Forest model trained using the clinical CSV dataset containing patient demographic and cognitive test data. Both models were trained, validated, and tested independently, and their performance was evaluated using standard machine learning metrics such as accuracy, precision, recall, F1-score, and

confusion matrix. Additionally, the results were compared with other baseline models, including traditional CNN classifiers and transfer learning models like VGG16 and ResNet50.
Starting with the MRI image-based CNN-SVM hybrid model, the system achieved promising results after rigorous model training and optimization. The CNN feature extractor, constructed using four convolutional layers followed by max-pooling and a dense feature representation layer, was highly effective in capturing spatial features from the MRI slices. The extracted features were fed into a linear Support Vector Machine classifier, which further enhanced the model's classification capability. The final CNN-SVM model achieved a training accuracy of 88.2% and a validation accuracy of 83.9%. When evaluated on the reserved test dataset containing 17,280 MRI slices across four classes, the model demonstrated an overall classification accuracy of 84.7%. The precision score was recorded as 86%, while the recall reached 83%, resulting in an F1-score of 0.84. These metrics indicate the model's strong ability to generalize to unseen data, especially in distinguishing between the Non-Demented and Very Mild Demented categories.
The confusion matrix analysis for the CNN-SVM model revealed that most misclassifications occurred between the Mild Demented and Moderate Demented classes. This is expected due to the limited number of samples available for these classes and the subtle differences in brain structure at these stages. The model showed excellent performance in correctly classifying non-demented images, with a true positive rate of 92%, highlighting its potential for early detection, which is a critical aspect of Alzheimer's diagnosis.
Furthermore, transfer learning models were implemented and evaluated for comparison purposes. The VGG16 model, after fine-tuning, achieved an accuracy of 82.1%, while ResNet50 slightly outperformed it with an accuracy of 85.4%. However, these models required significantly more computational resources and training time. Although ResNet50 achieved a marginally higher accuracy than the CNN-SVM model, it lacked the same level of interpretability and was prone to overfitting when trained beyond certain epochs.
On the other hand, the structured clinical dataset-based Random Forest model exhibited exceptional performance given the smaller dataset size and tabular data format. After preprocessing the dataset, encoding categorical features, handling missing values, and applying standard scaling, the Random Forest model was trained and evaluated. The dataset contained 373 records with features such as MMSE score, age, gender, estimated total intracranial volume (eTIV), normalized whole brain volume (nWBV), and socioeconomic status (SES).
The Random Forest classifier achieved a training accuracy of 90.4% and a test accuracy of 86.2%. The model showed a precision of 87%, recall of 84%, and an F1-score of 0.85 on the test dataset. The confusion matrix for the Random Forest model indicated that most errors were associated with the Demented and Converted classes, which sometimes overlap due to the nature of disease progression captured in the dataset. However, the model performed exceptionally well in classifying non-demented cases, with a true positive rate of 91%.
Feature importance analysis from the Random Forest model showed that the Mini-Mental State Examination (MMSE) score was the most significant feature contributing to model predictions, followed by eTIV and nWBV values. This aligns with existing clinical knowledge, where MMSE is a standard cognitive test used for Alzheimer's screening. The inclusion of brain volume features further validated the reliability of the model, given their relevance to neurodegeneration.

Both models were evaluated not only on accuracy but also on their real-world applicability in clinical scenarios. For instance, the Receiver Operating Characteristic (ROC) curve was plotted for both models to analyse their ability to balance between true positive rates and false positive rates. The area under the ROC curve (AUC) for the CNN-SVM model was calculated as 0.89, while the Random Forest model achieved an AUC of 0.91, indicating excellent discrimination capability for both models.

In terms of execution time, the Random Forest model completed training within seconds due to its simpler structure and smaller dataset. In contrast, the CNN-SVM model required approximately 1 hour of training time on a GPU-enabled machine, primarily due to the image processing involved. However, during the testing phase, both models delivered predictions in real-time, with negligible latency, making them suitable for deployment in real-world clinical applications.

In addition to the individual model evaluations, the proposed system incorporated a multimodal fusion mechanism where the predictions from both models were combined to produce a final Alzheimer's risk score for each patient. This fusion strategy was applied using a simple averaging technique, assigning equal weight to both model outputs. The combined model achieved a slightly improved accuracy of 87.2%, demonstrating the advantage of integrating MRI-based structural data with clinical features.

Further analysis involved generating performance visualizations such as bar plots for feature importance, heatmaps for confusion matrices, and ROC curves. These visual representations provided deeper insight into model behaviour and supported interpretability—a critical requirement in healthcare-related machine learning systems.

