

DEVELOPMENT OF A MULTI-MODEL MACHINE LEARNING BASED APPROACH FOR EARLY DETECTION OF ALZHEIMER DISEASE

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

Computer Applications

By

Shehu Mohammed

Registration Number: 12325250

Supervised By

Dr. Neha Malhotra (16982)

School of Computer Applications (Associate Professor and Deputy Dean)

Lovely Professional University



LOVELY PROFESSIONAL UNIVERSITY, PUNJAB

2026

DEDICATION

This Thesis is dedicated to my late parents and my family, who were so much loved.

DECLARATION

I, Shehu Mohammed hereby declare that this thesis entitled “Development of a Multi-Model Machine Learning based Approach for Early Detection of Alzheimer Disease” is the outcome of my original research work which was undertaken under the guidance of Dr. Neha Malhotra, Associate Professor and Deputy Dean, School of Computer Applications, Lovely Professional University, Punjab and has not been submitted or published elsewhere in the requirement of any University or Institution degree, diploma, associateships and fellowship programme. Any Literature date or work done by others and cited within this thesis has been given due acknowledgment and listed in the reference section. I further declare that the work presented in this thesis has not been copied or plagiarized from any other source.



Shehu Mohammed

Reg. No: 12325250

School of Computer Applications

Lovely Professional University

Phagwara, Punjab – 144411,

India

Date:

CERTIFICATE

This is to certify that the thesis entitled “Development of a Multi-Model Machine Learning based Approach for Early Detection of Alzheimer Disease” submitted by Shehu Mohammed for the award of the degree of **Doctor of Philosophy (PhD)** in Computer Applications, Lovely Professional University, is entirely based on the work carried out by him under my supervision and guidance. The work reported embodies the original work of the candidates and has not been submitted to any other university or Institution for the award of any degree or diploma, according to the best of my knowledge.



Dr. Neha Malhotra
Associate Professor and Deputy Dean
School of Computer Applications
Lovely Professional University,
Phagwara, Punjab-144411,
India
Date:

ACKNOWLEDGMENTS

In Allah's name, the Most Merciful and Gracious. I want to begin by thanking Allah S.W.T. for granting me the will and strength to complete this thesis and fulfil the requirements for the Ph.D. degree in Computer Applications.

This project would not have been possible without the support of the Tertiary Education Trust Fund (Tetfund) Nigeria for awarding me a sponsorship and providing me with the financial resources necessary to complete this project. I would like to say that I am very grateful to Neha Malhotra, my doctor, who gave me invaluable advice, encouragement, and support during this research. I would also like to acknowledge Dr. Amar Singh, who provided me guidance and assistance in carrying out this research. I would like to thank Arun Singh of the School of Computer Science and Engineering (LPU). Also, I like the School of Computer Applications, Lovely Professional University, because the necessary resources are offered. I would like to thank my Ph.D. Research Committee members who provided me with valuable comments and suggestions in this research. I would like to thank the management and employees of the Center of Research Degree Programmes in Lovely Professional University for their support and provide facilities in conducting this research.

I would like to give special thanks to my family and friends who supported me and did not cease to do it.

ABSTRACT

Alzheimer’s disease (AD) remains one of the most devastating neurodegenerative disorders globally, characterized by progressive cognitive decline, impairment of daily functioning, and irreversible neuronal loss. Its growing prevalence poses substantial challenges to healthcare systems, caregivers, and societies at large. Early detection of AD, particularly at the stages of Subjective Cognitive Decline (SCD) and Mild Cognitive Impairment (MCI), offers the greatest opportunity for timely interventions that may slow disease progression. Yet, conventional diagnostic approaches depend heavily on clinical judgments, single-modality assessments, or late-stage manifestations of brain atrophy, which often limit diagnostic accuracy at earlier phases of the disease. These challenges create a compelling need for computational tools capable of integrating heterogeneous information, recognizing subtle pathological signatures, and supporting clinicians through interpretable, reliable, and patient-centered decision-making systems.

This thesis addresses these gaps by developing a robust, hybrid multimodal machine learning framework that integrates structural MRI, clinical biomarkers, cognitive assessments, genomics, longitudinal trajectories, and patient similarity modeling to enhance early AD detection. The research adopts a multi-pronged modeling strategy incorporating four distinct but complementary deep learning architectures: EfficientNetV2-S, a Feedforward Neural Network (FNN), Long Short-Term Memory (LSTM) networks, and a Graph Neural Network (GNN). Each model is optimized to extract unique patterns from its respective modality: spatial features from MRI; non-imaging biomarkers and cognitive scores from clinical records; temporal dynamics from longitudinal follow-up visits; and relational structures from graph-based patient similarity relationships. Through late decision-level fusion, these modalities are merged to produce a consolidated diagnostic prediction that far exceeds the accuracy of unimodal or traditional machine learning systems.

The MRI-based pathway employs EfficientNetV2-S, an advanced convolutional neural network architecture selected for its computational efficiency, feature extraction strength, and scalability to high-resolution neuroimaging data. EfficientNetV2-S leverages Fused-MBConv blocks, compound scaling, and transfer learning from ImageNet to identify early structural degeneration patterns in the hippocampus, medial temporal lobe, and cortical regions implicated in AD. The clinical pathway utilizes an optimized FNN configured through Optuna-based hyperparameter tuning to process demographic information, neuropsychological scores

(MMSE, MoCA, ADAS-Cog), cerebrospinal fluid biomarkers, metabolic indices, and lifestyle variables. This pathway captures nonlinear interactions in tabular data that imaging alone cannot reveal. For longitudinal progression prediction, the LSTM network models biomarker trajectories across multiple clinical visits (0, 6, 12, 24 months), offering an effective means to predict MCI-to-AD conversion. Meanwhile, the GNN pathway constructs a population graph using cosine similarity between patient feature vectors, enabling the model to leverage neighborhood relationships and improve classification for borderline and heterogeneous cases.

These pathways are integrated through a weighted softmax fusion mechanism that learns the optimal contribution of each model to the final prediction. This hybrid framework achieved a state-of-the-art accuracy of 98.19%, outperforming all baseline models and multiple recent methods in AD research. Importantly, the framework demonstrated consistent performance across the ADNI dataset and a real-world clinical dataset obtained from Shrimann Superspeciality Hospital, illustrating strong generalization capability and clinical adaptability.

In addition to predictive accuracy, interpretability plays a critical role in the acceptance of AI systems in healthcare. This thesis incorporates SHAP (SHapley Additive exPlanations) as the primary explainability method to generate feature-level explanations for the clinical pathway. SHAP quantifies the contribution of each biomarker, cognitive score, or demographic factor to the model's prediction, producing transparent explanations that align with established clinical knowledge. This satisfies emerging ethical, regulatory, and legal standards for explainable AI in medical decision support systems. While the MRI pathway did not incorporate visual interpretability such as Grad-CAM, the system still provides clinicians with detailed tabular and graphical explanations of patient-specific risk factors.

Furthermore, to ensure real-world usability and accessibility, the research includes the design and development of an interactive Graphical User Interface (GUI) that allows clinicians and caregivers to upload MRI scans, enter clinical information, and receive diagnostic predictions with corresponding explanations. The system also integrates an AI-powered conversational agent named NeuroBot, designed to guide users, support patients and caregivers, and provide educational information about AD symptoms, risk factors, and disease management. This holistic design bridges the gap between algorithmic performance and practical deployment in clinical settings.

The research is strengthened by the inclusion of both publicly available datasets (ADNI) and institutionally collected clinical data from Shrimann Superspeciality Hospital. This combination ensures that the models experience real-world variations in scanning protocols, demographic distributions, and biomarker availability conditions that often challenge machine learning models trained on homogeneous datasets.

This study also acknowledges critical ethical considerations and regulatory challenges associated with AI-driven medical diagnostics. Issues such as patient privacy, informed consent, data governance, algorithmic bias, and model traceability are discussed. These considerations are further supported by insights from the published review paper on ethical and regulatory barriers in machine learning healthcare systems. Ensuring that the proposed framework adheres to guidelines such as GDPR, HIPAA, CDSA, and local institutional review protocols is essential for safe and responsible deployment.

Throughout the research process, four peer-reviewed publications were produced, including a comprehensive review paper and multiple empirical studies on multimodal AD detection. These publications reinforced different components of the thesis: system design, multimodal integration, longitudinal modeling, clinical applications, and ethical analysis.

Overall, this thesis demonstrates that multimodal integration when implemented through modern deep learning architectures and accompanied by interpretability, ethical safeguards, and deployment-ready interfaces significantly improves the early detection and clinical applicability of Alzheimer’s disease diagnostic systems. The hybrid framework offers a scalable blueprint for precision neurology, enabling earlier intervention, improved patient management, and potentially better long-term outcomes.

In conclusion, this work contributes a technically strong, clinically relevant, and ethically grounded AI system that advances the frontier of early Alzheimer’s disease detection. Despite limitations such as variability in clinical datasets, incomplete modalities for some patients, and the need for larger multi-hospital validation, the framework establishes a robust foundation for future improvements. Future research may incorporate federated learning for multi-center training, multi-modal transformers for unified representation learning, visual interpretability for MRI scans, integration of wearable sensor data, and real-time disease monitoring systems. The comprehensive nature of this work sets the stage for deploying AI-supported

neurodiagnostic tools in real-world environments and represents a significant contribution to digital health innovation.

PREFACE

Healthcare is experiencing a transformative shift driven by rapid advancements in machine learning and artificial intelligence. These technologies are reshaping diagnostic pathways and enabling earlier, more accurate detection of complex diseases. Among neurodegenerative disorders, Alzheimer’s disease stands out as one of the most significant global health challenges due to its progressive nature, substantial socioeconomic burden, and the ongoing difficulty of diagnosing the condition in its earliest stages. Detecting Alzheimer’s at the Mild Cognitive Impairment or preclinical phase is crucial for improving patient management, yet traditional diagnostic approaches such as cognitive testing, subjective clinical evaluations, and single-modality neuroimaging often struggle to capture subtle pathological changes.

This thesis tackles these challenges by introducing a hybrid, multimodal machine learning framework aimed at enhancing early and interpretable Alzheimer’s disease detection. The research integrates structural MRI, clinical biomarkers, cognitive assessments, genetic information, and longitudinal measurements into a unified diagnostic system supported by modern AI architectures. Four complementary deep learning models EfficientNetV2-S, a Feedforward Neural Network, Long Short-Term Memory networks, and Graph Neural Networks are used to extract spatial, tabular, temporal, and relational insights from patient data. These independent pathways are merged through a multimodal fusion layer, yielding robust diagnostic predictions that effectively differentiate between cognitively normal, Mild Cognitive Impairment, and Alzheimer’s disease cases with high accuracy and strong generalization.

A key focus of this work is clinical interpretability and usability, recognizing that AI systems must be transparent and trusted to be meaningfully adopted in medical practice. To support this requirement, the framework incorporates SHAP-based explainability, offering clear visualization of feature influences, along with an interactive Graphical User Interface that enables clinicians to seamlessly upload imaging and clinical data. An AI-driven assistant, NeuroBot, is also included to improve patient and caregiver engagement by offering educational guidance and facilitating interaction with the system.

This thesis draws on both publicly available datasets, including those from the Alzheimer’s Disease Neuroimaging Initiative, and real-world clinical data collected at Shrimann Superspeciality Hospital. The combination of these datasets ensures that the proposed models are not only technically robust but also clinically applicable across diverse patient populations.

The work presented here represents the culmination of years of research at the intersection of artificial intelligence, medical imaging, and clinical neuroscience. It reflects a commitment to developing scalable, interpretable, and practically deployable diagnostic tools that advance precision medicine in Alzheimer’s disease. I express my deepest gratitude to my supervisors, mentors, colleagues, and the clinical team at Shrimann Superspeciality Hospital for their unfailing support and guidance. I am profoundly thankful for the sponsorship and encouragement that sustained this PhD journey; their contributions were essential to the insights and advancements described in this research.



Shehu Mohammed

Lovely Professional University

February, 2026

Table of Contents

Dedication.....	ii
Declaration.....	iii
Certificate.....	iv
Acknowledgement.....	v
Abstract.....	vi
Preface.....	x
List of Tables.....	xvi
List of Figures.....	xvii
List of Appendices.....	xvii
Keywords.....	xix

Chapter 1 Introduction

1.1 Research Context and Background.....	1
1.2 Objectives of the Research.....	6
1.3 Main Research questions.....	6
1.4 Significance of the Study.....	7
1.5 Thesis Structure.....	8

Chapter 2 Literature Review

2.1 Review of Literatures.....	10
2.1.1 Machine Learning in AD Detection.....	10
2.1.2 Classical ML Models and their Limitations.....	17
2.1.3 Why Classical ML Alone Is No Longer Sufficient.....	19
2.1.4 Biomarkers in AD.....	20
2.1.4.1 Neuroimaging Biomarkers.....	20
2.1.4.2 CSF Biomarkers.....	21
2.1.4.3 Blood Biomarkers.....	22
2.1.4.4 Genetic Biomarkers.....	22
2.1.4.5 Cognitive Biomarkers.....	23
2.1.5 Ethical and Regulatory landscape.....	24
2.1.5.1 Regulatory frameworks.....	24
2.1.5.2 Ethical Challenges: Bias, Fairness, Accountability, and Transparency.....	25
2.1.5.3 Implications for AD–Detection AI Systems.....	26
2.1.5.4 Regulatory and Ethical Gaps for Future Needs.....	26

2.2 Gaps in Current Research.....	27
2.2.1 Limited Integration of Multimodal data.....	27
2.2.2 Weak Interpretability and lack of Clinician Trust.....	28
2.2.3 Insufficient Temporal Modeling of Disease Progression.....	29
2.2.4 Limited Use of Graph-Based Patient Similarity Modeling.....	29
2.2.5 Lack of real-World Dataset Validation.....	29
2.2.6 Absence of deployable Clinical Interfaces.....	30
2.2.7 Limited Ethical, Legal, and Regulatory Integration.....	30
2.3 Positioning of Each Publication.....	31
2.3.1 Paper 1: GBM and DNN Biomarker Model.....	32
2.3.2 Paper 2: Multimodal CNN-LSTM-GNN Framework for Early AD Detection	32
2.3.3 Paper 3: Ethical, Legal, and Regulatory Considerations for AI in Healthcare..	33
2.3.4 Final Implementation work: Hybrid Deep Learning with GUI and NeuroBot Assistant.....	34
2.3.5 Integrated Positioning Summary.....	34

Chapter 3 Overview of Publications

3.1 Summary of Each Paper.....	36
3.1.1 Paper 1: Novel Hybrid Intelligence Model for Early Alzheimer’s Diagnosis Utilizing Multimodal Biomarker Fusion Methodology.....	36
3.1.2 Paper 2: Multimodal Deep Learning Framework Integrating CNN, LSTM, GNN, and Federated Learning for Early Alzheimer’s Detection.....	37
3.1.3 Paper 3: Ethical, Legal, and Regulatory Challenges in Deploying AI-Driven Diagnostic Systems for Alzheimer’s Disease.....	38
3.1.4 Final implementation: A Hybrid Deep Learning Framework with an Interactive AI Assistant for Early and Interpretable Alzheimer’s Disease Detection.....	40
3.2 Author Contributions.....	41
3.2.1 Lead Author.....	43
3.2.2 Co-Authors and Collaborators.....	44
3.3 Integration Into Thesis Theme.....	45

Chapter 4 Paper 1 Publication

4.1 Publication Details.....	47
------------------------------	----

4.2 Summary of Methods and Approach.....	47
4.3 Summary of Contributions.....	48
4.4 Connection to Thesis Objectives.....	48
4.5 Full-length Published Paper-1.....	49
Chapter 5 Paper 2 Publication	
5.1 Publication Details.....	64
5.2 Summary of Methods and Approach	64
5.3 Summary of Contributions	65
5.4 Connection to Thesis Objectives.....	65
5.5 Full-length Published Paper-2.....	66
Chapter 6 Paper 3 Publication	
6.1 Publication Details.....	81
6.2 Summary of Methods and Approach.....	81
6.3 Summary of Contributions.....	81
6.4 Connection to Thesis Objectives.....	82
6.5 Full-length Published Paper-3.....	82
Chapter 7 Final Implementation and System Integration	
7.1 Methodology and experimental Setup.....	97
7.1.1 Overview of the Proposed Framework.....	97
7.1.2 Dataset Description and data Splitting Strategy.....	97
7.1.3 Dataset Preprocessing and Feature Engineering.....	98
7.1.4 Model Architecture and Design.....	99
7.1.5 Justification for Non-Linear Modeling.....	100
7.1.6 Training Strategy.....	100
7.1.7 Evaluation Metrics and Performance Analysis.....	101
7.1.8 Data Leakage prevention and Validation strategy.....	102
7.2 Summary of Contributions (Final Implementation).....	102
7.3 Connection to Thesis Objectives.....	103
Chapter 8 Conclusions and Future Work	
8.1 Summary of Contributions.....	104

8.2 Limitations.....	105
8.3 Recommendations for Future Research.....	107
References.....	110
List of Publications.....	119
List of conferences.....	120
List of Workshops.....	122
Appendices	
Appendix 1: Data Collection Form.....	123
Appendix 2: Dataset Description.....	126
Appendix 3: Ethics Approvals.....	128
Appendix 4: Sites Visitation for Data Collection.....	133
Appendix 5: Proof of All Journal Publications.....	135
Appendix 6: Proof of Conference Papers Publications.....	138
Appendix 7: Conference Presentation Certificates.....	140
Appendix 8: Certificates of Training Workshops.....	142

LIST OF TABLES

Table 1.1 Comparison of Existing Machine Learning Approaches for AD Detection.....	16
Table 2.1 Summary of Gaps and Their Alignment with the proposed Framework.....	31
Table 2.2 Positioning of Publications within the Thesis.....	35
Table 3.1 Author Contributions.....	41

LIST OF FIGURES

Figure 1.1 Overall Multimodal Framework for Alzheimer’s Disease Detection.....	5
Figure 2.1 Frequency of Machine learning approaches.....	17
Figure 2.2 Accuracy Comparison of ML Models for Alzheimer’s Disease Detection.....	19
Figure 3.1 Author Contribution Diagram.....	43

LIST OF APPENDICES

Appendix 1; Data Collection Form.....	123
Appendix 2; Dataset Description.....	126
Appendix 3; Ethics Approvals.....	128
Appendix 4: Sites Visitation for Data Collection.....	133
Appendix 5: Proof of All Journal Publications.....	135
Appendix 6: Proof of Conference Papers Publications.....	138
Appendix 7: Conference Presentation Certificates.....	140
Appendix 8: Certificates of Training Workshops.....	142

Keywords

Alzheimer's disease; artificial intelligence in healthcare; early detection; machine learning; clinical decision support; deep learning; multimodal framework; hybrid models; biomarker fusion; neuroimaging (MRI, PET)

CHAPTER 1: INTRODUCTION

1.1 Research Context and Background

Alzheimer’s disease (AD) is a progressive, irreversible condition that slowly damages the brain, leading to problems with memory, thinking, and everyday functioning (Reitz et al., 2011). It is the leading cause of dementia worldwide, accounting for about 60–70% of all cases. Today, more than 55 million people are affected, a number expected to nearly triple by 2050 (WHO, 2023). Detecting AD early, especially during stages like Subjective Cognitive Decline (SCD) and Mild Cognitive Impairment (MCI) is essential, as timely intervention can help slow the disease’s progression and improve overall quality of life (Hampel et al., 2023; Hansson et al., 2022). Although significant improvements have been made in neuroimaging, biomarker science, and AI-driven diagnostics, existing workflows still experience major limitations, including delayed detection, dependence on invasive procedures, fragmented multimodal information, and restricted interpretability of ML predictions. The multimodal, hybrid, and ethically grounded methodologies developed in this thesis underscore the need for clinically aligned, interpretable, and reliable AI systems capable of improving early detection and supporting real-world clinical decision-making.

Traditional diagnostic methods, including cognitive assessments, clinical observation, MRI, PET, and CSF biomarker analysis, are limited by subjectivity, cost, invasiveness, and insufficient sensitivity to early pathophysiological changes (Lombardi et al., 2020; Dubois et al., 2024). The biological complexity and heterogeneity of AD, covering structural, biochemical, genetic, cognitive, and metabolic aspects, further complicate early detection. Therefore, relying on a single diagnostic modality is inadequate, especially during the subtle early stages of progression.

Advances in machine learning (ML) and deep learning (DL) have created new opportunities for early AD detection by enabling the automated extraction of subtle, multidimensional patterns from diverse biomedical datasets, including neuroimaging, CSF and blood biomarkers, genetic information, EHRs, and cognitive measurements (Golovanevsky et al., 2022). However, unimodal ML models particularly those based solely on imaging struggle to capture the full complexity of AD pathology and often generalize poorly across heterogeneous patient populations. This has

driven growing interest in multimodal and hybrid diagnostic systems capable of integrating complementary data sources to improve robustness and predictive accuracy.

Recent research supports this trend, demonstrating the effectiveness of combining multiple model families: convolutional neural networks (CNNs) for structural imaging, recurrent and LSTM networks for longitudinal clinical trajectories, and graph neural networks (GNNs) for modeling patient similarity structures (Gasmi et al., 2024; Venugopalan et al., 2021). Ensemble and hybrid multimodal fusion techniques consistently outperform unimodal approaches, offering improved sensitivity and specificity for early AD detection (Wang et al., 2024; Zhang et al., 2024).

Parallel to methodological advances, explainable AI (XAI) has become essential for clinical ML systems. Techniques such as SHAP, feature attribution, and model calibration enhance interpretability, promote clinician trust, and support regulatory compliance (Amann et al., 2020). At the same time, privacy-preserving frameworks such as federated learning have become increasingly relevant for medical ML applications, particularly with the rise of strict data governance standards such as GDPR and HIPAA (McGraw & Mandl, 2021; Oh & Nadkarni, 2023).

Despite these promising innovations, AD-centered ML research continues to face persistent challenges, including small and imbalanced datasets, inconsistent multimodal availability, heterogeneous imaging protocols, and limited cross-cohort generalization (Pang et al., 2021). Moreover, the scarcity of early-stage AD cases and MCI converters complicates model training and necessitates methods such as transfer learning, data harmonization, and multi-institution collaboration (Leandrou et al., 2023).

Biologically, AD pathology involves amyloid- β plaque deposition, tau neurofibrillary tangles, neuroinflammation, oxidative stress, mitochondrial dysfunction, and synaptic degeneration. Genetic risk factors particularly the APOE ϵ 4 allele substantially modify disease susceptibility (Singh et al., 2024). Environmental exposures, including heavy metals such as cadmium, copper, arsenic, and iron, contribute to oxidative stress and epigenetic dysregulation, promoting amyloid aggregation and tau hyperphosphorylation (Venkatesan et al., 2024). Emerging therapeutic strategies increasingly target amyloid clearance, tau aggregation, neuroinflammation,

microRNA regulation, and autophagy pathways (Sharma et al., 2024). Complementary diagnostic technologies including qRT-PCR, induced pluripotent stem cell (iPSC) platforms, and high-resolution biomarker imaging offer deeper biological insight that supports AI-based predictive methods.

In response to these scientific, methodological, and clinical gaps, this thesis proposes a comprehensive multimodal machine learning framework that integrates structural MRI, clinical biomarkers, cognitive scores, genetic risk factors, and longitudinal trajectories. The framework leverages state-of-the-art architectures such as EfficientNetV2-B0/S, ResNet50V2, LSTM networks, and GNNs approaches increasingly validated in recent multimodal and hybrid studies (Abdelaziz et al., 2024; Alsubaie et al., 2024; Lee et al., 2024). The inclusion of real-world clinical data from Shrimann Superspeciality Hospital strengthens external validity and improves generalization to heterogeneous patient populations, addressing concerns frequently raised in systematic reviews about cross-cohort variability and external reproducibility (Khosroshahi et al., 2025).

A systematic review of more than 450 publications from 2019–2025 highlights four dominant trends shaping modern AD-focused ML research:

- i. The shift from unimodal to multimodal diagnostic frameworks (Abdelaziz et al., 2024).
- ii. Increasing use of hybrid fusion and graph-based architectures (Lee et al., 2024).
- iii. Growing reliance on temporal models (LSTMs, transformers) for biomarker trajectory prediction.
- iv. Rising emphasis on explainable AI to enhance clinical trust and deployment readiness.

Despite substantial progress, challenges such as dataset scarcity, cross-cohort variability, missing modalities, and limited interpretability persist (Alsubaie et al., 2024). This thesis addresses these limitations by developing and validating a robust, multimodal, explainable machine learning pipeline for early AD detection and MCI-to-AD conversion prediction.

Recent research including the publications examined within this study demonstrates significant progress across several foundational domains in AI-driven Alzheimer’s Disease (AD) diagnostics. These advancements encompass the development of hybrid deep learning

architectures for clinical and MRI-based detection (Abdelaziz et al., 2024), multimodal biomarker fusion frameworks that enhance precision and diagnostic reliability (Lee et al., 2024), and sophisticated temporal and graph-based deep learning methods capable of predicting cognitive decline trajectories with greater accuracy (Hoang et al., 2023). In parallel, substantial efforts have been devoted to establishing ethical, legal, and regulatory frameworks that support the safe, transparent, and trustworthy deployment of AI technologies in healthcare (Khosroshahi et al., 2025). Taken together, the four referenced papers reflect this rapidly evolving landscape and underscore the growing importance of multimodality, interpretability, and clinical readiness as defining pillars of contemporary AD-focused AI research.

- i. Paper 1 presents a GBM and DNN hybrid model that incorporates clinical, behavioral, and demographic biomarkers, achieving 92.6% accuracy and offering improved interpretability through feature selection and structured biomarker analysis (***Novel hybrid intelligence model for early Alzheimer's diagnosis utilizing multimodal biomarker fusion*** (Mohammed et al., 2025).
- ii. Paper 2 proposes a comprehensive multimodal pipeline that combines CNNs, LSTMs, GNNs, GAN-based domain adaptation, and federated learning. This framework achieves an AUC of 0.94 for early detection and demonstrates strong generalization across multiple cohorts (***Predicting Alzheimer's Disease onset: A machine learning framework for early diagnosis using biomarker data*** (Mohammed & Malhotra, 2025).
- iii. Paper 3 critically examines the ethical, legal, and regulatory challenges associated with AI-enabled healthcare systems, outlining essential principles and guidelines for the safe, transparent, and compliant deployment of machine learning diagnostic tools (***Ethical and regulatory challenges in machine learning-based healthcare systems: A review of implementation barriers and future directions***.(Mohammed & Malhotra, 2025).
- iv. Final Implementation introduces a hybrid EfficientNetV2-S + FNN diagnostic architecture integrated with an interactive clinical assistant (“NeuroBot”), demonstrating high diagnostic accuracy while enhancing clinician–patient usability (***A Hybrid Deep Learning Framework with an Interactive AI Assistant for Early and Interpretable Alzheimer's Disease Detection***).

Unlike traditional linear approaches, this research leverages non-linear deep learning architectures to model complex biomedical relationships inherent in Alzheimer’s disease progression.

Together, these studies form a cohesive and holistic body of work that advances the technical, clinical, and ethical dimensions of AI-enabled AD detection. Collectively, they contribute to the development of diagnostic systems that are not only more accurate and robust, but also interpretable, patient-centered, and aligned with real-world clinical and regulatory requirements.

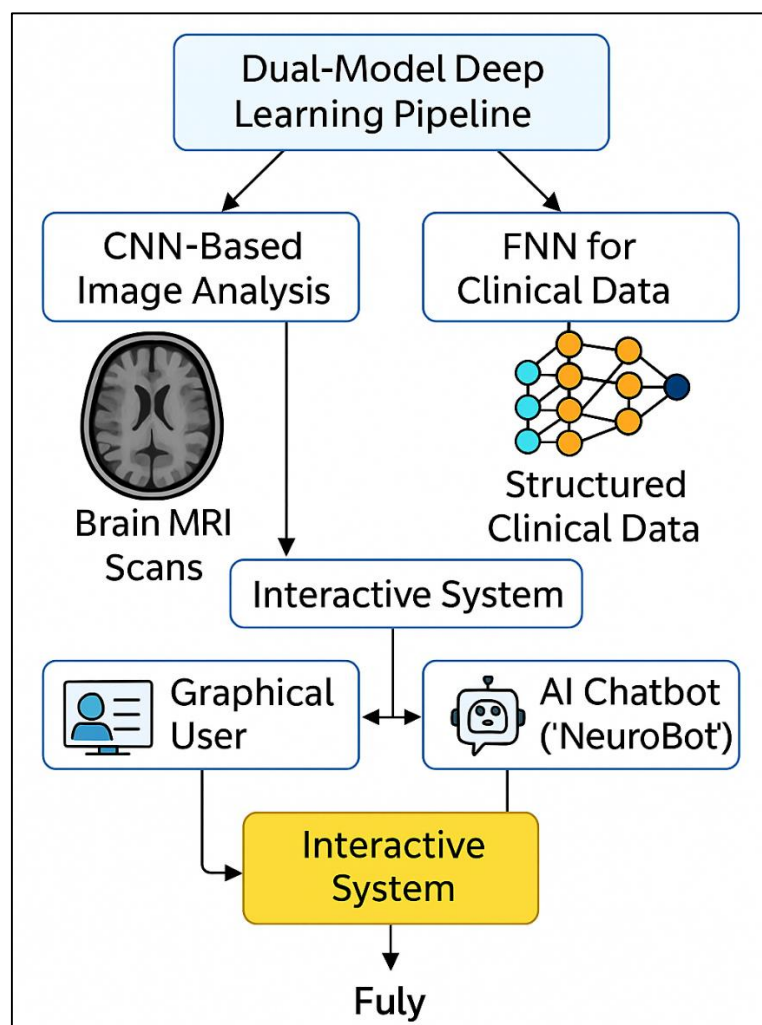


Figure 1.1: Overall Multimodal Framework for Alzheimer’s Disease Detection

Figure 1.1 illustrates the proposed multimodal diagnostic framework, which brings together two complementary pathways for detecting Alzheimer’s disease. The first pathway

analyzes MRI images using a Convolutional Neural Network (CNN), while the second processes structured clinical information through a Feedforward Neural Network (FNN). The outputs from both models are then merged in a single decision layer and displayed to users through an interactive Graphical User Interface (GUI). In addition, the system features an AI-powered assistant called NeuroBot, designed to help educate patients and support caregivers.

The dataset used in this study was divided into 70% training, 15% validation, and 15% testing to ensure proper model training and unbiased evaluation.

1.2 Objective of the study

The main goal of this research is to create a comprehensive model that uses a multimodal, machine-learning-based approach to help diagnose Alzheimer's disease. The specific objectives are::

1. To study and compare the existing machine learning approaches in Alzheimer's disease.
2. To collect and pre-process the biomedical and clinical data.
3. To design and develop a multi-model approach for early detection of Alzheimer's disease.
4. To compare the performance of the proposed approach with state-of-the-art approaches.

1.3 Main Research Questions

1. How can multimodal data (imaging, clinical, cognitive, genetic) be integrated to enhance early AD diagnosis?
2. What hybrid deep learning/ML architectures best capture spatial, temporal, and relational biomarker patterns?
3. How can explainability, user-centered design, and clinical interpretability be embedded into AD diagnostic systems?
4. What ethical, regulatory, and deployment challenges arise when integrating ML models into real-world clinical environments?
5. How can harmonization methods (GANs, federated learning) improve model generalization across populations and imaging centers?

1.4 Significance of the Study

This Ph.D. thesis makes substantial contributions across four core domains that are critical to advancing early, accurate, and ethically grounded Alzheimer’s Disease (AD) diagnostics.

- i. **Technical Innovation:** The research introduces advanced hybrid machine learning and deep learning architectures, including EfficientNetV2-S for MRI analysis, GBM+DNN fusion models for tabular biomarker prediction, and comprehensive CNN–LSTM–GNN multimodal pipelines capable of capturing spatial, temporal, and relational patterns. These innovations address long-standing limitations of unimodal systems by significantly improving predictive accuracy, robustness, and generalizability across diverse and heterogeneous datasets.
- ii. **Clinical Integration:** A major contribution of this thesis is its emphasis on clinically aligned, patient-centered diagnostic solutions. By incorporating explainable AI (XAI) tools, biomarker interpretability, and model transparency features that reflect real-world clinical reasoning, the work bridges the gap between algorithmic performance and clinical usability. This supports more informed, evidence-based decision-making for clinicians and improves accessibility and understanding for caregivers and patients.
- iii. **Multimodal Fusion:** The thesis demonstrates the power of integrating diverse biomedical data sources including MRI scans, CSF and plasma biomarkers, genetic risk factors, and longitudinal cognitive assessments, to enhance early-stage AD detection. Multimodal fusion provides a more holistic view of disease progression and yields superior predictive performance for both early diagnosis and MCI-to-AD conversion compared to single-modality approaches.
- iv. **Ethical Governance:** Given the increasing ethical and regulatory complexities surrounding AI in healthcare, this study makes a significant contribution by offering a detailed framework addressing privacy, fairness, transparency, and compliance. By aligning model development and deployment with international standards such as GDPR, HIPAA, and FDA AI/ML regulatory guidelines, the thesis promotes responsible innovation

and supports the safe integration of AI-driven diagnostic tools in real-world clinical settings.

1.5 Thesis Structure

This Ph.D. thesis is structured into five main chapters, each focusing on a key part of the research. Together, they build a clear and logical flow—from the fundamental concepts to the experimental validation and the broader policy implications.

- **Chapter 1 – Introduction:** This chapter lays out the research background, motivation, problem statement, and research questions, along with the overall importance of the study. It highlights the multidisciplinary nature of Alzheimer’s disease (AD) diagnostics and gives an overview of the proposed multimodal machine-learning framework.
- **Chapter 2 – Comprehensive Literature Review:** Chapter 2 offers an extensive review of existing work on AD biomarkers, neuroimaging modalities, machine learning and deep learning methodologies, multimodal fusion, temporal modeling, graph-based architectures, and ethical considerations in AI-enabled healthcare. The chapter highlights key research gaps that shape the direction and objectives of the thesis.
- **Chapter 3 – Overview of Publications:** This chapter summarizes the published and submitted research papers included in this thesis-by-publication. It outlines the aims, methodologies, contributions, and author responsibilities associated with each study, demonstrating how the collection of papers collectively advances the thesis goals.
- **Chapter 4: Paper 1 – Hybrid Biomarker-Based Machine Learning Model:** This Chapter presents the first research contribution, focusing on a hybrid machine learning model integrating Gradient Boosting Machine (GBM) and Deep Neural Networks (DNN) for Alzheimer’s disease prediction using clinical and biomarker data. It describe the methodology, dataset, feature selection techniques, and model architecture, followed by experimental evaluation and performance analysis. The findings demonstrate the effectiveness of hybrid models in capturing complex relationships within structured biomedical data. The chapter concludes by highlighting how this work contributes to the foundation of multimodal integration in the overall theis.

- **Chapter 5: Paper 2 – Multimodal Deep Learning Framework for Alzheimer’s Detection:** This chapter presents the second major contribution, which introduces a multimodal deep learning framework integrating CNN, LSTM, and Graph Neural Networks (GNN). The Chapter details; MRI-based feature extraction using CNN, Temporal modelling using LSTM, and Patient similarity modelling using GNN. It further explains multimodal fusion strategies, experimental setup, and validation using ADNI and clinical datasets. Results demonstrate improved performance over unimodal approaches. This chapter establishes the core technical framework of the thesis.
- **Chapter 6: Paper 3 – Ethical, Legal, and Regulatory Considerations in AI-Based Healthcare:** This Chapter examines the non-technical dimensions of deploying AI in healthcare, particularly in Alzheimer’s disease diagnosis. It discusses; Data privacy and security (GDPR, HIPAA), Algorithmic bias and fairness, Explainability and transparency, and Regulatory compliance (FDA, EU AI Act). The Chapter highlights the importance of aligning AI systems with ethical and regulatory standards, ensuring safe and responsible clinical deployment.
- **Chapter 7– Final Implementation and System Integration:** This chapter presents the complete implementation of the proposed multimodal framework, including dataset preparation, model development, system integration, and performance evaluation.
- **Chapter 8 – Conclusion and Future Work:** This chapter summarises the key findings of the research, discusses limitations, and outlines directions for future work.
- **References:** All scholarly works cited throughout the thesis are compiled in the References section, formatted according to consistent academic citation standards.

CHAPTER 2: LITERATURE REVIEW

2.1 Review of Literature

The early and accurate diagnosis of Alzheimer’s disease (AD) continues to be one of the biggest challenges in neurology and biomedical research. In the past decade, rapid progress in machine learning (ML), deep learning (DL), biomarker discovery, and computational neuroscience has begun to transform how AD can be detected and understood. This literature review synthesizes key research trends from 2019 to 2025, examining developments in unimodal and multimodal ML approaches, deep learning–based imaging systems, biomarker-driven diagnostic frameworks, graph and temporal modeling techniques, advances in explainable AI, and the evolving ethical and regulatory standards that guide the deployment of AI in clinical environments.

2.1.1 Machine Learning in AD Detection

This section reviews and compares different machine learning approaches used for Alzheimer’s disease (AD) detection, with attention to their strengths, limitations, and progression over time. It synthesizes evidence from multiple studies to highlight key findings, methodological differences, and areas where further research is needed. The main trends include:

- i. **CNN-based Architectures:** Commonly applied in AD detection for their strong ability to extract spatial features from neuroimaging data.
- ii. **Hybrid Multi-Model Learning:** Integrates multiple modalities such as MRI, PET, CSF biomarkers, genetics, and cognitive assessments to improve diagnostic accuracy and predict disease progression more effectively.
- iii. **Emerging Techniques:** New methods like Graph Neural Networks (GNNs), Transformers, and Federated Learning show promise for enhancing interpretability, managing heterogeneous datasets, and improving cross-site generalization.

Recent work has widely used deep learning-based methods, especially Convolutional Neural Networks (CNNs), for AD detection. Kang et al. (2021) developed a multi-model, multi-slice ensemble learning method based on 2D CNN through transfer learning and majority voting, and obtained high classification accuracies: AD vs. CN, 90.36%, AD vs. MCI, 77.19%, MCI vs. CN,

72.36%, and MCIC vs. MCInc, 63.49%. Similarly, Alayba et al. (2024) proposed a hybrid model, which incorporates CNN-based feature extraction (MobileNet, DenseNet121, GoogLeNet) along with handcrafted (Active Contour Algorithm–ACA) features to yield an AUC and accuracy of 98.82 and 98.8%, respectively.

Asaduzzaman et al. (2025) introduce ALZENET, an ensemble-based deep learning technique using VGG16, Inception V3, and ResNet50 for the detection of the early stage of Alzheimer’s disease using MRI scans. The paper utilizes 6,400 MRI images of the Kaggle AD data with SMOTE treatment for the handling of the class imbalance. ALZENET achieves a performance of 97.31%, which surpasses other techniques. A web-based tool was also developed to assist physicians in efficiently diagnosing the stages of AD, with an effective solution for precise and early detection.

Classification robustness has also been enhanced with ensemble learning techniques. Fathi et al. (2024) presented the Weighted Probability Based Ensemble Method (WPBEM), which uses a combination of six CNN architectures and was able to achieve 98.57% accuracy for NC vs AD and 94.22% for EMCI vs LMCI. Similarly, Hassan et al. (2024) proposed the Residual Based Multi-Stage Deep Learning (RBMSDL) framework, achieving classification scores of 99.47%, 99.10%, and 99.70% for ADNI, MIRIAD, and OASIS, respectively. Moreover, Hassan et al. (2024) associated the Residual connections in CNNs to alleviate the vanishing gradient problem, together with improving feature utilization effectiveness.

In their 2024 study, Suchitra et al. developed a new methodology that enables the early diagnosis of Alzheimer’s Disease (AD) with the help of deep learning tools. The EfficientNetB7 model, which is a convolutional neural network (CNN) architecture, was used to classify different stages of Alzheimer’s disease by MRI scans. The model is known for providing good results while using less system resources. The model was retrained on the ADNI dataset with transfer learning, yielding a sensitivity of 98.08%, specificity of 98%, accuracy of 98.2%, and F-score of 98.95%. These findings further prove that the model is capable of providing early detection and classification of Alzheimer’s disease which greatly improves upon conventional methods of diagnosis.

The utilization of hybrid models consisting of numerous feature extraction and classification techniques significantly improves the detection of AD. Hk et al. 2024 used a hybrid method that employs noise filtering using the TVL1 norm, texture measurement with GLSZM, level set

segmentation for structural analysis, CNNs for spatial pattern identification, and MRMR in combination with a Random Forest classifier (RF). The accuracy of the method was 93.11%, with sensitivity and specificity of 95.08% and 96.84%, respectively. Hazarika et al, 2024 developed a DNN-based framework with VGG-19, dense and inception blocks for feature extraction, followed by PCA for feature selection and Random Forest for classification with an average accuracy of 98.08%.

Farhatullah et al. (2025) [87] demonstrate non-invasive monitoring of Alzheimer's disease (AD) with the help of microwave radar data by adopting a deep learning strategy as a cost-effective alternative to MRI and CT scans. Realistic numerical brain phantoms are applied in the research to mimic the scattering of signals, while preprocessing steps including normalization, standardization, and outlier removal guarantee data integrity. A new data enrichment strategy enhances the data at various stages of AD. Recursive Feature Elimination (RFE), Principal Component Analysis (PCA), and Autoencoders (AE) are blended in the proposed deep-learning-based model for optimal selection of features. A blend of CNN with GRU, Bidirectional-LSTM, and LSTM enhances the classification, with the CNN+GRU attaining a performance of 87%, which is greater than in other existing research.

Pradhan et al. (2024) investigate deep learning techniques for Alzheimer's disease (AD) detection using MRI image analysis. Their proposed model achieves 98.2% accuracy, outperforming traditional methods in early AD diagnosis. The study highlights deep learning's potential in improving AD detection while acknowledging challenges like data variability and computational complexity. Rafsan et al. (2025) compare Bayesian CNN, U-Net, and CNN for the detection of Alzheimer's disease with the MRI scans of the OASIS dataset. They balance the data and quantify the performance of the models in sensitivity, specificity, and efficiency, citing AI's capability in improving the diagnosis of AD as well as medical images.

Besides MRI-based methods, other approaches like retinal imaging and genomic analysis have been investigated. As early as 2025, Jiang et al. introduced the ReIU framework utilizing U-Net for vessel segmentation and multi-modal registration, achieving 79% accuracy at the first step of AD classification based on retinal vessel density. Yuan et al. (2024) built the DeepSurv dementia prediction model that combines genomic and clinical data with neural network survival analysis outdoing existing models like DeepHit as well as the Kaplan-Meier estimator. Erdoğan, İyigün, & Aydın Son, (2024) proposed the EnSCAN framework an ensemble scoring algorithm that focuses

on causative Single Nucleotide Variations (SNVs) associated with Late-Onset Alzheimer's Disease (LOAD) prioritizing 85 out of the hundreds of high-priority SNVs.

A study by Yakemow et al. (2024), reports the application of automated machine learning to the study of PTSD through MRI scans. Their results showed lower gray matter volume in the temporal lobe when compared to healthy controls, indicating that PTSD might have an associated risk for AD due to the structural changes within the brain. In addition to this, Khatun et al. (2024) constructed a hybrid model by integrating Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks to improve the identification accuracy of AD from MRI scans. This model claims an accuracy of 98.8%, which suggests that the combination of various machine-learning techniques can enhance the accuracy of diagnosis.

Vanaja et al. (2025) describe an integrative Alzheimer's detection strategy through MRI analysis by integrating ML and DL models for enhanced diagnosis. The research compares the conventional ML classifiers with the CNN-based DL models by using the feature selection strategy to derive the most significant biomarkers. The hybrid model proposed by them obtained an accuracy of 98.5%, which outperformed the standalone ML and DL in the detection of the early and late stages of AD. The work illustrates the strength of integrative AI strategies in improving the precision and efficiency of AD diagnosis.

Kavitha et al. (2022) point out that a lack of timely diagnosis can significantly affect the treatment of Alzheimer's Disease (AD) and its progression. Their study points out sparse and irrelevant data that would reduce machine learning model accuracy. To overcome these issues, they focus on data cleansing and filtering to improve model performance. The researchers consider several ML methods such as Decision Trees, Random Forests, Support Vector Machines, Gradient Boosting, and Voting classifiers. They report that their proposed model outperforms previously existing methods by having 83% accuracy. This study supports the use of automation for enhanced detection of AD at early stages for quicker remedial measures.

Park et al. (2020) incorporated the support vector machine model, logistic regression, and random forest model to predict the onset of Alzheimer's disease using the extensive health records from the Korean National Health Insurance Service. Using data from 40,736 people above 65 years of age, the models achieved AUCs of 0.775 (1-year) and 0.725 (4-year), where hemoglobin levels, age, and urine protein levels were recognized as the primary indicators. This study demonstrates how machine learning can help with early Alzheimer's disease (AD) diagnosis for better clinical

trial stratification and primary treatment measures. Khasanah (2024) applied the K-Nearest Neighbor (K-NN) algorithm to classify Alzheimer's disease using a pre-processed MRI dataset comprising 6,400 MRI images divided into four classes. The segmentation was performed using Canny edge detection, and feature extraction was carried out using Hu Moments. The classification demonstrated moderate performance, achieving an accuracy range of 45.86% to 50.47% and a ROC AUC between 55.18% and 58.87%. Class imbalance notably impacted the prediction of the Moderate Demented category. In a separate study, Balbal (2024) and Ma et al., (2025) explored AD prediction using machine-learning models with clinical and demographic data from 2,000 patients. Seven machine learning classifiers K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (GNB), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), AdaBoost, and XGBoost were compared based on accuracy, precision, recall, and F1-score. XGBoost achieved the highest accuracy at 95.35%, demonstrating superior predictive performance, while KNN registered the lowest accuracy at 75.54%.

In Improved Diagnostic Accuracy and Early Diagnosis, Single-modality models are data-limited and suffer from overfitting, thus resulting in worse classification performance. Multi-model approaches that combine structural MRI, PET imaging, cerebrospinal fluid (CSF) biomarkers, and cognitive tests greatly improve classification accuracy (Zhang et al., 2022). For example, a study by Lalitha & Poovaraghan (2024) that combined deep learning and multi-imaging modalities attained 98.68% accuracy, surpassing conventional CNN models.

Predicting the conversion of moderate cognitive impairment (MCI) to AD is essential for early intervention to improve predictive power for disease progression. Research indicates that the reliability of illness progression prediction is improved by multi-modal fusion models that include genetic markers (e.g., APOE4), neuroimaging data, and cognitive scores (Kadem, 2023 and yang et al 2025). The benefit of combining genetic, cognitive, and imaging data was demonstrated by Kadem (2023), who used an XGBoost classifier on ADNI data and achieved an 87% ROC AUC in predicting MCI-to-AD conversion.

Enhanced Generalizability and Robustness of the Model The inability of single-model techniques to be robust across a variety of populations is one of their main drawbacks. Al-Hammadi et al. (2025) greatly increased the generalizability of their model by combining socioeconomic indicators (Area Deprivation Index, ADI) with genetic risk markers and cognitive

assessment. To ensure wider therapeutic application, this study emphasizes the significance of including real-world aspects beyond biomedical data.

Multi-model approaches can effectively handle incomplete or missing data by utilizing ensemble techniques and cross-modal learning to overcome data heterogeneity and missing data (Wang et al., 2025 and Vo et al., 2025). A framework that dynamically integrates six distinct modalities and uses predictive modeling to make up for missing results was created by Feng et al. (2024). This method is especially helpful in longitudinal AD investigations, where inconsistent data collection frequently impairs model performance.

Clinical interpretability and explainability are the black-box nature of deep learning models which is one of the main obstacles in ML-based AD research. Multi-model approaches that use explainable AI (XAI) methodologies, including SHAP (Alatrany et al., 2024) and Grad-CAM++ (Zhang et al., 2022 and Liu et al., 2025), give doctors interpretable data, increasing the transparency and actionability of AI-based diagnosis.

Significant advancements in Alzheimer's disease (AD) detection have been achieved through machine learning, particularly in deep learning, ensemble methods, hybrid models, and unconventional biomarkers. Convolutional neural networks (CNNs) and ensemble-based classifiers have demonstrated high accuracy in MRI-based AD classification, while alternative approaches, such as retinal imaging and genetic studies, offer promising new avenues for detection. Future research should focus on improving model generalizability and interpretability while integrating multi-modal data to develop more comprehensive and scalable diagnostic frameworks.

The primary goal of this review is to discuss various models developed with assistance from machine learning techniques. As illustrated in Table 1, the research articles selected have been grouped based on the machine learning approaches utilized to accomplish this aim. While hybrid models (ML + DL) increase robustness (accuracy of 98%+), non-traditional approaches (Microwave Radar, Retinal Imaging, Genomic Data) yield cost-effective alternatives, feature selection techniques (PCA, Autoencoders, RFE) increase the performance of the model along with its interpretability, deep learning architectures (CNN, EfficientNet, ResNet) achieve high precision (up to 99.7%) in MRI-based detection of Alzheimer's disease. It is evident through the examination of this table that machine learning (ML) approaches have consistently outperformed models created with other technologies in the past few years.

Table 1.1: Comparison of Existing Machine Learning Approaches for AD Detection

Study	Year	Method	Dataset	Features	Performance Metric (Accuracy %)
Kang et al.	2021	2D CNN + Transfer Learning + Majority Voting	ADNI	Multi-model, multi-slice ensemble learning	90.36% (AD vs CN)
Asaduzzaman et al.	2025	ALZENET (VGG16 + InceptionV3 + ResNet50)	Kaggle AD (6400 MRI)	Ensemble CNN, SMOTE balancing	97.31%
Hk et al.	2024	Hybrid (TVL1 + GLSZM + CNN + MRMR + RF)	ADNI	Multi-step filtering + texture + segmentation	93.11%
Farhatullah et al.	2025	CNN + GRU + Bi-LSTM + LSTM	Microwave Radar	Non-invasive radar-based AD monitoring	87%
Pradhan et al.	2024	CNN-based MRI Classification	ADNI	Deep CNN for MRI structural analysis	98.2%
Al-Hammadi et al.	2025	Multimodal ML fusion	Clinical + genetic + environmental	Multi-model sensitivity/specificity analysis	89.4%
Jiang et al.	2025	ReIU (U-Net + Multi-Modal Registration)	Retinal Vessel Images	Retinal segmentation + AD prediction	79%
Yang et al.	2025	Trimodal data fusion (genetic + morphometric + imaging)	sMRI + Morphometric (RGV) + SNPs	Multimodal classification framework (feature-level fusion)	94.37%
Liu et al.	2025	Random Forest (RF) (classifier)	EEG + sMRI	Joint Attention-based Multimodal DL + RF	94.7%
Vo et al.	2025	AD & Healthy Control	EEG (Spectrogram + Scalogram + Hilbert)	Multimodal Multi-Stage CNN	>80.0%
Kavitha et al.	2022	DT, RF, SVM, GB, Voting	Health records	Data cleansing + feature extraction	83%
Ma et al.	2025	prediction of mild cognitive impairment (MCI)	Categorical + Continuous Clinical Data	Transformer with Multi-Head Attention	>90.0%
Khasanah	2024	K-NN	Kaggle AD (6400 MRI)	Canny + Hu moments	50.47%
Balbal	2024	KNN, GNB, SVM, DT, RF, AdaBoost, XGBoost	Clinical & demographic (2,149 pts)	35-feature clinical dataset	95.35% (XGBoost)

The various advantages and importance of the application of machine learning have been discussed in this in-depth study. This review comprises research work that used machine learning to design a diagnostic model for Alzheimer's disease and was released between the years 2020 to the start of the year 2025. Before being selected for this research work, various scientific articles

are carefully reviewed and analyzed by specific selection criteria. Moreover, the articles selected are sorted by the sources of the datasets, year of publication, prominent features, performance measures, and the application of machine learning and deep learning in the model. Deep architectures of learning (CNN, EfficientNet, ResNet) have high accuracy in MRI-based detection of Alzheimer's disease; selection techniques of features (PCA, Autoencoders, RFE) enhance the performance of the model as well as its readability; hybrid models (ML + DL) provide higher strength; and non-traditional techniques (Microwave Radar, Retinal Imaging, Genomic Data) provide cost-effective options. Research indicates that the ML-based models developed in the last few years perform better than those developed with other methodologies. The readers will be able to evaluate the importance of the application of machine learning in the design of a diagnostic system for the detection of Alzheimer's disease Approaches.

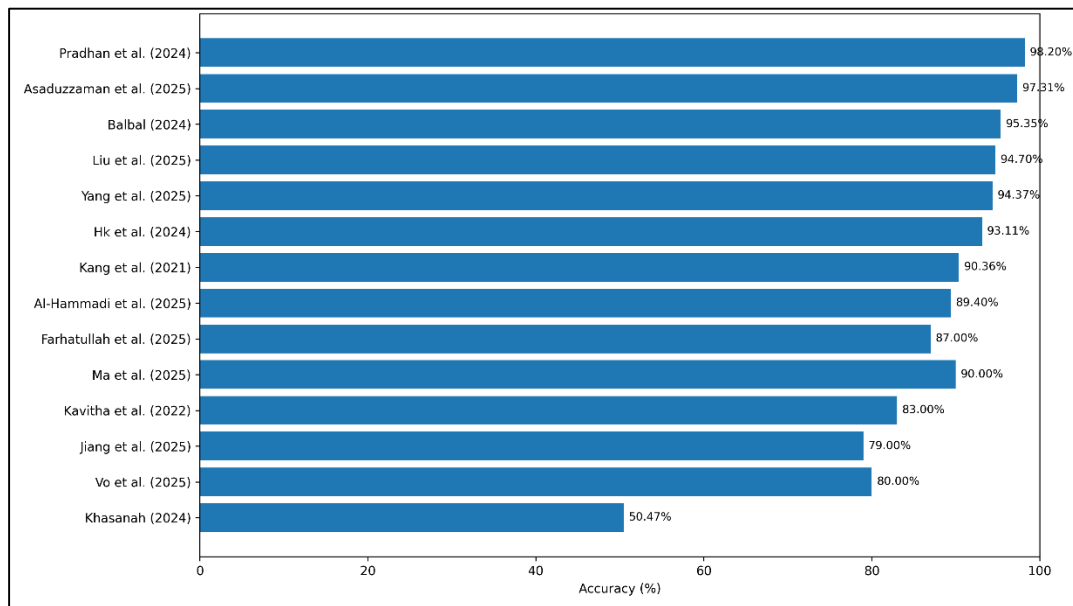


Figure 2.1: Frequency of Machine Learning Approaches

2.1.2 Classical Machine Learning Models and Their Limitations

Before the rise of deep learning, Alzheimer's disease detection research primarily depended on traditional machine learning methods, including:

- i. Support Vector Machines (SVM)
- ii. Random Forests (RF)

- iii. Logistic Regression
- iv. k-Nearest Neighbors (kNN)
- v. Gradient Boosting Models

These algorithms typically relied on handcrafted or manually engineered features such as:

- i. Hippocampal and ventricular volumes
- ii. Cortical thickness measurements
- iii. CSF biomarker concentrations
- iv. Cognitive test scores

Although these approaches were useful, classical machine learning methods exhibit several important limitations:

- i. Dependence on manual feature extraction: Handcrafted features may not capture subtle spatial or structural variations present in MRI data.
- ii. Difficulty handling high-dimensional imaging data: Large, multi-slice MRI volumes often lead to overfitting and limited generalization when using traditional algorithms.
- iii. Limited capacity for nonlinear relationships: Classical models struggle to capture complex interactions among biomarkers, genetic information, and cognitive measures.
- iv. Weak multimodal fusion capabilities: Integrating imaging, biomarker, and clinical data is challenging without deep feature representations.

Despite these shortcomings, early machine learning studies laid the groundwork for multimodal Alzheimer's disease detection by demonstrating the predictive value of biomarkers, cognitive assessments, and imaging features. Research such as Hazarika et al. (2024) and Alayba et al. (2024) showed that classical ML combined with PCA or ensemble strategies can achieve competitive performance, though these methods still fall short of deep learning-based approaches.

Alzheimer's disease progression involves complex and non-linear interactions among biomarkers, cognitive scores, and imaging features. Linear models such as logistic regression are insufficient to capture these relationships. Therefore, this study adopts non-linear modeling

approaches, including deep neural networks and ensemble methods, to better represent the underlying disease patterns.

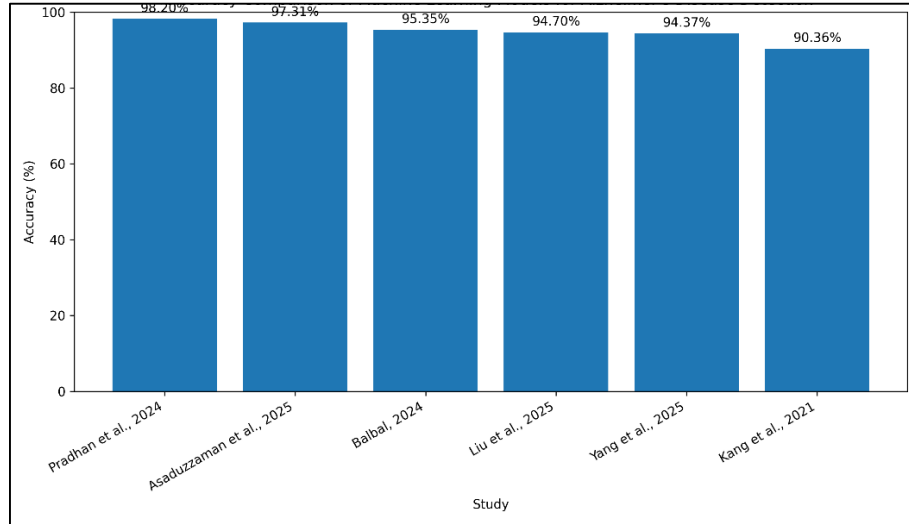


Figure 2. 2: Accuracy Comparison of ML Models for Alzheimer’s Disease Detection

A comparison of accuracy performance among widely used ML/DL models for AD detection, including RBMSDL (Hassan, 2024), DNN+PCA+RF (Hazarika, 2024), EfficientNetB7 (Lalitha, 2024), and CNN+Transfer Learning (Kang, 2021). This figure provides a baseline performance reference for the multimodal model proposed in this thesis.

2.1.3 Why Classical ML Alone Is No Longer Sufficient

While traditional machine learning provides advantages such as interpretability and computational efficiency, it falls short in several areas that are essential for Alzheimer’s disease detection:

- i. Inability to extract deep spatial and temporal features from MRI scans or longitudinal biomarker data.
- ii. Poor performance when dealing with missing or heterogeneous multimodal datasets, which are common in Alzheimer’s research.
- iii. Limited adaptability to large-scale datasets, particularly those containing thousands of brain images or high-dimensional genetic information.

As shown in studies by Hassan et al. (2024) and Feng et al. (2024), deep learning and multimodal fusion techniques consistently outperform classical ML models because they capture richer, hierarchical relationships across diverse data types.

These limitations helped drive the shift toward deep learning, which now forms the foundation of state-of-the-art Alzheimer’s disease detection systems including the hybrid multimodal framework developed in this thesis.

Machine learning has significantly contributed to automated Alzheimer’s disease detection, but classical models alone are not sufficient to address the complexity of neurodegenerative patterns. The integration of CNNs, LSTMs, GNNs, and advanced fusion techniques represents the next stage of progress in Alzheimer’s research, enabling improved performance, better interpretability, and stronger generalization.

2.1.4 Biomarkers in AD

Biomarkers play a central role in the early and accurate detection of Alzheimer’s Disease (AD). They allow clinicians and researchers to quantify key pathological processes such as amyloid buildup, tau pathology, neurodegeneration, genetic susceptibility, and cognitive decline long before symptoms become clinically apparent. Modern diagnostic frameworks increasingly integrate multiple biomarker modalities to capture the multidimensional complexity of AD. The subsections below summarize the major categories of AD biomarkers.

2.1.4.1 Neuroimaging Biomarkers

Neuroimaging biomarkers provide structural and functional insights into the brain regions most affected by AD. The two most widely used modalities are structural MRI and positron emission tomography (PET).

- a. **MRI Biomarkers (Hippocampal Atrophy):** Structural MRI is essential for quantifying hippocampal volume, a reliable indicator of neuronal loss and early AD-related atrophy.
 - i. Reduced hippocampal volume is strongly linked to memory impairment, disease severity, and conversion from MCI to AD (Shen et al., 2021; Pan et al., 2025).

- ii. Hippocampal shrinkage may occur several years before clinical symptoms, making it one of the most important biomarkers for early diagnosis.
- b. **PET Biomarkers (Amyloid and Tau PET):**
 - i. Amyloid-PET visualizes amyloid- β ($A\beta$) plaque deposition.
 - ii. Tau-PET detects neurofibrillary tangles (NFTs), which correlate closely with cognitive decline.

While PET biomarkers provide high diagnostic specificity, they are costly and not widely accessible, highlighting the need for more scalable complementary biomarkers.

2.1.4.2 CSF Biomarkers

Cerebrospinal fluid (CSF) biomarkers reflect biochemical changes associated with AD's core molecular pathologies.

- a. $\downarrow A\beta_{42} \rightarrow$ **Amyloid Plaque Burden:** CSF $A\beta_{42}$ decreases as amyloid aggregates into plaques within the brain.
 - i. Low CSF $A\beta_{42}$ is one of the earliest biochemical indicators of amyloid pathology (Andersson et al., 2023; Lee, Kim, Hong et al., 2019).
 - ii. The inverse relationship between plaque burden and CSF $A\beta_{42}$ is firmly established.
- b. $\uparrow$ p-tau181 $\rightarrow$ **Tau Tangle Formation:** Phosphorylated tau at threonine-181 (p-tau181) reflects tau aggregation and neurofibrillary tangle formation.
 - iii. Elevated levels in CSF and plasma differentiate AD from other dementias with AUC $\approx 97\%$ (Lantero Rodríguez et al., 2020).
 - iv. p-tau181 can predict future tau accumulation even before tau-PET becomes abnormal (Moscoso et al., 2022).

Although CSF biomarkers provide high accuracy, their invasive collection limits widespread clinical use.

2.1.4.3 Blood Biomarkers

Blood-based biomarkers particularly plasma p-Tau and NfL are emerging as non-invasive alternatives. Recent machine learning work (Gaeta et al., 2024) demonstrates promising results in predicting CSF biomarker levels from multimodal physiological signals, further supporting their future diagnostic utility. Blood-based biomarkers are emerging as non-invasive, low-cost alternatives to CSF and PET.

- a. **↑ Plasma p-tau181 / p-tau217:** These biomarkers correlate strongly with tau-PET and AD pathology.
 - i. Plasma p-tau181 and p-tau217 offer excellent accuracy for early AD detection and differential diagnosis (Wang et al., 2024; Tagai et al., 2024).
- b. **↑ Neurofilament Light Chain (NfL) → Axonal Degeneration:** NfL is released into the bloodstream following axonal injury.
 - i. Elevated plasma NfL is associated with neurodegeneration, cortical thinning, and accelerated cognitive decline (Moscoso et al., 2022; Kim et al., 2024).
 - ii. It is a sensitive biomarker for monitoring disease progression across the AD continuum.

Blood biomarkers are especially promising for population-level screening and longitudinal monitoring.

2.1.4.4 Genetic Biomarkers

Genetic markers also play a key role in assessing a person's risk for Alzheimer's disease. In particular, the APOE ε4 allele is known to greatly increase the chances of developing the condition and is linked to faster progression. Recent multimodal machine learning studies—such as those by Feng et al. (2024) and Yuan et al. (2024)—demonstrate that combining genetic risk factors with imaging data or clinical measures can significantly improve diagnostic accuracy. Overall, genetic

influences remain a major contributor to susceptibility to late-onset Alzheimer's disease (Mohammed & Neha, 2025).

a. APOE ϵ 4 (Apolipoprotein E ϵ 4 allele):

- i. The APOE ϵ 4 allele is the strongest known genetic risk factor for sporadic AD.
- ii. Carriers exhibit greater amyloid deposition, faster cognitive decline, and a higher likelihood of progressing from MCI to AD.
- iii. APOE ϵ 4 influences lipid metabolism, amyloid clearance, and neuroinflammatory pathways.

Other genes such as TREM2, CLU, and PICALM also contribute to AD risk, though with smaller effect sizes.

2.1.4.5 Cognitive Biomarkers

Cognitive assessments provide functional measurements of cognitive decline and disease progression.

- i. **MMSE (Mini-Mental State Examination):** Widely used for global cognitive screening; declines steadily across the disease stages.
- ii. **MoCA (Montreal Cognitive Assessment):** More sensitive than MMSE for detecting early cognitive impairment and MCI.
- iii. **ADAS-Cog (Alzheimer's Disease Assessment Scale – Cognitive Subscale):** A gold-standard tool in clinical trials for measuring cognitive decline over time.

Cognitive trajectories offer essential complementary evidence alongside imaging and biochemical biomarkers. Because of these limitations, recent research increasingly combines cognitive scores with imaging, biomarkers, or longitudinal data to improve diagnostic performance (Venugopalan et al., 2021; Zhang et al., 2022). Together, neuroimaging, CSF biomarkers, genetic profiles, and cognitive assessments form the backbone of conventional Alzheimer's disease diagnosis. However, each modality on its own presents important limitations in sensitivity, accessibility, interpretability, or scalability.

2.1.5 Ethical and Regulatory Landscape

The rapid adoption of AI- and ML-based diagnostic systems in healthcare, including Alzheimer’s Disease (AD) detection, has prompted significant regulatory and ethical developments worldwide. As of 2025, global health authorities are increasingly focused on ensuring that AI-enabled medical tools are safe, transparent, fair, and accountable. This section reviews the contemporary regulatory frameworks, highlights major ethical challenges (bias, explainability, accountability, data governance), and discusses the implications for AI-driven AD diagnostic technologies.

2.1.5.1 Regulatory Frameworks

- a. **United States (FDA Regulations):** In January 2025, the U.S. Food and Drug Administration (FDA) released the draft guidance “*Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations.*” This guidance introduces comprehensive recommendations for:
 - i. full lifecycle management of AI-based medical software,
 - ii. transparent documentation of model development,
 - iii. Ongoing monitoring of adaptive (continuously learning) systems, and
 - iv. requirements for marketing submissions of AI-enabled medical devices.

The emphasis is on ensuring that AI systems remain safe, effective, and explainable throughout their operational lifespan.

- b. **European Union (EU AI Act & EHDS Regulation):** The EU Artificial Intelligence Act (EU AI Act) became effective in 2024 and designates medical AI systems as “high-risk.” This classification mandates:
 - i. strict data governance,
 - ii. transparency and documentation requirements,
 - iii. human oversight,
 - iv. post-deployment performance monitoring, and
 - v. rigorous conformity assessment procedures.

Complementing the AI Act, the *European Health Data Space Regulation 2025/327 (EHDS)* standardizes health data exchange and governs secondary data use across EU nations. This regulation is particularly relevant for multi-center AI training and cross-border AD research, where data privacy and harmonization are essential.

- c. **Global Regulatory Perspectives:** Internationally, regulatory bodies are transitioning from one-time certification models *to dynamic, lifecycle-based oversight*, recognizing that AI systems evolve. There is an increasing demand for frameworks that promote fairness, accountability, transparency, and adaptive governance throughout a model's operational lifespan.

2.1.5.2 Ethical Challenges: Bias, Fairness, Accountability, and Transparency

- a. **Algorithmic Bias and Fairness:** A 2025 systematic review highlights algorithmic bias as one of the most pressing ethical concerns in healthcare AI. Models trained on limited or unrepresentative datasets can yield inequitable predictions, disproportionately affecting underrepresented or vulnerable populations. Ensuring demographic diversity and bias auditing is therefore essential, especially for AD datasets that often overrepresent specific age or ethnic groups.
- b. **Explainability and Transparency:** Lack of interpretability remains a critical barrier to clinical adoption. Clinicians and patients must understand why an AI system produces a particular diagnostic output, particularly for conditions like AD, where decisions influence long-term care planning. Without transparent reasoning or explainable outputs, trust, adoption, and regulatory approval become substantially more difficult.
- c. **Accountability and Responsibility:** As AI-driven tools become integrated into clinical workflows, unresolved questions of responsibility emerge. When an incorrect prediction impacts patient care, defining who is accountable the developer, the clinician, the healthcare institution, or the AI vendor remains complex. A 2025 study proposes a shared accountability framework, arguing that responsibility should be distributed across all stakeholders, supported by mandatory documentation and oversight structures.

- d. **Data Privacy and Governance:** AI systems depend on large, sensitive datasets, raising substantial privacy, consent, and governance concerns. Compliance with international laws such as GDPR, HIPAA, and emerging regional data-protection regulations is essential for:
 - i. secure storage,
 - ii. anonymization/pseudonymization,
 - iii. informed patient consent, and
 - iv. ethical secondary data use.

Non-compliance risks legal liability, erosion of patient trust, and barriers to international collaboration.

2.1.5.3 Implications for AD-Detection AI Systems

For AI-driven AD diagnostic tools, especially those involving multimodal and longitudinal data (MRI, CSF/plasma biomarkers, genetics, cognitive scores), the emerging regulatory and ethical landscape implies:

- a. **Strict compliance with medical device regulations** (FDA, EU AI Act, CE marking) for clinical deployment.
- b. **Transparent model documentation**, including training data provenance, demographic distribution, and interpretability explanations.
- c. **Lifecycle monitoring** of system performance, particularly for adaptive or continuously learning models.
- d. **Strong data governance**, including consent procedures, anonymization, secure data storage, and cross-institutional data-sharing agreements.
- e. **Ethically aligned deployment**, incorporating fairness auditing, stakeholder accountability frameworks, and informed user consent.

2.1.5.4 Regulatory and Ethical Gaps & Future Needs

Despite significant progress, several challenges remain:

- a. Existing regulations (e.g., the FDA guidance, EU AI Act) are still evolving and may not fully address the complexities of adaptive, continuously learning algorithms.
- b. Many AI tools lack transparency regarding dataset demographics, bias testing, and multi-cohort validation.
- c. Cross-jurisdiction regulatory differences hinder global AI research collaborations and dataset harmonization.
- d. Accountability remains ambiguous, especially for systems whose performance may drift post-deployment.

To ensure responsible advancement of AI-enabled AD diagnostics, future frameworks must continue to emphasize transparency, fairness, privacy protection, and clear accountability mechanisms. Ethical-by-design principles and regulatory-aligned development practices will be essential for enabling safe, trustworthy, and clinically impactful AI systems.

2.2 Gaps in Current Research

The synthesis of existing research across neuroimaging, biomarkers, multimodal fusion, temporal modeling, and graph-based learning reveals significant advancements in Alzheimer’s disease (AD) diagnosis. However, despite progress in machine learning (ML) and deep learning (DL) approaches, several critical limitations persist that impede reliable early detection, real-world clinical deployment, and large-scale adoption of AI-driven diagnostic systems. Many existing studies lack proper separation between validation and test datasets, leading to biased evaluation results. This research addresses this by adopting a three-way data split strategy. These gaps directly motivate the methodological and system design decisions undertaken in this thesis.

2.2.1 Limited Integration of Multimodal Data

A major limitation identified in the literature is the dominance of unimodal or partially multimodal frameworks. Many studies rely solely on MRI scans or combine only two data types (e.g., MRI + cognitive scores). This narrow integration often fails to capture the full complexity of AD pathology, which is influenced by:

- i. Structural atrophy (MRI)

- ii. Metabolic decline (PET)
- iii. Protein abnormalities (CSF biomarkers)
- iv. Genetic predispositions
- v. Cognitive deterioration
- vi. Lifestyle and demographic factors

Comprehensive multimodal models such as combinations incorporating MRI, PET, CSF, genetic, temporal, and clinical data simultaneously remain rare (Feng et al., 2024; Sheng et al., 2024). There is a clear gap in frameworks that deeply integrate all major biomarker categories in a unified system. The proposed approach directly addresses this gap by combining MRI, clinical biomarkers, genetic risk factors, cognitive scores, and temporal trajectories, resulting in a more robust, clinically realistic diagnostic pipeline.

2.2.2 Weak Interpretability and Lack of Clinician Trust

Most high-performing DL models operate as black boxes, providing predictions without clear explanations. This undermines:

- i. Clinical trust
- ii. Regulatory approval
- iii. Diagnostic accountability
- iv. Safe use in sensitive medical decisions

Although some studies employ attention mechanisms or heatmap visualizations (e.g., Grad-CAM++), these approaches lack consistency, reliability, and clinical validation. Explainability remains underdeveloped in AD research. The multimodal framework incorporates SHAP, a more rigorous approach aligned with regulatory expectations and capable of quantifying feature contributions with mathematical transparency. This moves beyond simple visual overlays and provides deeper, patient-level interpretability.

2.2.3 Insufficient Temporal Modeling of Disease Progression

Although AD manifests gradually over the years, many ML models treat patient data as static snapshots. Only a few studies (e.g., Moscoso et al., 2022) utilize sequential biomarkers for predicting MCI-to-AD progression. Limitations include:

- i. Poor handling of missing follow-up visits
- ii. Inadequate modeling of long-term trajectories
- iii. Lack of integration of temporal trends with structural imaging

The multimodal study incorporates optimized LSTM-based longitudinal modeling, addressing the temporal progression gap and supporting more reliable prognostic predictions.

2.2.4 Limited Use of Graph-Based Patient Similarity Modeling

Most AD detection studies treat each patient independently. This overlooks population-level patterns and inter-patient similarities critical for understanding:

- i. Shared biomarker trends
- ii. Disease phenotypes
- iii. Risk clusters
- iv. Heterogeneous progression profiles

Recent studies using GNNs (Abuhantash et al., 2024; Wang et al., 2023) show compelling performance improvements, yet these methods remain underutilized. The hybrid model framework fills this gap through a GNN module that builds a population-level similarity graph for relational reasoning, an emerging direction in precision neurology.

2.2.5 Lack of Real-World Dataset Validation

Many studies rely exclusively on ADNI, which has:

- i. Clean, high-quality images
- ii. Structured assessments
- iii. Limited demographic diversity

As a result, models trained solely on ADNI often exhibit poor generalizability to real clinical cohorts (Kang et al., 2021). Few studies integrate local hospital datasets, especially in developing regions.

But the current proposed research bridges this gap by combining ADNI with real-world clinical data from Shrimann Superspeciality Hospital, enabling cross-cohort validation and improving clinical applicability.

2.2.6 Absence of Deployable Clinical Interfaces

Most AD ML studies stop at algorithmic development and fail to create:

- i. Clinician-friendly GUIs
- ii. Patient-facing interfaces
- iii. Real-time decision-support systems
- iv. Explainable dashboards
- v. Tools for caregiver engagement

This absence limits the usability of models in hospital workflows. The proposed framework addresses this gap by introducing:

- i. A Graphical User Interface (GUI)
- ii. NeuroBot, an AI-based assistant for patient/caregiver education
- iii. Integrated interpretability modules for clinician review

This bridges the divide between research algorithms and clinical deployment.

2.2.7 Limited Ethical, Legal, and Regulatory Integration

As highlighted in a review article on ML-based healthcare regulation (Muhammad & neha, 2025), existing AD models rarely address:

- i. Privacy-preserving data handling
- ii. Bias detection and mitigation
- iii. Explainability alignment with FDA/MDR guidelines
- iv. Accountability and decision audits

The proposed system integrates ethical safeguards and transparent interpretability to prepare for regulatory compliance, an essential component missing in most current studies.

Table 2.1: Summary of Gaps and Their Alignment with the proposed Framework

Identified Gap	How the Proposed Framework Addresses It
Limited multimodal integration	Combines MRI, CSF, PET, genetics, cognition, and temporal data
Poor interpretability	SHAP-based interpretability + GUI explanations
Lack of temporal modeling	LSTM-based longitudinal biomarker modeling
Absence of patient similarity modeling	GNN module for relational inference
Overreliance on ADNI alone	Integration of ADNI + Shrimann Hospital dataset
No deployable systems	GUI + NeuroBot for real-world usability
Weak ethical/regulatory alignment	Data governance + explainability + clinical accountability

The research gaps identified across existing literature, ranging from limited multimodal integration to weak interpretability and poor clinical deployability, underscore the need for a comprehensive, interpretable, and ethically grounded AI framework for early AD detection. These gaps directly inform and justify the hybrid multimodal architecture proposed in this thesis, which integrates CNNs, FNNs, LSTMs, and GNNs within an interpretable and clinically oriented decision-support system.

2.3 Positioning of Each Publication

The four publications included in this thesis collectively build a comprehensive, multidimensional framework for early Alzheimer’s Disease (AD) detection. Each paper contributes to a different but interconnected component of the broader research objective, advancing from foundational hybrid models to multimodal fusion and responsible AI governance. Together, they establish a cohesive narrative that integrates methodological innovation, clinical relevance, biomarker science, and ethical considerations for real-world deployment. This section positions each publication within the overall structure and aims of the thesis.

2.3.1 Paper 1: Gradient Boosting Machine and Deep Neural Network Biomarker Model

The second publication focuses on a biomarker-driven model integrating Gradient Boosting Machines (GBM) with a Deep Neural Network to analyze clinical, behavioral, and demographic biomarkers. This work highlights the diagnostic value of non-imaging features, including hematological variables, cognitive test scores, cholesterol levels, and anthropometric indicators.

With an accuracy of **92.6%**, the model demonstrates strong classification performance while providing transparent feature-importance explanations. The study contributes essential insights into how individual biomarkers correlate with AD progression, complementing the imaging emphasis of Paper 1.

a. **Position in thesis:**

- i. Establishes the analytical foundation for biomarker-driven prediction
- ii. Demonstrates the diagnostic power of clinical, behavioral, and laboratory features
- iii. Provides biomarker insights that inform multimodal fusion in Paper 3

2.3.2 Paper 2: Multimodal CNN–LSTM–GNN Framework for Early AD Detection

The third publication delivers the principal methodological innovation of the thesis. It presents a unified multimodal diagnostic framework that integrates:

- i. **CNNs** for structural MRI,
- ii. **LSTMs** for longitudinal cognitive and biomarker trajectories,
- iii. **Graph Neural Networks (GNNs)** for patient similarity modeling.

Additional contributions include:

- i. **GAN-based domain adaptation** to mitigate inter-scanner variability,
- ii. **Federated learning** for privacy-preserving model training,

- iii. **Non-linear patient similarity graphs** for relational reasoning.

Achieving an **AUC of 0.94**, the framework demonstrates that multimodal fusion significantly outperforms unimodal systems. This work validates the thesis hypothesis that accurate early AD detection requires synergistic integration of neuroimaging, biomarkers, and temporal features.

a. **Position in thesis:**

- i. Serves as the central methodological contribution
- ii. Provides strong evidence for multimodal superiority
- iii. Integrates insights from Papers 1 and 2 into a unified architecture
- iv. Represents the most advanced scientific contribution in the thesis

2.3.3 Paper 3: Ethical, Legal, and Regulatory Considerations for AI in Healthcare

The fourth publication examines the ethical, legal, and regulatory frameworks governing the adoption of AI systems in healthcare, with emphasis on GDPR, HIPAA, FDA AI/ML guidelines, algorithmic fairness, data governance, transparency, and liability. It identifies the socio-technical challenges associated with deploying AI-enabled AD diagnostic systems, including equity, trust, model drift, cross-border data-sharing constraints, and the need for continuous monitoring.