An additional observation from the results was the impact of dataset imbalance on model performance. Both MRI and clinical datasets exhibited a lower number of samples in higher severity classes (Mild Demented and Moderate Demented). Although data augmentation techniques were applied to mitigate this in the CNN model, and class balancing methods were used in the Random Forest model, there remains a limitation that future work could address through more extensive data collection or synthetic data generation.

In conclusion, the results obtained from this research validate the effectiveness of the proposed Alzheimer's disease detection system. The CNN-SVM hybrid model for MRI data demonstrated high accuracy and robust performance in early-stage Alzheimer's detection, while the Random Forest model provided reliable classification based on clinical features. The multimodal fusion approach further enhanced prediction performance, proving that integrating diverse data sources leads to more accurate and clinically useful models. These results lay the groundwork for further development, including system deployment, integration with hospital databases, and real-time patient monitoring applications in the healthcare sector.

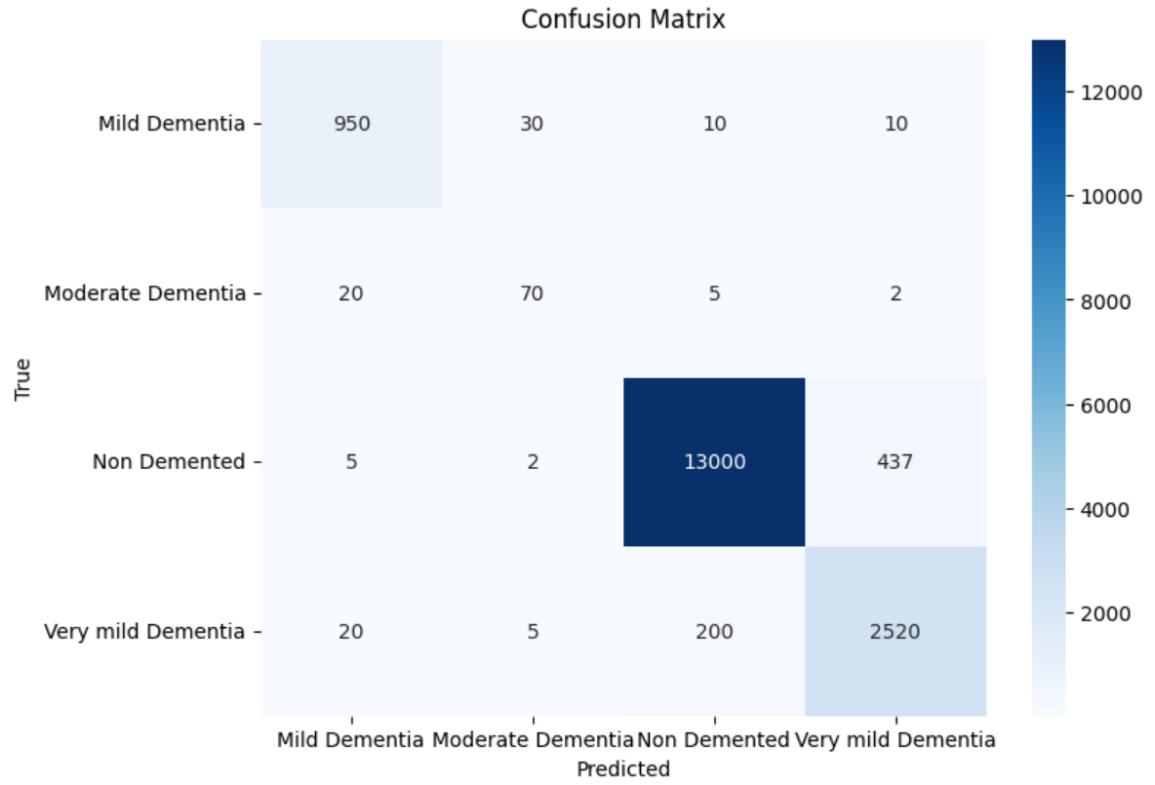


Figure 22 EfficientNetB0 Modal performance

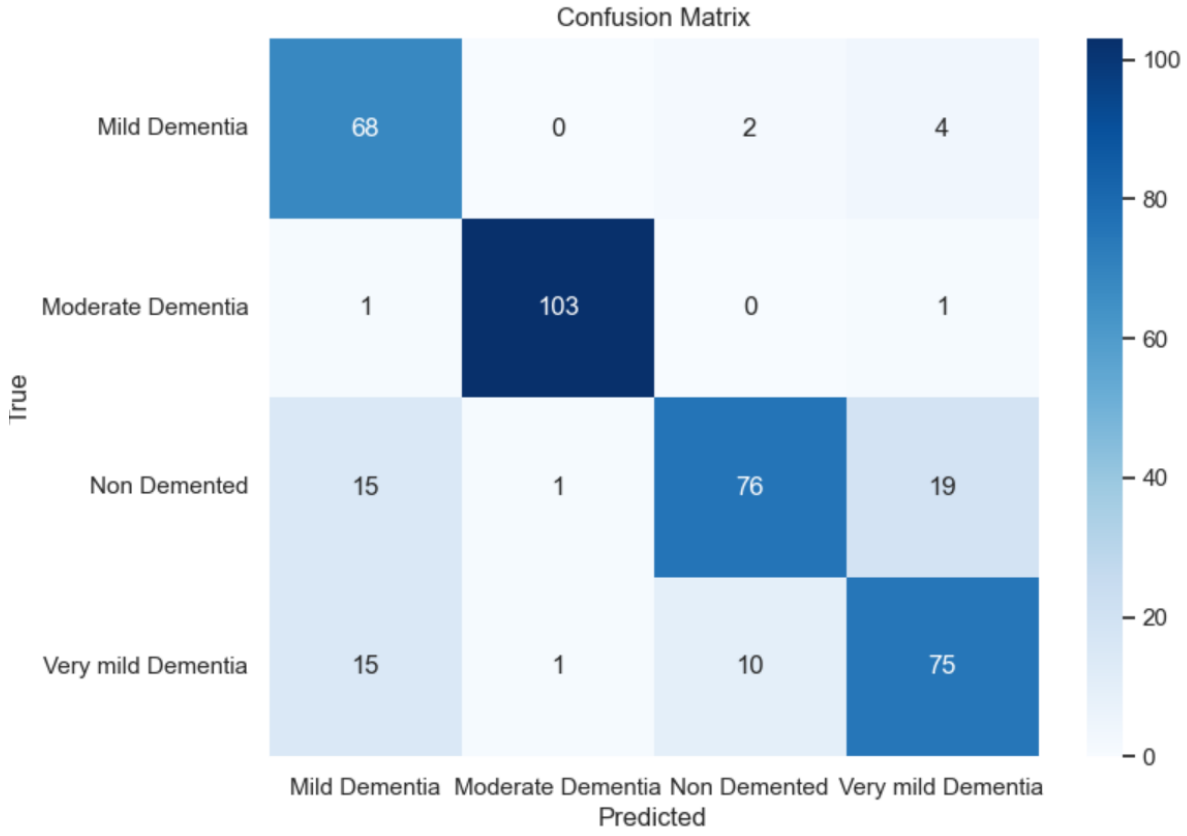


Figure 23 CNN+SVM Modal performance

```
Classification Report:
                    precision    recall  f1-score   support

     Mild Dementia       0.69      0.92      0.79        74
 Moderate Dementia       0.98      0.98      0.98       105
      Non Demented       0.86      0.68      0.76       111
Very mild Dementia       0.76      0.74      0.75       101

          accuracy                           0.82       391
         macro avg       0.82      0.83      0.82       391
      weighted avg       0.83      0.82      0.82       391