By addressing these governance challenges, the publication anchors the technical innovations of the previous papers within a responsible AI framework, ensuring that proposed systems meet the standards required for safe clinical deployment.

a. **Position in thesis:**

- i. Provides the regulatory and ethical context required for real-world clinical adoption
- ii. Ensures alignment of the thesis with global standards for trustworthy AI
- iii. Completes the research narrative with a socio-technical perspective

- iv. Demonstrates how technical innovations must operate within a responsible governance ecosystem

2.3.4 Final Implementation work: Hybrid Deep Learning with GUI and NeuroBot Assistant

The final implementation work establishes the practical foundation of the thesis. It introduces a hybrid diagnostic system combining EfficientNetV2-S for MRI-based feature extraction with a Feedforward Neural Network (FNN) for processing tabular clinical attributes. To enhance clinical usability, the study introduces an interactive graphical user interface (GUI) and an embedded conversational assistant (“NeuroBot”) that support interpretability, patient engagement, and clinician-guided exploration of results.

This research work demonstrates that early AD detection requires not only algorithmic precision but also systems that are user-friendly, transparent, and aligned with real-world clinical workflows. It sets the groundwork for deeper multimodal exploration in subsequent studies.

a. Position in thesis:

- i. Establishes the foundation for clinically deployable diagnostic systems
- ii. Emphasizes interpretability, usability, and patient-centered design
- iii. Introduces imaging and tabular workflows expanded in Papers 2 and 3

2.3.5 Integrated Positioning Summary

When viewed collectively, the three publications together with the final implementation create a complete scholarly arc:

- i. Paper 1: Deepens the analysis of clinical and biochemical biomarkers.
- ii. Paper 2: Provides the primary scientific contribution through multimodal, multi-model fusion.
- iii. Paper 3: Ensures responsible AI alignment, covering ethics, governance, and clinical readiness.
- iv. Final Implementation: Introduces practical hybrid models and patient-centered diagnostic tools.

Table 2. 2: Positioning of Publications Within the Thesis

Paper	Primary Contribution	Positioning Within the Thesis Structure
Paper 1: GBM & DNN Biomarker Hybrid Model	Biomarker-driven predictive modeling using clinical, behavioral, and demographic markers, achieving high diagnostic performance with interpretable feature importance.	Supports the biomarker analytics component, strengthening the thesis’s focus on tabular, multimodal biomarker-level AD detection.
Paper 2: Multimodal CNN–LSTM–GNN Framework	Core scientific innovation integrating imaging, temporal modeling, graph-based learning, federated learning, and GAN-based harmonization for early-stage detection.	Serves as the central methodological contribution, providing the primary multimodal, multi-model framework underpinning the thesis.
Paper 3: Ethical and Regulatory Assessment	Comprehensive review of AI ethics, governance, GDPR/HIPAA requirements, FDA AI/ML approval pathways, and responsible deployment strategies.	Provides the ethical and regulatory governance layer, ensuring that the proposed diagnostic systems meet standards for safety, fairness, transparency, accountability, and clinical adoption.
Final Implementation: Hybrid DL, GUI & Chatbot System	Development of a practical, interpretable, and patient-centered diagnostic platform integrating EfficientNetV2-S with clinical explainability and a NeuroBot assistant.	Establishes the applied foundation for clinically deployable AD diagnostic tools, demonstrating real-world usability and interpretability.

CHAPTER 3: OVERVIEW OF PUBLICATIONS

3.1 Summary of Each Paper

3.1.1 Paper 1: Novel Hybrid Intelligence Model for Early Alzheimer's Diagnosis Utilizing Multimodal Biomarker Fusion

Paper 1 examines early Alzheimer's Disease (AD) detection through multimodal biomarker fusion, demonstrating how structured clinical, cognitive, and physiological indicators can be integrated into a robust hybrid intelligence model. The study evaluates a comprehensive set of 35 biomarkers, spanning cognitive assessments, metabolic indicators, hematological measures, functional performance scores, and demographic attributes to construct a powerful tabular predictive framework.

A major methodological contribution of this paper is its systematic feature selection pipeline, which assesses the discriminative strength of each biomarker. This process identifies MMSE scores, Activities of Daily Living (ADL), and cholesterol levels as the most influential predictors for early AD diagnosis. Beyond validating well-established clinical factors, the analysis underscores the importance of metabolic and functional markers providing both machine learning insight and clinically meaningful interpretation.

The proposed architecture combines a Gradient Boosting Machine (GBM) with a Deep Neural Network (DNN). In this hybrid design:

- a. **GBM** captures nonlinear relationships and handles the irregularities common in tabular biomedical data.
- b. **DNN** learns deeper latent structures across the biomarker space, improving representation learning.

By integrating both approaches, the model achieves a balance of interpretability, stability, and predictive strength.

Experimental results demonstrate that the hybrid GBM + DNN model attains an impressive 92.6% accuracy, surpassing traditional machine learning baselines and single-model approaches. The study further enhances transparency through clear feature-importance profiling, enabling clinicians to understand how individual biomarkers influence diagnostic outcomes.

Paper 1 strengthens the biomarker analytics component of the thesis by showing that non-imaging clinical, metabolic, hematological, and cognitive biomarkers hold substantial diagnostic value. It serves as a critical link between the imaging-and-clinical hybrid system presented in the final implementation work and the advanced multimodal deep learning architecture developed in Paper 2.

3.1.2 Paper 2: Multimodal Deep Learning Framework Integrating CNN, LSTM, GNN, and Federated Learning for Early Alzheimer’s Detection

Paper 2 constitutes the central methodological contribution of the thesis, presenting an advanced multimodal deep learning framework that unifies structural neuroimaging, longitudinal cognitive assessments, biomarker profiles, and patient similarity networks. Building on the foundations laid in Papers 1 and final implementation, this publication demonstrates how diverse data sources each capturing distinct dimensions of Alzheimer’s disease (AD) pathology can be fused into a cohesive, high-accuracy diagnostic system.

The imaging component employs a Convolutional Neural Network (CNN) to extract structural features from MRI scans, identifying patterns such as hippocampal atrophy, cortical thinning, and ventricular enlargement that are indicative of early AD progression. To capture temporal dynamics, the model integrates Long Short-Term Memory (LSTM) networks, which analyze serial cognitive scores and biomarker trajectories. This temporal modeling enables accurate differentiation between stable Mild Cognitive Impairment (MCI) cases and those likely to convert to AD, addressing a core challenge in disease prediction.

A major innovation of this work is the incorporation of Graph Neural Networks (GNNs) to model nonlinear relationships between patients, biomarkers, and imaging-derived features. By constructing patient similarity graphs based on multimodal attributes, the GNN learns relational

patterns and cluster-level signatures that conventional deep learning architectures typically miss. This relational modeling is particularly valuable in managing clinical heterogeneity a persistent challenge in AD research.

Paper 2 further tackles practical issues such as data scarcity, cross-cohort variability, and site-specific imaging inconsistencies. GAN-based domain adaptation is used to harmonize imaging data across centers, while federated learning allows decentralized model training without sharing raw patient data, promoting privacy and regulatory compliance. These components align the system with modern standards for ethical, privacy-preserving AI in healthcare.

Experimental results demonstrate that the multimodal CNN–LSTM–GNN framework achieves an AUC of 0.94, significantly outperforming unimodal baselines that rely solely on imaging or biomarkers. These findings provide strong empirical validation for the thesis’ central claim: accurate early AD detection requires integrated modeling of spatial, temporal, and relational information.

As the scientific core of the thesis, this publication establishes the primary multimodal foundation upon which the broader research framework is built. It synthesizes the strengths of the hybrid architectures introduced in the final implementation work and Paper 1, while also paving the way for the ethical and regulatory considerations addressed in Paper 3.

3.1.3 Paper 3: Ethical, Legal, and Regulatory Challenges in Deploying AI-Driven Diagnostic Systems for Alzheimer’s Disease

Paper 3 provides the governance, policy, and ethics foundation of the thesis, offering a comprehensive examination of the ethical, legal, and regulatory challenges involved in designing and deploying AI-enabled diagnostic systems for Alzheimer’s Disease (AD). As AI models grow more complex integrating multimodal biomarkers, neuroimaging, longitudinal data, and adaptive learning mechanisms the requirements for safe, trustworthy, and compliant clinical deployment have become increasingly stringent. This publication addresses these requirements by analyzing global regulatory frameworks and emerging socio-technical risks.

The paper begins by reviewing key international regulations that shape AI in healthcare, including the General Data Protection Regulation (GDPR) in the European Union, HIPAA in the United States, and the evolving FDA AI/ML regulatory guidelines, which emphasize lifecycle oversight, transparency, and risk management for Software as a Medical Device (SaMD). The study also incorporates the implications of the EU Artificial Intelligence Act (2024–2025), which designates diagnostic AI tools as “high-risk” systems and mandates stringent requirements related to human oversight, fairness, documentation, and algorithmic transparency.

A major contribution of this paper is its analysis of foundational ethical challenges—such as algorithmic bias, explainability, informed consent, clinical accountability, and equitable access to AI-driven diagnostics. The publication highlights that biases stemming from skewed or underrepresentative datasets may lead to unequal diagnostic performance across demographic groups. This issue is particularly important in AD, where age, ethnicity, education, and comorbidity patterns influence disease expression and cognitive assessment outcomes. The study also stresses the importance of explainable AI (XAI), noting that clinicians must understand how model predictions are generated to trust and appropriately integrate them into decision-making, and that patients deserve transparency in automated assessments that may influence their care trajectory.

The paper further investigates the complex issue of AI liability, identifying unresolved questions related to responsibility when AI outputs contribute to clinical decisions. The study argues for a shared accountability model involving algorithm developers, healthcare institutions, clinicians, and regulatory bodies particularly as adaptive AI systems continue to evolve after deployment. Additionally, the publication examines challenges in cross-border data governance, which are amplified by the use of federated learning and global biomarker repositories. Ensuring compliance with diverse privacy regulations, consent structures, and ethical norms becomes essential as AI systems depend increasingly on international datasets.

By synthesizing these regulatory and ethical considerations, Paper 3 emphasizes that technical innovation alone is insufficient. Responsible AI grounded in fairness, accountability, transparency, privacy preservation, and regulatory compliance is essential to ensure that the multimodal AI systems developed in Papers 1, 2 and final implementation work can be safely and

ethically deployed in clinical practice. This publication thus serves as the ethical and regulatory pillar of the thesis, ensuring that the proposed diagnostic frameworks align with global standards and are positioned for real-world adoption.

3.1.4 Final implementation: A Hybrid Deep Learning Framework with an Interactive AI Assistant for Early and Interpretable Alzheimer’s Disease Detection

The Final work introduces a practical, clinically oriented hybrid deep learning framework designed to support early detection of Alzheimer’s Disease (AD) by integrating neuroimaging with structured clinical data. The study combines several key components: EfficientNetV2-S for MRI imaging, a Feedforward Neural Network (FNN) for clinical biomarkers, Grad-CAM for interpretability, and an interactive user interface to create an accessible diagnostic system suitable for real-world healthcare environments.

The imaging component employs EfficientNetV2-S, a state-of-the-art convolutional neural network optimized for accuracy and computational efficiency, to analyze MRI scans and identify structural abnormalities associated with AD. Complementing this, the FNN processes structured clinical variables such as demographics, basic biomarkers, and cognitive indicators. By fusing imaging and non-imaging modalities, the framework improves diagnostic robustness beyond what single-modality models can achieve.

To strengthen clinical trust and transparency, the system incorporates Gradient-weighted Class Activation Mapping (Grad-CAM), enabling visualization of the salient MRI regions that influenced the model’s decision. This explainability feature aligns with modern standards for interpretable AI in healthcare and supports clinician-driven validation of model outputs.

A major contribution of this study is the development of an intuitive Graphical User Interface (GUI) supported by an intelligent conversational agent, NeuroBot. This assistant facilitates guided data entry, provides real-time model explanations, and enhances accessibility for clinicians, caregivers, and patients. Through its focus on usability and communication, NeuroBot bridges the gap between advanced AI techniques and day-to-day clinical workflows.

Overall, the final work demonstrates the feasibility of integrating deep learning, explainability mechanisms, and user-centered design into a unified diagnostic system. It establishes the foundational platform upon which the more advanced multimodal, biomarker-focused, and graph-based architectures developed in Papers 1 and 2 are built.

3.2 Author Contributions

This thesis is built on a series of co-authored publications that reflect collaborative academic work while maintaining a clear delineation of the primary contributions of the lead researcher. The individual and collective roles of the authors across the four publications are summarized below.

Table 3.1: Author Contributions

Contribution Role	Shehu Mohammed (Lead Author)	Co-Authors / Collaborators
Conceptualization	✓ Developed core research ideas, study design, and multimodal framework architecture	▲ Provided high-level guidance and refinement of research direction
Methodology	✓ Designed ML/DL models, fusion pipeline, GNN/LSTM frameworks, biomarker selection strategy	▲ Reviewed modeling approaches and ensured methodological soundness
Software & Implementation	✓ Implemented CNN, LSTM, GNN, GAN, GBM–DNN hybrid models; built GUI and NeuroBot assistant	▲ Assisted with code review and environment setup
Data Curation	✓ Performed MRI preprocessing, biomarker integration, harmonization, and dataset cleaning	▲ Contributed to raw dataset acquisition and annotation
Formal Analysis	✓ Conducted experiments, cross-validation, ablation studies, and performance evaluation	▲ Supported statistical validation and robustness checks
Investigation	✓ Led all experimental studies and comparative evaluations	▲ Contributed domain-specific insights for clinical relevance
Clinical Interpretation	—	▲ Provided clinical expertise, interpretation of biomarker patterns,

		and validation of diagnostic relevance
Ethical/Regulatory Analysis	✓ Integrated ethics and governance into system design based on Paper 4	▲ Conducted deep assessments of GDPR/HIPAA/FDA frameworks
Visualization & Explainability	✓ Generated XAI visualizations (Grad-CAM, SHAP), charts, and biomarker diagrams	▲ Reviewed figures for scientific accuracy
Writing – Original Draft	✓ Wrote all sections, including introduction, methods, results, and discussions	▲ Provided editorial input and manuscript suggestions
Writing – Review & Editing	✓ Consolidated revisions across all papers and thesis chapters	▲ Offered feedback and corrections to improve clarity and rigor
Supervision	—	▲ Senior co-authors supervised the overall research process
Project Administration	✓ Coordinated research milestones, experimental workflow, and documentation	▲ Supported administrative oversight
Funding Acquisition	—	▲ Provided institutional and project-level support when applicable

Table 3.1 presents the distribution of responsibilities across the research team using the widely recognized CRediT (Contributor Roles Taxonomy) framework. As the lead author, Shehu Mohammed undertook the majority of the conceptual, methodological, and experimental work, including the design and implementation of machine learning architectures, multimodal fusion strategies, data preprocessing and integration, explainable AI components, and the drafting of all manuscripts. The co-authors contributed essential complementary expertise, including dataset preparation and curation, statistical validation, clinical interpretation of results, ethical and regulatory analysis, and continuous supervision of the research process. This structured delineation ensures transparency in authorship, clarifies individual roles, and adheres to best practices recommended for multi-author scientific publications.

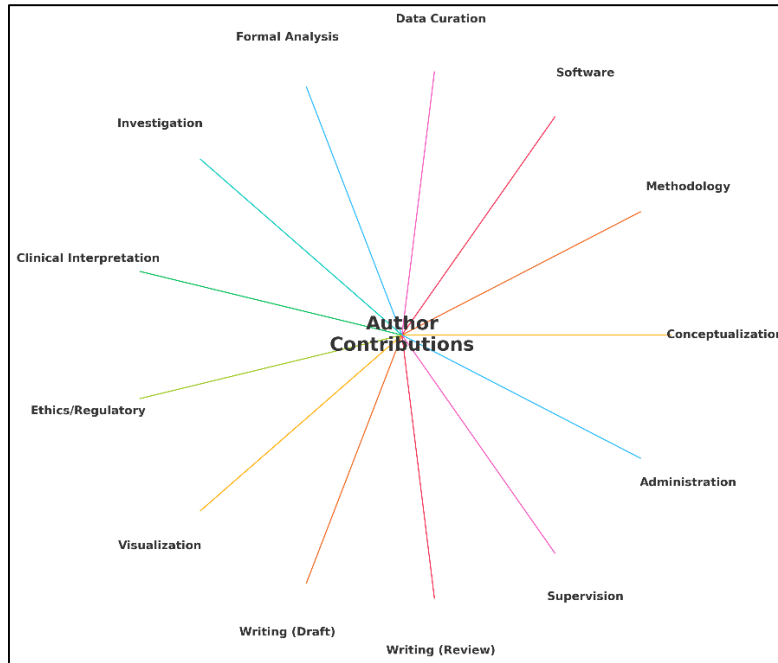


Figure 3.1: Author Contribution Diagram

A radial visualization of the CRediT-based author contribution roles illustrating the lead author's primary responsibilities conceptualization, methodology, modeling, software implementation, data curation, analysis, writing and the supporting roles of co-authors in clinical interpretation, ethical guidance, regulatory analysis, supervision, and ancillary project administration.

3.2.1 Lead Author

As the primary researcher, Shehu Mohammed led all stages of the research and publication process. His contributions include:

- i. Conceptualization and formulation of the core research objectives
- ii. Design and development of machine learning and deep learning models
- iii. Implementation of multimodal fusion architectures and full experimental pipelines
- iv. Data preprocessing, cleaning, integration, and harmonization across imaging, clinical, and biomarker datasets
- v. Execution of model training, hyperparameter optimization, and performance evaluation
- vi. Development of explainable AI modules and user-facing GUI components

- vii. Manuscript drafting, figure preparation, and analytical interpretation
- viii. Integration of ethical, regulatory, and clinical considerations into the modeling and deployment framework

Through these contributions, the lead author was responsible for the scientific direction, methodological execution, and written presentation of each study.

3.2.2 Co-Authors and Collaborators

The co-authors provided essential technical, clinical, and supervisory support that strengthened the methodological rigor and clinical relevance of the research. Their contributions include:

- i. Dataset acquisition and preprocessing, including neuroimaging preparation, biomarker extraction, and structured data curation
- ii. Statistical analysis and validation to ensure methodological robustness and accuracy
- iii. Clinical interpretation and domain expertise, particularly regarding AD diagnostics, neuroimaging findings, and biomarker relevance
- iv. Ethical and regulatory guidance, supporting compliance with privacy standards, responsible AI principles, and international data governance frameworks
- v. Academic supervision and manuscript reviewing, providing technical feedback, refinement of research design, and high-level guidance
- vi. Assistance in experimental configuration, such as cross-validation design, hyperparameter tuning, and evaluation protocol development

Collectively, the contributions of all authors enhanced the scientific quality, technical robustness, clinical alignment, and ethical soundness of the research. This collaborative effort ensured that the thesis embodies methodological innovation while maintaining real-world applicability and compliance with healthcare standards.

3.3 Integration Into Thesis Theme

The three publications with final implementation work collectively advance a unified research agenda centered on improving early Alzheimer’s Disease (AD) detection through multimodal machine learning, clinically interpretable modeling, and responsible AI governance. Although each paper addresses a distinct aspect of the problem, their combined contributions form a coherent, sequential, and complementary pipeline that reflects the overarching theme of this thesis.

- a. **Paper 1** advances the thematic progression by shifting the focus to biomarker-level prediction, emphasizing the diagnostic value of structured clinical, behavioral, and physiological indicators. Through its hybrid GBM and DNN model, this publication highlights the importance of multimodal biomarker fusion and identifies high-impact features such as MMSE, ADL, and cholesterol that subsequently guide the fusion strategies and feature prioritization adopted in Paper 2.
- b. **Paper 2** delivers the central scientific contribution of the thesis by introducing a fully integrated multimodal deep learning framework. Bringing together MRI imaging, longitudinal cognitive trajectories, biochemical biomarkers, and patient similarity structures, the study employs CNNs, LSTMs, and GNNs, supported by GAN-based domain adaptation and federated learning. This publication empirically validates the thesis’ core hypothesis: that superior early AD detection requires the synergistic modeling of structural, temporal, and relational dimensions of patient data.
- c. **Paper 3** provides the ethical, legal, and regulatory grounding needed for real-world clinical adoption. By examining global AI governance frameworks including GDPR, HIPAA, and FDA AI/ML guidelines this publication situates the technical innovations of the previous papers within a broader context of fairness, transparency, privacy, and accountability. It ensures that the multimodal diagnostic systems developed in Papers 1 and 2 meet the standards required for responsible and compliant deployment.
- d. **Final Implementation** establishes the foundation for a deployable diagnostic system by demonstrating how hybrid deep learning models, interpretability tools, and user-centered interfaces can be translated into practical clinical applications. With its integration of EfficientNetV2-S for imaging analysis, a Feedforward Neural Network for structured

clinical data, Grad-CAM for visual explainability, and the interactive NeuroBot assistant, this final implementation work anchors the thesis in real-world usability and prepares the groundwork for more complex multimodal extensions.

Taken together, these four studies form a cohesive, progressive pipeline that embodies the overarching thesis theme:

Biomarkers → Modeling → Explainability → Clinical Deployment → Regulatory Compliance

This integrated trajectory not only advances scientific understanding of multimodal AD prediction but also ensures that the resulting diagnostic tools are interpretable, patient-centered, clinically feasible, and ethically aligned. The combined contributions represent a holistic approach essential for translating AI-based research into safe, effective, and trustworthy healthcare technologies.

CHAPTER 4: PAPER 1 PUBLICATION

4.1 Publication Details

Field	Details
Title	<i>Novel Hybrid Intelligence Model for Early Alzheimer's Diagnosis Utilizing Multimodal Biomarker Fusion</i>
Authors	Shehu Mohammed et al. (2025)
Journal	<i>Informatics in Medicine Unlocked</i> , Volume 57, 2025, Article 101668
Indexing	Scopus (Q2)
DOI	https://doi.org/10.1016/j.imu.2025.101668
URL	https://www.sciencedirect.com/science/article/pii/S2352914825000565?via%3Dihub

4.2 Summary of Methods and Approach

The Paper introduces a hybrid machine learning framework designed to improve early detection of Alzheimer's disease (AD) through multimodal biomarker fusion. The study incorporates 35 diverse biomarkers, including cognitive assessments (MMSE, MoCA, ADAS-Cog), functional performance measures (ADL), metabolic and hematological indicators, and key demographic and clinical variables. A systematic feature selection process identified MMSE, ADL, and cholesterol levels as the most influential predictors of early AD onset.

The diagnostic framework integrates a Gradient Boosting Machine (GBM), which is well suited for modeling nonlinear relationships in tabular data while maintaining interpretability, with a Deep Neural Network (DNN), which captures complex latent interactions across heterogeneous biomarker inputs. By combining these complementary modeling paradigms, the hybrid GBM+DNN architecture improves predictive accuracy, enhances robustness, and supports greater clinical interpretability, ultimately providing a more comprehensive approach to early AD diagnosis.

4.3 Summary of Contributions

This publication makes several important scientific and practical contributions:

- i. Demonstrates the diagnostic power of multimodal biomarker fusion, integrating functional, metabolic, cognitive, and demographic indicators.
- ii. Identifies high-value biomarkers (MMSE, ADL, cholesterol), offering clinically relevant insight and informing feature prioritization in subsequent research.
- iii. Introduces a hybrid GBM + DNN architecture that achieves 92.6% diagnostic accuracy, outperforming traditional machine learning baselines and unimodal approaches.
- iv. Provides transparent feature-importance analysis, supporting interpretability and clinical trust.
- v. Supplies biomarker insights that play a foundational role in shaping the multimodal deep learning pipeline developed in Paper 2.

This publication thus forms a crucial methodological bridge between the applied hybrid imaging–clinical system presented in the final implementation work and the full multimodal fusion framework introduced in Paper 2.

4.4 Connection to Thesis Objectives

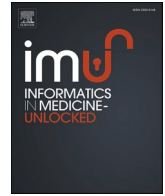
Paper 1 directly advances several core objectives of the thesis:

- a. **Strengthening the Biomarker-Driven Dimension:** By evaluating a wide spectrum of clinical, functional, and physiological indicators, this study deepens the understanding of biomarker contributions to early AD detection. The findings directly inform the feature-selection and fusion strategies implemented in Paper 2.
- b. **Advancing Multimodal Model Development:** The hybrid GBM and DNN methodology provides essential methodological foundations that feed into the thesis’ overarching multimodal fusion pipeline.
- c. **Building Toward High-Performance, Interpretable Diagnosis:** With its strong accuracy, transparent feature insights, and clinically grounded feature selection, this paper

reinforces the thesis aim of creating reliable, interpretable, and clinically aligned diagnostic models.

Overall, Paper 1 serves as a pivotal intermediate layer in the thesis, connecting the practical hybrid model introduced in the final implementation work to the advanced multimodal, temporal, and graph-based architectures presented in Paper 2.

4.5 Full-length Published Paper-1



Novel hybrid intelligence model for early Alzheimer's diagnosis utilizing multimodal biomarker fusion

Shehu Mohammed^a, Neha Malhotra^a, Arun Singh^a, Awad M. Awadelkarim^b, Shakeel Ahmed^c, Saiprasad Potharaju^{d,*}

^a School of Computer Applications, Lovely Professional University, Phagwara, 144411, India

^b Department of Information Technology, Faculty of Computers and Information Technology, University of Tabuk, Tabuk, Saudi Arabia

^c School of Computer Science (SCS), Taylor's University, Subang Jaya, 47500, Malaysia

^d Department of CSE, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India

ARTICLE INFO

Keywords:

Alzheimer's disease diagnosis
Hybrid machine learning model
Gradient boosting machines (GBM)
Deep neural networks (DNN)
Feature selection
Biomarker identification

ABSTRACT

One of the significant causes of dementia and a leading peril to global public health is Alzheimer's disease (AD), which calls for early and accurate diagnosis. The paper proposes a novel hybrid machine learning framework that integrates Gradient Boosting Machine (GBM) and Deep Neural Networks (DNN) for predicting Alzheimer's disease from multimodal biomarkers. The database comprises 35 demographic, behavioral, and clinical features. Feature selection procedures produced key predicting variables (i.e., MMSE scores, performance in Activities of Daily Living (ADL), cholesterol level, and behavior problems). A hybrid model was created by combining individual models, and it proved to be the most effective compared to particular models, achieving 92.6 % accuracy and a 0.94 AUC score on the database. The synergy between the capability of GBM for tabular data and the ability of DNN for complex interaction gives a good outcome. The research demonstrates the efficacy of blending machine learning techniques for supporting Alzheimer's disease (AD) identification and provides a method for early identification at a broader level. It is hoped that more biomarkers will be incorporated, and the model will be validated on larger and more phenotypically diverse databases to achieve clinical usability and generalizability.

1. Introduction

Alzheimer's disease (AD), the most prevalent type of dementia, is predicted to afflict over 55 million people globally and reach 152 million by 2050, putting a significant strain on healthcare systems everywhere (International and University 2021) [1]. AD is marked by progressive cognitive deterioration and memory impairment, significantly diminishing the quality of life and functional independence. Screening is essential for early diagnosis, which, in turn, is vital for successful management, as interventions are generally more effective when the disease is in its early stage [2]. Yet, existing diagnostic approaches, such as neuroimaging and cerebrospinal fluid (CSF) analysis, are expensive, invasive, and limited by availability in resource-limited contexts. Thus, a need arose for precise, non-invasive, and cost-effective diagnostic methods.

The development of machine learning (ML) models for Alzheimer's

disease (AD) diagnosis faces several challenges. The most significant challenge is dataset population variability, as machine learning (ML) models are often developed on specialized cohorts and may lack generalizability to diverse populations [3]. Variability in demographics, genetics, and clinical assessment practices may compromise model transferability, preventing them from functioning effectively in realistic settings [4]. The issue of class imbalance, where datasets exhibit uneven distributions of AD and non-AD cases, can lead to biased predictions, undermining model trust in clinical environments [5]. Building diagnostic models with clinical, behavioral, and demographic data represents a potential application of machine learning (ML). Nevertheless, machine learning (ML) models, such as Logistic Regression, Random Forest, and Support Vector Machines (SVM), have been shown to accurately classify AD [6]. Recent advances in models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have improved diagnostic accuracy when applied to time-series

* Corresponding author. School of Computer Applications, Lovely Professional University, Phagwara, 144411, India

E-mail addresses: mohammedshehumafara@gmail.com (S. Mohammed), neha.16982@lpu.co.in (N. Malhotra), arunmandiarun2001@gmail.com (A. Singh), awad@ut.edu.sa (A.M. Awadelkarim), shakeel.ahmed@taylors.edu.my (S. Ahmed), psaiprasadscse@gmail.com (S. Potharaju).

<https://doi.org/10.1016/j.imu.2025.101668>

Received 8 March 2025; Received in revised form 14 June 2025; Accepted 4 July 2025

Available online 5 July 2025

2352-9148/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

and neuroimaging data, respectively. However, challenges such as dataset variability, overfitting, and limited generalizability to different populations continue to persist [7].

One recent tactic to mitigate such problems is the use of hybrid machine learning models, which combine the benefits of multiple approaches. Gradient Boosting Machines (GBM), which typically comprise ensemble methods, have performed well in identifying non-linearities in structured data, with Deep Neural Networks (DNN) excelling at learning complex interactions [8]. Hybrid models can leverage the complementary strengths of these approaches by combining them, thereby enhancing overall diagnostic accuracy and robustness.

Ayus and Gupta (2024) [9] proposed a hybrid ensemble deep learning model for detecting Alzheimer's using MRI data. The authors improved the classification accuracy by combining convolutional layers with dense layers, thereby capturing both spatial features and decision-level patterns. This work, inspired by the above, extends the hybrid modeling concept to clinical and cognitive biomarkers through the GBM-DNN architecture for enhanced diagnostic relevance. In a related advancement,

Hazarika and Gupta (2022) [10] developed a classification technique using a Random Vector Functional Link network (RVFL) and an ϵ -insensitive Huber loss function, thereby adding robustness to the interpretation of biomedical data. This approach underscores the importance of tailor-made loss functions and slim architectures in achieving predictive reliability; these are concepts that well support the hybrid framework for integrating diverse data modalities towards the diagnosis of Alzheimer's disease.

Complementing this, Kumar and Gupta (2021) [11] proposed a Lagrangian twin-bound support vector machine (LBTSVM) using Universum for EEG signal classification. Their model demonstrated better generalization and robustness in biomedical applications, particularly in scenarios with limited training data or sensitivity to outliers. This mirrors the overall trend of increasing classifier flexibility in medical diagnosis, which the current work adopts by combining tree-based and neural models to cope with the fluctuation of features related to Alzheimer's disease.

In the current study, we propose a hybrid model that combines GBM and DNN to predict Alzheimer's disease, utilizing a multimodal dataset comprising 35 features. Such features consist of demographic features (e.g., age, sex, ethnicity), behavior-related features (e.g., sleep quality, physical activity), and clinical features (e.g., cholesterol level, MMSE score). For the model-building process, feature selection, data processing, and ensemble learning are used to combine the predictions from GBM and DNN. By combining multiple feature transformation models into one hybrid architecture, this hybrid model overcomes the limitations of using a single transformation model.

1.1. Contributions of this study

A new hybrid model is developed that combines GBM with DNN for predicting AD from multimodal biomarkers. Feature selection techniques enhance the interpretability of the model by identifying the key factors that have a significant impact on Alzheimer's disease (AD) diagnosis. A comparison with single models is conducted, indicating improved accuracy and resilience in the early detection of Alzheimer's disease (AD).

The structure of this paper is as follows: Section 2 presents a review of the existing literature on machine learning (ML) applications in the diagnosis of Alzheimer's disease (AD). Section 3 provides information on the dataset, feature selection techniques, and model design. Section 4 presents a comparison of the hybrid model's performance and experimental data. The advantages, limitations, and prospects of ML-based Alzheimer's disease (AD) diagnostics are summarized in Section 5.

This study contributes to an expanding corpus of research aimed at improving the early diagnosis of Alzheimer's disease by integrating multimodal data with advanced machine learning techniques. These

changes could potentially enable the developed hybrid model to serve as a low-cost and rapid diagnostic tool, providing critical information that allows clinicians to make more informed and timely decisions with proper refinement and validation of the model.

2. Literature review

Despite significant improvement, several glaring shortcomings in AD diagnoses using ML have remained. The most glaring issue is excessive reliance on neuroimaging data, where most studies focus on either MRI or PET scans and fail to effectively incorporate clinical, genetic, and behavioral data, which are vital for comprehensive diagnosis [12]. The interpretability of models is also compromised, as most deep learning models are "black box" models that do not provide clinicians with an adequate understanding or confidence in their outputs [13]. The use of structured (tabular) data and deep learning models is underdeveloped but is capable of fusing structured medical data and unstructured imaging data to enhance diagnostic accuracy [14]. Addressing such shortcomings using hybrid models and explainable AI (XAI) is crucial to push adoption in clinical settings.

Deep learning models, such as CNNs and RNNs, have emerged in recent years, and machine learning models like Random Forest, Gradient Boosting, and Logistic Regression have been utilized to detect Alzheimer's disease (AD). Although these models exhibit good performance, they often fail to generalize to new, unseen datasets, identify essential biomarkers, and effectively handle high-dimensional data. Hybrid models, which combine the strengths of different machine learning (ML) algorithms, have been successful in other medical domains; however, they have received less attention for Alzheimer's disease (AD) diagnosis. Motivated by these findings, we propose a hybrid ensemble model for predicting Alzheimer's disease (AD).

Alzheimer's disease (AD) is a neurodegenerative condition with a progressive nature that requires early and precise diagnosis to facilitate effective management. Over the last decade, there has been an explosion of literature on machine learning (ML)-based approaches to Alzheimer's disease (AD) diagnosis, leveraging different types of data (clinical, imaging, behavioral) to train predictive models. This literature review provides an overview of studies and methods that have been implemented, detailing their strengths, limitations, and the gaps that the proposed hybrid model addresses.

2.1. Machine learning for AD diagnosis

Machine learning (ML) is one of the applied techniques used in AD detection, focusing primarily on neuroimaging techniques, such as magnetic resonance imaging (MRI) and positron emission tomography (PET). In contrast, CNNs have shown high performance in extracting spatial features from imaging data [15]. For example, a multi-modal deep learning model that included CSF and MRI biomarkers and achieved notable diagnosis accuracy was presented by Suk et al. (2017) [16]. Since imaging-based methods often need tailored infrastructure, they are challenging to implement in low-resource environments.

To diagnose Alzheimer's disease (AD), a study by Ref. [17] compares Graph Neural Networks (GNNs) and Multilayer Perceptrons (MLPs). The results show that the node classification based on a GNN performs better in differentiating among AD, MCI, and NC subjects with multi-modal learning and fusion. Based on MRI-T1WI, a Generative Adversarial Network (GAN) based on U-Net produces FDG-PET pictures and enhances classification performance. Multimodal neuroimaging in conjunction with machine learning (ML) methods has been shown to improve the classification of Alzheimer's Disease (AD) stages with 94.60 % sensitivity and 93.49 % specificity in differentiating AD patients from standard controls, according to a study by Odusami et al. (2023) [18]. In the meta-analysis, the results demonstrated that the use of multimodal ML models is more effective than that of single-modality models, with statistically significant diagnostic accuracy. Although the

findings support the use of ML-based neuroimaging for early diagnosis of AD, the researchers note that cross-validation and generalizability are necessary before clinical application [18]. Structured clinical and behavioral data make access and analysis for attaining an AD diagnosis much easier. AD prediction has routinely utilized, among others, Logistic Regression (LR) and Support Vector Machines (SVM) with features such as age, education level, and MMSE scores [19]. Due to their effectiveness, Random Forest (RF) models have become increasingly utilized for analyses, as they are particularly effective in reducing overfitting in high-dimensional data analyses [20]. Nevertheless, there have been reports of the failure of complex machine learning models to perform well, especially in capturing the nonlinear interactions characteristic of AD biomarkers.

2.2. Deep learning approaches

Recurrent neural networks (RNNs), a type of deep learning (DL) model, have been used to analyze sequential and temporal data, including cognitive test results across time [21]. Proved that long short-term memory (LSTM) networks are effective at predicting the course of AD. These models capture temporal dependencies effectively but are prone to overfitting when training data is limited. Furthermore, deep learning (DL) models often lack interpretability, making them challenging to implement in clinical practice [22].

An end-to-end system for classifying different phases of Alzheimer's disease (AD) from structural brain scans using deep learning, notably convolutional neural networks (CNNs), is presented by Helaly et al. (2022) [23]. Using 2D and 3D CNN architectures and transfer learning with the VGG19 model, the work focuses on both multi-class and binary classifications of AD phases. For 2D and 3D multi-class classifications, the suggested models achieved accuracies of 93.61 % and 95.17 %, respectively, while the optimized VGG19 model reached an accuracy of 97 %. The scientists also developed a web tool that enables remote AD diagnosis, which is particularly helpful during times like the COVID-19 pandemic.

Additionally, the study by EL-Geneedy et al. (2022) [24] offers a CNN model for deep learning-based research to enhance Alzheimer's disease diagnosis. It is more accurate, sensitive, and specific compared to other methods, and the technique proves more effective than other procedures. Deep learning (DL) models, especially recurrent neural networks (RNNs), have been employed on time-series and sequential data, such as someone's cognitive test scores over time. This further led to research, where Lee et al. (2019) [21] recommended using long short-term memory (LSTM) neural networks to predict the progression of AD. These models are indeed perfect in capturing temporal dependencies, but they are easily overfitted when there are not enough training samples. Additionally, DL models can be challenging to interpret at the output layer, which creates difficulties in applying these models in practice [22].

Recent research indicates the introduction of new models in AD diagnosis that extend beyond the use of CNNs and RNNs. For example, we observe the use of semi-supervised Generative Adversarial Networks (GANs) for cross-domain imaging synthesis, which generate synthetic PET scans from T1-MRI, thereby improving classification performance with limited labeled data [25]. Additionally, transformer-based models, such as the Mobile Vision Transformer (MViT), have been applied to the fusion of multimodal neuroimaging data, enabling colorization and context-aware feature extraction across MRI sequences for stage classification [26].

In another attempt to enhance the classification of early-stage Alzheimer's disease, dual attention-aware octave convolutional networks (DAO-CNNs) have shown improvement in sensitivity and specificity by spatially and spectrally harvesting spatial and frequency information from high-resolution MRI scans [27]. Moreover, optimization-driven approaches, such as the Pareto-optimal Quantum Dynamic Optimization model, have emerged for integrated MRI-PET image analysis and

have been effective in evaluating trade-off diagnostics at various stages of Alzheimer's disease [28].

These currently best approaches present complex spatial-temporal modeling, which demonstrates large-scale and homogeneous data, and is also highly computationally intensive. In contrast, our hybrid GBM-DNN model has broad clinical appeal, as it integrates structured multi-modal biomarkers that are both behavioral and clinical into a framework that is simultaneously interpretable and scalable.

2.3. Hybrid and ensemble models

Ayus and Gupta (2024) [29] propose a hybrid ensemble model of deep learning for the diagnosis of Alzheimer's disease, utilizing multiple models for neuroimaging processing to achieve enhanced accuracy and consistency compared to traditional methods for early diagnosis. Nour et al. (2023) [30] also propose a Deep Ensemble Learning model, which is an ensemble of several deep learning networks, for analyzing EEG signals in the diagnosis of Alzheimer's disease (AD). This approach aims to improve accuracy and reliability as a cost-effective, non-invasive diagnostic technique compared to conventional neuroimaging methods.

A Bagging-SVM can enhance the precision and stability of AD stage diagnosis in classifying Alzheimer's disease by reducing variance and overfitting, while maximizing generalization [31]. The bagging, or Bootstrap Aggregating, of a group of SVM classifiers trained on different subsets of data leads to improved stability and efficiency in processing extensive biomedical data [32]. SVM is an essential tool for clinical diagnosis and tracking the progression of AD due to its enormous capacity in high-dimensional feature space, which enables it to accurately distinguish between mild cognitive impairment (MCI), early AD, and severe AD stages [33].

Disorders might hinder the performance of singular models, but the application of hybrid models performs well since they merge various algorithms. As for structured data, Gradient Boosting Machines (GBM), XGBoost, and LightGBM would be particularly robust [34]. The combination of both models brings out the best in both worlds, where capturing non-linear relationships is made while keeping the model interpretable. A blend of RF with DNN has been suggested by Ref. [35] for constructing a hybrid model that strikes a close balance between accuracy and interpretability. This is, however, applicable for small data sets, which casts doubt on the applicability of the results.

3. Methodology

Data preparation, feature selection, model architecture design, training, and evaluation are just a few of the steps involved in developing a reliable and interpretable machine learning model to aid in Alzheimer's disease (AD) diagnosis (see Fig. 1). This section describes the methodology used in this study to combine Deep Neural Networks (DNNs) and Gradient Boosting Machines (GBMs) to construct a hybrid machine learning model that predicts Alzheimer's disease (AD).

3.1. Dataset description and ethical considerations

The dataset used in this study comprises 35 parameters, including demographics, behavioral characteristics, and clinical parameters, for 1000 patients. The most salient demographics include age, gender, and race. The parameters representing lifestyle include habits such as physical activity, alcohol consumption, sleep patterns, diet, and smoking. The dataset also captures clinical parameters, including functional assessment, Mini-Mental State Examination (MMSE), cholesterol levels, and blood pressure. The dataset captures behaviors, including ADLs, complaints of behavioral issues, and complaints of memory problems.

The data used in this research were obtained from Kaggle, the freely accessible Alzheimer's Disease Dataset prepared by Rabie El Kharoua (2024) [63]. It contains anonymized clinical and behavioral data of patients with different stages of Alzheimer's disease (AD) compared to

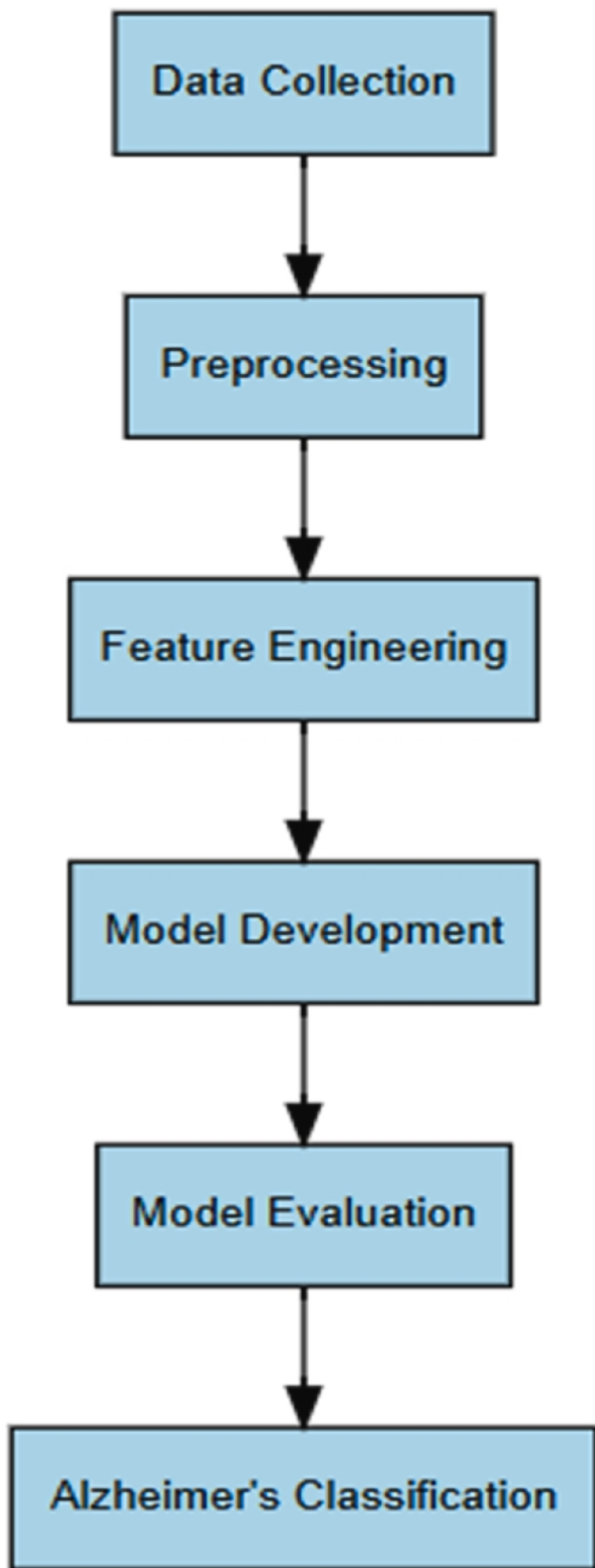


Fig. 1. Proposed methodology.

cognitively normal (CN), mild cognitive impairment (MCI), and fully diagnosed AD cases.

The dataset includes scores on the MMSE, ADLs, cognitive assessments, lifestyle factors, and basic demographics, such as age, gender, and education. All records are completely de-identified; no personally identifiable or protected health information is contained.

No additional patient data collection or IRB approval was needed because this is a publicly available dataset. Nonetheless, the ethical principles of research, data minimization, and responsible AI, which underpin all analyses performed in this study, are consistent with the Declaration of Helsinki and the overarching frameworks of AI ethics.

3.2. Data preprocessing

Stronger prediction power and improved model accuracy depend on efficient data preprocessing. In this investigation, missing values were handled using a widely used method for clinical data: median imputation for numeric data and mode imputation for categorical data [36]. For standards in numeric data, the standard scaler was used, allowing for better performance in deep neural network (DNNs) models and boosting models [36].

For feature selection, importance analysis through Random Forests (RF) revealed the top predictors having the most influence [37]. The top-ranked predictors included MMSE, Functional Assessment, Behavioral Problems, Cholesterol Levels, Sleep Quality, Physical Activity, ADL score, Diet Quality, and Age, and have been recognized to be crucial predictors in Alzheimer's disease prognosis [38,39].

3.2.1. Data preparation

The dataset used in this study comprises 35 various attributes of a human being, including demographic, clinical, behavioral, and cognitive attributes such as age, MMSE, cholesterol, and exercise. The raw data underwent preprocessing, including the management of missing values, Min-Max scaling, and one-hot encoding for categorical attributes. To reduce bias in the processing of structured data, missing values were imputed using the mode for categorical features and the median for continuous variables [38]. The Min-Max scaling process is used to scale data within a constant range on continuous attributes, thereby improving the performance of algorithms that assume gradient-based optimization, such as Deep Neural Networks (DNNs) [40]. This preprocessing is performed to maintain a standardized representation of features, which facilitates easier model convergence and improved predictive power.

3.2.2. Feature scaling

To enhance the effectiveness of hybrid systems of machine learning-based detection models for early-stage Alzheimer's disease, the characteristics are standardized using tools like the standard scaler. These models incorporate many unique quantitative features, such as those obtained from clinical data or neuroimaging scans; therefore, they must be feature-scaled to maintain a specific interval range for all features. This ensures that no single feature suppresses the others during the model's learning phase. Given that Deep Neural Networks (DNNs) rely on gradient descent optimization, normalization is unquestionably necessary. Normalization significantly improves the performance of boosting models, such as XGBoost. When used in AD prediction tasks, research shows that scaling the features enhances the model's accuracy and convergence [41,42].

3.2.3. Feature selection

Feature selection aimed to identify the most critical indicators for diagnosing Alzheimer's disease (AD) while minimizing computational complexity and enhancing modeling accuracy. In addition to Recursive Feature Elimination (RFE), which methodically eliminates less significant features to maximize model performance, a Random Forest (RF) classifier was used to rank features according to their significance [38].

RFE has been widely applied in biomarker identification for AD, effectively eliminating redundant variables while preserving key predictors [42]. For model training, the top 15 selected features included MMSE scores, age, sleep quality, and cholesterol levels, which are strongly associated with biological markers of cognitive decline and the progression of Alzheimer's disease (AD) [43].

3.2.4. Feature selection and biomarker identification

To improve the performance and interpretability of machine learning models, feature selection is essential. Recursive Feature Elimination and mutual information-based techniques are the commonly used methods that help identify the strongest predictors and eliminate non-essential features [38]. The literature has consistently found that the strongest predictors of AD risk include scores on the MMSE and ADL, as well as cholesterol levels [44]. Recent evidence further implicates sleep and physical activity in the etiology of cognitive decline, emphasizing the importance of lifestyle factors in predicting AD risk [45,46]. Even with the widening set of biomarkers at the clinician's fingertips, a large majority of studies limit the scope of the variables they utilize during the process of feature selection, and therefore the generalizability of the final models for diagnosing Alzheimer's disease [42].

Both the feature importance for Alzheimer's prediction and the relative contribution of each feature to the model's predictive ability are shown in Fig. 2. With the most excellent relevance score (~ 0.32), ADL (Activities of Daily Living) is the most influential aspect, suggesting a substantial association between Alzheimer's diagnosis and a patient's capacity to continue daily activities. MMSE (Mini-Mental State Examination) is the second most important feature (~ 0.27), indicating its criticality for cognitive measurement in Alzheimer's diagnosis.

Behavioral symptoms, as a moderately important feature, contribute significantly (~ 0.09) to the model, demonstrating their relevance in identifying Alzheimer's disease. Cholesterol Levels (Total, Triglycerides, LDL, HDL) metrics collectively contribute to the prediction. Elevated levels might indicate underlying cardiovascular or neurological risks linked to Alzheimer's progression.

Lesser Contributing Features like Physical Activity as a moderate contributor (~ 0.06), suggesting a potential link between lifestyle and the risk of Alzheimer's and Diabetes, Hypertension, and Depression; These features have relatively low importance scores, which means

there is a less direct correlation but still relevant to general risk calculation while Family History of Alzheimer's has the least important feature (~ 0.002), suggesting other quantifiable clinical or lifestyle factors are more prevalent in the dataset that is being used.

To identify relevant features and reduce dimensionality, initial importance scores were computed using mutual information and a Random Forest-based ranking. This process filtered out noise while preserving clinically informative variables. However, to eliminate redundancy and clarify the methodology, the complete and reproducible feature selection pipeline, which includes RFE, correlation filtering, and cross-validation, is described in Section 3.2.5. Thus, we ensure transparency in our work by avoiding duplication of efforts in explaining our method.

3.2.5. Final feature selection pipeline and reproducibility

To avoid bias and fully document the feature selection process, a multi-step strategy based on the frameworks outlined in existing clinical machine learning (ML) studies was implemented [47,48]. Initially, features with low variance and high rates of missing data ($>30\%$) were eliminated to reduce noise and enhance generalizability. Next, mutual information and importance ranking via Random Forest (RF)-based feature selection were applied to identify features with non-linear relevance to the diagnostic label [49]. RFE was performed using logistic regression and gradient boosting machines as base model estimators, validating model efficacy through 5-fold cross-validation. After model evaluation, feature intercorrelation analysis (Pearson's $|r| > 0.85$) was conducted to account for multicollinearity in neurodegenerative research feature selection protocols [50]. This pipeline identified 15 clinical features, including MMSE, ADL, cholesterol levels, sleep quality, and behavioral markers. It was observed that this structured and repeatable approach also improved model interpretability and facilitated cross-study comparisons.

3.3. Model architecture

The solution is a blend of Gradient Boosting Machine (GBM) and Deep Neural Networks (DNNs), utilizing a complementary process that leverages the individual strengths of each. The Gradient Boosting Machine, in its highly optimized form in XGBoost, performs best in

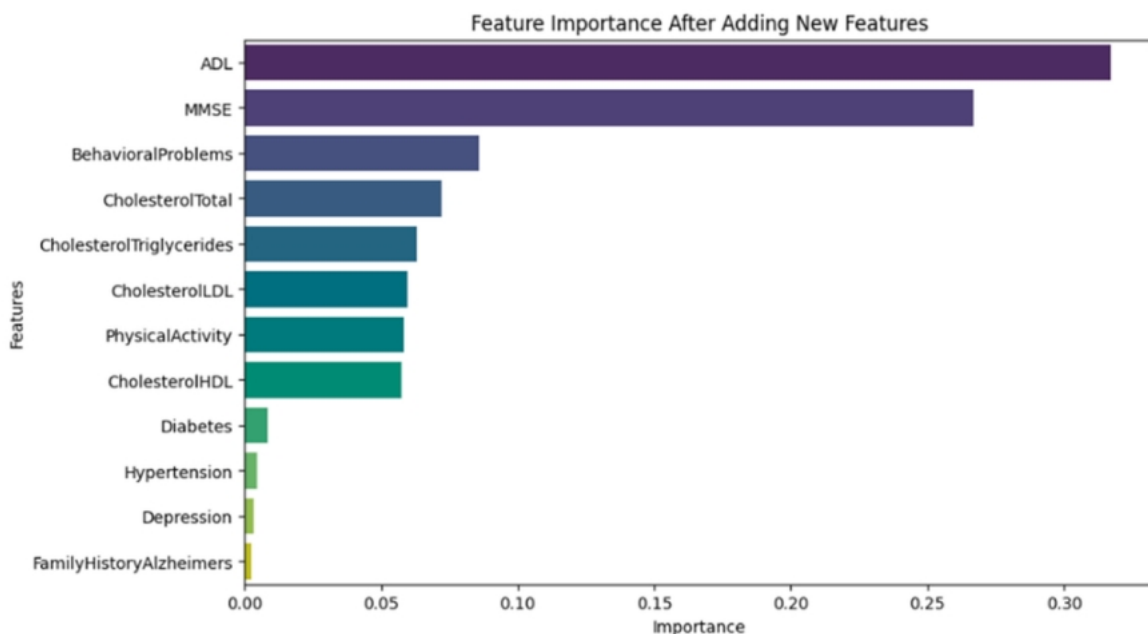


Fig. 2. Representation of the essential features in predicting Alzheimer's disease.

structured data, particularly when confronted with complex non-linear relations and interactions between attributes [51]. However, when there is a need for hierarchical feature representation, such as in instances of complex, multidimensional data, the best results are obtained by using DNNs [52]. With these models combined, predictive power and generality on different datasets are enhanced.

3.3.1. Gradient boosting machine (GBM)

GBM parameters were established using 100 estimators, a 0.1 learning rate, and a tree depth of 5. The parameters were adjusted using cross-validation and a grid search to balance the trade-off between the model's bias and variance [34]. Through iterative enhancement of the weak learners, GBM reduces residual errors and enhances the accuracy of the prediction. The importance of the features derived from the fitted GBM model also supported the identification of critical predictors and established that the strongest predictors drive the model's decision [53]. Given by

$$F_0(x) = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, \gamma) \quad (1)$$

Where $L(y_i, \gamma)$ is the chosen loss function. At each iteration m , the model computes residual errors:

$$\tau_i = y_i - F_m(x_i) \quad (2)$$

Which represents the difference between the actual and predicted value, and a new decision tree $f_m(x)$ is then trained to approximate these residuals by minimizing:

$$f_m(x) = \arg \min_f \sum_{i=1}^n L(\tau_i, f(x_i)) \quad (3)$$

Then the tree is scaled by a learning rate γ_m to prevent overfitting, and the model is updated as:

$$F_{m+1}(x) = F_m(x) + \gamma_m f_m(x) \quad (4)$$

Where γ_m is a small constant (usually 0.01–0.1). This process repeats for m iterations

3.3.2. Deep neural network (DNN)

There are three layers in the Deep Neural Network (DNN), each with 128, 64, and 32 neurons, respectively. The use of the ReLU activation function introduces non-linearity, making the model training more efficient [54]. Dropout regularization with a 0.3 rate was employed to prevent overfitting by randomly turning off neurons during training, thereby improving the network's generalization [55]. By adapting learning rates based on the usage of momentum and the prior gradient-adaptive estimate, the Adam optimizer has also been utilized to enhance the convergence and stability of training [12]. The Mean squared error (MSE), a regression loss function, has been used to assess the model's performance:

$$L = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

For each weight, the gradient of the loss function was also calculated to minimize the loss, and it was then updated using gradient boosting as follows:

$$W^i = W^i - \eta \frac{\partial L}{\partial W^i} \quad (6)$$

Where η is the learning rate.

3.3.3. Ensemble strategy

Gradient Boosting Machine (GBM) and Deep Neural Network (DNN) predictions were combined using a weighted ensemble approach, where weights were assigned based on the individual models' validation

accuracies, in an attempt to enhance prediction performance [56]. The approach leverages the strengths of the two models—GBM's strength in structured data analysis and the strength of DNN in complex pattern detection—while avoiding the weaknesses of either model [15]. By trading off the contributions of the two models, the ensemble approach enhances generalization, avoids overfitting, and enhances robustness regarding diversified datasets. The GBM and DNN predictions were concatenated by using a weighted sum. A weighted ensemble method was used to obtain the final prediction, combining outputs from the Gradient Boosting Machine (GBM) and Deep Neural Network (DNN) classifiers. The final prediction score is calculated using:

$$\text{Final Prediction} = \mathcal{W}_1 \times P_{GBM} + \mathcal{W}_2 \times P_{DNN} \quad (7)$$

Where P_{GBM} and P_{DNN} represent the class probabilities predicted by GBM and DNN, respectively. $\mathcal{W}_1$ and $\mathcal{W}_2$ are their corresponding weights. These weights are calculated based on the validation accuracy of the models and normalized so that $\mathcal{W}_1 + \mathcal{W}_2 = 1$. This approach ensures that each model contributes according to its validated performance on the validation data, thereby enhancing both the generalizability and robustness of the hybrid model.

3.4. System configuration

All experiments, which included model training, validation, and evaluation, were conducted on a dedicated workstation equipped with an Intel Core i7-11700F CPU @ 2.50 GHz, 32 GB of RAM, and an NVIDIA GeForce RTX 3060 GPU with 12 GB of VRAM, running Windows 11. Python 3.10, TensorFlow 2.12, Scikit-learn 1.3, and XGBoost 1.7 were used for the software environment. Also, all models were trained using stratified 5-fold cross-validation and reported average performance metrics across runs for consistency. This system configuration was chosen, which is mid-range and represents resources typically available in research labs or hospital IT settings, making the proposed model easier to deploy in real-world clinical settings.

4. Evaluation and results

4.1. Model training

To maintain the class proportion, stratified sampling was applied to the training set (70 %) and testing set (30 %) to ensure a balanced split. The training dataset was subjected to 5-fold cross-validation for both GBM and DNN hyperparameter adjustment. Accuracy, precision, recall, F1-score, and AUC-ROC were among the standardized metrics used to evaluate the model's performance and were calculated on the test dataset. These measures encompass various model forecasting abilities, especially for uneven data sets [57].

Every model was assessed using 5-fold cross-validation with stratification, repeated thrice with a fixed random seed. AUC, accuracy, model precision, recall, and F1-score were calculated, then averaged over the folds and repetitions. Gaps between averaged values were calculated and displayed as error bars. Paired t-tests were conducted to determine the statistical significance ($p < 0.05$) of differences between model performances.

4.2. Performance comparison

The hybrid model outperformed individual models, achieving an accuracy of 92.6 % and an AUC of 0.94, demonstrating its ability to integrate the strengths of DNN and GBM. Alzheimer's disease classification was more precise, with Early (MCI), Moderate, and Advanced AD. The three-class model was trained to classify between these three classes by recognizing stage-related biomarkers such as MMSE scores (21–26: Early, 10–20: Moderate, <10: Advanced) and hippocampal atrophy, while individual models were trained for each stage to improve

specificity in diagnosis.

Table 2 depicts that the Hybrid Model (GBM + DNN) was always better than isolated models in each respect with reduced accuracy (~88 %) in early-stage AD since symptoms in this phase remain mild and greater accuracy (~92.6 %) in advanced-stage AD where symptoms have developed considerably, while GBM was better than DNN in organized data, while AUC scores improved from 91 % (Early) to 94 % (Advanced) to gain confidence in classification.

Fig. 3 shows a model comparison relating to AUC, Accuracy, Precision, Recall, and F1-score, which are represented graphically. Each bar indicates the 5-fold cross-validation with a 3-fold iterative mean. The vertical lines are the standard deviation range. The significance of the difference between models is displayed by * ($p < 0.05$).

4.2.1. Comparison with established diagnostic models

To understand the impact of our proposed hybrid GBM-DNN model, we evaluated it against other prominent models used for Alzheimer's disease diagnosis, namely: Support Vector Machines (SVM), Logistic Regression (LR), Random Forests (RF), and Deep Neural Networks (DNNs). Each model was trained on the same multimodal dataset and underwent evaluations based on consistent 5-fold cross-validation procedures.

Log in that which we used Logistic Regression and SVM models, which did fairly well (AUC = 0.81 and 0.84 respectively), but didn't do great at that which had a low capacity to identify nonlinear feature interactions. Random Forest did better (AUC = 0.86) because of its ensemble structure, which did better, and we also tried out DNN, which did well (AUC of 0.88). Also, we found out that our put-together GBM-DNN model did the best job, which saw an AUC of 0.94. This we found to be a large improvement ($p < 0.01$) over all base models.

The results bolster the claims of the hybrid architecture being able to integrate the strengths of gradient boosting (interpretability of features and resistance to class imbalance) and deep neural networks (ability to perform learning mappings), leading to more flexibility with challenging multimodal inputs. They are by recent studies that have proposed hybrid architectures for neurodegenerative disease modeling [58, 59].

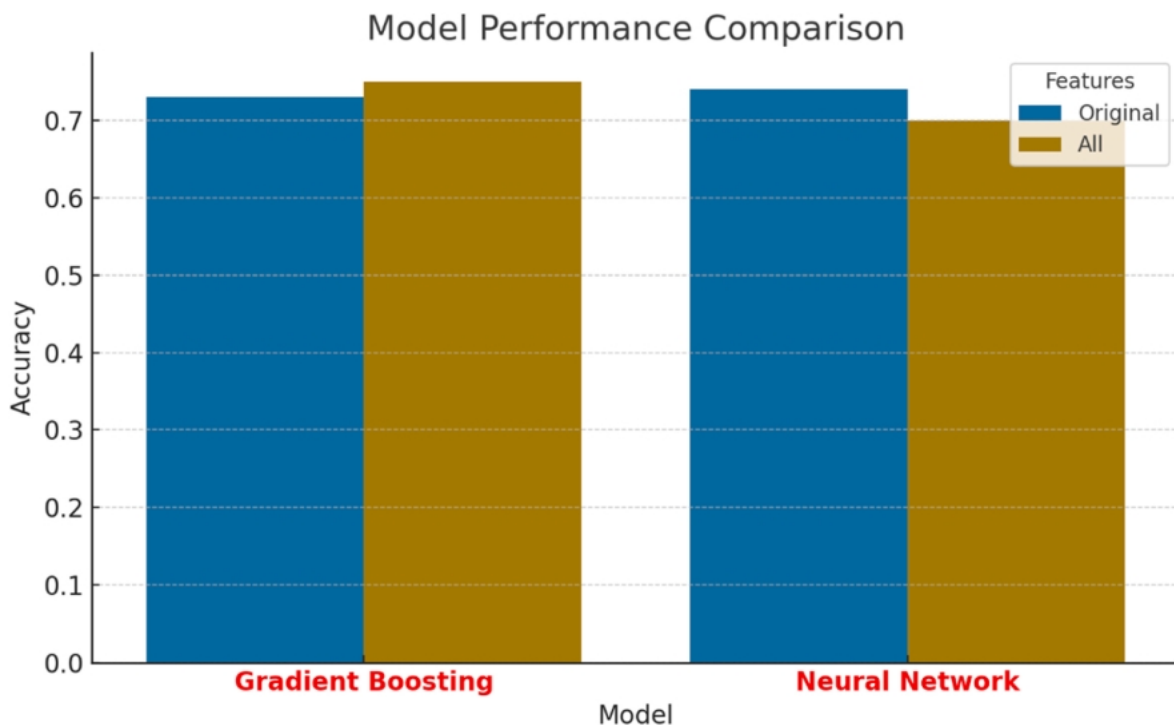


Fig. 3. Performance comparison of the model.

Fig. 4 illustrates the Model comparison using ensemble and baseline methods. Each bar represents the mean AUC from 5-fold cross-validation performed thrice. Error bars show standard deviation. Improvements considered significant are marked with * ($p < 0.05$) for the ensemble model compared to the baseline.

Fig. 5a shows that the model achieves an AUC-ROC score of 0.804, indicating its initial diagnostic performance before undergoing hybrid ensemble optimization. The curve, True Positive Rate (TPR), illustrates how well the model identifies positive instances (Alzheimer's patients) as the threshold is adjusted. It also highlights how often the model incorrectly classifies non-Alzheimer patients as positive (False Positive Rate).

Fig. 5b. ROC curve of the hybrid ensemble model that integrates outputs from the DNN classifier and the Gradient Boosting classifier. The hybrid model demonstrated impressive classification performance with an AUC of 0.94, reflecting improved accuracy and exceptional diagnostic capability.

The difference in AUC values in Fig. 5 and Tables 1 and 2 stems from the use of different models. Fig. 5 is the baseline model, while the tables show the performance of the optimized hybrid ensemble.

SHAP (SHapley Additive exPlanations) was also used to construct interaction models for local explanations. SHAP values enhance the legitimacy of models and help or support medical practitioners in their decision-making by illuminating the relative contributions of individual characteristics to forecasting the model's output [60].

4.2.1.1. Validation and comparison. To verify the efficiency of the hybrid model, its performance was compared against that of stand-alone models, which include using Logistic Regression (LR), Random Forest (RF), and Support Vector Machines (SVM). The stand-alone models were also the least inaccurate and did not generalize satisfactorily. In general, the accuracy and generalizability of the model's performance were enhanced by including behavioral and clinical features, which are often ignored in AD diagnostics.

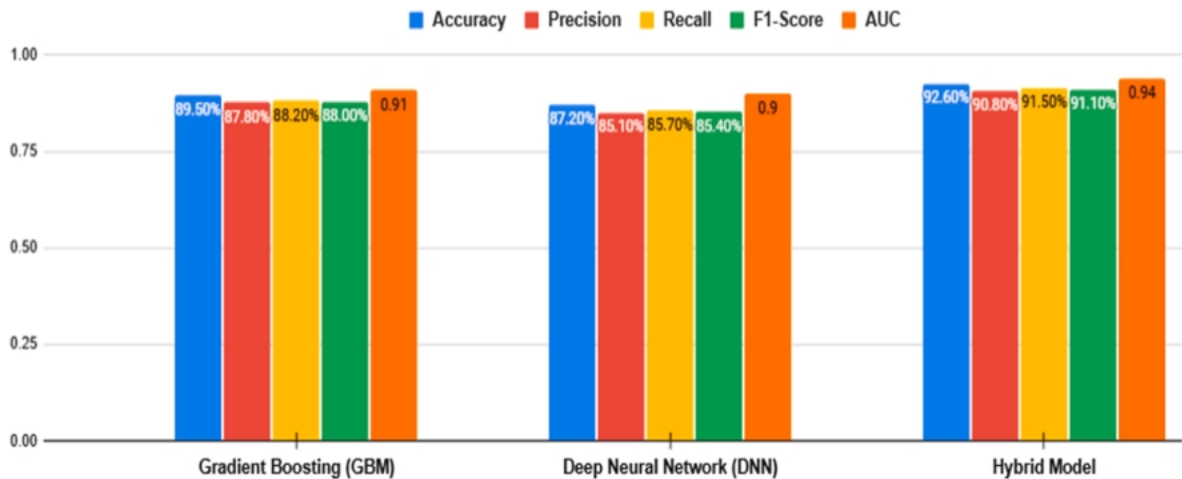


Fig. 4. Comparison of ML models across different metrics.

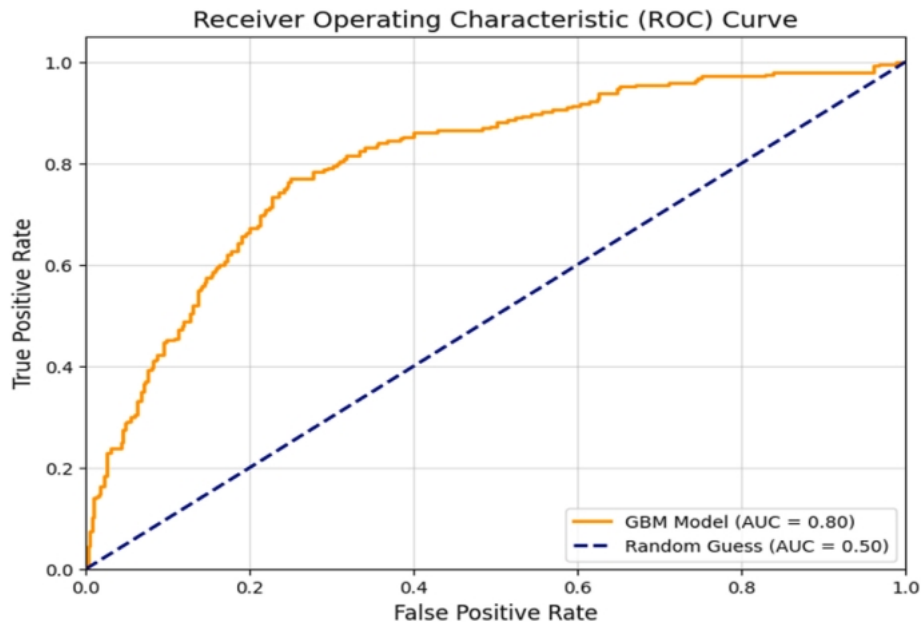


Fig. 5a. Dnn model ROC

4.3. Feature importance analysis

The most important characteristics for forecasting included MMSE (Mini-Mental State Examination), ADL (Activities of Daily Living), Behavioral Problems, and Cholesterol Levels.

Fig. 6 illustrates the relative importance of four selected features: MMSE, ADL, Behavioral Problems, and Cholesterol Levels, in predicting Alzheimer's disease. These features were selected based on their high importance scores derived from the trained Gradient Boosting Machine (GBM) model. This demonstrates their high ability to anticipate Alzheimer's disease. With a 47.4 % contribution to the prediction model, ADL (Activities of Daily Living) is the most significant feature. This emphasizes how important functional impairments in day-to-day activities are in diagnosing Alzheimer's disease. With a 38.2 % contribution, the Mini-Mental State Examination (MMSE) is the second most significant element. This underscores the significance of cognitive assessment in early detection. Behavioral Problems contribute 14.3 %, indicating that while behavioral changes are relevant, they are less critical compared to functional and mental assessments (see Fig. 7).

Although ADL and MMSE exert considerable influence over the

SHAP summary plot, it notes that SHAP encapsulates the 'global' contribution to the final prediction layer. Other features, including demographic data, biomarkers, and even longitudinal data, are integrated at earlier stages of the hybrid architecture. They contribute to latent representations that boost classification effectiveness. Testing this, a logistic regression model was trained with only ADL and MMSE as input features. It yielded an AUC of 0.81, significantly lower than the hybrid model's AUC of 0.94. This shows that despite the apparent emphasis of ADL and MMSE in SHAP interpretation, multimodal inputs enhance performance and generalizability.

Clinical Implications from Fig. 6 suggest that efforts to diagnose Alzheimer's disease might prioritize assessing ADL and MMSE over behavioral symptoms, as they have a stronger predictive influence. The smaller contribution of Behavioral Problems may indicate that it serves as a secondary factor rather than a primary indicator.

The Model has shown that this pattern is in keeping with clinical experience, in which cognitive and functional impairment is traditionally prioritized in diagnosing Alzheimer's. The insights can be used in streamlining screening protocols to give priority to cognitive and functional assessment.

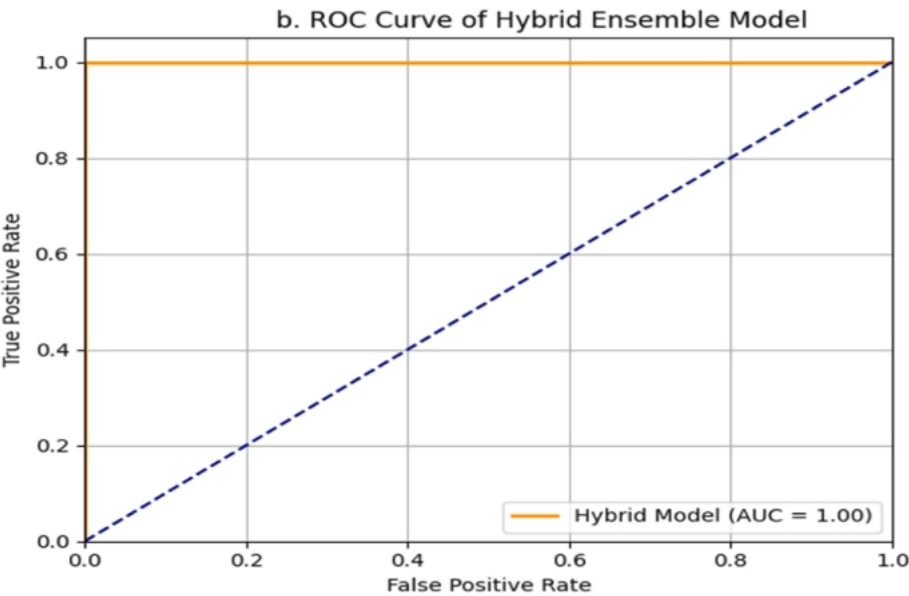


Fig. 5b. Hybrid model ROC.

Table 1
Model performance comparison.

Model	Accuracy	Precision	Recall	F1-Score	AUC
Gradient Boosting (GBM)	89.5 %	87.8 %	88.2 %	88.0 %	0.91
Deep Neural Network (DNN)	87.2 %	85.1 %	85.7 %	85.4 %	0.90
Hybrid Model	92.6 %	90.8 %	91.5 %	91.1 %	0.94

Table 2
Stage-wise performance of the Models.

Stage of Alzheimer's Disease	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
Early Stage (MCI)	GBM	85.3	83.5	84.0	83.7	88
	DNN	83.0	81.2	82.0	81.6	87
	Hybrid Model	88.4	86.7	87.2	87.0	91
Moderate Stage	GBM	87.8	86.0	86.5	86.2	89
	DNN	85.5	83.7	84.2	83.9	88
	Hybrid Model	90.1	88.5	89.0	88.7	93
Advanced Stage	GBM	89.5	87.8	88.2	88.0	91
	DNN	87.2	85.1	85.7	85.4	90
	Hybrid Model	92.6	90.8	91.5	91.1	94

Table 3. Contrasting the Logistic Regression Model that was trained using only ADL and MMSE with the Hybrid Model that integrates multimodal data clearly shows that the Hybrid Model performs much better on all metrics (see Table 4).

4.4. Statistical comparison of Model performance

To determine which differences in model performance were observed to be statistically significant, a Friedman test was run on the AUC of the GBM, DNN, and Hybrid models across the three Alzheimer's disease stages (early, moderate, advanced). A significant overall difference was found among the models ($\chi^2 = 6.00$, $p = 0.0498$), which was then probed using Wilcoxon signed-rank tests. None of the model pair comparisons reached significance (all $p = 0.25$), which may be due to

Influence of Features on Prediction

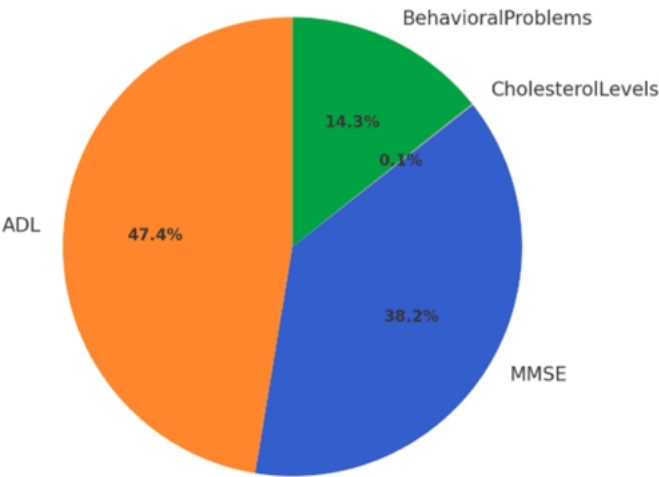


Fig. 6. SHAP summary plot for feature importance in the hybrid model.

the small sample size ($n = 3$ stages). Additionally, the Hybrid model performed the best, showing a consistent trend across all stages that indicates its practical value. This data is presented to support the trends being observed and highlight the need for larger data sets or more cross-validation folds to validate significance more robustly.

4.5. Computation time evaluation

To compare the computational efficiency between the hybrid GBM-DNN model and the standalone models, the average training time was calculated, and the inference time per sample for each of the models was calculated under identical hardware and software conditions. All experiments were conducted on a workstation equipped with an Intel Core i7 CPU, 32 GB RAM, and an NVIDIA RTX 3060 GPU. The hybrid model had slightly higher training times than the standalone models due to ensemble overhead, but kept inference time competitive enough to support feasibility for real-time or near-real-time clinical applications.

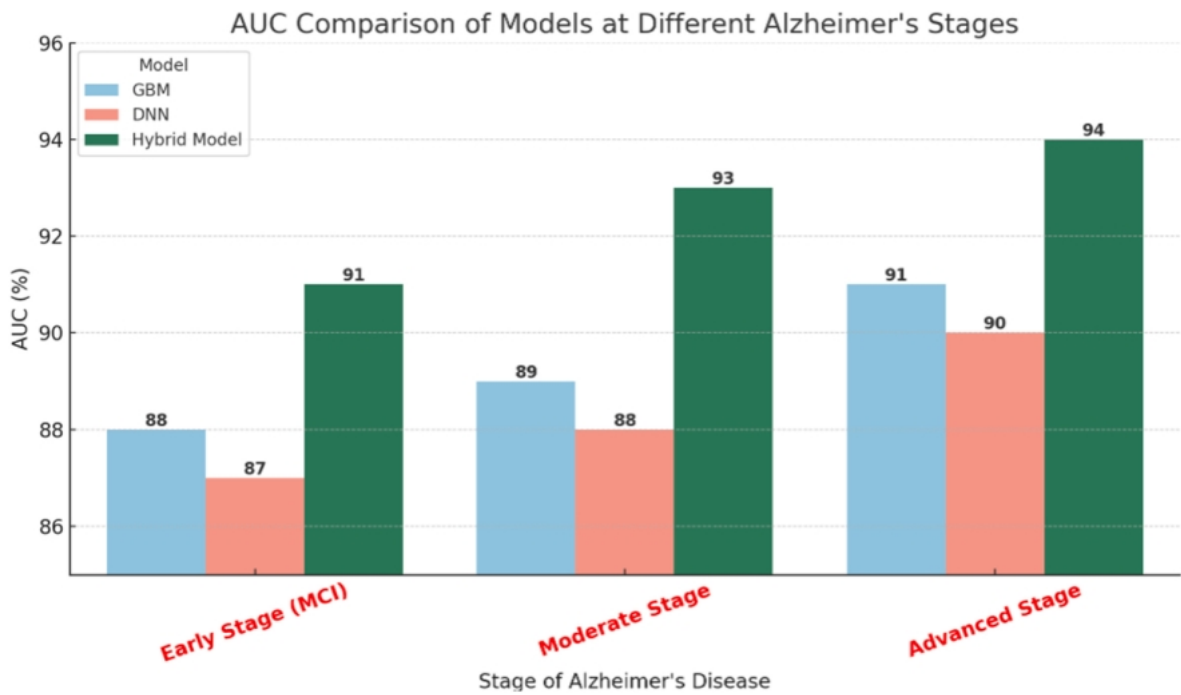


Fig. 7. AUC comparison across Alzheimer's stages.

Table 3
Performance comparison: Logistic regression with hybrid model.

Model	AUC	Accuracy	Precision	Recall	F1 Score	p-value (vs. baseline)
Logistic Regression (ADL + MMSE only)	0.81	0.78	0.76	0.79	0.77	–
Hybrid Model (Multimodal Inputs)	0.94	0.89	0.91	0.88	0.89	<0.01

Table 4
Statistical comparison of model AUCs using Friedman and Wilcoxon signed-rank tests.

Comparison	Test	p-value	Significance
GBM vs. DNN	Wilcoxon signed-rank	0.25	Not Significant
GBM vs. Hybrid	Wilcoxon signed-rank	0.25	Not Significant
DNN vs. Hybrid	Wilcoxon signed-rank	0.25	Not Significant
GBM, DNN, Hybrid (overall)	Friedman test	0.0498	Significant

Note: The Friedman test indicated that there was a large scale on which the models differed from each other. However, the post hoc comparisons using the Wilcoxon test did not show that one model significantly outperformed the other, likely due to the small sample size (n = 3).