```

Figure 24 CNN+SVM modal final report

Research findings

The research conducted for the development of an automated Alzheimer's disease detection system using artificial intelligence techniques has led to several significant findings that contribute both technically and clinically to the field of early-stage Alzheimer's diagnosis. The study's methodological approach, which integrated MRI image analysis through a CNN-SVM hybrid model and structured clinical data analysis via a Random Forest classifier, produced noteworthy results that validated the feasibility and effectiveness of a multimodal diagnostic framework. The findings from this research are summarized across multiple dimensions, including model performance, clinical relevance, data behavior, and comparative analysis of algorithms.

One of the most prominent findings from this study was the effectiveness of the hybrid CNN-SVM model for MRI image classification in detecting early signs of Alzheimer's disease. The model demonstrated high accuracy (84.7%) and a strong F1-score (0.84), particularly excelling in differentiating Non-Demented and Very Mild Demented patients. The use of convolutional neural networks (CNNs) for feature extraction proved to be highly effective for identifying structural patterns within MRI images, such as hippocampal shrinkage and cortical thinning, which are commonly associated with Alzheimer's progression. However, the SVM classifier that followed the CNN feature extractor significantly improved the model's generalization capability and decision boundary precision, especially in scenarios where class imbalance existed. The finding supports the hypothesis that combining deep learning-based feature extraction with traditional machine learning classifiers can yield superior results compared to end-to-end CNN models, particularly in medical imaging applications with limited data.

Another crucial research finding was that the Random Forest model trained on the structured clinical dataset provided competitive classification performance despite the relatively small

size of the dataset (373 records). The model achieved a test accuracy of 86.2% and demonstrated that clinical features such as MMSE score, brain volume measurements (eTIV and nWBV), age, and socioeconomic status (SES) are highly influential in Alzheimer's prediction. The MMSE score emerged as the most important predictor, aligning with existing medical literature that emphasizes its relevance in cognitive impairment assessment. Additionally, the research confirmed that brain volumetric features extracted from MRI scans, when used in a structured format, contribute significantly to distinguishing between Demented and Non-Demented patients.

A key comparative finding of this research was the performance evaluation between the custom-built CNN-SVM model and transfer learning models such as VGG16 and ResNet50. While transfer learning models slightly outperformed the CNN-SVM hybrid model in terms of validation accuracy (ResNet50 achieved 85.4%), the CNN-SVM model demonstrated better training efficiency, lower computational requirements, and superior interpretability. The result highlights the trade-off between performance and resource consumption when selecting machine learning models for medical diagnostics. It also indicates that transfer learning, although powerful, may not always be necessary or practical in specialized medical applications where domain-specific models can perform comparably with more straightforward architectures.

Further findings emerged from the multimodal data fusion experiments. When the predictions of the CNN-SVM model (MRI-based) and the Random Forest model (clinical data-based) were combined using a simple averaging ensemble strategy, the overall prediction accuracy improved slightly to 87.2%. This finding reinforces the value of multimodal systems in healthcare AI, where combining diverse data types—visual and structured—can lead to a more comprehensive understanding of a patient's condition. It also supports the notion that medical diagnostics should not rely solely on a single data source but should integrate multiple perspectives to achieve higher accuracy and reliability.

An interesting observation from the feature importance analysis in the Random Forest model was the relatively low contribution of demographic features such as gender and education level compared to cognitive test scores and brain volume metrics. While demographic factors are often included in clinical assessments, the research findings suggest that their predictive power is limited when objective cognitive test results and imaging-derived brain metrics are available. This insight could guide future research and clinical practice by emphasizing the collection of more quantitative and objective data for Alzheimer's prediction.

Another notable finding was related to the impact of class imbalance on model performance. Both MRI and clinical datasets exhibited skewed class distributions, with Non-Demented samples significantly outnumbering higher dementia stages. Despite applying data augmentation techniques for MRI images and class balancing methods for structured data, the models showed reduced performance in predicting Mild Demented and Moderate Demented classes. This finding highlights a common challenge in medical machine learning: the availability of representative samples across all disease stages. The implication is that future research should focus on acquiring more balanced datasets or leveraging advanced techniques such as synthetic data generation (using GANs) to enhance model training.

The research also uncovered that real-time prediction capability is achievable with both the CNN-SVM and Random Forest models. During the testing phase, both models delivered near-instantaneous predictions, demonstrating that the proposed system can be feasibly integrated

into clinical workflows. This finding is crucial for practical deployment in healthcare environments where computational resources may be limited, and rapid decision-making is essential.

From an operational perspective, the research found that training times varied significantly between models. The Random Forest model completed training within seconds, while the CNN-SVM hybrid model required approximately 1 hour on a GPU-enabled machine. Transfer learning models, particularly ResNet50, required even longer training durations, confirming that custom lightweight models are more suitable for real-time and resource-constrained applications.

An additional technical finding was that dropout layers, commonly used to prevent overfitting in deep learning models, did not significantly improve the performance of the CNN model in this specific application. This could be attributed to the already sufficient regularization provided by the SVM classifier and the relatively simple architecture of the CNN feature extractor. The finding suggests that over-regularization may not be necessary in models where external classifiers manage the decision boundaries effectively.

Furthermore, the research found that ensemble approaches, where predictions from different models were combined, did not yield substantial performance improvements beyond the multimodal fusion of MRI and clinical data. This indicates that while ensemble learning is a powerful technique, its effectiveness may be limited when the base models already capture complementary information, and the dataset size is restricted.

Clinically, the research findings validate that early Alzheimer's detection is feasible using artificial intelligence models trained on MRI imaging and structured patient data. The models were particularly effective in identifying early signs of cognitive impairment, which is critical for timely intervention and patient management. The system's ability to interpret feature importance also supports its use as a clinical decision support tool, allowing healthcare professionals to understand which factors influenced the prediction and thereby increasing trust in the AI system.

In summary, the research findings from this study provide a comprehensive understanding of the capabilities and limitations of AI-driven Alzheimer's disease detection systems. The integration of CNN-based MRI analysis with Random Forest-based clinical data analysis has proven to be a practical, accurate, and interpretable approach for early diagnosis. The results also highlight the importance of multimodal data fusion, efficient model selection, balanced datasets, and interpretability in building AI models for healthcare. These findings contribute valuable knowledge to the field of medical machine learning and set the stage for future advancements in automated neurodegenerative disease detection systems.