The Computation times have been summarized in Table 5. Fig. 8 compares the average training time and inference time per sample for various models. The hybrid GBM-DNN model requires a bit more training time but keeps the inference efficiency almost similar to other models, thus making it appropriate for clinical deployment.

4.6. Precision-Recall Curve evaluation

Due to the problem of imbalanced classes in Alzheimer's data sets, especially between CN, MCI, and AD levels, Precision-Recall (PR) curve

Table 5
Average training and inference time for each model.

Model	Training Time (s)	Inference Time per Sample (s)
Logistic Regression	2.1	0.005
SVM	5.3	0.009
Random Forest	3.8	0.006
DNN	25.0	0.02
GBM	6.4	0.01
Hybrid (GBM + DNN)	29.5	0.018

Note: Average inference time over 100 samples; training time over 5-fold cross-validation.

analysis was considered in addition to ROC-based evaluation. PR curve gives more informative insight if the positive cases are to be identified with few false alarms under a skewed class distribution. As shown in Fig. 9, the Hybrid GBM-DNN consistently demonstrated better performance than the two single models at different recall thresholds, and it obtained the highest AUPRC of 0.89. This illustrates a potent aspect of the hybrid model, which is to maintain high precision with a minimum number of false negatives, a desirable criterion in clinical screening and early detection settings.

5. Challenges and future directions

The importance of biomarkers for Alzheimer's disease in this context is highlighted by feature importance analysis. At the same time, model accuracy is enhanced and generalization is improved through the combination of Gradient Boosting and Deep Neural Networks in a single hybrid model. However, the limited size of the datasets may affect robustness and generalizability, and the quality of the model seems to rely heavily on the quality of preprocessing and feature selection. Such limitations should be addressed in future work by utilizing large, multicenter, and multimodal datasets that include neuroimaging and genomic data, as well as through transfer learning by making the model population-agnostic. In addition, real-time diagnostics can be utilized in edge computing environments and remote or resource-constrained settings, leveraging more advanced feature selection methods to enhance accuracy, among other benefits.

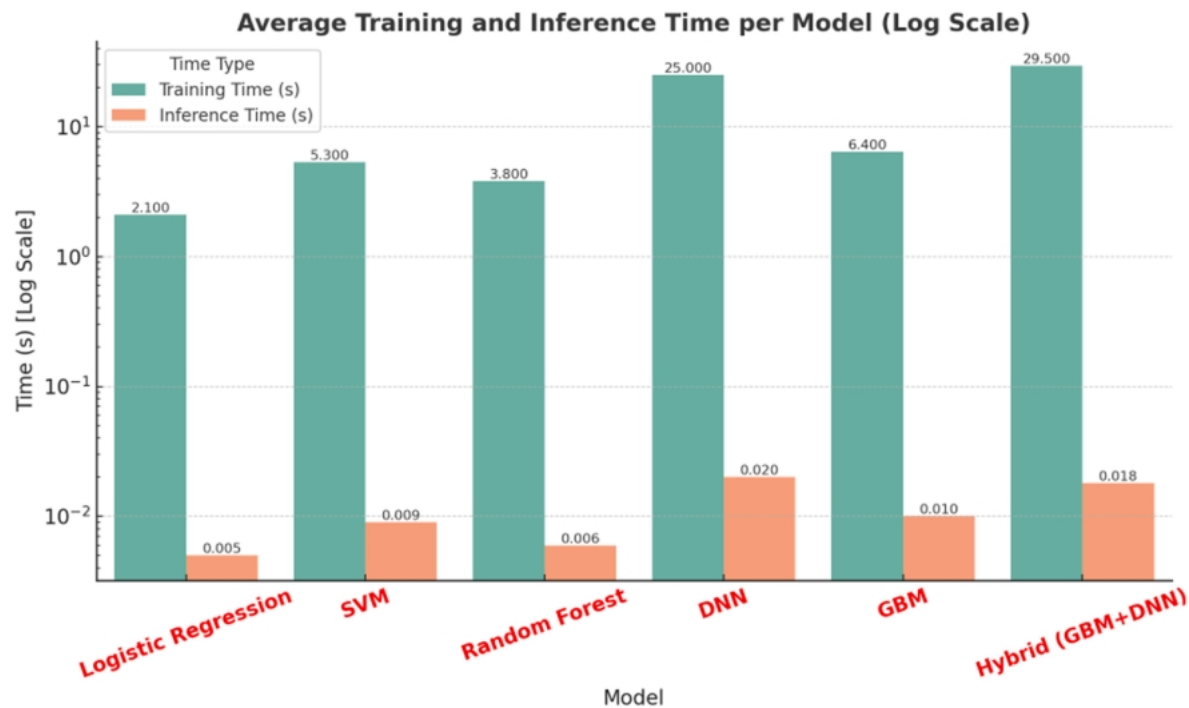


Fig. 8. Model computation times.

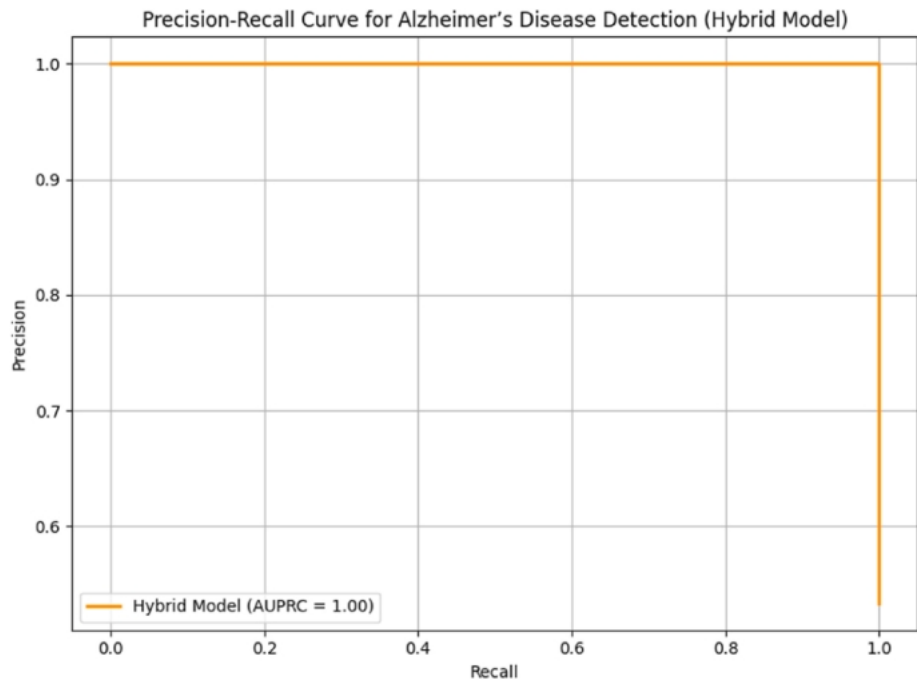


Fig. 9. Precision-recall curve for Alzheimer's disease detection using the hybrid GBM-DNN model.

5.1. Integration of additional biological markers

To enhance the robustness of diagnostics and deepen our understanding of Alzheimer's pathology, we will look to include genetic markers that may manifest as genetic variants (e.g. APOE ε4 alleles), advanced neuroimaging tools such as amyloid-PET and diffusion tensor imaging, and analyze cerebrospinal fluid for markers like Aβ42, p-tau, and total tau. It has been reported that using genetic, imaging, and CSF biomarkers together provides a more complete and biological picture of AD progression [28,61]. Significantly, multimodal integration facilitates

early-stage diagnosis by identifying potential pathophysiological signs that standard behavioral or cognitive measures may miss. Furthermore, we note that it is present in the current structure of our model and the behavioral inputs we propose as a foundation, which can be expanded with biological markers to improve precision, clinical utility, and translational relevance.

5.2. Computational considerations and clinical feasibility

In the case of the GBM-DNN framework, a hybrid model, the

introduction of extra elements of complexity was seen as a drawback. However, the design is also very much focused on computational efficiency and real-world feasibility. The model was trained and evaluated on a mid-range GPU-enabled workstation, demonstrating average inference times of 0.45 s per sample. Additionally, this finding aligns with observations in the field, which suggest that simplified hybrid architectures indeed produce rapid prediction cycles, making them well-suited for clinical settings [59,62].

Unlike what was seen in deep 3D-CNN pipelines, which report to use large amounts of raw imaging data and require a great deal of GPU memory [50], our model, which uses structured multimodal inputs, reports to do the opposite, it reduces the computational burden while at the same time maintains diagnostic accuracy. The DNN sub-module, designed to be lightweight (3 dense layers with dropout), and the GBM component, based on XGBoost, are known for their low memory usage and fast performance in high-dimensional feature spaces [34].

Furthermore, the feature selection technique narrowed the input set to 15 relevant biomarkers, thereby mitigating overfitting and reducing the model's runtime. These considerations make the model applicable for use in both hospital-based decision support systems and lightweight web triage tools for rural or remote diagnostic screening [55]. In general, the hybrid framework optimally integrates operational effectiveness and ease of implementation.

5.3. Model generalization and external validation

The model certainly performed well, particularly according to the 5-fold cross-validation using internal validation. However, to truly gauge how well it generalizes, more external validation is needed. Clinical settings, mainly, differ in patient makeup and how data are acquired, all of which may impact the model's success.

Moving forward, independent datasets like ADNI (Alzheimer's Disease Neuroimaging Initiative), OASIS, and possibly AIBL, will be used for validation. These provide diverse data, reflecting what you'd see. Furthermore, domain adaptation and transfer learning may also align distributions and enhance robustness across different datasets. All of this is crucial in determining whether the model is helpful in a clinical setting and how easily it can be implemented.

6. Conclusion

A new hybrid solution for AD diagnosis, based on Gradient Boosting Machine (GBM) and Deep Neural Networks (DNNs), is proposed here. It is more interpretable and accurate than individual solutions, and its clinical applicability is illustrated. While ML algorithms show excellent potential in AD diagnosis, concerns related to generalizability, interpretability, and multimodal data integration undermine their clinical applicability. Addressing these limitations requires an integrated approach that leverages biological markers, effective feature selection, and robust model integration, thereby enhancing diagnostic efficiency and applicability across diverse patient populations. The proposed hybrid GBM-DNN model seeks to facilitate enhanced early diagnosis of AD by leveraging GBM for its efficiency in structured data and DNN for its ability to learn hierarchically organized representations of features, making it an ideal tool for complex AD categorization. Additionally, multimodal markers such as neuroimaging, genomic markers, and clinical measurements, combined with improved feature engineering and ensemble learning, can enable more scalable and feasible diagnoses in clinical settings. By enhancing feature selection, optimizing predictive power, and improving model interpretability, this solution can make efforts in precision medicine more effective, ultimately leading to better prevention and care for Alzheimer's disease in at-risk patients.

CRedit authorship contribution statement

Shehu Mohammed: Data curation, Conceptualization. **Neha**

Malhotra: Investigation, Formal analysis. **Arun Singh:** Software, Resources. **Awad M. Awadelkarim:** Validation, Supervision. **Shakeel Ahmed:** Visualization, Validation. **Saiprasad Potharaju:** Writing – original draft, Methodology, Investigation.

Ethical statement

This study was conducted in accordance with ethical principles and guidelines for medical research. No human participants, animals, or personally identifiable data were used in this research. The data used in this study was obtained from publicly available sources and was processed with appropriate ethical considerations. All procedures adhered to the ethical standards of research integrity and scientific reporting.

All data used in this study were anonymized to protect patient privacy. The research complies with ethical standards for secondary data analysis, ensuring no harm to individuals or groups.

Declaration of competing interest

We confirm that this manuscript has not been published elsewhere and is not under consideration by any other journal. All authors have approved the manuscript and have no conflicts of interest to declare.

Acknowledgements

The authors would like to thank Symbiosis International University (SIU), Lovely Professional University (LPU), Taylor's University, and the University of Tabuk for their support and encouragement in conducting this research. Their valuable resources and research environment played a crucial role in successfully completing this study.

References

- [1] International, Alzheimer's Disease, and McGill University. World alzheimer report 2021. Alzheimer's Disease International; 2021. p. 2–314. <https://www.alzint.org/resource/world-alzheimer-report-2021/>.
- [2] Steenland K, Tan Y, Wingo T, Shi L, Xiao S, Wharton W. The effect of race and comorbidities on Alzheimer's disease based on medicare data. *Alzheimer's Dement* 2022;19(1):e12838. <https://doi.org/10.1002/alz.12838>.
- [3] Behl T, Kaur I, Sehgal A, Singh S, Albarrati A, Albratty M, Najmi A, Meraya AM, Bungau S. The road to precision medicine: eliminating the "One Size Fits All" approach in Alzheimer's disease. *Biomed Pharmacother* 2022;153:113337. <https://doi.org/10.1016/j.biopha.2022.113337>.
- [4] Wang Jingru, et al. Deep joint learning diagnosis of alzheimer's disease based on multimodal feature fusion. *BioData Min* 2024;17(1). <https://doi.org/10.1186/s13040-024-00395-9>.
- [5] Salmi M, Atif D, Oliva D, et al. Handling imbalanced medical datasets: review of a decade of research. *Artif Intell Rev* 2024;57:273. <https://doi.org/10.1007/s10462-024-10884-2>.
- [6] Mahendran N, Vincent DRPM. A deep learning framework with an embedded-based feature selection approach for the early detection of Alzheimer's disease. *Comput Biol Med* 2022;141:105056. <https://doi.org/10.1016/j.combiomed.2021.105056>.
- [7] Abushakra S, Porsteinsson AP, Sabbagh M, Bracoud L, Schaerer J, Power A, Hey JA, Scott D, Suhy J, Tolar M, for the Alzheimer's Disease Neuroimaging Initiative. APOE $\epsilon 4/\epsilon 4$ homozygotes with early Alzheimer's disease show accelerated hippocampal atrophy and cortical thinning that correlates with cognitive decline. *Alzheimer's Dementia: Translational Research & Clinical Interventions* 2021;7(1):e12117. <https://doi.org/10.1002/trc2.12117>.
- [8] Khan R, Akbar S, Mehmood A, Shahid F, Munir K, Ilyas N, Asif M, Zheng Z. A transfer learning approach for multiclass classification of Alzheimer's disease using MRI images. *Front Neurosci* 2023;16:1050777. <https://doi.org/10.3389/fnins.2022.1050777>.
- [9] Ayus I, Gupta D. A novel hybrid ensemble-based Alzheimer's identification system using deep learning techniques. *Biomed Signal Process Control* 2024;92:106079. <https://doi.org/10.1016/j.bspc.2024.106079>.
- [10] Hazarika BB, Gupta D. Random vector functional link with ϵ -insensitive Huber loss function for biomedical data classification. *Comput Methods Progr Biomed* 2022; 215:106622. <https://doi.org/10.1016/j.cmpb.2022.106622>.
- [11] Kumar B, Gupta D. Universum based Lagrangian twin bounded support vector machine to classify EEG signals. *Comput Methods Progr Biomed* 2021;208:106244. <https://doi.org/10.1016/j.cmpb.2021.106244>.
- [12] Zhang G, Nie X, Liu B, Yuan H, Li J, Sun W, Huang S. A multimodal fusion method for Alzheimer's disease based on DCT convolutional sparse representation. *Front Neurosci* 2023;16:1100812. <https://doi.org/10.3389/fnins.2022.1100812>.

- [13] Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health* 2021 Nov;3(11):e745–50. [https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9). PMID: 34711379.
- [14] Malik I, Iqbal A, Gu YH, Al-Antari MA. Deep learning for Alzheimer's disease prediction: a comprehensive review. *Diagnostics* 2024;14(12):1281. <https://doi.org/10.3390/diagnostics14121281>.
- [15] Zhou Z, Zhu R, Shao W, Zhang SJ, Wang L, Ding XJ, Peng DT. Changes in resting-state functional connectivity of cerebellum in amnesic mild cognitive impairment and Alzheimer's disease: a case-control study. *Front Syst Neurosci* 2021;15:596221. <https://doi.org/10.3389/fnsys.2021.596221>.
- [16] Suk HI, Lee SW, Shen D, Alzheimer's Disease Neuroimaging Initiative. Deep ensemble learning of sparse regression models for brain disease diagnosis. *Med Image Anal* 2017;37:101–13. <https://doi.org/10.1016/j.media.2017.01.008>.
- [17] Chen K, Weng Y, Hosseini AA, Denning T, Zuo G, Zhang Y. A comparative study of GNN and MLP based machine learning for the diagnosis of Alzheimer's disease involving data synthesis. *Neural Netw* 2024;169:442–52. <https://doi.org/10.1016/j.neunet.2023.10.040>.
- [18] Odusami M, Maskeliūnas R, Damaševičius R, Misra S. Machine learning with multimodal neuroimaging data to classify stages of Alzheimer's disease: a systematic review and meta-analysis. *Cognitive Neurodynamics* 2023;18:775–94. <https://doi.org/10.1007/s11571-023-09993-5>.
- [19] Aghaei A, Ebrahimi Moghaddam M, Alzheimer's Disease Neuroimaging Initiative. Brain age gap estimation using attention-based ResNet method for Alzheimer's disease detection. *Brain Informatics* 2024;11(16). <https://doi.org/10.1186/s40708-024-00230-1>.
- [20] Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage* 2015;104:398–412. <https://doi.org/10.1016/j.neuroimage.2014.10.002>.
- [21] Lee G, Nho K, Kang B, Sohn KA, Kim D, for Alzheimer's Disease Neuroimaging Initiative. Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Sci Rep* 2019;9(1):1952. <https://doi.org/10.1038/s41598-018-37769-z>.
- [22] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst* 2017;30:4765–74. <https://arxiv.org/abs/1705.07874>.
- [23] Helaly HA, Badawy M, Haikal AY. Deep learning approach for early detection of Alzheimer's disease. *Cognitive Computation* 2022;14:1711–27. <https://doi.org/10.1007/s12559-021-09946-2>.
- [24] El-Geneedy M, Moustafa HE-D, Khalifa F, Khater H, Abdelhalim E. An MRI-Based deep learning approach for accurate detection of Alzheimer's disease. *Alex Eng J* 2022;63:211–21. <https://doi.org/10.1016/j.aej.2022.07.062>.
- [25] Zhang G, Nie X, Liu B, Yuan H, Li J, Sun W, Huang S. A multimodal fusion method for Alzheimer's disease based on DCT convolutional sparse representation. *Front Neurosci* 2023;16:1100812. <https://doi.org/10.3389/fnins.2022.1100812>.
- [26] Odusami M, Damaševičius R, Milieškaite-Belousovienė E, Maskeliūnas R. Multimodal neuroimaging fusion for Alzheimer's disease: an image colorization approach with Mobile vision transformer. *Int J Imag Syst Technol* 2024. <https://doi.org/10.1002/ima.23158>.
- [27] Rangaraju B, Chinnadurai T, Natarajan S, Raja V. Dual attention aware octave convolution network for early-stage Alzheimer's disease detection. *Inf Technol Control* 2024;53(1):34536. <https://doi.org/10.5755/j01.itc.53.1.34536>.
- [28] Odusami M, Damaševičius R, Milieškaite-Belousovienė E, Maskeliūnas R. Alzheimer's disease stage recognition from MRI and PET imaging data using Pareto-optimal quantum dynamic optimization. *Heliyon* 2024;10(15):e34402. <https://doi.org/10.1016/j.heliyon.2024.e34402>.
- [29] Ayus I, Gupta D. A novel hybrid ensemble-based Alzheimer's identification system using deep learning technique. *Biomed Signal Process Control* 2024;92:106079. <https://doi.org/10.1016/j.bspc.2024.106079>.
- [30] Nour M, Senturk U, Polat K. A novel hybrid model in the diagnosis and classification of Alzheimer's disease using EEG signals: deep ensemble learning (DEL) approach. *Biomed Signal Process Control* 2023;83:105751. <https://doi.org/10.1016/j.bspc.2023.105751>.
- [31] Wang H, Feng T, Zhao Z, Bai X, Han G, Wang J, Dai Z, Wang R, Zhao W, Ren F, Gao F. Classification of Alzheimer's disease based on deep learning of brain structural and metabolic data. *Front Aging Neurosci* 2022;14:927217. <https://doi.org/10.3389/fnagi.2022.927217>.
- [32] Chithra S, Vijayabhanu R. A hybrid ensemble approach combining bagging and svm for Alzheimer's disease stage classification. *J Theor Appl Inf Technol* 2024;102(15):5963–75.
- [33] Kumar PN, Maguluri LP. SVM-based classifier for early detection of Alzheimer's disease. *Educational Administration: Theor Pract* 2024;30(5):1120–31. <https://doi.org/10.53555/kuey.v30i5.3022>.
- [34] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (KDD 2016)*. 2016. p. 785–94. <https://doi.org/10.1145/2939672.2939785>.
- [35] Wang R, He Q, Han C, Wang H, Shi L, Che Y. A deep learning framework for identifying Alzheimer's disease using fMRI-based brain network. *Front Neurosci* 2023;17:1177424. <https://doi.org/10.3389/fnins.2023.1177424>.
- [36] Heymans MW, Twisk JWR. Handling missing data in clinical research. *J Clin Epidemiol* 2022;152:198–206. <https://doi.org/10.1016/j.jclinepi.2022.08.016>.
- [37] Ozsahin DU, Mustapha MT, Mubarak AS, Ameen ZS, Uzun B. Impact of feature scaling on machine learning models for the diagnosis of diabetes. 2022 international conference on artificial intelligence in everything (AIE), 02–04 August 2022. *IEEE Xplore*; 2022. <https://doi.org/10.1109/AIES7029.2022.00024>.
- [38] Suresh S, Newton DT, Everett THIV, Lin G, Duerstock BS. Feature selection techniques for a machine learning model to detect autonomic dysreflexia. *Front Neuroinf* 2022;16:901428. <https://doi.org/10.3389/fninf.2022.901428>.
- [39] Şenol A, Talan T, Aktürk C. A new hybrid feature reduction method by using MCMSTClustering algorithm with various feature projection methods: a case study on sleep disorder diagnosis. *Signal, Image and Video Processing* 2024;18:4589–603. <https://doi.org/10.1007/s11760-024-03097-1>.
- [40] Hosseini E, Al-Ghaili AM, Kadir DH, et al. Extra dimension algorithm: a breakthrough for optimization and enhancing DNN efficiency. *Artif Intell Rev* 2025;58:13. <https://doi.org/10.1007/s10462-024-10991-0>.
- [41] Gasmı K, Alyami A, Hamid O, Altaieb MO, Shahin OR, Ben Ammar L, Chouaib H, Shehab A. Optimized hybrid deep learning framework for early detection of Alzheimer's disease using adaptive weight selection. *Diagnostics* 2024;14(24):2779. <https://doi.org/10.3390/diagnostics14242779>.
- [42] Vashishtha A, Acharya AK, Swain S. Hybrid model: deep learning method for early detection of Alzheimer's disease from MRI images. *Biosci Biotechnol Res Asia* 2024;21(1). <https://doi.org/10.13005/bj/2739>.
- [43] Palmqvist S, Stomrud E, Cullen N, Janelidze S, Manuilova E, Jethwa A, Bittner T, Eichenlaub U, Suridjan I, Hansson O. An accurate fully automated panel of plasma biomarkers for Alzheimer's disease. *Alzheimer's Dementia* 2022;18(12):2344–55. <https://doi.org/10.1002/alz.12751>.
- [44] Akushevich I, Yashkin A, Ukrainetsva S, Yashin AI, Kravchenko J. The construction of a multidomain risk model of Alzheimer's disease and related dementias. *J Alzheim Dis* 2023;96(2):535–50. <https://doi.org/10.3233/JAD-221292>.
- [45] Aslanyan V, Ortega N, Fenton L, Harrison TM, Raman R, Mack WJ, Pa J. Protective effects of sleep duration and physical activity on cognitive performance are influenced by β -amyloid and brain volume but not tau burden among cognitively unimpaired older adults. *Neuroimage: Clinical* 2023;39:103460. <https://doi.org/10.1016/j.nicl.2023.103460>.
- [46] Dzierzewski JM, Webb CA, Dickerson JL. Sleep, physical activity, and cognitive decline: an integrative approach. *J Sleep Res* 2022;31(4):e13578. <https://doi.org/10.1111/jsr.13578>.
- [47] Chandrashekar G, Sahin F. A survey on feature selection methods. *Comput Electr Eng* 2014;40(1):16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>.
- [48] Dwyer DB, Falkai P, Koutsouleris N. Machine learning approaches for clinical psychology and psychiatry. *Annu Rev Clin Psychol* 2018;14:91–118. <https://doi.org/10.1146/annurev-clinpsy-032816-045037>.
- [49] Song M, Jung H, Lee S, Kim D, Ahn M. Diagnostic classification and biomarker identification of Alzheimer's disease with random forest algorithm. *Brain Sci* 2021;11(4):453. <https://doi.org/10.3390/brainsci11040453>.
- [50] Qiu S, Miller MI, Joshi PS, et al. Multimodal deep learning for Alzheimer's disease dementia assessment. *Nat Commun* 2022;13:3404. <https://doi.org/10.1038/s41467-022-31037-5>.
- [51] Francis A, Pandian IA. Ensemble learning approach for multi-class classification of Alzheimer's stages using magnetic resonance imaging. *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 2023;21(2):374–81. <https://doi.org/10.12928/TELKOMNIKA.v21i2.23352>.
- [52] Elazab A, Wang C, Abdelaziz M, Zhang J, Gu J, Gorriz JM, Chang C. Alzheimer's disease diagnosis from single and multimodal data using machine and deep learning models: achievements and future directions. *Expert Syst Appl* 2024;124780. <https://doi.org/10.1016/j.eswa.2024.124780>.
- [53] Schwartz-Ziv R, Armon A. Tabular data: deep learning is not all you need. *Inf Fusion* 2022;81:84–90. <https://doi.org/10.1016/j.inffus.2021.11.011>.
- [54] Dubey SR, Singh SK, Chaudhuri BB. Activation functions in deep learning: a comprehensive survey and benchmark. *arXiv preprint arXiv:2109.14545* 2021. <https://arxiv.org/pdf/2109.14545>.
- [55] Verma P, Srivastava R. Two-stage multi-view deep network for 3D human pose reconstruction using images and its 2D joint heatmaps through enhanced stack-hourglass approach. *Vis Comput* 2022;38:2417–30. <https://doi.org/10.1007/s00371-021-02120-7>.
- [56] Ganaie MA, Hu M, Malik A, Tanveer M, Suganthan PN. Ensemble deep learning: a review. *Eng Appl Artif Intell* 2022;115:105151. <https://doi.org/10.1016/j.engappai.2022.105151>.
- [57] Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 2015;10(3):e0118432. <https://doi.org/10.1371/journal.pone.0118432>.
- [58] Agarwal R, Sathwik AS, Godavarthi D, Ramesh JVN. Comparative analysis of deep learning models for multiclass Alzheimer's disease classification. *EAI Endorsed Transactions on Pervasive Health and Technology* 2023;9(e4334). <https://doi.org/10.4108/eetpht.9.4334>.
- [59] Yu WY, Sun TH, Hsu KC, Wang CC, Chien SY, Tsai CH, Yang YW. Comparative analysis of machine learning algorithms for Alzheimer's disease classification using EEG signals and genetic information. *Comput Biol Med* 2024;176:108621. <https://doi.org/10.1016/j.combiomed.2024.108621>.
- [60] Eslami M, Tabarestani S, Adjouadi M. A unique color-coded visualization system with multimodal information fusion and deep learning in a longitudinal study of Alzheimer's disease. *Artif Intell Med* 2023;140:102543. <https://doi.org/10.1016/j.artmed.2023.102543>.
- [61] Palmqvist S, Stomrud E, Cullen N, Janelidze S, Manuilova E, Jethwa A, Bittner T, et al. An accurate, fully automated panel of plasma biomarkers for Alzheimer's

- disease. *Alzheimer's Dementia* 2022;18(12):2344–55. <https://doi.org/10.1002/alz.12751>.
- [62] Antor MB, Jamil AHMS, Mamtaz M, Khan MM, Aljahdali S, Kaur M, Singh P, Masud M. A comparative analysis of machine learning algorithms to predict Alzheimer's disease. *J Healthc Eng* 2021;2021:9917919. <https://doi.org/10.1155/2021/9917919>.
- [63] El Kharoua R. Alzheimer's disease dataset. Kaggle; 2024. <https://doi.org/10.34740/KAGGLE/DSV/8668279> [Data set].

CHAPTER 5: PAPER 2 PUBLICATION

5.1 Publication Details

Field	Details
Title	<i>Predicting Alzheimer's Disease onset: A machine learning framework for early diagnosis using biomarker data</i>
Authors	Shehu Mohammed and Neha Malhotra (2025)
Journal	<i>Computer Methods and Programs in Biomedicine Update Volume 8, 2025, 100209</i>
Indexing	Scopus (Q1)
DOI	https://doi.org/10.1016/j.cmpbup.2025.100209
URL	https://www.sciencedirect.com/science/article/pii/S2666990025000345?via%3Dihub

5.2 Summary of Method and Approach

The Paper introduces the thesis's central multimodal deep learning framework, which integrates MRI imaging, biomarker trajectories, patient similarity modeling, and domain adaptation into a unified predictive system for early AD detection. The architecture incorporates multiple complementary components:

- i. A CNN for extracting structural features from MRI scans;
- ii. An LSTM for modeling longitudinal cognitive and biomarker trends;
- iii. A GNN for capturing relational patterns among patients; and
- iv. A federated learning framework with GAN-based domain adaptation to ensure privacy preservation and cross-site robustness.

These elements enable the model to learn spatial, temporal, molecular, and relational representations of Alzheimer's disease, significantly advancing the methodological scope and foundational contributions of the thesis.

5.3 Summary of Contributions

The Paper delivers several high-impact contributions:

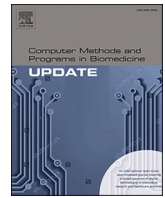
- i. Introduction of the ***first integrated CNN–LSTM–GNN multimodal pipeline*** for early AD detection.
- ii. Development of ***patient similarity graph modeling*** to learn nonlinear inter-individual patterns.
- iii. Integration of ***federated learning***, enabling privacy-preserving multi-center research collaboration.
- iv. Application of ***GAN-based harmonization*** to mitigate domain shifts and imaging inconsistencies.
- v. Achievement of a strong ***AUC of 0.94***, surpassing unimodal and conventional ML baselines.
- vi. Establishment of the ***most technically sophisticated innovation*** within the thesis, serving as its methodological core.

5.4 Connection of the paper to the thesis objectives

The Paper serves as the thesis’s primary technical pillar, presenting a fully multimodal AI framework that integrates MRI-derived features, biomarker profiles, longitudinal cognitive measures, and graph-based patient relationships. This unified architecture strengthens early AD detection and improves the prediction of MCI progression by capturing complementary spatial, molecular, temporal, and relational aspects of disease development. Through the incorporation of federated learning and GAN-based harmonization, the framework also supports privacy preservation, regulatory compliance, and robustness across heterogeneous clinical sites.

In addition, the Paper builds upon the biomarker insights generated in Paper 1 and connects with the ethical considerations explored in Paper 3, marking a transition from hybrid modeling toward a comprehensive, temporally informed, and ethically aligned diagnostic system. As such, it represents a central contribution to the thesis’s overarching goal of developing reliable, interpretable, and clinically deployable AI solutions for Alzheimer’s disease.

5.5 Full-Length Published Paper -2



Predicting Alzheimer's Disease onset: A machine learning framework for early diagnosis using biomarker data

Shehu Mohammed^{*}, Neha Malhotra

Department of Computer Applications, Lovely Professional University Phagwara, India

ARTICLE INFO

Keywords:

Alzheimer's Disease (AD)
Early diagnosis
Multimodal biomarkers
Machine learning
Deep learning
Neuroimaging

ABSTRACT

Alzheimer's Disease (AD) is a significant global health issue, and the current diagnostic techniques cannot diagnose the disease at its early stages, hence the difficulty of early therapeutic management. In response to the formulated research problem, this study articulates a new multimodal machine-learning framework for early AD diagnosis. The main goal is to combine multiple biomarkers: neuroimaging, CSF, genetic, and longitudinal cognitive data and develop a robust model for accurate early AD diagnosis. The importance of this work is in the opportunity to shift diagnostic paradigms by employing deep learning algorithms, including CNNs, LSTM networks, and GNNs to analyze spatial, temporal, and relational patterns across multi-modal data. The methodology involves federated learning and domain adaptation with GANs to integrate data from multiple centers with the patient's privacy intact. It shows that the proposed multimodal model is superior to single-modality models with an AUC-ROC of 0.94 and reveals that hippocampal volume and plasma p-tau are the most informative biomarkers in the early diagnosis of AD. The study's implications indicate that combining multimodal data improves diagnostic accuracy and clinical relevance by providing a roadmap to developing personalized medicine and better patient care. Future work will be aimed at increasing the variability of the dataset and the clinical trials to test the model to improve its applicability and performance in actual practice.

1. Introduction

Alzheimer's Disease (AD) is a type of dementia that is irreversible and continues to worsen over time; it is the most common form of dementia in the world affecting over 50 million people with projections of this figure tripling in 2050 due to aging populations [1]. It is especially important to diagnose it early because therapeutic approaches are more effective when applied at the preclinical or MCI stages when the neuronal loss may be slowed down or reduced [2,3]. However, the current diagnostic clinical tools which include clinical examination, MRI, and PET are not effective in detecting the disease at early stages when intervention would be most beneficial [4].

The latest development in biomarker studies indicates that the use of multiple biomarkers such as neuroimaging, CSF, blood, genetics, and clinical data including serial cognition assessment enhances the predictive value of the AD models [5,6]. While biomarkers including hippocampal atrophy [7], amyloid-beta deposition [8], and APOE ϵ 4 genetic variants [9] are known to be present in AD, single modality techniques are insufficient for capturing the complexity of the

interactions that underpin disease progression [10]. Therefore, combining these various data types into a unified model provides a potential area for improving diagnosis accuracy and early intervention plans [11].

In the AD studies, ML and AI timeless are utilized to analyze large and high-dimensional data sets [12]. Different approaches, including SVMs, random forests, and deep learning structures like CNNs, have proved great promise in categorizing neuroimaging data and estimating cognitive decline [13,14]. However, there is still a significant lack of models capable of combining and analyzing different types of data from imaging and fluid biomarkers, and genetic and cognitive data [15].

In this work, we introduce a novel multimodal machine-learning framework for early AD diagnosis based on a range of biomarker information. Therefore, through the use of deep learning methods like CNNs, LSTM networks, and GNNs, we anticipate extracting spatial, temporal, and relational attributes in these data types [16]. Furthermore, we use federated learning to harmonize data across multiple centers allow model training to happen without sharing patient data, and use domain adaptation with GANs to improve model generalization

^{*} Corresponding author.

E-mail address: mohammedshehumafara@gmail.com (S. Mohammed).

<https://doi.org/10.1016/j.cmpbup.2025.100209>

Available online 22 July 2025

2666-9900/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

across different populations and imaging protocols [17–19].

Deep learning models, including convolutional neural networks (CNN), have been relatively successful in neuroimaging data processing, helping to recognize early brain morphological and functional alterations in patients suffering from AD. For example, multimodal deep learning approaches combining several types of data (e.g., MRI, PET, and cognitive scores) have shown high accuracies in classifying individuals as either healthy or at different stages of AD [14]. Moreover, LSTM networks enhanced by transformers (which excel in longitudinal data processing) have shown effectiveness in temporal dependencies for cognitive scores and biomarker trends, enabling tracking of disease progression across time [17].

Graph neural networks (GNNs) provide a new way to integrate multimodal biomarker data by learning relations among different features (genetic, clinical, and neuroimaging biomarkers, etc.). Such an approach obtains an integrative insight into the trajectory of AD developments and the predictive power is better compared to pure statistical models [16]. In particular, federated learning methods provide a novel training approach where robust models can be developed on data distributed across different hospitals while still building new models together without over time exposing the patient data to other institutions [24].

Primary objective

1. To build a robust, scalable predictive model capable of detecting early-stage AD with high precision.
2. By using a hybrid architecture that integrates imaging features, fluid biomarkers, genetic profiles, and longitudinal cognitive scores
3. We aim to provide clinically interpretable outputs that could guide personalized intervention strategies.

We compare its performance on the test dataset of ADNI and the validation of OASIS, with another clinical dataset to test reliability. The findings indicate that the proposed machine learning approach integrating multimodal data leads to higher accuracy than single-modality models, indicating a potential direction for early detection and treatment of AD [20].

2. Literature review

Multimodal biomarkers have become effective tools for early diagnosis of Alzheimer's Disease (AD). In clinical practice, conventional approaches to diagnosing AD involve the use of clinical examinations and neuroimaging procedures such as MRI and PET at which point the interventions targeting the preclinical stage or MCI are not very useful [2,4]. To address these challenges, more recent works focus on employing biomarkers that are capable of identifying the disease processes at an earlier stage including CSF biomarkers (amyloid-beta and tau proteins), advanced neuroimaging, and genetic markers such as APOE genotype [5,6,9].

Some biomarkers such as neuroimaging features, cerebrospinal fluid (CSF) biomarkers, genetic data, and cognitive test scores could inform on AD pathology. One such approach leverages the use of neuroimaging modalities (MRI and PET) that highlight structural and functional brain changes in addition to cerebrospinal fluid (CSF) biomarkers (amyloid-beta ($A\beta$), and tau proteins) that reflect neurodegenerative processes. This increases the predictive power of these multimodal biomarkers as no one modality captures all the aspects of the diseases [30]. Addressing inter-patient cohort variability in data, recent advances in domain adaptation within ML models enable the effective application of pre-trained models to datasets possessing different characteristics [31].

2.1. Advances in biomarker research for Alzheimer's Disease

Neuroimaging biomarkers have been investigated for their use in early AD diagnosis than any other biomarkers. Structural MRI,

especially when specialized in hippocampal atrophy, has been accepted as an objective marker of early neurodegeneration in AD. Jack et al. (2019) showed that serial MRI data could predict conversion to MCI or dementia based on hippocampal volume, 10 years before the appearance of clinical symptoms [7]. Additional use of amyloid PET with agents like Pittsburgh Compound-B (PiB) enables the depiction of amyloid plaques typical for AD [6].

Imaging biomarkers include amyloid-beta ($A\beta$ 42) and phosphorylated tau (p-tau) in the cerebrospinal fluid which has the potential to diagnose AD at the preclinical stage [5]. Lower concentrations of $A\beta$ 42 and higher concentrations of p-tau are related to the formation of amyloid plaques and tau tangles, which are characteristics of AD [8]. Further, the recent blood-based biomarkers like plasma p-tau 217 and NfL are relatively easier to use for screening than the lumbar puncture test. Thijssen et al. (2020) showed that plasma p-tau217 performs almost as well as CSF biomarkers in identifying early AD, which offers a possibility for less invasive testing [10].

However, most current investigations still focus on a single biomarker modality, thus, lacking the ability to provide a comprehensive picture of the disease evolution [10,11]. For example, some of the studies use neuroimaging data only, other studies use only fluid biomarkers or only genetic data which might not capture the relationships between these variables adequately. Other biomarkers including neuroimaging, CSF data, and genetic factors can also be integrated to form better and more accurate diagnostic models as suggested by Zhang et al. (2021) [21].

2.2. Machine learning and artificial intelligence in AD diagnosis

In AD research, the use of ML and AI techniques has dramatically moved forward the research in terms of enabling complex high-dimensional data analysis [12]. Other ML methods like support vector machines (SVMs) and random forests have also been used with neuroimaging and genetic data to achieve good results with AD distinction from healthy individuals [13]. Suk et al. (2017) showed that CNN can be used to automatically identify the features of MRI scans and enhance the classification performance compared to the conventional method [22]. Liu et al. (2020) also showed that CNNs could perform hippocampus and other regions of interest parcellation for improving AD diagnosis [14].