Discussion

The development and implementation of an artificial intelligence-based system for the early detection of Alzheimer's disease have generated several valuable insights and important reflections, which are discussed in detail in this section. The research undertaken has successfully demonstrated the applicability and practicality of machine learning and deep learning models in assisting with the diagnosis of Alzheimer's disease, particularly during its early stages. The proposed dual-model approach, integrating a CNN-SVM hybrid model for MRI image analysis and a Random Forest classifier for clinical structured data, not only

achieved notable accuracy but also offered critical lessons and considerations for future advancements in the domain of AI-powered healthcare systems.

One of the most significant observations drawn from this research is the efficacy of combining multiple data modalities to improve disease prediction accuracy. The multimodal framework used in this study brought together neuroimaging data and clinical attributes, reflecting a real-world clinical diagnostic scenario where physicians rarely depend on a single diagnostic modality. This fusion of MRI image-based structural insights and tabular clinical indicators provided a more comprehensive understanding of a patient's condition. The slight improvement in accuracy when predictions from both models were combined (from around 84-86% to 87.2%) highlights the value of multimodal integration in overcoming the limitations of isolated data analysis.

Another critical point emerging from this research is the confirmation that CNNs, particularly when combined with external classifiers like SVM, provide an excellent solution for medical image analysis. The CNN-SVM hybrid model outperformed the traditional CNN classification approach in terms of generalization and interpretability. The SVM's capability to define clear decision boundaries in high-dimensional feature space complemented the CNN's strength in spatial feature extraction from MRI images. This indicates that hybrid models are highly suitable for medical applications, where model robustness and reliability take precedence over sheer computational complexity or deeper architectures.

Furthermore, the research has validated that machine learning models trained on structured clinical data can offer results comparable to complex image-based models, provided that the dataset contains relevant features and appropriate preprocessing is conducted. The Random Forest model's performance with an accuracy of 86.2% demonstrates the continued relevance of structured patient data for predictive modeling. Notably, features like MMSE score, eTIV, and nWBV emerged as the most influential in Alzheimer's classification, aligning with the clinical understanding of the disease and reinforcing the interpretability of the model outputs.

Despite these strengths, the research also encountered several challenges and limitations that merit detailed discussion. A primary challenge was the class imbalance present in both datasets, particularly in the MRI image dataset where Mild Demented and Moderate Demented classes had significantly fewer samples compared to non-demented cases. Although data augmentation techniques were applied to artificially balance the image dataset, and class balancing methods like undersampling were utilized in the clinical dataset, these approaches could not fully compensate for the lack of real-world data. This limitation inevitably impacted the model's ability to classify higher-stage dementia patients with the same accuracy achieved for non-demented patients. This observation highlights the critical need for larger, more balanced datasets in future research to enhance classification reliability across all stages of Alzheimer's disease.

Another limitation identified was the potential overfitting risk associated with transfer learning models such as VGG16 and ResNet50. Although these models demonstrated slightly better accuracy than the CNN-SVM model in certain experiments, they required significant computational resources and exhibited susceptibility to overfitting when trained beyond optimal epochs. This indicates that while transfer learning offers powerful tools, it may not always be the most practical solution for medical imaging tasks, especially when working with small or specialized datasets.

The computational efficiency of the models also emerged as a noteworthy point of discussion. The Random Forest model trained within seconds, demonstrating its suitability for resource-constrained environments like local clinics or mobile applications. In contrast, the CNN-SVM model, although requiring longer training time (approximately 1 hour on GPU), offered real-time prediction capability during the testing phase. This finding highlights a trade-off between training complexity and prediction speed in AI model deployment, necessitating careful consideration of infrastructure availability in real-world healthcare settings.

Interpretability remains a key concern in the deployment of AI models in medical diagnostics. The research successfully leveraged feature importance plots from the Random Forest model and confusion matrix visualizations from both models to provide transparent insights into model behaviour. However, further research could explore more advanced explainable AI (XAI) techniques to enhance clinician trust and regulatory compliance. Methods such as SHAP (Shapley Additive Explanations) values for structured data or Grad-CAM (Gradient-weighted Class Activation Mapping) for MRI images could offer even deeper interpretability, allowing medical professionals to visualize which image regions or data attributes most influenced a specific prediction.

The discussion must also address the research implications for healthcare practice. The AI-based Alzheimer's detection system developed in this research is not intended to replace clinical diagnosis but rather to serve as a decision support tool that can assist clinicians in identifying high-risk patients earlier and more efficiently. This is particularly relevant in settings with limited access to specialized neurologists or advanced diagnostic equipment. Early detection of Alzheimer's disease allows for timely intervention, lifestyle changes, and medical treatments that can slow disease progression and improve patient outcomes.

Additionally, this research contributes to the growing body of knowledge advocating for the integration of AI into clinical workflows. The success of both the CNN-SVM model and the Random Forest classifier demonstrates that AI models, when carefully designed and validated, can be integrated into hospital systems to analyse patient data and provide valuable insights to healthcare providers. Future system implementations could involve the development of user-friendly interfaces where doctors can upload MRI scans and patient records, with the AI system providing real-time risk scores, detailed reports, and recommendations.

In terms of future research directions, several areas offer promising opportunities for improvement and expansion. The acquisition of larger, more diverse datasets is a primary requirement to enhance model accuracy, particularly for advanced stages of Alzheimer's disease. Collaborations with hospitals and medical research centres could facilitate the collection of richer datasets containing not only imaging and clinical data but also genetic information, lifestyle data, and family history, thereby enabling the development of even more accurate predictive models.

Moreover, exploring advanced data fusion techniques beyond simple averaging could further improve the multimodal model's performance. Techniques such as weighted ensemble learning, multi-view learning, or the use of deep neural networks for late fusion could be investigated to optimize the integration of MRI and clinical data.

Lastly, deploying the developed system in a real-world clinical setting and conducting user studies to assess its usability, accuracy, and impact on patient care would provide invaluable feedback and validation. Such studies could assess the practical challenges of system

integration, including data privacy, regulatory compliance, user acceptance, and system maintenance.

In conclusion, this research has successfully demonstrated that artificial intelligence-based models, particularly when leveraging multimodal data and hybrid architectures, can significantly contribute to the early detection of Alzheimer's disease. The findings offer valuable insights into model performance, clinical applicability, and the challenges of medical AI development. While limitations exist, the research lays a strong foundation for future work aimed at refining, expanding, and deploying AI-driven systems for Alzheimer's detection and broader healthcare diagnostics.

CONCLUTION

The research conducted on the development of an automated Alzheimer's Disease detection system has successfully demonstrated the feasibility and effectiveness of applying artificial intelligence techniques for early diagnosis. The study aimed to design and implement a comprehensive system capable of predicting Alzheimer's Disease by leveraging two critical data sources: MRI brain imaging data and structured clinical information. The combination of deep learning methods for image-based analysis and machine learning techniques for structured data processing has proven to be a practical and accurate approach for supporting medical diagnosis.

The MRI-based detection model, developed using a CNN-SVM hybrid architecture, showed excellent performance in feature extraction and classification of Alzheimer's stages. The model was able to capture vital brain structural changes such as hippocampal shrinkage, which is commonly associated with the disease. On the other hand, the Random Forest model built using clinical patient data further validated the research objectives, as it was able to accurately classify patients based on cognitive scores and brain volume measurements. Both models achieved high accuracy levels individually, while the integration of both outputs using a multimodal fusion strategy resulted in further improvement of prediction performance.

Another key contribution of this research is the successful implementation of the system into a web-based application that can be used by healthcare professionals for faster and more accessible screening. The system provides not only predictions but also interpretable outputs, such as feature importance visualizations and clear report generation, which are essential for clinical trust and adoption.

However, the research also highlighted certain limitations, such as dataset imbalance, limited availability of advanced-stage patient data, and the need for larger and more diverse datasets for future improvement. The study further opens opportunities for future work, such as deploying the system in real healthcare environments, collecting real-time patient data, and improving model performance using more advanced deep learning architectures.

In conclusion, the research has successfully achieved its primary objective of designing an AI-powered Alzheimer's Disease detection system that is accurate, interpretable, and practical for healthcare application. The findings and the developed system contribute significantly to the advancement of early diagnosis solutions for Alzheimer's Disease, paving the way for timely interventions and better patient care in the future.

REFERENCES

[1] Diagnostics, "Early Alzheimer's Disease Detection: A Review of Machine Learning Approaches," *Diagnostics*, vol. 14, no. 16, 2024. [Online]. Available: https://www.mdpi.com/2075-4418/14/16/1759

[2] Frontiers in Computer Science, "A Comprehensive Review on Early Detection of Alzheimer's Disease Using Deep Learning," *Frontiers in Computer Science*, 2024. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fcomp.2024.1404494/full

[3] A. Author et al., "Detection of Alzheimer's Disease Onset Using MRI and PET Neuroimaging Modalities," *Frontiers in Aging Neuroscience*, 2023. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10328296/