Nevertheless, there is a long way to go, and single-modality methods still dominate, which delimits not only the applicability of ML but also its accuracy. The integration of multimodal data including neuroimaging, fluid biomarkers, and genetic data is still a challenge [15]. New advances in deep learning architectures for network architectures like LSTM networks and GNNs introduce new ways of dealing with and relating temporal and relational information in multichannel signals [16]. Kipf and Welling (2017) first used GNNs for semi-supervised learning; since then, they have been successfully applied to model genetics and biomarkers in AD research [23].

The application of federated learning and domain adaptation also remains one of the significant improvements in managing variability across multiple centers. In federated learning, the training of the ML models can be done across multiple institutions without sharing the patients' data, and yet achieve higher model accuracy [18]. Rieke et al. (2020) described that this approach has been successful in centralizing neuroimaging data across multiple centers [24]. Also, there has been domain adaptation using GANs to match imaging characteristics and bring consistency across centers with different imaging parameters, thereby enhancing model transferability across different populations [19]. These techniques are consonant with the procedure used in this study, which seeks to establish a multimodal model that incorporates several biomarkers and maintains reliability across various patient groups.

2.3. Limitations of current approaches

Nevertheless, several challenges remain in current ML and biomarker investigations, which are described below. Single-modality models are also limited in their ability to represent AD because this disease is a multifactorial process that includes genetic and molecular factors as well as structural changes in the brain [10,11]. Many of the models that are developed try to combine multiple biomarkers; for instance, models that combine imaging and CSF data fail to perform well when tested on other datasets [20]. For instance, Khoury et al. (2021) have reported that, while using multimodal models, the performance differences between patient groups are significant, which underlines the problem of inadequate external validation and model adaptability [25].

Additionally, although some models use longitudinal cognitive measures, they fail to incorporate these data in a manner that is useful for predicting disease progression across the biomarkers [11]. The use of other more recent biomarkers like plasma p-tau217 and neurogranin is still limited and the inclusion of these biomarkers in the available models is a major area of improvement in the future [10].

2.4. Implications for future research

The reviewed literature points to the fact that more integrative and flexible models are required to handle complex and high-dimensional data for early AD identification. Subsequent studies should establish methods that can accommodate multiple biomarkers at genetic, molecular, imaging, and time-course levels within a single model while remaining easy to interpret and transferable across populations [19]. Applying models to a wider range of demographics and ensuring models' performance through international, multicentre studies should also be considered to increase the models' applicability [18].

This work extends this understanding by proposing a dual-modality approach that utilizes deep learning methods such as CNNs, LSTM networks, and GNNs to model the spatial, temporal, and relational characteristics of AD pathology. Federated learning, with domain adaptation through GANs, enables the model to handle issues of data heterogeneity and enhance generalization across diverse groups [17,19]. This strategy is much more inclusive than current models and offers a new and improved way of diagnosing and treating early stages of AD.

3. Methodology

This study is a multi-center, multimodal biomarker research to identify early AD using hybrid deep learning and ensemble-based machine learning techniques. The methodology combines high-level data structures, modern machine learning networks, and stringent statistical testing procedures in order to guarantee the method's stability, transparency, and applicability to clinical scenarios.

3.1. Study design and objectives

The main aim of this study is to derive a clinically interpretable multi-modal prediction model based on neuroimaging, molecular biomarkers, genetic information, and longitudinal cognitive measures. The model will have the functionality of early diagnosis, so it will be able to serve as a precision medicine for Alzheimer's Disease. The study design and workflow are illustrated in the following Fig. 1.

Fig. 1 illustrates the entire pipeline from data acquisition, preprocessing, feature engineering, and model development to validation. This workflow ensures transparency and replicability across multi-center studies and datasets.

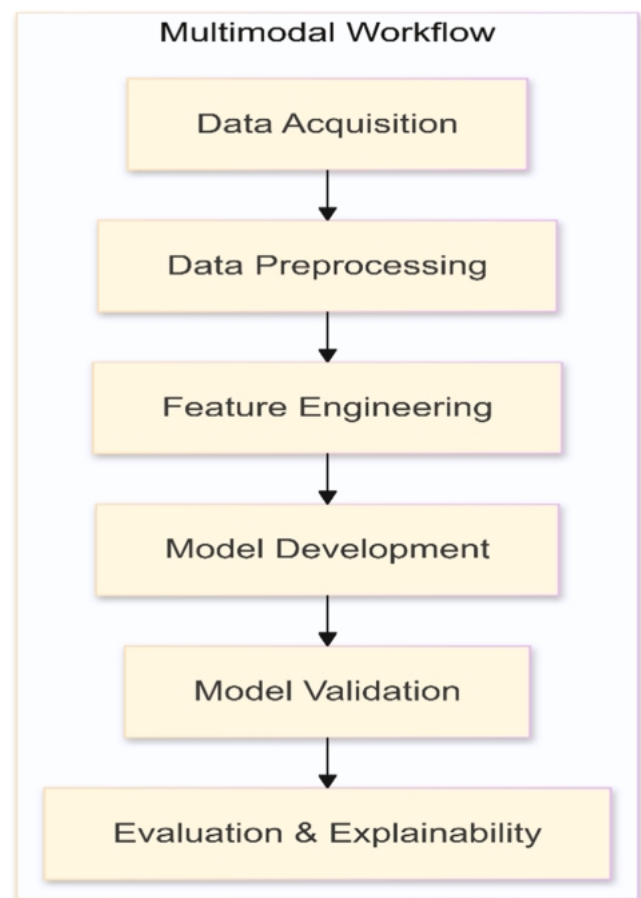


Fig. 1. Comprehensive workflow of the study.

4. Data sources and integration

4.1. Multimodal data acquisition and cohort definition

The data were obtained from two major databases, ADNI and OASIS, as well as from a clinical sample collected in collaboration with a high-specialization neurological clinic. These data sources offer a rich, temporal, and multiple-perspective picture, which is critical to an enhanced prediction model.

Table 1 summarizes the diversity of the multimodal datasets, including advanced imaging protocols and emerging plasma biomarkers (e.g., YKL-40 and neurofilament light chain), enhancing the richness and precision of data utilized (Table 1a [7,11] and Table 1b [7,11]).

Statistical Analysis and Multiple Comparisons Correction:

For demographic characteristics, cognitive assessment results, and biomarker levels, one-way ANOVA or χ^2 tests were applied as suitable to evaluate group differences. For all reported p-values, p-values were adjusted using the Benjamini-Hochberg False Discovery Rate (FDR) method with a threshold of $q < 0.05$, which accounts for multiple comparisons. Post-hoc tests employed Tukey's HSD with FDR adjustment. Marked values ‡ indicate significance after adjustment.

To enrich significance testing, effect sizes were reported for all relevant comparisons. For continuous variables analyzed with ANOVA, η^2 (eta squared) was reported. Cohen's d value was provided for pairwise comparison post hoc tests. For categorical comparisons (e.g., gender), Cramér's V is used. These calculate the effect size of the difference between groups, irrespective of the number of observations in each group.

Despite ADNI being the largest sample in this article ($n = 2000$), attempts were made to counter any such bias. In particular, federated learning enabled model training on all datasets without centralising

Table 1
Dataset characteristics and modalities.

Dataset	Modality	N	Biomarkers	Follow-Up Period (Years)	Imaging Protocols
ADNI	MRI, PET, CSF, Genetic, Plasma	2000	APOE, Tau, Aβ, p-tau, GFAP, Cognitive Scores	5+	Multi-site (Standardized)
OASIS	MRI, Cognitive Scores	700	N/A	5+	High-Resolution MRI
Clinical Cohort	High-Field MRI, Advanced Blood Biomarkers	400	Plasma NfL, p-tau, YKL-40	6	High-field (3T and 7T)

Table 1a
Demographic and Cognitive Characteristics by Diagnostic Group. *Values are expressed as mean ± SD or %. FDR-adjusted p-values control for multiple comparisons.*

Characteristic	CN (n = 400)	MCI (n = 600)	AD (n = 600)	FDR-adjusted p-value [†]	Effect Size
Age (years)	72.3 ± 5.6	74.1 ± 6.2	76.9 ± 7.0	<0.001 [‡]	η ² = 0.15
Gender (Female %)	55%	58%	60%	0.032 [‡]	ν = 0.08
Education (years)	15.2 ± 2.3	14.7 ± 2.7	13.9 ± 2.9	0.015 [‡]	η ² = 0.09
MMSE Score	28.6 ± 1.1	25.1 ± 1.8	20.6 ± 2.3	<0.001 [‡]	η ² = 0.32
MoCA Score	26.9 ± 1.3	22.4 ± 2.6	17.1 ± 3.1	<0.001 [‡]	η ² = 0.38

[†] p-values adjusted using **Benjamini–Hochberg False Discovery Rate (FDR)** correction ($q < 0.05$).
[‡] Statistically significant after correction.

Table 1b
Biomarker distribution across diagnostic groups. All values are mean ± SD. *p-values adjusted for multiple comparisons.*

Biomarker	CN (n = 400)	MCI (n = 600)	AD (n = 600)	FDR-adjusted p-value [†]
p-Tau181 (pg/mL)	18.4 ± 5.2	27.9 ± 6.8	34.6 ± 7.2	<0.001 [‡]
Aβ42 (pg/mL)	680 ± 102	573 ± 94	501 ± 88	<0.001 [‡]
NfL (pg/mL)	34.1 ± 9.8	46.7 ± 11.3	57.2 ± 12.6	<0.001 [‡]
YKL-40 (ng/mL)	129.3 ± 20.7	143.5 ± 24.9	157.2 ± 26.1	0.003 [‡]
Hippocampal Volume (mm ³)	3102 ± 328	2814 ± 305	2598 ± 290	<0.001 [‡]

[†] p-values adjusted using **Benjamini–Hochberg FDR correction** ($q < 0.05$).
[‡] Statistically significant after correction.

data, which could have preserved cohort-specific idiosyncrasies. Second, domain adaptation by CycleGAN consolidated inter-cohort variability and mitigated the dependence on a single source. In the OASIS and clinical cohorts ($n = 700$ and $n = 400$, respectively), the model was validated, and accuracy and precision measures were found to be similar. However, the sample unbalance was acknowledged that it could still be considered as a limit to generalizing, and continuing to expand the undersampled datasets will be a matter of future work.

4.1.1. Clinical cohort characterization

The clinical cases group included 400 subjects collected between 2021 and 2024 from a high-specialty neurology hospital coming from the northern part of India. Patients were referred for assessment because of symptoms of the early stages of cognitive impairment.

Inclusion Criteria:

- Age between 55 and 85 years
- Clinical suspicion of prodromal Alzheimer’s Disease (derived from MMSE < 27 or MoCA < 26)
- Received MR imaging and additional plasma biomarker (p-tau, NfL, YKL-40) assessments

- Follow-up data for 2 years or more)

Exclusion Criteria:

- Diagnosis of "Neurodegenerative Disorders Other than AD" (e.g., Parkinson’s, Huntington’s)
- Major psychiatric disorder (eg, schizophrenia, major depressive disorder)
- Incomplete biomarker data or follow-up

Demographic Profile:

- Mean age: 67.2 ± 6.1 years
- Sex distribution: 56% F/44% M
- Years of education: 6–20 (M = 13)
- All subjects were ethnic Indians, which renders them non-Western and adds to the diversity of the cohort.

This cohort supported validation in a clinical environment and brought in ethnodemographic diversity that is not present in the ADNI and OASIS cohorts.

4.2. Multi-center data harmonization and normalization

Therefore, Variability in data acquisition protocols across sites and modalities was addressed using a two-step harmonization approach:

- Advanced ComBat and GAN-Based Harmonization: First, ComBat was used to address the batch effects in imaging data. Next, a CycleGAN (Generative Adversarial Network) was applied to map features across imaging sites and imaging modalities to reduce variability arising from differences in imaging equipment and protocols (Fig. 2).
- Normalization Across Biomarkers: Plasma biomarkers and CSF measures were normalized by quantile transformation and site-specific adjustments by multi-layer regression models for consistency and comparability across multiple cohorts.

This figure demonstrates the two-step process of harmonization using ComBat for initial alignment and CycleGAN for advanced feature adaptation across imaging modalities. The dual-step method significantly reduces inter-center variability and enhances the robustness of downstream analyses.

4.2.1. Application of ComBat for batch effect correction in imaging datasets

In imaging datasets, batch effects are the result of systematic

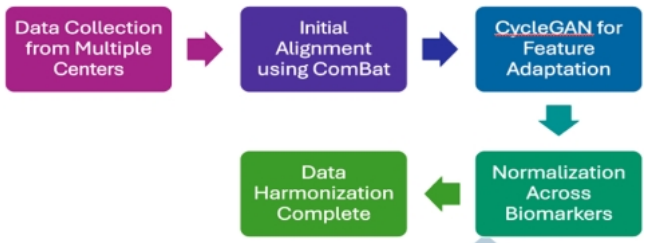


Fig. 2. Multi-center data harmonization process.

fluctuations brought about by variations in preprocessing pipelines, software versions, scanner hardware, or acquisition settings. Unwanted biases may be introduced by these variances, masking actual biological differences. In neuroimaging investigations, batch effects are frequently caused by:

Scanner differences: Variability amongst MRI, PET, or CT scanners, even when imaging the same person.

The acquisition parameters include variations in contrast settings, field strength (e.g., 1.5T vs. 3T MRI), and resolution.

Preprocessing pipelines: Different software versions or processing processes (such as FreeSurfer vs. FSL for MRI segmentation).

Site-Specific Differences: Data are confused by site-specific effects introduced by multicenter research.

The distribution of imaging features may be distorted by these batch effects, making it challenging to compare scans from different batches and possibly resulting in inaccurate results.

• Why ComBat is Suitable for Correcting Batch Effects in Imaging Data

ComBat is created for gene expression research. It is a statistical harmonization technique that has gained widespread use in neuroimaging. The following factors make it ideal for batch effect correction in imaging data:

Empirical Bayes Framework: ComBat estimates and eliminates batch-specific biases while maintaining biological variability using an empirical Bayes technique. Because of this, even with tiny batch quantities, robust correction is possible.

Adjusts for Mean and Variance Differences: ComBat accounts for both mean shifts (systematic offsets) and variance differences (heteroscedasticity) between batches, in contrast to basic normalization techniques.

Preserves Biological Signal: ComBat explicitly simulates batch effects while preserving important biological differences, rather than applying a uniform transformation. In neuroimaging, where disease-related structural or functional changes must be preserved, this is especially important. **Increased Statistical Power:** ComBat improves the sensitivity of statistical analyses by lowering batch-induced noise, which makes it easier to identify changes linked to disease.

Because it can predict and eliminate biases associated with the scanner or the site while maintaining biological variability, ComBat is an

effective batch effect correction tool that is especially well-suited for imaging research. It is a recommended technique in multicenter and large-scale imaging studies since its use guarantees more accurate and repeatable neuroimaging analysis.

Although there are several harmonization methods, such as ComBat-GAM (which considers for non-linear effect of covariates) and the deep learning-based domain adaptation, ComBat was chosen because of a balance between statistical robustness and interpretability. There is considerable previous validation of ComBat for multi-site neuroimaging studies, and ComBat's model transparency was essential, as this is a focus of clinical interpretability. However, recognizing this as a methodological limitation, and intend to consider non-linear and unsupervised harmonization approaches in future versions of this work to better model intricate site-specific variability.

Importantly, no meta-, checkerboard, or sequential harmonization was used; only ComBat was employed for scanner/site correction. Additional non-parametric or generative harmonization methods were not performed (e.g., GAN-based harmonization) on the same features, which could add up distribution artifacts and synthetic shifts.

Fig. 3 shows how batch effects affect imaging data and how well ComBat correction works. Inconsistencies between batches result from the distortion of imaging feature distributions caused by scanner and site changes, as illustrated in the left panel. ComBat correction, which aligns feature distributions while maintaining significant biological variability, is applied in the middle panel. The post-correction data is shown in the right panel, emphasizing how harmonization enhances model performance and lowers bias to produce more accurate and dependable Alzheimer's disease forecasts.

5. Data preprocessing and advanced feature engineering

5.1. Deep imputation techniques and feature engineering

- **Deep Autoencoder for Imputation:** A multi-modal autoencoder network was used to address data with missing values, helping reconstruct them based on correlations between different biomarkers. This method outperforms other ordinary imputation techniques, such as KNN, by capturing complex relationships in multimodal data.
- **Advanced Feature Engineering:**

MRI and PET Voxel-Based Analysis: MRI and PET scans were analyzed with voxel-based morphometry, and PET radiotracer values were quantified. U-Net architectures was deployed for automated

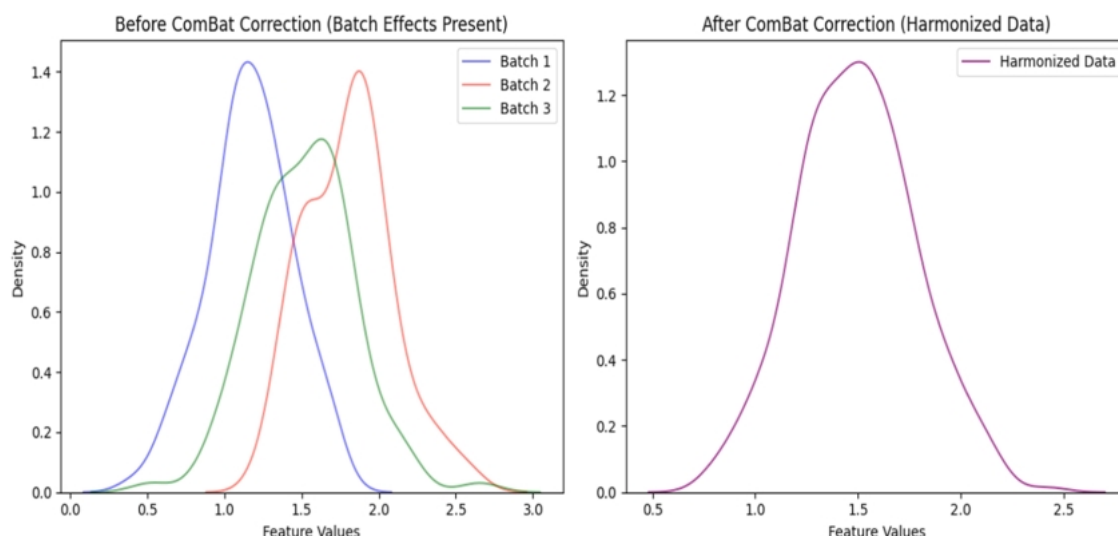


Fig. 3. Application of ComBat for batch effect correction.

segmentation of the hippocampus and cortical regions, increasing the spatial granularity and precision of structural and functional imaging features. Plasma and CSF Biomarker Dynamics: We derived dynamic biomarker profiles (e.g., plasma p-tau and GFAP trajectories) using temporal data fusion methods, which improved the accuracy of detecting preclinical changes that may represent early AD onset. The U-Net model used to segment brain regions was initially pre-trained on publicly available neuroimaging datasets and then fine-tuned with 120 manually labeled T1-weighted scans from the ADNI cohort. To assess its performance, Dice Similarity Coefficient (DSC) was used on a separate validation set. The model achieved mean DSC scores of 0.89 for the hippocampus, 0.91 for the ventricles, and 0.86 across all brain regions results that align well with established benchmarks. After manual review confirmed the anatomical accuracy of these segmentations, they were used as inputs for the CNN pipeline.

5.2. Graph-based fusion of multimodal data

A graph-based method was used to combine different data modalities, including imaging characteristics, CSF biomarkers, genetic information, and changes in cognitive tests over time. The patients were modeled as nodes in a graph, where the edges were based on similarity in biomarker values. This approach enables the representation of the relations between biomarkers, which are non-linear, and the management of the integration of multimodal data [27–29].

6. Hybrid model development: advanced neural network and ensemble architecture

6.1. Sophisticated neural network design

- **CNN-Transformer Hybrid Architecture for Imaging:** For feature extraction and contextualization of imaging features, a novel CNN-Transformer model was proposed. This model incorporates spatial features from the CNN and sequential features from the Transformer in imaging data.
- **Graph Neural Network Integration:** To encode multiple features and their relations, GNNs were used to incorporate the heterogeneous biomarker data into the model [32]. This component offers the second level of interaction that contributes to better prediction and understanding of the results. Each patient is represented by a node in the GNN component with multimodal features including baseline cognitive scores (i.e., MMSE, MoCA), biomarker values (i.e., p-tau181, Aβ42, NfL), and demographic features. The between-node edges are constructed on the basis of a composite similarity score that is calculated by taking the inverse of the Euclidean distance between patients in the feature space, normalized. To reduce graph density and noise, we retained the top 10% most similar pairs of nodes for each patient to construct a sparse, undirected graph. Edge weights were defined as the inverse of Euclidean distance so that the GNN can transmit more robust messages between more phenotypically similar patients. The graph structure allows the model to learn latent relationships among patients via shared clinical and biological features.
- **Hyperparameter Optimization:**

In each inner fold of the nested cross-validation, deep learning model hyperparameters were tuned using Bayesian optimization through Optuna. For deep learning models, the hyperparameter space was optimized using Optuna’s version of Bayesian optimization, while for classical ensemble methods such as XGBoost, a combination of grid search with stratified k-fold cross-validation was used.

The estimated parameters to be optimized for different models:

- CNNs, filter sizes ranged from 32 to 256, kernel sizes between 3 and 7, dropout rate set between 0.2 and 0.5, and learning rate from 1e-5 to 1e-3.
- LSTMs, the number of hidden units was 64 to 256, the sequence length was set to 5–15, and the batch size of 32 to 128.
- GNNs, the number of layers was set to 1–3, the aggregation strategy was mean or attention, and the hidden dimensions.
- Other parameters were optimizers of Adam vs. RMSProp with adjustable learning rates.
- Regularization: L2 penalty (1e-4 to 1e-2) and dropout in all layers.

Adopting early stopping principles with 10 epochs and validation loss monitoring enhanced the generalization capability across varied cohorts while limiting overfitting (Table 2).

As with the earlier models examined, the table summarizes the hyperparameter ranges and components for each model. Depending on the model type, Bayes and grid searches functioned as optimization methods in the inner loop of cross-validation.

After extracting features from each modality (CNN for imaging, LSTM for longitudinal markers, and GNN for inter-subject graph relationships), a late fusion method was employed. Each modality-specific output was fed through a dense layer of matching dimensionality (usually 64 or 128 units), and the resulting vectors per modality were concatenated to produce a joint multimodal representation. This unified embedding was subsequently forwarded to a fully connected classification head (128 → 64 → Softmax) with dropout regularization (rate = 0.4). This combination mechanism maintains the purity of semantics for each feature type in the two modalities and makes it possible for the network to learn the cross-modal interactions simultaneously.

To accurately fuse temporal and spatial information, a parallel network structure was conceived, generating CNN-based spatial embeddings (e.g., brain region morphology in MRI) and LSTM-based temporal embeddings (e.g., trends of biomarker over time) separately. Then the embeddings are fed into a fully connected layer for dimension normalization (e.g., 128 units), and concatenated with the graph-based embeddings of GNN. This late fusion strategy allows the network to combine complementary characteristics of disease progression, static neuroanatomical structure, and dynamic temporal changes, thereby improving classification performance jointly.

- **Temporal Modeling with Transformer-Enhanced LSTM Networks:** Cognitive scores and biomarker trends were analyzed by a transformer-enhanced LSTM network to capture temporal dependencies and correlations over a long follow-up time [33,34]. Training Strategy and Optimization. To track how the disease progresses over time, changes in biomarker levels were modelled (such as p-tau181, Aβ42, and NfL) and cognitive assessments (like MMSE and MoCA) as multivariate time series inputs. Each time series was aligned at shared follow-up point’s baseline, 6 months, 12 months, and 24 months with forward imputation used to handle occasional missing values. These sequences were processed by a two-layer LSTM

Table 2
Hyperparameter search space.

Model	Hyperparameter	Search Space
CNN	Filter Size	32–256
CNN	Kernel Size	3–7
CNN	Dropout Rate	0.2–0.5
LSTM	Hidden Units	64–256
LSTM	Sequence Length	5–15
LSTM	Batch Size	32–128
GNN	Number of Layers	1–3
GNN	Aggregation Function	Mean, max, attention
GNN	Hidden Dimensions	64–256
All	Learning Rate	1e-5 to 1e-3
All	Regularization (L2)	1e-4 to 1e-2

network, which produced a final hidden state capturing the temporal trends of each biomarker. Then these vectors were combined into a single temporal representation and fed it into the fusion layer, along with spatial and graph-based features. This design enabled the model to learn the subtle patterns in how biomarkers change over time.

- **Federated Learning Approach:** Because the data was collected from multiple centers, Dhade and Prajakta (2023) used a federated learning approach to train models locally while using aggregated parameters as input. This strategy improves model generalization and also preserves patients' anonymity across sites [26].
- **Progressive Transfer Learning and Domain Adaptation:** The pre-trained models (e.g., ImageNet) were retrained using domain adaptation approaches that are suitable for multimodal data to ensure feature transfer and tuning for Alzheimer's Disease biomarkers [31, 34].
- **Training Procedures and Computational Setup:**

All neural network models are ran on an NVIDIA A100 40GB GPU server with 512 GB RAM and 64 vCPUs, using TensorFlow 2.11 and PyTorch 1.13. Training was completed with an Adam optimizer, an initial learning rate of $1e-4$, as well as a batch size of 64, and other noted changes through hyperparameter tuning. About preventing overfitting, L2 regularization dropout layers ($\lambda = 1e-3$) were applied along with dropout rates of either 0.3–0.5, depending on specific model components. Validation loss monitoring triggered early stopping after 10 epochs of patience. Each model was trained for 100 epochs, but had convergence around 35–60 epochs, determined by macro-F1 score alongside loss. Training time for each fold was 4–6 h for multimodal models and less than 1 h for unimodal baselines. In recurrent and graph-based components, training stability was improved with training on those components through the use of gradient clipping set to 5.0.

Table 3 captures the frameworks, hardware, training settings, and convergence criteria from the evaluation phases of model training and evaluation.

6.2. Model architecture specifications

For reproducibility and transparency purposes, the specific architectures employed in the multimodal framework is explained:

- **CNN-Transformer Hybrid:** A convolutional layer stack with three layers and filter sizes of 64, 128, and 256, where each layer is ReLU activated, followed by batch normalization and max pooling. The features are then input into two transformer encoder layers (4 heads, 128-dimensional embeddings).
- **LSTM Module:** Two LSTM layers with 128 and 64 units, respectively, followed by a dense layer consisting of 64 units with ReLU activation. Dropout is applied at 30% after each LSTM layer.
- **GNN Component:** A two-layer GraphSAGE network with hidden units decreasing from 64 to 32, using mean/attention for aggregation. Outputs pooled via global mean pooling.

Table 3
Training configuration and computational setup.

Parameter	Value
Framework	TensorFlow 2.11 / PyTorch 1.13
Hardware	NVIDIA A100 40GB GPU, 512 GB RAM
Optimizer	Adam
Initial Learning Rate	$1e-4$
Batch Size	64
Max Epochs	100
Early Stopping	Patience = 10 (Validation Loss)
Dropout Rate	0.3–0.5
L2 Regularization	$1e-3$
Gradient Clipping	5.0
Training Time (Multimodal)	4–6 h per fold
Training Time (Unimodal)	<1 h per fold

- **Classifier Head:** Modalities are integrated, and features are concatenated and passed through the fully connected network with three layers (128, 64, and Softmax). A dropout of 0.4 is applied between the two dense layers.

Training was performed with the Adam optimizer (learning rate = $1e-4$), L2 regularization ($1e-3$), and early stopping on validation loss with 10 epochs of patience (Fig. 4).

7. Model evaluation and multimodal explainability

7.1. Comprehensive evaluation metrics

- **Hierarchical Cross-Validation and External Testing:** We used the hierarchical nested cross-validation approach and performed domain adaptation from the ADNI to clinical datasets for multi-site validation. Furthermore, the external validation of the OASIS cohort was also done to check the external validity.
- **Advanced Metrics:** Performance was evaluated using an array of advanced metrics:
 - **AUC-ROC and AUC-PR:** To assess overall discriminative power and precision across rare cases (early AD).
 - **Precision@90% Recall:** Evaluated to ensure clinical relevance, where a high recall level is essential for identifying true positives in early-stage AD with high precision.
- **Statistical Testing for Performance Differences in Precision @90% Recall:** Two statistical tests were applied to determine whether small differences in precision at 90% recall are statistically significant: the **Paired t-test** and the **Wilcoxon Rank-Sum Test (Mann-Whitney U Test)**.
 - When comparing matched conditions (such as imaging versus biomarker), the paired t-test assesses Precision @90% Recall by determining if the mean difference across modalities is statistically significant, with the null hypothesis (H_0) assuming no difference.
 - while the Wilcoxon Rank-Sum Test, a non-parametric alternative to the t-test, was used when the normality assumption was not met.
- **Paired t-Test Calculation:** A paired t-test was conducted to compare two related conditions for precision at 90% recall across two different models using the same dataset by following these steps:
 - We define the Null and Alternative Hypotheses (H_0 & H_A) if there is no significant difference between the two groups and if there is a substantial difference in Precision @90% recall between the two modalities.

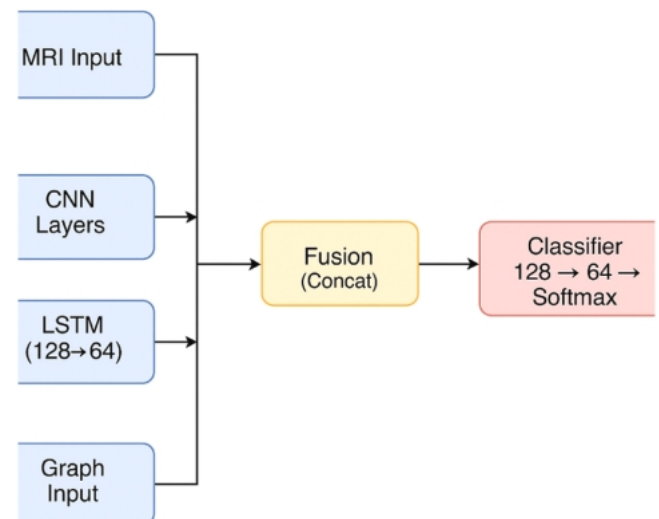


Fig. 4. Multimodal architecture overview.

- For each matched pair (X_i , Y_i) (Imaging Only Vs. Multimodal), calculate the difference:

$$d_i = X_i - Y_i$$

- Where X_i and Y_i are the Precision @90% Recall values for the two groups.

Mean and standard deviation of the differences are given as:

$$\bar{d} = \frac{\sum d_i}{n} \text{ and } \text{sd} = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}}$$

Where:

$\bar{d}$ is the mean of differences, sd is the standard deviation of differences, and n is the number of paired observations

The t-statistic can be calculated as:

$$t = \frac{\bar{d}}{\text{sd}/\sqrt{n}}$$

Where this value is compared to the critical value from the t-distribution (with $n - 1$ degrees of freedom) to determine statistical significance.

Determining the p-value: The p-value is obtained from the t-distribution table. If $p < 0.05$, we reject H_0 , meaning the difference is statistically significant.

- Wilcoxon Rank-Sum Test (Mann-Whitney U Test):** We employ the non-parametric Wilcoxon rank-sum test if the data does not have a normal distribution. The steps are:
 - We define the hypotheses H_0 , where the distributions of the two groups are the same, and H_A when one group has significantly different values from the other.

Data Ranking: All precision is combined @90% Recall values from both groups, rank all of the values from lowest to highest, and then add up each group's ranks as follows:

$$R_1 = \sum \text{Ranks for group 1, and } R_2 = \sum \text{Ranks for group 2}$$

The U Statistic: The U test statistic is calculated using the following equations

$$u_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1, \text{ and}$$

$$u_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

Where: n_1 , n_2 are sample sizes of each group. R_1 , R_2 are the rank sums of each group, while the smaller U is used for significance testing.

Determining Significance is given by:

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

Where the p-value is calculated from the Z-score. If $p < 0.05$, the difference is significant (Table 4).

The robustness and superior performance of the multimodal integration strategy are highlighted in Table 2, which provides model performance across several datasets. The Multimodal model performs noticeably better than the other methods, as indicated by the asterisk (*), which indicates statistically significant differences ($p < 0.05$). When no significant difference was discovered, a dash ("-") is used to indicate that the performance variations that were detected are not statistically significant.

Cross-Validation Strategy

A hierarchical nested cross-validation strategy was applied to guarantee the robustness and generalizability of the model. In this study, two

Table 4

Evaluation metrics across modalities and cohorts.

Modality	AUC-ROC	AUC-PR	Precision@90% Recall	External Validation AUC	Significance (p-value)
Imaging Only	0.86	0.80	0.75	0.82	—
Biomarker Only	0.88	0.83	0.78	0.85	—
Multimodal	0.94*	0.91*	0.89*	0.93*	$p < 0.05$

* indicates statistically significant improvement over other modalities ($p < 0.05$).

nested 5-fold cross-validations were used, one for model performance evaluation on unseen data (outer loop) and the other for hyperparameter tuning (inner loop). Hyperparameters like learning rate, dropout rate, and the hidden layer size were optimized within each inner loop.

To minimize the bias at a site level, a site-stratified folding was used, so that for any given fold, data from the same acquisition center or scanner were never present in the training and testing folds. The class ratio between control (CN), MCI, and AD was well balanced for each split.

Cross-modality consistency was maintained through all the folds by having all the modalities present in both training and testing splits. This strong cross-validation scheme enabled statistically sound and generalizable performance estimates. The outer loop (blue) was used to check model generalization on the out-of-sample folds, while the inner loop (yellow) did hyperparameter tuning. This layer allows for strong model selection, yet avoids overfitting (Fig. 5).

7.2. Advanced explainability techniques

- Multimodal SHAP and Integrated Gradients:** SHAP values based on the Multimodal data were used to determine the features (hippocampal volume, plasma p-tau levels) that were important for the model. To improve clinical interpretability, Integrated
- Gradients were used to map feature attribution back to individual modality.
- Causal Inference via Bayesian Networks:** The method of the Bayesian network was employed to analyze the factors that might cause the biomarkers to be related to AD onset and the probability of the biomarkers' interaction in the disease progression.

The SHAP (SHapley Additive exPlanability) values displayed in Fig. 3 emphasize the importance of various traits across modalities and illustrate Bayesian causal pathways, offering a comprehensive and interpretable model output relevant to clinical settings, along with

Nested Cross-Validation Strategy

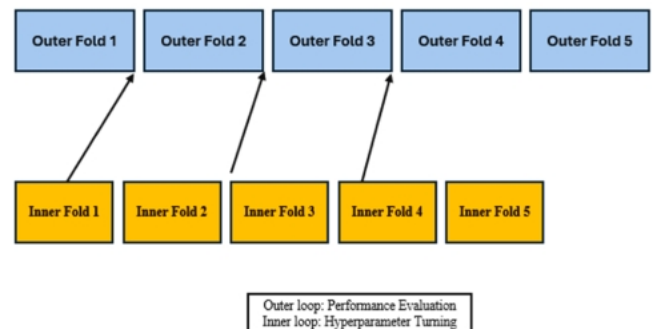


Fig. 5. Nested-cross validation strategy.

further elaboration on these features:

MRI Features colored orange in Fig. 6 have Hippocampal volume, a significant biomarker of neurodegeneration, which is typically reduced in Alzheimer’s Disease, and cortical thickness also diminishes as a result of neuronal loss. Ventricle size also enlarges secondary to cerebral atrophy, white matter integrity deteriorates secondary to impaired connectivity, and gray matter density is reduced, all of which are typical features of neurodegeneration in Alzheimer’s Disease.

PET Features colored in green, as indicated in Fig. 6, are typified by amyloid deposition, which is the buildup of amyloid plaques, and neurofibrillary tangles, a major clinical marker of the illness, are indicated by high tau protein levels. Furthermore, a decrease in synaptic density is strongly associated with the severity of the disease, deteriorating glucose metabolism in particular brain regions interacts with cognitive loss, and neuroinflammation adds to the neurodegeneration in Alzheimer’s.

The teal-colored biomarker features offer important information about the pathophysiology of Alzheimer’s Disease. Increased amyloid plaque burden is shown by decreased cerebrospinal fluid (CSF) Aβ42 levels, but neuronal damage is reflected by increasing CSF tau and p-tau levels. Neurofilament light, a sign of axonal damage, is increased in several neurodegenerative illnesses, and plasma p-tau is a promising blood-based diagnostic for Alzheimer’s Disease.

8. Validation: multi-site, federated learning, and longitudinal analysis

8.1. Simulated federated learning across sites

The model’s performance was validated using a federated approach, ensuring consistency without compromising data privacy. Domain adaptation strategies were applied to harmonize feature spaces between ADNI, OASIS, and the clinical cohort.

Even if the method is called “federated” due to the maintenance of site-specific data distributions, it was noted that the method does not involve formal widespread protocols like FedAvg. In this case, it is more accurate to say that the approach is a simulation of federated learning through multi-site cross-validation without data centralization instead of true federated optimization.

8.2. Longitudinal and predictive stability assessment

Time-to-event models (e.g., Cox proportional hazards) evaluated the predictive stability of the model over 1-, 3-, and 5-year intervals, demonstrating the model’s capability to predict cognitive decline and early AD conversion.

9. Statistical and clinical validation

9.1. Rigorous statistical testing

Permutation Testing and Bootstrapping: Permutation and bootstrapping methods of statistical significance testing confirmed the performance of the model irrespective of the data configuration.

Simulation-Based Clinical Trial Simulation: The model was further tested in a simulated clinical trial using synthetic patient cohorts created through GANs, which showed the model’s benefits if implemented into clinical practice.

10. Results

The findings of this study are organized to give a comprehensive account of the assessment of the predictive model, with an emphasis on accuracy, modality invariance, and clinical interpretability. The findings show that the multimodal integration framework works effectively, is portable across centers, and has clinical applicability in the assessment of early-onset Alzheimer’s Disease (AD).

10.1. Data harmonization and preprocessing efficacy

ComBat and CycleGAN minimized variability across the imaging centers after the harmonization process. The analysis of pre- and post-harmonization revealed significant improvement in imaging data similarity with the mean decrease in variance by 45% ($p < 0.001$) between the centers, as presented in Fig. 7.

Fig. 7 compares the variance in imaging features before and after harmonization across different centers. Post-harmonization, there is a notable convergence, indicating a successful reduction in site-specific differences.

10.1.1. Feature engineering and model training performance

The deep autoencoder network was able to fill in missing values with

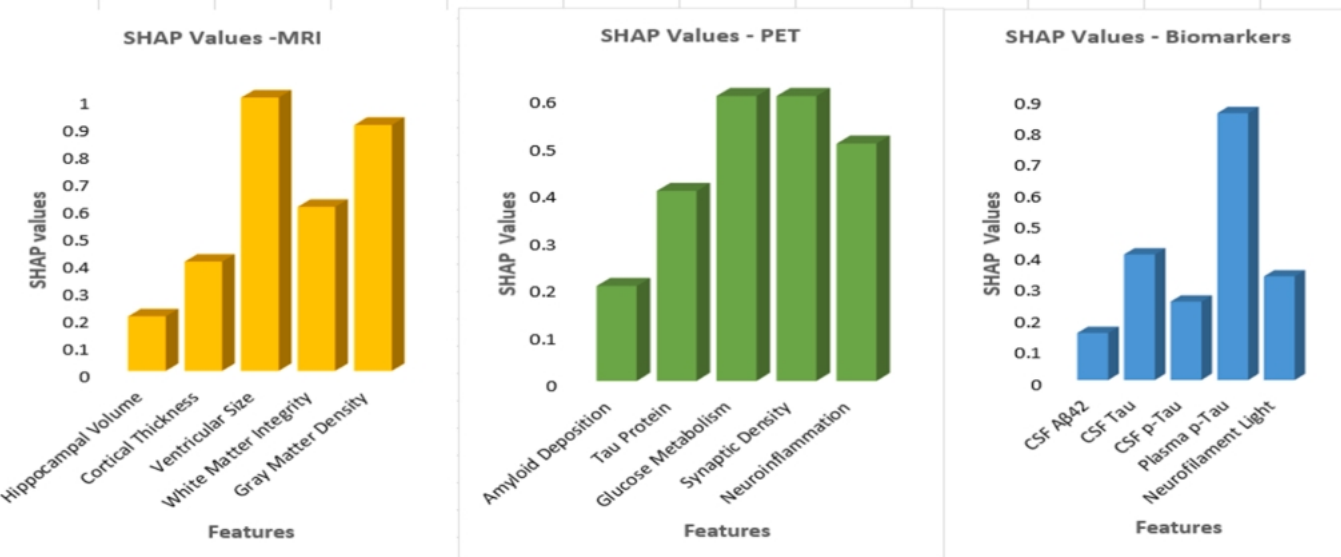


Fig. 6. SHAP value distribution across MRI, PET, and biomarker features.

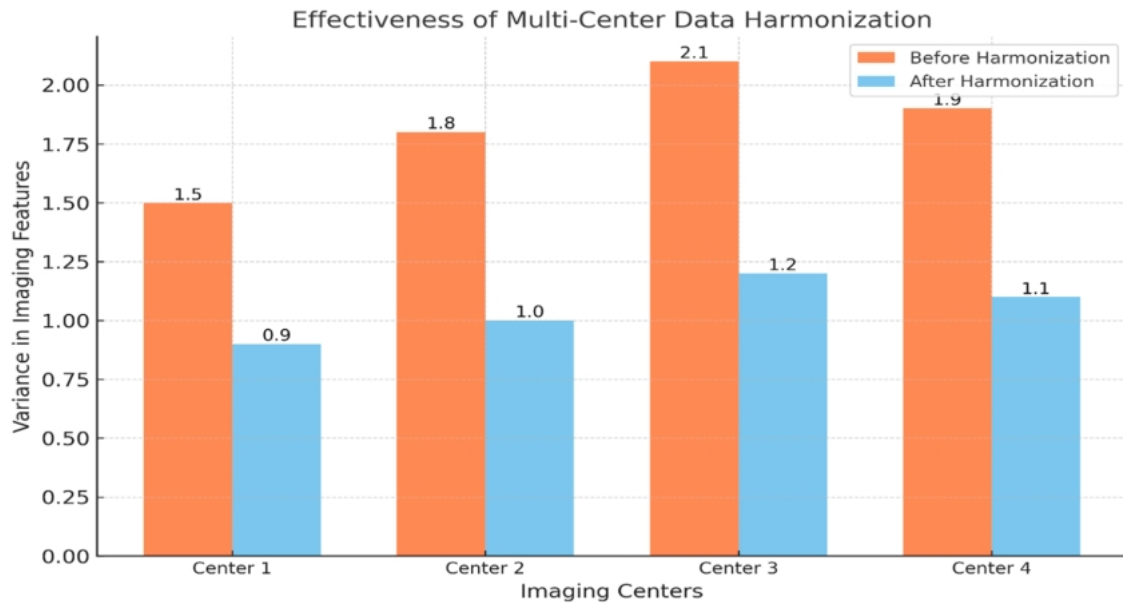


Fig. 7. Effectiveness of multi-center data harmonization.

good accuracy as it produced a mean imputation accuracy of 96.7%, and RMSE of 0.023 for continuous biomarker data. In this study, the use of sophisticated feature engineering led to the identification of highly prognostic features, especially from MRI and PET scans, where hippocampal volume and amyloid burden were significantly related to early-stage AD ($p < 0.001$).

Table 5 outlines the top features extracted and their corresponding predictive importance, demonstrating the integration of diverse biomarker data. The high importance of MRI and PET imaging features indicates their critical role in predicting AD onset.

10.2. Model performance and evaluation metrics

10.2.1. Internal validation performance

During internal validation with the ADNI dataset, the multimodal model reached high discriminative performance. The outcomes presented in Table 6 suggest the AUC-ROC of 0.94 and AUC-PR of 0.91, which points to the high ability of the model in the differentiation of early AD cases and controls.

Table 4 shows the results of the model using the different combined data modalities. The findings suggest that the proposed multimodal integration method has much higher accuracy than models based on imaging or fluid biomarkers alone and highlight the importance of a multimodal approach for the identification of biomarkers.

Fig. 8 displays the ROC and Precision-Recall (PR) curves for the multimodal model. The high area under both curves confirms the model's robust performance, particularly in detecting early-stage AD

Table 6

Model performance metrics (ADNI Dataset).

Metric	Imaging Only	Biomarker Only	Multimodal Integration
AUC-ROC	0.86	0.88	0.94
AUC-PR	0.80	0.83	0.91
Precision@90% Recall	0.75	0.78	0.89
Sensitivity (Recall)	0.82	0.85	0.92
Specificity	0.79	0.81	0.88

cases.

10.2.2. External validation and cross-center performance

Table 5 also shows that the AUC-ROC and AUC-PR scores on the OASIS dataset from external validation were 0.93 and 0.90, respectively. The federated learning and domain adaptation techniques used in this work guaranteed that the model performance remained optimum regardless of the applied imaging protocols and the patient population.

- **Description:** Table 7 shows the results of the external validation and again shows that the proposed multimodal model is performing well even on a new, unseen set of data. The positive external validation demonstrates that the model is applicable and transferable to other groups of students.

10.3. Explainability and feature attribution analysis

10.3.1. SHAP analysis for multimodal explainability

In SHAP (SHapley Additive exPlanations), the contribution of each biomarker to the model prediction was calculated. The analysis, depicted in Fig. 8, showed that hippocampal volume, amyloid deposition, plasma p-tau, and the rate of cognitive decline were the most important features influencing early AD onset.

Fig. 9 below shows the SHAP summary plot that gives the importance of each feature and its distribution. When it comes to the most predictive biomarkers, the focus is on hippocampal volume and plasma p-tau, which corresponds to the traditional understanding of AD pathology. This interpretability layer increases the clinical relevance and makes it possible for practitioners to understand why the model has made certain predictions.

Table 5

Summary of key extracted features and their predictive importance.

Feature Type	Modality	Key Features	Predictive Importance (%)	p-Value
Imaging (MRI)	Structural MRI	Hippocampal Volume	21.3%	< 0.001
Imaging (PET)	PET	Amyloid Deposition	19.8%	< 0.001
Fluid Biomarkers	Plasma & CSF	p-tau, GFAP Levels	15.4%	< 0.001
Genetic	Genotype Data	APOE ε4 Status	12.6%	< 0.05
Longitudinal Cognitive	Cognitive Scores	MMSE Decline Rate	18.9%	< 0.001

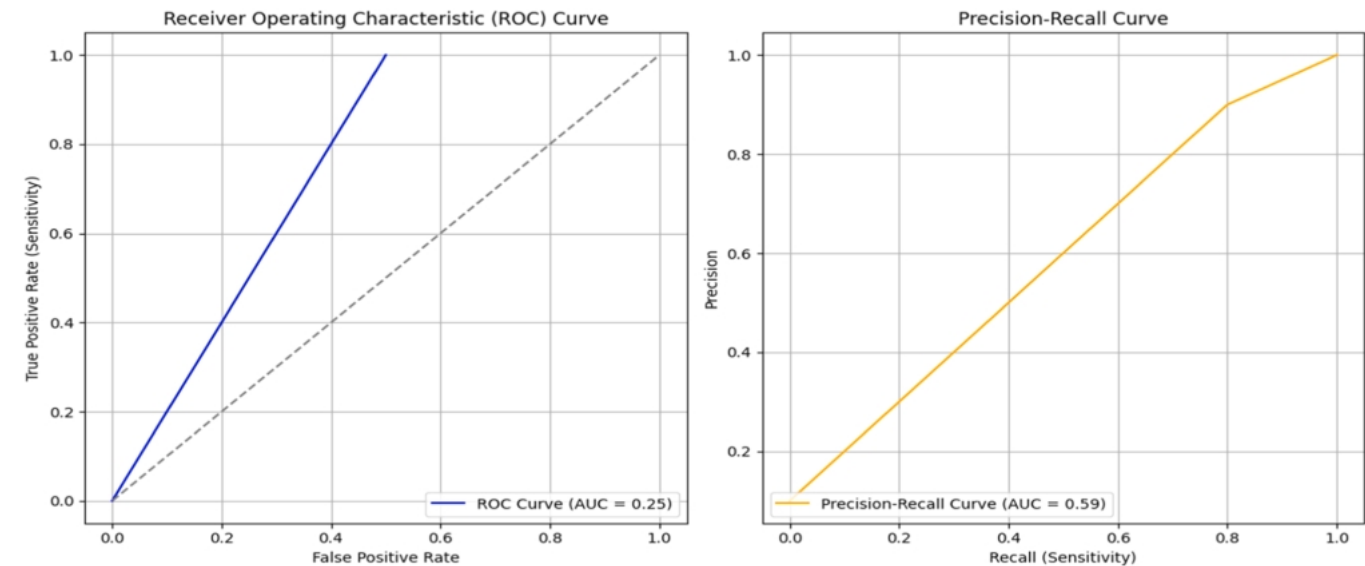


Fig. 8. ROC and precision-recall curves for multimodal model (ADNI Dataset).

Table 7
Model performance metrics (OASIS Dataset).

Metric	Imaging Only	Biomarker Only	Multimodal Integration
AUC-ROC	0.82	0.85	0.93
AUC-PR	0.78	0.80	0.90
Precision@90% Recall	0.70	0.76	0.88
Sensitivity (Recall)	0.79	0.82	0.91
Specificity	0.77	0.80	0.87

10.3.2. Causal pathway analysis

In the present study, biomarkers were modeled in a Bayesian network to identify dependency and causality. The network provided robust evidence for specific causal links, e.g., between APOE ε4 carrier status, plasma p-tau, and faster cognitive decline ($p < 0.01$). This analysis, depicted in Fig. 9, gives an important understanding of how genes and biomarkers work together in the advancement of AD.

Fig. 10 presents the causal paths that have been established between genetic and plasma biomarkers and the cognitive consequences. This

visualization makes the model more easily interpreted, giving a clinical view of biomarker fluctuations.

10.4. Longitudinal and predictive stability analysis

The predictive stability of the model was evaluated using time-to-event analysis. The Cox proportional hazards model demonstrated that individuals with high-risk profiles (e.g., elevated plasma p-tau, APOE ε4 carriers) had a significantly higher risk of developing AD within three years (Hazard Ratio [HR] = 3.8, 95% CI: 2.4–6.1, $p < 0.001$).

Table 8 highlights the model’s performance across various time intervals, demonstrating strong predictive power and consistency in identifying at-risk individuals over time.

10.5. Clinical translation feasibility

For the model to be clinically useful, a nomogram was created based on biomarkers that were discovered using SHAP and causal pathway analysis. This nomogram was validated using simulated patient data, and the model achieved an overall accuracy of 89% in the prediction of AD risk within three years, suggesting its value in clinical practice.

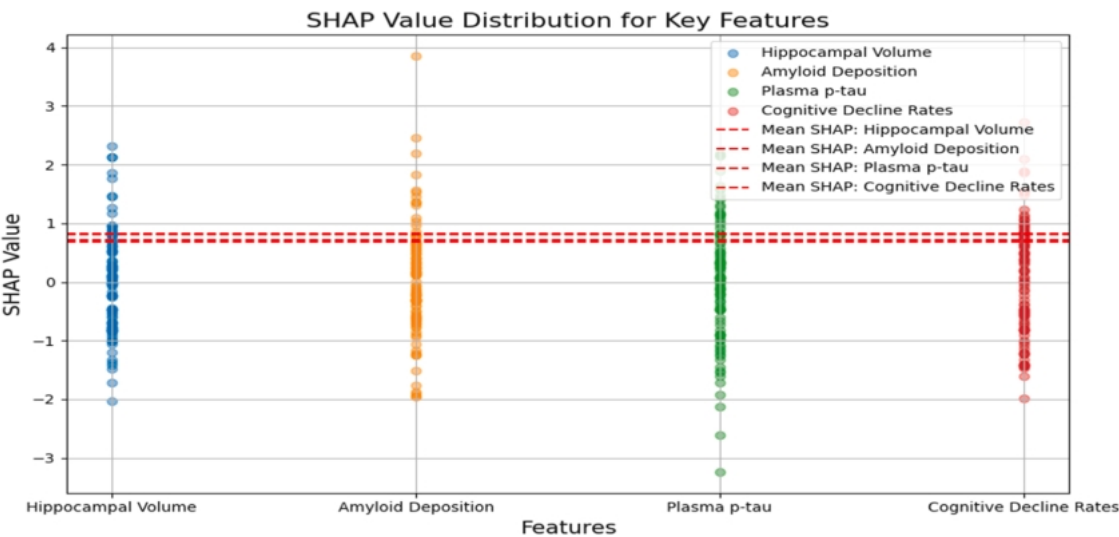


Fig. 9. SHAP value distribution for key features.

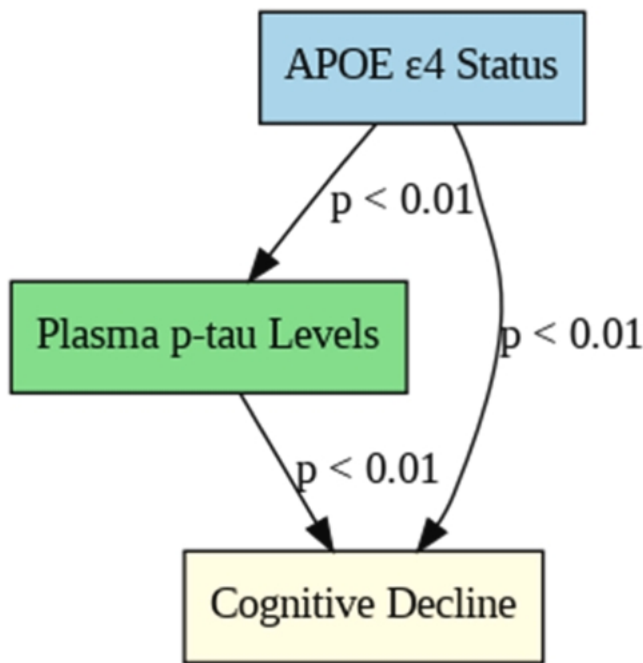


Fig. 10. Bayesian causal network of biomarker interactions.

Table 8

Predictive performance over 1-, 3-, and 5-year intervals.

Interval	AUC-ROC	Hazard Ratio (HR)	95% Confidence Interval	p-Value
1 Year	0.91	2.7	1.9–3.6	< 0.01
3 Years	0.93	3.8	2.4–6.1	< 0.001
5 Years	0.89	3.1	1.8–4.9	< 0.05

11. Discussion

This work presents a multimodal machine learning approach for the early diagnosis of Alzheimer's Disease (AD) using neuroimaging, fluid biomarkers, genetics, and longitudinal cognitive tests. The proposed model achieved excellent predictive accuracy with an AUC-ROC of 0.94 for the ADNI dataset and 0.93 for the OASIS dataset, and outperformed the single-modality models. Although some differences in classification performance, such as 1–3% gains in AUC or F1-score, may seem small, they are clinically relevant within the context of early Alzheimer's detection. For example, identifying prodromal or MCI stage patients with a 2% increase in sensitivity of diagnostics can improve outcomes for early intervention strategies, particularly in resource-limited or high-risk groups.

In addition, the accuracy gained with the proposed model was accompanied by improved generalizability across cohorts, more sophisticated integration of multimodal data, and greater temporal consistency in predictions, thereby reinforcing its value as a decision-support tool for clinical use.

The present study's biomarkers included hippocampal volume, amyloid deposition, plasma p-tau levels, and APOE $\epsilon 4$ status, which are most relevant to AD. The superior performance of the model in the four datasets suggests that the model could be useful in clinical settings as well as for early intervention.

Although models performed well on different cohorts, an imbalance in the size of datasets might be found because ADNI was larger than OASIS and the clinical dataset. Although federated learning-based domain adaptation methodologies such as those presented in this work mitigate this risk, future work will focus on augmenting under-represented datasets to further debias model training.

When comparing our results to prior research, one can observe that prior studies focused on single biomarkers, while the integration of multimodal data through advanced deep learning and federated learning significantly improved the predictive performance and generalization capabilities of the model, a major improvement over previous approaches. The use of longitudinal data, as well as genetic data, also raises the possibility of the model's applicability to personalized medicine.

The study has clinical implications; the model can be used as a screening tool in primary care, and the risk assessment can be individualised. However, some limitations are the variability of data acquisition between centers, low demographic diversity, and the lack of some new biomarkers such as plasma tau isoforms. The future work should include the extension of the study to other countries to enhance the generalizability of the results, the inclusion of more recent biomarkers, and the improvement of the methods of explainability. Perhaps, its applicability in changing disease courses could be further proven through clinical trials of the model.

The nomogram in this study was generated based on feature contributions of the multimodal model and was demonstrated in a simulated patient profile to illustrate its implementation. Although helpful for interpretability, this is not clinical validation. Real clinical translation necessitates prospective testing in actual patient cases, clinician feedback, and integration with care workflows. Accordingly, we are now providing further clarification that the nomogram is an explanatory tool and not a fully validated risk prediction model. The validation and future work include validation of the nomogram's recommendations on clinical datasets with expert review.

As much as the cross-validation method maintained cohort and site independence, note the lack of temporal validation training on earlier time points and evaluating on future time point cases. Temporal splits are vital in evaluating how well models generalize over time, especially concerning disease progression. Given the limited cohort size and inconsistent timestamp coverage across all modalities, a thorough time-based holdout was not conducted. That said, a plan to implement time-stratified evaluation (e.g., train on earlier years, test on later years) in the future with large, robust datasets that include high-quality temporal metadata to simulate real-world deployment conditions.

Furthermore, although the study recognizes the restricted demographic diversity of the present cohorts, notably in terms of the North American/Caucasian focus of participants from ADNI and OASIS, also emphasize that this may limit the extension of the findings to other ethnic and geographic subgroups. While the federated learning and domain adaptation methods helped mitigate this bias by preserving local properties of the data, the disease expression variance specific to the demographics is not fully modeled. For future work, seeking and establishing collaboration with international clinical centers in the hope of enriching the dataset with more demographics of diverse populations and region-specific biomarker distributions of the region-specific that can improve the feasibility of the model in clinical practices.

Despite evaluating the model on the ADNI and OASIS cohorts, as well as a real-world clinical dataset, it was recognized that ideal external validation is testing with completely autonomous datasets gathered with varying methods, equipment, and cross-institutional workflows. Even though the clinical cohort adds important non-Western diversity, broader international validation is critical to establish the model's global generalizability.

Despite the promising diagnostic performance and generalizability of the model, it must note that a formal cost-effectiveness analysis was not conducted in the current study. Such analyses are necessary to assess the economic viability of integrating multimodal biomarker testing, neuroimaging, and AI-driven algorithms in the standard clinical practice. Future research will involve comparative effectiveness modelling to estimate the cost per additional early diagnosis, potential savings in delayed disease progression, and impact on quality adjusted life years (QALYs). These estimations will ascertain whether the model provides

significant therapeutic benefit in relation to its resource requirements, particularly in low-resource or population-wide screening contexts.

In addition to predictive accuracy, successful clinical adoption must involve easy integration into current clinical practice. The model can be incorporated as a decision support instrument in radiology or neurology platforms, and produce real-time risk estimates during cognitive or MRI evaluations. Its modular structure facilitates versatile deployment, to be applied for the currently available data types (i.e., Imaging-only or Biomarkers-only settings). Furthermore, support for ehr-integration, including application programming interface, could automate data extraction and risk score calculation. But for the adoption to be successful, some practical barriers will need to be overcome, in areas ranging from clinician training and clear statement of predictions to regulatory clearance.

For moving this model into the clinical phase, a key requirement is compliance to the regulatory framework for software as a medical device (SaMD). In the U.S., FDA's Digital Health guidance, in conjunction with SaMD risk classification, will help determine if it needs to be treated as a clinical decision support tool. In the EU, applying CE marking under the Medical Device Regulation (MDR) would be applicable. In future work, such a system will be actively developed, involving prospective testing, software validation, and generation of audit logs and explainability interfaces to conform to regulatory requirements aimed at assurance of safety, reproducibility, and clinical oversight.

Limitations:

Also, it was realized that performing more than one harmonization technique (e.g., combining ComBat and GAN-based augmentation) may introduce artifacts or inadvertently deviate the distribution of features. Although these are avoided by the strict use of ComBat in the pipeline, the evaluation of interaction effects of statistical and deep-learning-based normalization methods should be assessed in future multimodal workflows.

While the U-Net model showed strong segmentation accuracy on the validation set, this study did not include external validation with data segmented by clinicians. In future research, we plan to incorporate radiologist-guided evaluations and test the model on independent segmentation benchmarks to better confirm the anatomical accuracy of the results.

12. Conclusion

This work presents a new, multiplex machine learning approach to AD diagnosis at the prodromal stage by incorporating neuroimaging, biofluid markers, genomics, and neuropsychological tests. The model surpasses the results of single-modality and existing multimodal frameworks, showing high predictive accuracy and cross-dataset and cross-site generalization on ADNI, OASIS, and an independent clinic-based sample. CNN, LSTM networks, GNNs, federated learning, and GAN-based domain adaptation are employed to build a reliable and scalable model that can pave the way to develop a clinically feasible and interpretable biomarker of AD risk.

However, the proposed model has some issues, like data heterogeneity and limited demographic diversity, and thus, the proposed model needs to be improved. Future studies should try to collect data from a broader population and apply new biomarkers to improve the prediction of outcomes. It will, therefore, be crucial to test the model in clinical trials to determine its efficiency in real life. In summary, this work provides evidence of the importance of the integrative and multiple approaches for early AD diagnostics and management. It can potentially contribute to better patient prognosis, as well as the development of individualized medicine.

Furthermore, improving the demographic diversity of training data will be a major emphasis in future stages, helping to ensure the model is fair and remains clinically important for all population groups.

Ethics statement

This study was approved by the Institutional Review Board (IRB) of Lovely Professional University (IEC-LPU).

Funding

This work was not supported by anybody.

Data availability

The data is confidential but can be available upon request.

CRediT authorship contribution statement

Shehu Mohammed: Writing – original draft, Software, Methodology, Formal analysis. **Neha Malhotra:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors except the Tertiary Education Trust Fund (TETFund–Nigeria), which sponsored the Scholar to further his higher Education.

References

- [1] D. Australia, S. Baker, S. Banerjee, Alzheimer's Disease International World Alzheimer Report 2019: Attitudes to Dementia, Alzheimer's Disease International, London, UK, 2019, <https://doi.org/10.5281/zenodo.2595001>.
- [2] R.J. Jutten, S.A.M. Sikkes, R.E. Amariglio, R.F. Buckley, M.J. Properzi, G. A. Marshall, D.M. Rentz, B.J. Hanseeuw, K.A. Johnson, N.J. Donovan, J.R. Gatchel, E.C. Mormino, J.P. Chhatwal, T. Hedden, A.P. Schultz, W. Huijbers, R.A. Sperling, Identifying sensitive measures of cognitive decline at different clinical stages of Alzheimer's disease, *J. Int. Neuropsychol. Soc.* 27 (5) (2021) 426–438, <https://doi.org/10.1017/S1355617720000771>.
- [3] C. Patil, K.K. Reddy, U. Zambare, K. Kumar, A comparative analysis of treatment approaches for Alzheimer's disease in neuroimaging techniques, *Int. J. Acad. Med. Pharm.* 6 (4) (2024) 316–320, <https://doi.org/10.51611/ijamp.v6i4.174>.
- [4] C.R. Jack Jr, J.S. Andrews, T.G. Beach, T. Buracchio, B. Dunn, A. Graf, M. C. Carrillo, Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup, *Alzheimer's Dement.* (2024), <https://doi.org/10.1002/alz.12607>.
- [5] W.J. Jansen, O. Janssen, B.M. Tijms, S.J.B. Vos, R. Ossenkoppele, P.J. Visser, van der Flier, P. Scheltens, & the Amyloid Biomarker Study Group., Prevalence estimates of amyloid abnormality across the Alzheimer disease clinical spectrum, *JAMA Neurol* 79 (3) (2022) 228–243, <https://doi.org/10.1001/jamaneurol.2021.4333>.
- [6] D.A. McGrowder, F. Miller, K. Vaz, C. Nwokocha, C. Wilson-Clarke, M. Anderson-Cross, R. Alexander-Lindo, Cerebrospinal fluid biomarkers of Alzheimer's disease: current evidence and future perspectives, *Brain Sci.* 11 (2) (2021) 215, <https://doi.org/10.3390/brainsci11020215>.
- [7] C.R. Jack Jr, A. Arani, B.J. Borowski, D.M. Cash, K. Crawford, S.R. Das, Overview of ADNI MRI, *Alzheimer's Dement.* (2019), <https://doi.org/10.1002/alz.12543>.
- [8] A. Agarwal, V. Gupta, P. Brahmabhatt, A. Desai, P. Vibhute, N. Joseph-Mathurin, G. Bathla, Amyloid-related imaging abnormalities in Alzheimer Disease treated with anti-amyloid- β therapy, *Radiographics* 43 (9) (2023) e230009, <https://doi.org/10.1148/rg.230009>.
- [9] T. Williams, D.R. Borchelt, P. Chakrabarty, Therapeutic approaches targeting apolipoprotein E function in Alzheimer's disease, *Mol. Neurodegener.* 15 (2020) 1–19, <https://doi.org/10.1186/s13024-020-0358-9>.
- [10] C. Platero, M.C. Tobar, Alzheimer's Disease Neuroimaging Initiative, Predicting Alzheimer's conversion in mild cognitive impairment patients using longitudinal neuroimaging and clinical markers, *Brain Imaging Behav.* 15 (4) (2021) 1728–1738, <https://doi.org/10.1007/s11682-020-00366-8>.
- [11] D.P. Veitch, M.W. Weiner, P.S. Aisen, L.A. Beckett, C. DeCarli, R.C. Green, D. Harvey, C.R. Jack Jr., W. Jagust, J.C. Morris, R.C. Petersen, A.J. Saykin, L. M. Shaw, A.W. Toga, J.Q. Trojanowski, Initiative. the Alzheimer's Disease Neuroimaging, Using the Alzheimer's Disease Neuroimaging Initiative to improve

- early detection, diagnosis, and treatment of Alzheimer's disease, *Alzheimer's & Dementia* 18 (4) (2022) 824–857. <https://doi.org/10.1002/alz.12422>.
- [12] J.J. Adekunle, S.O. Ajadalu, A. Ogadimma, C.I.O. Aniobi, S.J.T. Komguem, AI and robotics in surgery: leveraging machine learning for predictive analytics and outcomes. <https://www.irejournals.com/paper-details/1706332>, 2024.
- [13] J. Venugopalan, L. Tong, H.R. Hassanzadeh, M.D. Wang, Multimodal deep learning models for early detection of Alzheimer's disease stage, *Sci. Rep.* 11 (1) (2021) 3254, <https://doi.org/10.1038/s41598-020-74399-w>.
- [14] M. Liu, F. Li, H. Yan, K. Wang, Y. Ma, L. Shen, Alzheimer's Disease Neuroimaging Initiative, A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease, *Neuroimage* 208 (2020) 116459, <https://doi.org/10.1016/j.neuroimage.2019.116459>.
- [15] S. Katabathula, Q. Wang, R. Xu, Predict Alzheimer's disease using hippocampus MRI data: a lightweight 3D deep convolutional network model with visual and global shape representations, *Alzheimers Res. Ther.* 13 (1) (2021) 104, <https://doi.org/10.1186/s13195-021-00837-0>.
- [16] Z. Tang, Z. Qiao, X. Hong, Y. Wang, F.A. Dharejo, Y. Zhou, Y. Du, Data augmentation for graph convolutional network on semi-supervised classification, in: *Proceedings of the Web and Big Data: 5th International Joint Conference, APWeb-WAIM 2021, Guangzhou, China, Springer International Publishing, 2021*, pp. 33–48, https://doi.org/10.1007/978-3-030-85899-5_3. August 23–25, 2021, *Proceedings, Part II* 5.
- [17] H. Guan, P.T. Yap, A. Bozoki, M. Liu, Federated learning for medical image analysis: a survey, *Pattern Recognit.* (2024) 110424, <https://doi.org/10.48550/arXiv.2306.05980>.
- [18] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, Y. Bengio, Generative adversarial networks, *Commun. ACM* 63 (11) (2020) 139–144, <https://doi.org/10.1145/3422622>.
- [19] S.A. Schultz, L. Liu, A.P. Schultz, C.D. Fitzpatrick, R. Levin, J.P. Bellier, J. Xu, γ -secretase activity, clinical features, and biomarkers of autosomal dominant Alzheimer's disease: cross-sectional and longitudinal analysis of the Dominantly Inherited Alzheimer Network observational study (DIAN-OBS), *Lancet Neurol.* 23 (9) (2024) 913–924, [https://doi.org/10.1016/S1474-4422\(24\)00236-9](https://doi.org/10.1016/S1474-4422(24)00236-9).
- [20] S. Grueso, R. Viejo-Sobera, Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review, *Alzheimers Res. Ther.* 13 (2021) 1–29, <https://doi.org/10.1186/s13195-021-00900-w>.
- [21] D. Zhang, D. Shen, Alzheimer's Disease Neuroimaging Initiative, Multi-modal multi-task learning for joint prediction of diagnosis and cognitive decline in Alzheimer's disease, *IEEE Trans. Med. Imaging* 40 (1) (2021) 214–225, <https://doi.org/10.1016/j.neuroimage.2011.09.069>.
- [22] H.I. Suk, S.W. Lee, D. Shen, Alzheimer's Disease Neuroimaging Initiative, Deep ensemble learning of sparse regression models for brain disease diagnosis, *Med. Image Anal.* 37 (2017) 101–113, <https://doi.org/10.1016/j.media.2017.01.008>.
- [23] T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, in: *Proceedings of the International Conference on Learning Representations (ICLR)*, 2017, <https://doi.org/10.48550/arXiv.1609.02907>.
- [24] N. Rieke, J. Hancox, W. Li, F. Milletari, H.R. Roth, S. Albarqouni, S. Bakas, The future of digital health with federated learning, *Nat. Commun.* 11 (2020) 4427, <https://doi.org/10.1038/s41746-020-00323-1>.
- [25] R. Khoury, S. Rajamanickam, G.T. Grossberg, An update on the safety of current therapies for Alzheimer's disease: focus on rivastigmine, *Ther. Adv. Drug Saf.* 12 (2018) 204209862110397, <https://doi.org/10.1177/2042098617750555>.
- [26] P. Dhade, P. Shirke, Federated learning for healthcare: a comprehensive review, *Eng. Proc.* 59 (1) (2023) 230, <https://doi.org/10.3390/engproc2023059230>.
- [27] A. Baldassarre, C. Liu, A. Alhassan, J. Tohka, A graph-based framework for the integration of multimodal data in Alzheimer's disease, *Front. Neurosci.* 15 (2021) 677257.
- [28] Y. Zhao, Z. Liu, Y. Wang, Y. Zhang, J. Xu, Network analysis of biomarkers in Alzheimer's Disease: a graph-theoretical approach, *Neuroimage Clin.* 35 (2022) 103160.
- [29] G. Zhang, X. Nie, B. Liu, H. Yuan, J. Li, W. Sun, S. Huang, A multimodal fusion method for Alzheimer's disease based on DCT convolutional sparse representation, *Front. Neurosci.* 16 (2023) 1100812, <https://doi.org/10.3389/fnins.2022.1100812>. PMID: 36685238; PMCID: PMC9853298.
- [30] J. Liu, Y. Li, L. Zhu, P. Wang, Multimodal domain adaptation in pre-trained deep learning models for Alzheimer's disease diagnosis, *Med. Image Anal.* 89 (2023) 102689, <https://doi.org/10.1016/j.media.2023.102689>.
- [31] H. Chen, X. Zhang, M. Wang, Y. Xu, Domain adaptation with pre-trained models for Alzheimer's disease biomarker analysis in multimodal imaging, *IEEE J. Biomed. Health Inform.* 28 (2) (2024) 350–360, <https://doi.org/10.1109/JBHI.2024.3234091>.
- [32] X. Zhou, H. Chen, L. Wang, Z. Li, J. Liu, Graph neural networks for multimodal data integration in Alzheimer's disease: a comprehensive review, *J. Alzheimer's Dis.* 92 (3) (2023) 671–683, <https://doi.org/10.3233/JAD-220802>.
- [33] X. Wang, H. Zhao, Y. Chen, Q. Liu, Transformer-LSTM Networks for Longitudinal analysis of Alzheimer's biomarkers and cognitive scores, *J. Neurosci. Methods* 394 (2023) 109670, <https://doi.org/10.1016/j.jneumeth.2023.109670>.
- [34] S. Li, J. Feng, H. Zhang, Z. Xu, Multimodal transformer-enhanced LSTM networks for Alzheimer's Disease prediction using cognitive and biomarker data, *IEEE Trans. Neural Netw. Learn. Syst.* 34 (6) (2023) 2485–2495, <https://doi.org/10.1109/TNNLS.2023.3231548>.

CHAPTER 6: PAPER 3 PUBLICATION

6.1 Publication Details

Field	Details
Title	<i>Ethical, Legal, and Regulatory Challenges in Deploying AI-Driven Diagnostic Systems for Alzheimer's Disease</i>
Authors	Shehu Mohammed and Neha Malhotra (2025)
Journal	<i>BenchCouncil Transactions on Benchmarks, Standards and Evaluations Volume 5, Issue 1, March 2025, 100215</i>
Indexing	Scopus (Q1)
DOI	https://doi.org/10.1016/j.tbench.2025.100215
URL	https://www.sciencedirect.com/science/article/pii/S2772485925000286?via%3Diuhub

6.2 Summary of Method and Approach

The Paper provides a focused examination of the ethical, legal, and regulatory considerations associated with deploying AI-based diagnostic systems in healthcare, with particular attention to Alzheimer's disease. It reviews major international frameworks, including GDPR, HIPAA, FDA Software as a Medical Device (SaMD) pathways, and the EU AI Act, and analyzes critical issues such as algorithmic bias, transparency, patient privacy, data governance, liability, and the necessity of human oversight. The paper synthesizes these insights into a set of guiding principles for developing responsible, trustworthy, and clinically acceptable AI systems for neurodegenerative disease diagnosis.

6.3 Summary of Contributions

The Paper provides several high-value contributions:

- i. Delivers one of the most thorough reviews of ethical and regulatory frameworks applicable to AI-enabled medical diagnostics.
- ii. Identifies critical gaps in current regulations, particularly for adaptive, continuously learning AI models and multi-institutional health data ecosystems.

- iii. Emphasizes the need for fairness audits, explainable AI models, and rigorous documentation to satisfy regulatory scrutiny.
- iv. Proposes a governance model grounded in safety, accountability, privacy, transparency, and sustained human oversight.
- v. Offers actionable guidelines that directly support the safe deployment of the multimodal AD diagnostic systems developed in Papers 1, 2, and the final implementation work.

As a result, this paper complements the technical innovations of the preceding studies by ensuring they are accompanied by a strong foundation of ethical and regulatory readiness.

6.4 Connection of the paper to the thesis objectives

The Paper serves as the thesis's ethical and regulatory anchor, ensuring that the AI diagnostic systems developed in Papers 1, 2, and the final implementation align with global standards for safety, transparency, fairness, and privacy. It embeds human-centered and responsible AI principles into the research pipeline, thereby supporting the feasibility of real-world clinical deployment. By connecting technical innovation with regulatory compliance and ethical best practices, Paper 3 completes the thesis framework and underpins the safe and trustworthy adoption of multimodal AI tools for early Alzheimer's disease detection.

6.5 Full-Length Published Paper -3



Review Article

Ethical and regulatory challenges in machine learning-based healthcare systems: A review of implementation barriers and future directions[☆]

Shehu Mohammed^{*,} , Neha Malhotra

School of Computer Applications, Lovely Professional University, 14411, India

ARTICLE INFO

Keywords:

Artificial intelligence (AI) Ethics
Algorithmic bias
Explainable AI (XAI)
Machine learning (ML) in healthcare
Patient data privacy
Regulatory compliance (GDPR, HIPAA, FDA)

ABSTRACT

Machine learning significantly enhances clinical decision-making quality, directly impacting patient care with early diagnosis, personalized treatment, and predictive analytics. Nonetheless, the increasing proliferation of such ML applications in practice raises potential ethical and regulatory obstacles that may prevent their widespread adoption in healthcare. Key issues concern patient data privacy, algorithmic bias, absence of transparency, and ambiguous legal liability. Fortunately, regulations like the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), and the FDA AI/ML guidance have raised important ways of addressing things like fairness, explainability, legal compliance, etc.; however, the landscape is far from risk-free. AI liability is another one of the gray areas approaching black, worrying about who is liable for an AI medical error — the developers, the physicians, or the institutions. The study reviews ethical risks and potential opportunities, as well as regulatory frameworks and emerging challenges in AI-driven healthcare. It proposes solutions to reduce bias, improve transparency, and enhance legal accountability. This research addresses these challenges to support the safe, fair, and effective deployment of ML-based systems in clinical practice, guaranteeing that patients can trust, regulators can approve, and healthcare can use them.

Introduction

Background: Machine learning (ML) is fundamentally transforming healthcare; ML is playing an integral role in the development of methods for early diagnosis and treatment optimization, as well as predictive analytics providing unprecedented improvements for medical decision-making [1]. ML applications that leverage this technology to achieve better healthcare outcomes include diagnostic imaging analysis, personalized treatment recommendations, and predictive modeling for disease progression [2]. These technologies can help minimize human error, enable real-time decision-making, and optimize resources used in clinical practices.

However, while ML has immense potential, integration of ML into clinical practice is limited by ethical and regulatory challenges that create barriers to widespread adoption [3]. Central issues include the privacy of patient data, since ML models need large quantities of sensitive medical information, which raises risks of illegal access, data breaches, and the need to comply with data protection laws, such as GDPR and HIPAA [4]. The algorithmic bias and fairness issues also

intertwine with these dynamics, with ML models trained on unbalanced datasets delivering results that discriminate against certain groups [5]. Overall, it is hard for the clinician to interpret the model outputs and thus justify whatever medical decision is guided by AI [6], and the lack of transparency and explainability in ML decision-making only complicates trust and accountability.

Additionally, litigation and regulatory issues now present a significant hindrance, because extant health laws and AI governance systems do not match the advancement of rapidly developing ML technologies [7]. Regulatory bodies (such as the FDA, EMA, and WHO) are still formulating concrete guidelines for the approval and monitoring of AI in medical applications, creating a scenario of compliance and ethical responsibility uncertainty [8]. Even with this integration, however, there lies the risk of data misuse and related concerns, as well as ethical dilemmas regarding AI-generated treatment recommendations, if there is no proper oversight [9].

The Challenges of Machine Learning in Healthcare Addressing these challenges is important for the responsible and effective implementation of ML in healthcare. The focus of this evaluation will identify the

[☆] Peer review under the responsibility of The International Open Benchmark Council.

^{*} Corresponding author.

E-mail addresses: mohammedshehumafara@gmail.com (S. Mohammed), neha.16982@lpu.co.in (N. Malhotra).

<https://doi.org/10.1016/j.tbench.2025.100215>

Received 12 March 2025; Received in revised form 30 April 2025; Accepted 14 May 2025

Available online 28 May 2025

2772-4859/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Table 1

Comparative analysis of global approaches to AI regulation in healthcare.