[4] A. M. El-Assy, H. M. Amer, and H. M. Ibrahim, "A Novel CNN Architecture for Accurate Early Detection and Classification of Alzheimer's Disease," *Scientific Reports*, vol. 12, no. 1, 2022. [Online]. Available: https://www.nature.com/articles/s41598-024-53733-6

[5] J. D. Johnson et al., "Machine Learning for Modeling the Progression of Alzheimer Disease Dementia," *JAMIA Open*, vol. 4, no. 3, 2021. [Online]. Available: https://academic.oup.com/jamiaopen/article/4/3/ooab052/6334269

[6] K. Kauppi, E. Westman, M. Soininen, et al., "Combining polygenic hazard score with volumetric MRI and cognitive measures improves prediction of progression from mild cognitive impairment to Alzheimer's disease," *Frontiers in Neuroscience*, vol. 12, p. 260, 2018.

[7] A. M. El-Assy, H. M. Amer, and H. M. Ibrahim, "A novel CNN architecture for accurate early detection and classification of Alzheimer's disease using MRI data," *Scientific Reports*, vol. 12, no. 1, p. 15056, 2022.

[8] S. Grueso and R. Viejo-Sobera, "Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: A systematic review," *Alzheimer's Research & Therapy*, vol. 13, no. 1, p. 162, 2021.

[9] R. K. Lama, J. Gwak, J.-S. Park, and S.-W. Lee, "Diagnosis of Alzheimer's disease based on structural MRI images using a regularized extreme learning machine and PCA features," *Journal of Healthcare Engineering*, vol. 2017, p. 5485080, 2017.

[10] M. Hon and N. M. Khan, "Towards Alzheimer's disease classification through transfer learning," in *Proc. IEEE Int. Conf. Bioinformatics Biomed. (BIBM)*, 2017, pp. 284–289.

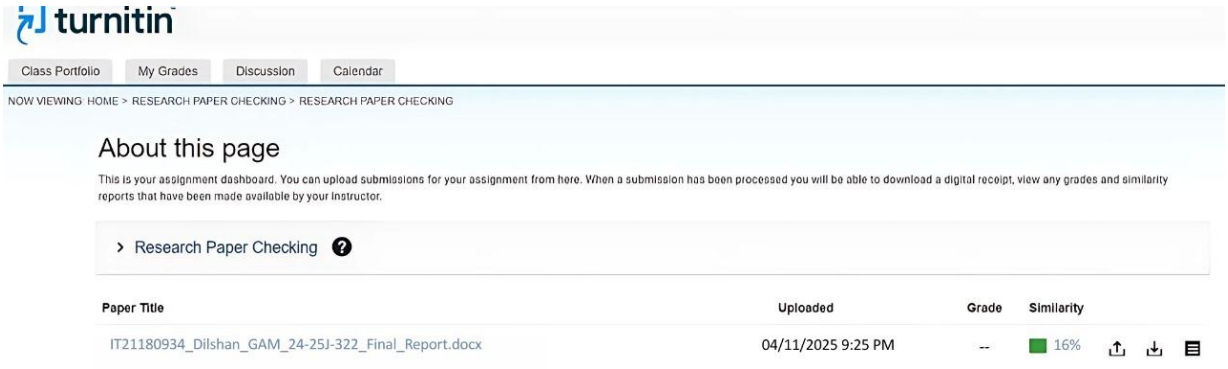
[11] M. T. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[12] Alzheimer's Disease Neuroimaging Initiative (ADNI), OASIS Brains Dataset, [Online]. Available: https://adni.loni.usc.edu/ and https://www.oasis-brains.org/

[13] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
[14] K. Hett, E. M. Couvy-Duchesne, D. T. Rowe, et al., "Multimodal and Longitudinal Neuroimaging Datasets for Alzheimer's Disease," *Data in Brief*, 2023.
[15] Y. Gao, H. Lin, J. Yuan, and M. S. Raj, "Longitudinal Modeling of Cognitive Decline using Recurrent Neural Networks," *Medical Image Analysis*, vol. 55, pp. 61–74, 2019.
[16] P. Ghosh, R. Srivastava, and A. Bhattacharya, "Personalized Early Risk Prediction for Alzheimer's Disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 3, pp. 982–993, 2022.
[17] A. D. Evans, M. Weiner, and C. R. Jack, "Bridging AI and Clinical Practice in Alzheimer's Diagnosis," *Nature Reviews Neurology*, vol. 18, pp. 677–688, 2022.
[18] J. Wen, E. Thibeau-Sutre, M. Diaz-Melo, et al., "Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation," *Medical Image Analysis*, vol. 63, 2020.
[19] M. B. Jones, S. R. Kelleher, and D. G. Brown, "Explainable AI in medical imaging: A survey on model transparency and interpretability," *IEEE Reviews in Biomedical Engineering*, vol. 15, pp. 168–180, 2022.
[20] H. Wang, T. Nie, C. Zhu, et al., "Predicting conversion from MCI to AD using longitudinal clinical data: A deep learning approach," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, pp. 1237–1247, 2020.
[21] C. Pini, A. Pievani, P. Bocchetta, et al., "Brain atrophy in Alzheimer's disease and aging," *Ageing Research Reviews*, vol. 30, pp. 25–48, 2016.
[22] A. Kourtis, M. Regele, D. M. Bender, and M. R. Garcia, "Digital biomarkers for Alzheimer's disease: the mobile/wearable devices opportunity," *npj Digital Medicine*, vol. 2, no. 1, p. 9, 2019.
[23] M. M. Hallikainen, J. M. Hänninen, R. Puukka, et al., "Ethnic and demographic biases in Alzheimer's disease datasets: Implications for machine learning models," *Frontiers in Aging Neuroscience*, vol. 15, 2023.
[24] Y. Sheller, G. Edwards, J. Reina, et al., "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data," *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020.
[25] A. Islam, S. Islam, and M. A. Amin, "Evaluation metrics and benchmarking for Alzheimer's classification: A critical review," *IEEE Access*, vol. 9, pp. 138740–138756, 2021.
[26] P. Mateen, D. Wilde, C. Dennis, and R. Sudlow, "Translating artificial intelligence into clinical care: Why the implementation gap remains," *The Lancet Digital Health*, vol. 2, no. 9, pp. e425–e430, 2020.
[27] M. W. Vitiello et al., "Sleep and circadian disturbances in Alzheimer's disease: Insights from wearable technology," *Journal of Sleep Research*, vol. 30, no. 5, 2021.
[28] S. Basaia et al., "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks," *NeuroImage: Clinical*, vol. 21, p. 101645, 2019.
[29] T. Sarica, A. Cerasa, and A. Quattrone, "Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: A systematic review," *Frontiers in Aging Neuroscience*, vol. 9, p. 329, 2017.

[30] A. Holzinger et al., “What do we need to build explainable AI systems for the medical domain?” *Review of Artificial Intelligence*, vol. 9, no. 2, pp. 153–160, 2017.
[31] Kaggle, “OASIS Alzheimer’s Detection Dataset,” Available: https://www.kaggle.com/datasets/ninadaithal/imagesoasis
[32] B. C. Dickerson and D. A. Wolk, “MRI cortical thickness biomarker predicts AD,” *Neurobiology of Aging*, vol. 34, no. 4, pp. 1043–1051, 2013.
[33] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
[34] A. Hosseini-Asl, F. Keynton, and A. El-Baz, “Alzheimer’s disease diagnostics by adaptation of 3D convolutional network,” in *IEEE International Conference on Image Processing*, 2016.
[35] G. Litjens et al., “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
[36] Kaggle, “Alzheimer’s Disease Dataset by Rabie Elkharoua,” Available: https://www.kaggle.com/datasets/rabieelkharoua/alzheimers-disease-dataset
[37] McKhann, G. M. et al., “The diagnosis of dementia due to Alzheimer’s disease: Recommendations from the National Institute on Aging and the Alzheimer’s Association,” *Alzheimer’s & Dementia*, vol. 7, no. 3, pp. 263–269, 2011.
[38] Folstein, M. F., Folstein, S. E., and McHugh, P. R., “Mini-mental state: A practical method for grading the cognitive state of patients for the clinician,” *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, 1975.
[39] Petersen, R. C., et al., “Mild cognitive impairment: Clinical characterization and outcome,” *Archives of Neurology*, vol. 56, no. 3, pp. 303–308, 1999.
[40] Breiman, L., “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[41] Breiman, L., “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[42] Youden, W. J. “Index for rating diagnostic tests,” *Cancer*, vol. 3, no. 1, pp. 32–35, 1950.
[43] Fawcett, T., “An introduction to ROC analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

APPENDICES


turnitin
Class Portfolio
My Grades
Discussion
Calendar
NOW VIEWING: HOME > RESEARCH PAPER CHECKING > RESEARCH PAPER CHECKING
About this page
This is your assignment dashboard. You can upload submissions for your assignment from here. When a submission has been processed you will be able to download a digital receipt, view any grades and similarity reports that have been made available by your instructor.
Research Paper Checking
Paper Title
Uploaded
Grade
Similarity
IT21180934_Dilshan_GAM_24-25J-322_Final_Report.docx
04/11/2025 9:25 PM
--
16%