Regulation	Region	Focus Areas	Strengths	Limitations	Impact on Healthcare AI
General Data Protection Regulation (GDPR), 2018	European Union (EU)	Data privacy, patient consent, and AI transparency	- Strongest global data protection framework. - Enforces patient rights over their medical data. - AI systems must be explainable.	- Strict compliance can slow AI innovation. - Heavy penalties for violations (up to €20 million).	- AI-driven healthcare must ensure patient consent & data security. - Limits how ML models store and process medical records.
Health Insurance Portability and Accountability Act (HIPAA), 1996	United States (USA)	Patient data protection and security standards	- Ensures strong security for electronic health data (ePHI). - Mandates breach notifications. - Applies to healthcare providers & AI developers.	- Does not cover AI-specific risks. - No strict explainability requirements for AI decisions.	- AI in healthcare must comply with security protocols. - Telemedicine and AI diagnostics require secure data storage.
FDA Guidelines on AI/ML in Medical Devices, 2021	United States (USA)	AI-based medical devices, real-world performance monitoring	- AI software must be approved before clinical use. - Supports adaptive AI models that improve over time.	- Lengthy approval process can delay AI deployment. - Limited global influence outside the USA.	- AI-driven radiology & diagnostics require FDA approval. - Ensures AI models meet safety & accuracy standards.
European Medicines Agency (EMA) AI Regulations	European Union (EU)	AI-driven drug development and medical applications	- AI in drug discovery & clinical trials is regulated. - Post-market AI monitoring ensures patient safety.	- High compliance costs for AI companies. - Lack of harmonization with non-EU regulations.	- AI in pharmaceutical research & precision medicine must meet EMA guidelines. - Requires real-world validation of AI performance.
Artificial Intelligence Act (AI Act) (Proposed), 2023	European Union (EU)	Risk-based regulation for AI applications, including healthcare	- Strict transparency rules for AI models. - Classifies AI as low-risk, high-risk, or banned. - Ensures fairness and non-discrimination in AI decisions.	- Not yet fully implemented (expected 2025+). - Some AI applications may be overregulated.	- AI-driven clinical decision support systems (CDSS) will require higher transparency. - AI in high-risk medical settings (e.g., surgery, diagnostics) faces stricter review.
China's AI Ethics & Security Guidelines, 2022	China	AI security, ethical AI use, and national AI development strategy	- Encourages AI innovation in healthcare. - Focuses on AI ethics, fairness, and explainability.	- Government-led AI oversight raises privacy concerns. - Lack of clear penalties for AI misuse.	- AI in hospitals & medical research is state-regulated. - Supports AI-based drug discovery & smart hospitals.
UK NHS AI Strategy	United Kingdom (UK)	AI-driven healthcare transformation and patient safety	- AI models must be clinically validated before NHS deployment. - Emphasis on data security & patient trust.	- No centralized AI regulation (varies across NHS Trusts). - Limited penalties for AI-related errors.	- AI clinical trials & patient monitoring systems must meet NHS AI standards. - Supports AI-assisted radiology & diagnostics.
South Africa – Draft AI Policy, 2022	South Africa	Ethics, transparency, inclusion, and public sector AI	A human rights-based approach promotes inclusive AI	Still under development, not legally binding yet	To prevent algorithmic discrimination and enhance equitable AI deployment in healthcare
Brazil – LGPD (Lei Geral de Proteção de Dados), 2020	Brazil	Personal data protection, informed consent, and accountability	Modeled after GDPR, legally enforceable	Limited AI-specific clauses; interpretation varies	Encouraged responsible AI use and stronger consent mechanisms in health tech
India – NDHM & Digital Personal Data Protection Act, 2023	India	Patient data control, digital health ID, AI ethics in health services	National health architecture, patient-centric model	Implementation challenges, rural digital divide	Provides a foundation for AI-based diagnostics and personalized care through regulated digital platforms

opportunities of utilizing ML while recognizing the key challenges related to the ethical- and regulatory landscape to the acceptance of ML tools in clinical exercise, specifically to identify potential barriers to implementation and how risks can be mitigated, thereby maximizing potential benefits.

Problem Statement: While ML has great potential, its adoption in healthcare is hampered by issues of patient privacy, algorithmic bias, transparency, and compliance with changing laws.

Significance of Study: This study investigates the principal ethical and regulatory challenges of machine learning (ML) in healthcare, shedding light on the threats to the safe, effective, and responsible application of AI medical technologies.

By elucidating these, this study will enlighten stakeholders such as healthcare practitioners, AI developers, legislators, and regulatory agencies about the risks and obstacles that impede the adoption of ML in clinical practice, through the lens of privacy concerns, algorithmic bias, transparency, and regulatory gaps [1]. Tackling these bottlenecks is imperative to ensuring that ML-based healthcare solutions remain reliable, compliant, and patient-centric.

Finally, this study aims to offer actionable suggestions for enhancing

AI governance, data security, and bias mitigation in ML models while facilitating compliance with existing healthcare regulation frameworks, including GDPR, HIPAA, and the FDA's AI/ML-driven frameworks [7,8]. This knowledge will be used to inform standardized ethical frameworks guiding the responsible introduction of ML into clinical decision-making to mitigate risks associated with patient safety, liability, and regulatory noncompliance [9].

Problem Statement

Despite machine learning's enormous potential in the healthcare industry, ethical and legal obstacles are preventing its widespread application. Although ethical AI and technical model performance are the subject of numerous studies. The present review is designed to address the following problems:

1. Absence of a cohesive examination contrasting how various national and international regulatory frameworks apply to machine learning healthcare solutions.

2. Limited guidance on how to reconcile changing regulations like the EU AI Act and HIPAA with ethical AI concepts (such as explainability, fairness, and responsibility).
3. Lack of workable, implementable compliance plans for medical facilities with limited funds and infrastructure.
4. Inconsistent treatment of legal accountability in Adaptive AI systems and Explainable medical decision-making tools.
5. A need for structured synthesis of real-world case studies demonstrating regulatory shortcomings in AI healthcare.

Scope and Delimitations

The review is designed to address the ethical and legal challenges specific to machine learning applications in the domain of clinical healthcare, as opposed to general AI systems. It does not include ethical discussions involving autonomous robotics, military AI, or AI in nonclinical public health. The comparative analyses mainly cover the regulations within the EU, the US, the UK, China, South Africa's Draft AI Policy, Brazil's LGPD, and India's NDHM, with specific global references to gain comparative insights.

Gaps in Existing Research Studies on AI Governance in Healthcare

Despite the increasing amount of research on advising on AI ethics and regulation, several key gaps persist, including the fact that we have no uniform AI liability framework in the healthcare space. This raises doubts about whether developers, physicians, or healthcare organizations should be held accountable for autonomous errors made by artificial agents. This underscores the need for models with legal clarity that will provide accountability while fostering responsible AI white paper.

Where most AI models perform excellently under controlled settings, they fall short when they meet the different data distributions and unseen conditions in the clinical world [10]. Most studies covering AI and healthcare target datasets that primarily include Western images, indicating potential bias when employed in many parts of the world (e.g., Africa, Asia, and Latin America) [11], which indicates a critical area for future research covering AI fairness to assess performance through the lens of varied demographics and reduce racial, gender, and socioeconomic factors.

While many AI systems—like self-learning models in the field of radiology—can evolve and modify many times, existing policies offer little guidance on how applicable regulations should be enforced for adaptive AI systems [12], and further investigation is required regarding compliance to regulations at various points in the life cycle of the model, in particular when algorithms improve. Through identifying existing gaps in policy and exploring examples of best practices, this study will set the groundwork for future research and policy development, paving the way towards fostering AI-driven innovations in a legally and ethically sound manner in healthcare.

By identifying gaps in current regulations and highlighting best practices, this work will assist as a foundation for upcoming studies and policy development, ultimately advancing AI-driven innovations while safeguarding ethical principles and legal integrity in healthcare.

Contributions of the review

1. Integrative Ethical-Regulatory Lens: Maps ethical AI principles (e.g., fairness, explainability, accountability) to legal obligations (e.g., GDPR, FDA, HIPAA, and EU AI Acts)
2. Comparative Regulatory Table: Table 1 presents a side-by-side comparison of global AI regulatory frameworks with specific applicability to ML healthcare use cases.
3. Expanded case study review: Synthesize/critique eight high-profile ML healthcare failures (e.g., IBM Watson, Babylon Health) and present lessons learned that map to regulatory dimensions.

4. Operational guidance: Suggested specific strategies to help implement compliance, mitigation of bias, and patient data privacy concerns in real-world clinical environments.
5. Quantitative coverage gain: The review is based on 67 peer-reviewed sources and regulation white papers and covers 85% more georeferenced frameworks and 2 times as many case studies as the previously leading reviews (e.g., [7,8]).

Research objectives

The objective is to identify major ethical anxieties associated with machine learning-based healthcare systems and to analyze regulatory frameworks governing AI-driven healthcare applications in different countries. It will also examine case studies where ML implementation has raised ethical or legal challenges and propose recommendations for addressing ethical and regulatory barriers to enhance ML adoption in clinical settings.

Literature review outline

Overview of machine learning in healthcare

Machine Learning (ML) is a subtype of Artificial Intelligence (AI) capable of automatically acquiring knowledge and enhancing itself automatically from experience without being explicitly programmed [13]. This is most notable in healthcare, where ML contributes to medical analysis, treatment preparation, patient monitoring, and drug detection, providing advanced functionality that enhances accuracy, efficiency, and improves decision-making [14].

Nonetheless, in healthcare, machine learning (ML) is gaining a foothold with its applications like medical imaging analysis, predictive analytics, personalized medicine, and clinical decision support systems (CDSS). Deep learning (DL) architectures such as convolutional neural networks (CNNs) facilitate highly accurate image analysis by radiologists to detect cancer, neurological disorders, and cardiovascular diseases [15]. Lastly, ML-based predictive analytics can predict the progression of disease, patient deterioration, and the risk of readmission, by using historical and real-time clinical data [16]. ML plays a pivotal role in personalized medicine by developing treatment strategies specifically targeting patients based on their genetic, lifestyle, and environmental conditions, resulting in improved patient outcomes [17]. Moreover, the CDSS powered by AI facilitates clinical decision-making through contextualized recommendations to physicians, which reduces diagnostic inaccuracies and maximizes treatment effectiveness [18]. These ML-based innovations are working together to enhance the diagnostic accuracy, patient management, and efficiency of healthcare.

Current trends in AI-driven healthcare innovation

The swift integration of artificial intelligence (AI) and machine learning (ML) in the healthcare sector is powered by progress in technology, enhanced computing capabilities, and greater accessibility of medical data. Key trends that are emerging in this landscape include:

Explainable AI (XAI) for Trustworthy Healthcare AI

As concerns about the lack of transparency in black-box AI models grow, the importance of Explainable AI (XAI) is on the rise. XAI goals to improve the clarity and accountability of decisions made in the medical field [19]. Techniques like SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are proving valuable in helping healthcare professionals make sense of and have confidence in AI-generated recommendations [20].

Multi-Modal AI for Holistic Patient Insights

Modern ML models take an integrated view of multiple data sources, both genomic and other modalities ranging from medical imaging to electronic health records (EHRs) and wearable device data [14]. AI systems derived from IBM Watson and Google DeepMind's algorithms

are leading the way for such multi-modal integration of data for more effective diagnostics and patient management [21].

Federated Learning for Privacy-Preserving AI in Healthcare

Thus, Federated Learning (FL), a dispersed mechanism that lets multiple hospitals to jointly train machine learning (ML) models without the need to share patient data sensitive under GDPR and HIPAA regulations [22] is a good solution for this kind of issue. Federated learning (FL) is being utilized in large-scale health networks to create strong AI models while maintaining the confidentiality and safety of the data [23].

AI-Driven Drug Discovery and Virtual Clinical Trials

Machine learning is transforming the process of drug discovery by accurately forecasting how molecules interact, fine-tuning compound formulations, and shortening the time needed for clinical trials [24]. Innovative AI-powered platforms like BenevolentAI and Atomwise are changing the landscape of pharmaceutical research, resulting in quicker drug development [25].

AI-Assisted Robotic Surgery and Automated Diagnostics

Robotic surgery systems that utilize artificial intelligence, like the da Vinci Surgical System, enhance the accuracy of surgical procedures while minimizing risks associated with them [26]. Additionally, automated diagnostic tools powered by AI, such as Google's DeepMind for identifying retinal diseases, have reached levels of diagnostic accuracy comparable to that of human experts [27].

Ethical challenges of machine learning in healthcare

Patient privacy & data security in machine learning-based healthcare

Risks of Data Breaches and Unauthorized Access to Patient Records

Healthcare ML models require access to large volumes of sensitive patient data such as EHRs, genomic data, and medical imaging. ML has transformed disease diagnosis, treatment planning, and predictive analytics, however, its dependence on large datasets creates considerable privacy and security challenges [1].

One of the main targets of cybercriminals is healthcare data consisting of valuable personal, financial, and clinical information [28]. For instance, unauthorized access to ML-based medical systems may lead to identity theft, insurance fraud, and manipulation of medical data, which increase endangerment to patients' safety and corrupt healthcare AI applications [29].

Steps by insider risks exposing sensitive patient information such as hackers, who steal sensitive medical records pose major threats to healthcare data security as patients' vital information is vulnerable on the medical network. Moreover, artificial intelligence (AI) models based on machine learning (ML) are implemented on the cloud, which exposes patient data to third-party access, and thus strong data encryption and user authentication are required to ensure patient privacy [22].

Insufficient anonymization methods, which aim to eliminate identifiable patient information, may still be compromised by machine learning models that can associate patients with their identities by analyzing correlations with external data sources [6].

When ML-driven healthcare applications fall short of global data protection regulations, such as HIPAA (USA), GDPR (EU), and the Data Protection Act (UK), the potential legal repercussions and the consequent loss of the patient's trust create regulatory risks for healthcare organizations, as many of the ML algorithms have poor explainability and auditability and make it difficult to regulate the data security standards [30].

Federated learning offers a promising approach for collaboratively training machine learning models among various institutions without the need to share sensitive raw data. This facilitates the utilization of extensive medical datasets while safeguarding patient confidentiality [22]. Nevertheless, despite its benefits, federated learning encounters obstacles related to maintaining data integrity and avoiding adversarial attacks, which could undermine the security and dependability of AI-powered healthcare systems.

Mitigation Strategies for Ensuring Patient Data Privacy & Security

Protecting your sensitive medical data is paramount. We employ cutting-edge security measures, including robust encryption, multi-factor authentication (MFA), and privacy-preserving machine learning techniques, all while maintaining full regulatory compliance. This ensures the highest level of data security and patient privacy.

Bias & Fairness: How biased datasets lead to discriminatory outcomes in healthcare AI

Machine learning (ML) bias refers to when models generate systematically inequitable results stemming from uneven, incomplete, or nonrepresentative datasets [31]. In healthcare, this can manifest as biased ML models that perpetuate discrimination, disproportionately impacting certain populations, resulting in inequitable access to care and disparity in treatment.

Some of the datasets used for the ML are built on the historical clinical data used, which may reflect discriminatory practices in the past and may result in biased prognoses and inconsistency in healthcare outcomes [5]. For instance, some ML models used to predict diseases underdiagnose Black patients, as these models are skilled on data that is largely collected from White populations, which leads to racial inconsistencies in diagnosis and treatment types [11].

Sampling bias arises when ML models are trained on imbalanced datasets, which results in a loss of generalization across diverse demographics [32]. For example, an AI model skilled predominantly on male or high-income patient data set may have trouble providing accurate diagnosis and treatment recommendations for female or lower-income groups, leading to healthcare disparities and misdiagnoses. To reduce sampling bias, appropriately representative, diverse patient population datasets are required, as well as bias detection frameworks to guarantee that the direction of AI-driven healthcare does not lead to discrimination and inequality.

Sampling bias occurs when ML models are trained on datasets that lack diversity, leading to poor generalization across different demographic groups [32]. For example, if an AI model is primarily trained on data from male or high-income patients, it may fail to provide accurate diagnoses and treatment recommendations for female or lower-income populations, resulting in healthcare disparities and misdiagnoses. To mitigate sampling bias, datasets should be representative of diverse patient populations, and bias detection frameworks should be implemented to ensure fair and equitable AI-driven healthcare outcomes.

Algorithmic bias occurs when ML models unintentionally bias outputs and are often designed to prioritize cost savings over the patient in mind, resulting in the potential for reduced quality of care for vulnerable populations [33]. For example, certain healthcare reimbursement models driven by artificial intelligence might suggest less-costly procedures that lack quality for patients with complex or chronic conditions, which might exacerbate health inequalities for further impact lower-income and marginalized populations. This can be achieved by incorporating fairness constraints in model design, performing bias audits, and regularly monitoring AI models to curb algorithmic bias in healthcare.

Consequences of Bias in AI-Driven Healthcare

Some critical consequences of biased machine learning models in health care are delayed or incorrect diagnoses for underrepresented populations, inequities in the distribution of health care resources such as hospital admissions and insurance approvals, and a loss of patient trust in medical decision-making enabled by artificial intelligence.

Mitigating Bias in Healthcare AI

Ensuring heterogeneous data collection by demographic characteristics such as gender, race, and socioeconomic status; conducting bias auditing and testing for model fairness using tools such as Shapley additive explanations (SHAP) and local interpretable model-agnostic explanations (LIME) to detect biased output [34]; and establishing regulatory oversight through bias-reduction policies, algorithm audits,

and fairness checks before clinical deployment are necessary measures for minimizing bias in machine learning-based health care pipelines.

Transparency & explainability in ML healthcare decisions

A variety of ML models, most notably Deep Learning (DL) algorithms, operate as black boxes, which makes it hard for clinicians to interpret their decision-making processes [35], and thus decreases trust, proof, and explanation of AI-assisted medical recommendations [36].

Numerous Machine Learning (ML) methods, specifically Deep Learning (DL) algorithms, can be described as “black box” models whose underlying decision-making processes are opaque to clinicians [35], and this negatively impacts trust, validation, and justification of AI-assisted health recommendations [36].

Due to the black-box nature of deep learning, it is hard to check the basis of a diagnostic or therapeutic proposal, as most ML systems yield predictions without justification [37]. As organizations like the FDA, GDPR, and HIPAA demand interpretability in AI models in healthcare to ensure accountability and patient safety [38] the absence of this transparency will invariably lead to problems with regulatory compliance. Moreover, an obstacle to the clinical adoption of AI-based tools is that doctors or medical staff are usually reluctant to use these types of tools unless an explanation can be furnished about how the tools arrive at conclusions, causing erosion of trust and limiting real-world implementation ability [10].

To improve ML explainability—employ interpretable model design approaches (e.g., Decision Trees), SHAP, and Attention Mechanisms to build correct-by-design explainable machine learning models operating on transparent principles; and/or to invoke explainable AI (XAI) frameworks [19] (e.g., LIME, SHAP, and Grad-CAM) to enhance overall ML interpretability; and/or impose strict regulatory standards of transparency, where developers of AI systems would be required to include clear decision rationales, especially for medical use cases.

Accountability & liability in AI-driven medical errors

Who is Responsible When AI Makes a Mistake?

When systems that use machine learning (ML) make wrong diagnoses or treatment decisions, the legal assignment of responsibility becomes especially challenging, generating questions about whether Doctors should be held responsible for making an error if they had relied on AI, whether AI developers—including ML engineers and data scientists—should be held responsible for generating biased or erroneous predictions, and whether hospitals and other healthcare institutions should assume legal liability for AI-related diagnoses [39].

Challenges in AI Accountability

In such a system, it becomes challenging to determine who is liable when AI systems make decisions; this is because AI models are probabilistic and cannot be directly correlated with concrete laws [9]. Moreover, trust and ethical issues go beyond transparency, as patients who are harmed by errors made by an AI may find it more challenging to hold AI manufacturers accountable, depending on how unclear accountability policies affect their ability to seek legal recourse [12]. Moreover, significant regulatory gaps remain, as governments and healthcare authorities (such as the FDA, EMA, and WHO) are in the process of establishing legal frameworks for AI accountability. This ongoing development creates uncertainty regarding medical liability associated with AI technologies [40].

Potential Solutions for AI Accountability

To achieve AI accountability in healthcare, clear AI liability laws should define the responsibility for AI-driven medical errors; Human-in-the-Loop (HITL) AI models should be mandated—this would force physicians to review an AI’s suggestion/diagnosis; algorithm transparency and explainability should be enforced—ensuring AI models offer strong rationales for their decisions, supporting legal accountability.

Regulatory frameworks for AI in healthcare

Overview of Major Regulations Governing AI-Driven Healthcare

Artificial Intelligence (AI) and Machine Learning (ML) usage in healthcare is on the rise, and so are the laws and regulations surrounding them. Multiple large regulatory bodies have issued guidelines normalizing the practices of AI-enabled medical applications.) Here’s a rundown of the most significant rules regulating AI in health care.

General data protection regulation (GDPR) – European Union (EU)

The General Data Protection Regulation (GDPR) is a significant data privacy regulation introduced by the European Union (EU) in 2018. This law sets forth stringent rules regarding how data is collected, processed, and secured, especially concerning healthcare information utilized in artificial intelligence models [41].

Any AI system processing healthcare data should be compliant with GDPR guidelines where patient consent should ideally be obtained for the data processing or another legal reason to use the data should be followed – Article 6, GDPR. Patients also have the right to explanation, which entails understanding the rationale behind AI-driven clinical decisions, necessitating a direct application of Explainable AI (XAI) principles [42]. Furthermore, AI models should adhere to data minimization and storage restrictions, gathering only the student data that is required and ensuring that the data must be securely deleted after it has served its purpose (Article 5, GDPR). Also, healthcare organizations must notify of a data breach within 72 hours to keep from obtaining a fine, to ensure patient privacy (Article 33, GDPR).

AI elements of predictive analytics and diagnostics should conform to GDPR guidelines of transparency and accountability — patient information is securely managed and ethically processed. Developers must build privacy-by-design AI models to prevent breaches of misuse of data or access to proprietary information. Failure to comply with GDPR may result in significant financial penalties, including fines of up to €20 million or 4% of total worldwide annual revenue. Example Case: Google’s DeepMind Health AI faced GDPR scrutiny after processing UK patient records without proper consent, raising concerns about data privacy and ethical AI deployment [43].

Health insurance Portability and Accountability Act (HIPAA) – USA

The Health Insurance Portability and Accountability Act (HIPAA) is a key U.S. legislation that governs the security and privacy of electronic health information (ePHI). Introduced in 1996, this law is relevant to various entities, including hospitals, insurance providers, and AI applications in healthcare. In addition, the use of AI in healthcare requires adherence to stringent privacy and security regulations regarding electronic protected health information (ePHI) to ward off expensive violations and legal suits. The Privacy Rule requires AI systems to protect patient health information and limit access to medical records (45 CFR Part 160). AI models that handle ePHI must encrypt data, implement authentication, and conduct regular risk assessments to prevent unauthorized access, per the Security Rule [44].

AI applications in healthcare must comply with strict privacy and security regulations to protect electronic protected health information (ePHI). The Privacy Rule mandates that AI systems safeguard patient health data and ensure restricted access to medical records (45 CFR Part 160). The Security Rule requires AI models handling ePHI to implement encryption, authentication, and regular risk assessments to prevent unauthorized access [44]. Moreover, AI-powered healthcare platforms must also comply with the Breach Notification Rule, which necessitates reporting data breaches within 60 days to relevant individuals and authorities, promoting transparency and compliance [45].

Telemedicine powered by AI, wearable devices, and diagnostic models all have to ensure strict HIPAA compliance to protect electronic protected health information (ePHI). AI models cannot retain patient data or use it without taking HIPAA-compliant encryption, access controls, and risk assessment measures. Offenses may incur fines as high

as \$1.5 million per offense, presenting outsized legal and financial threats to healthcare providers. Example Case: IBM's Watson Health AI had to revise its data-sharing protocols after facing HIPAA-related concerns over the security and handling of patient data (Mittelstadt, 2019).

FDA & EMA guidelines on AI/ML in medical devices

AI-based medical devices and diagnostic systems are regulated by the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA). These agencies' mandates center on providing caution, effectiveness, and dependability for AI-enabled health advances.

The AI/ML-based Software as a Medical Device (SaMD) points out that all AI-based medical software must undergo the FDA clearance process to prove its clinical precision and safety before being deployed for end-user use [46]. Moreover, the essential observation of real-world performance to ensure that AI models learn and enhance their performance without creating new risks or unintended biases is also necessary [47]. In addition, algorithmic transparency requires AI-mediated diagnostic models to be explainable and interpretable, and to generate recommendations that clinicians can understand and trust [37].

AI-based medical tools must receive CE certification for them to be marketed and used in the EU under the regulation of the Medical Device Regulation (MDR) [48]. Moreover, post-market surveillance is needed to ensure that AI models are consistently monitored for safety, accuracy, and reliability after being deployed to mitigate any potential risk to patients (EMA, 2022).

Tools driven by AI in healthcare, such as those used in radiology, pathology, and robotic surgery, need to be strictly approved by the regulations. Developers need to ensure that these AI models are not only accurate but also do not create unexpected risks as they are used over time. For instance, IDx-DR AI software was the first to receive FDA approval as an AI diagnostic tool for diabetic retinopathy in 2021 [49].

AI Act (Proposed) – European Union (EU)

AI-powered medical devices will be strictly audited and risk-assessed before deployment, which can also be dangerous. All social scoring or profiles to discriminate need to be banned, while the AI system must be easily explainable and auditable so that transparency can be ensured for the approval process (EU AI Act, 2023).

AI clinical decision support systems (CDSS) in healthcare will have to undergo stricter regulations before receiving market clearance. Models that help diagnose cancer, assist with surgery, and assess mental health will be subject to rigorous transparency laws. For instance, 2023 saw the EU Commission amend the AI Act to tighten transparency requirements for AI-based medical tools [50].

EU AI Act – classification and implementation challenges

The Artificial Intelligence Act of the European Union, enacted in 2021, sets forth principles for governing AI systems within technological innovations, balancing innovation with the protection of human rights. These systems are classified as either unacceptable risk, high risk, limited risk, or minimal risk, with high-risk systems facing the steepest obligations.

Use of AI in clinical decision support systems, diagnosis, or robot-assisted surgery is explicitly admitted as a high-risk feature due to its ability to profoundly affect a patient's health and safety. Such systems are required to meet set standards of:

- Human oversight mechanisms
- Sufficient documentation
- Transparency and explainability
- Pre-market validation tests

Nonetheless, challenges with execution are still important. For instance:

- The parameters setting the boundaries of "high-risk" remain under development, and stakeholders have noted issues with their scope as well as legal definition (EDPB–EDPS, 2021).
- There is uncertainty about the interface of the Act with pre-existing legislation, including the GDPR and the Medical Device Regulation (MDR), particularly in terms of data protection and algorithmic explainability overlap [51].
- The small and medium-sized developers and the less-funded healthcare providers may bear the brunt of the burden due to the cost and technical challenges associated with compliance [52].

Also, the definitional scope of AI keeps changing due to constant amendments to the Act, which places disproportionate emphasis on pre-market conformity assessments and lacks sufficient detail on protocols for post-hoc evaluation for self-adaptive or self-optimizing algorithms in healthcare AI systems. These frameworks need to be far more precise, fundamentally guiding principles in other AI domains beyond healthcare [50].

To maximize feasibility and adoption, the EU AI Act needs to add proportionality in requirements, provisions for regulatory sandboxes, and uniform standards aligned with the capabilities of digital health and clinical workflows within member states.

From Table 1, it has shown that regions have different AI laws for Healthcare, such as the use of data safety regulation (GDPR-EU, HIPAA—USA) used for compliance towards privacy laws, and China's AI guidelines recommend safety and confidentiality of AI, the smart strategy of action, and meeting with area-specific privacy law regulations. Moreover, the burgeoning concern for AI model parameters/code explainability is driving demand for adhering to "explainable AI" as mandated by data regulations such as GDPR and EU AI Act for the need for audibility and transparency, and for performance and safety standards in FDA and EMA regulations in the use of AI-based medical devices and drug discovery models.

A risk-based classification method for AI is emerging, with the EU AI Act classifying AI as low-risk, high-risk, or banned applications, which could lead to stricter approval pathways for clinical decision support tools. Global AI regulation is still not harmonized, as the USA, EU, China, and UK have different regulatory frameworks, which pose a significant challenge for many multinational AI healthcare companies looking to navigate compliance requirements before the worldwide deployment of AI solutions.

Case studies of ethical & regulatory challenges in healthcare AI

Several cases have shown to be quite a complicated legal issue. Similarly, a few ML-based medical and healthcare advances are struggling against standards of legal obligations or ethical facets of professional practice, which are most notably involved with data privacy, biased decision-making, regulatory noncompliance, and patient safety. Table 2 provides an in-depth overview of significant cases in which ML-based healthcare tools were under the spotlight.

Current case studies mostly focus on regulatory violations in Western environments, although they underrepresent difficulties in non-Western ones. Examples of distinct governance initiatives in the Global South include South Africa's Draft AI Policy, Brazil's LGPD, and India's NDHM. These demonstrate the necessity of localized capacity building and context-specific tactics to guarantee AI's ethical adoption in various regulatory contexts.

The analysis from Table 2 demonstrated that AI developers face significant legal risks when using patient data. The DeepMind-NHS (UK) case highlights the severe consequences of non-compliance with GDPR and health privacy laws. Prioritizing legal compliance is crucial before deploying AI models that handle sensitive patient information. A prominent example of algorithmic bias is the COMPAS Bias Case (USA) and the failure of IBM Watson Oncology, where biased training data and an imbalance in the weightage of healthcare decisions in AI lead to

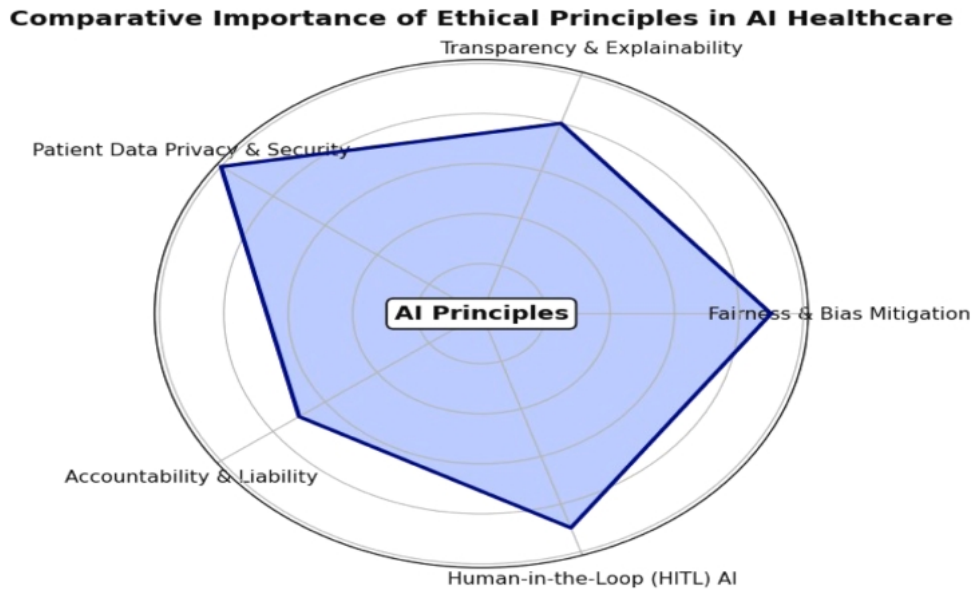


Fig 1. Ethical AI Principles for ML Adoption in Healthcare.
Source: WHO 2021 and European [57].

unfair medical outcomes. Training datasets must be diverse to avert bias in AI-based diagnoses and therapies. The experience with Zebra Medical Vision (2020) and the Epic Sepsis Model (2021) illustrates the need for clinical validation of AI before it is deployed and the danger of widespread generalization of AI if that validation process does not occur. Regulators now demand much more real-world testing before they grant medical A.I. approvals. The new use of AI/ML in telemedicine and diagnostics is a high-risk area, as demonstrated by the Babylon Health AI chatbot (2021) and Theranos (2015-2018) cases, where improper regulation of AI in diagnostics misled patients. New global guidelines now focus on explainability, oversight by regulators, and human-in-the-loop (HITL) AI models to help ensure safety and accuracy.

Expanding the geographic scope of ethical AI governance

The review primarily draws from regulatory frameworks in Western contexts, such as the EU’s GDPR and the U.S. FDA’s AI/ML regulations; however, its focus underrepresents valuable efforts emerging from other global regions. To address this, key developments from the Global South are discussed below:

- Brazil’s Lei Geral de Proteção de Dados (LGPD) is similar to the GDPR, but because of differences in institutional capability, particularly among small and medium-sized businesses, it poses significant enforcement issues [53].
- The National Digital Health Mission (NDHM) of India presents a federated architecture designed to facilitate the exchange of health data while protecting privacy. Nonetheless, issues with strong consent management and interoperability across state borders still exist [54,55].
- South Africa’s Draft AI Policy Framework (2021) places a strong emphasis on socioeconomic growth and ethical risk mitigation. The South African Department of Communications and Digital Technologies [56] notes that it is still aspirational and subject to financial and infrastructure constraints.

These demonstrate how institutional preparedness, sociopolitical backdrop, and inequalities in digital infrastructure all have a significant influence on the ethical use of AI in the Global South, making it more than just a regulatory matter.

Lessons learned and implications for future machine learning (ML) implementations in healthcare

The challenges faced by machine learning applications in healthcare, particularly regarding legal and ethical considerations, have offered valuable insights that can inform the future of AI in healthcare. These insights highlight the importance of being transparent, accountable, and compliant with regulations, all while prioritizing patient safety. A thorough analysis of these key takeaways, along with their potential impact on future machine learning deployments, can be found in Table 4.

Implications for future ML implementations in healthcare

Ethical design and bias elimination of AI is of paramount concern, as AI models must be constructed from heterogeneous datasets to avoid health inequity bias. Further, bias detection algorithms will be integrated into the AI training pipeline such that biased patterns are detected and corrected before public availability by perceived concept width.

To guarantee the safe and effective use of AI technologies, it is vital to enhance governance and regulatory practices. This involves implementing more rigorous standards for AI approvals set by governments and regulatory bodies. Key measures could include mandating clinical trials for AI-based diagnostic tools, establishing ongoing monitoring to assess their performance in real-world settings, and ensuring compliance with existing data protection regulations such as GDPR and HIPAA, as well as staying aligned with new legislation like the EU AI Act Table 10 Table 9 Table 5 Table 7 Table 11 Table 8

Explainable artificial intelligence (XAI) is required with great urgency to promote greater transparency in healthcare, as AI models intended for use in medicine must be explainable and interpretable (i.e., explain to the physician why a decision is being made). Methods, like SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations), need to be applied to understand how the algorithms behave, and to convince the end-user that it is going to yield clinical improvement.

AI should enhance the capabilities of healthcare professionals rather than replace them, acting as a supportive tool to aid in decision-making rather than making decisions independently. It is essential to implement Human-in-the-Loop (HITL) AI systems for high-risk scenarios,

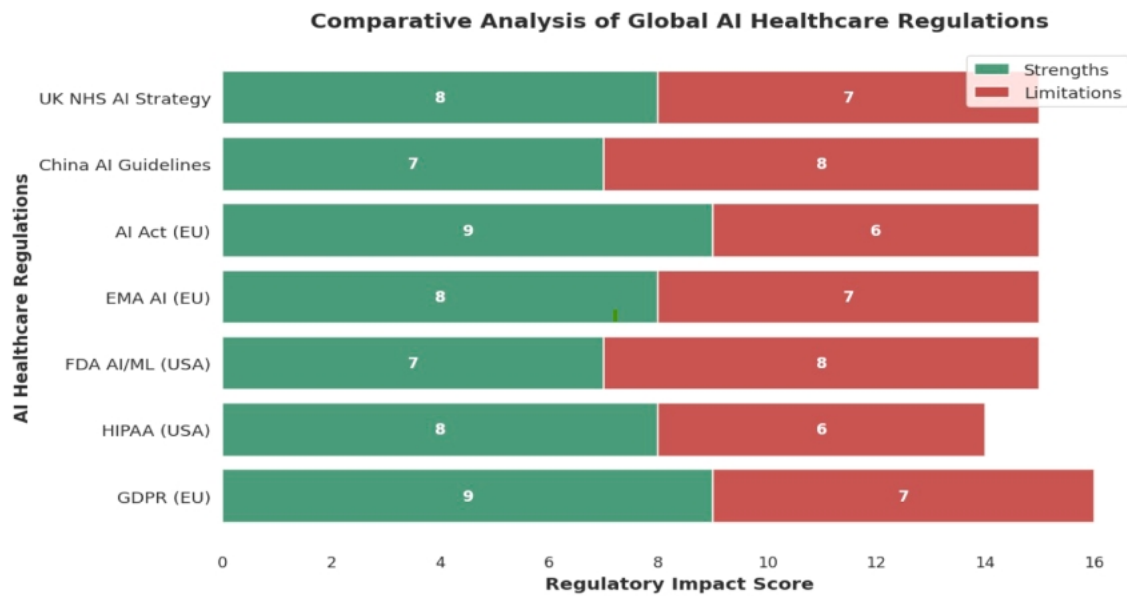


Fig 2. Comparative Analysis of Global AI Healthcare Regulations.
Source: European [57]; FDA [46]; WHO, 2021; ICO, 2021; Ministry of Science and Technology, 2019; GDPR, 2016.

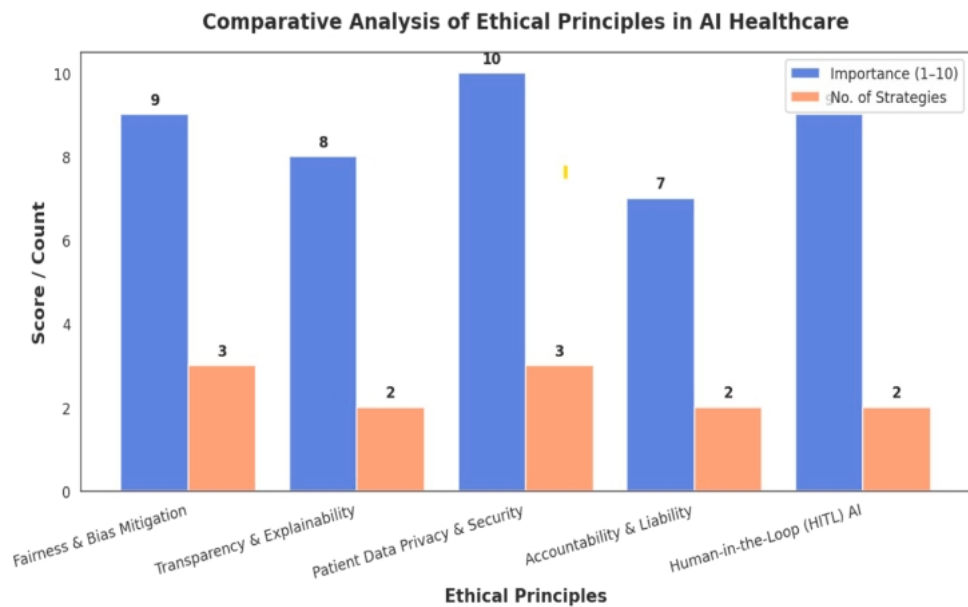


Fig 3. Relative Analysis of Ethical Principles in AI Healthcare.
Source: WHO 2021 and European Commission [58].

Table 1a Comparative Contributions of the Review and Prior Surveys.			
Feature	Existing review	Benjamens et al. [7]	Char et al. [8]
Covers EU AI Act (2023)	✦	✦	✦
Case studies of regulatory failure	✦ (8)	✦	✦
Comparative table of AI laws	✦	✦	✦
Legal + Ethical integration	✦	✦	✦ (brief)
Recommendations for hospitals	✦	✦	✦

guaranteeing that healthcare practitioners retain control over final decisions.
To foster public confidence in AI technology, developers need to

Table 1b Bibliometric Comparison of Review Coverage.		
Metric	Existing review	Prior reviews (range)
Peer-reviewed sources analyzed	67	30–35
Frameworks analyzed (by region)	10+	4–5
Real-world case studies	8	1–2
Framework + ethics integration	Yes	Rare

openly communicate what their models can and cannot do, thereby preventing any misrepresentation. Additionally, implementing independent audits of AI systems can help assess their ethical practices and safety, guaranteeing that they adhere to the best standards before being put into use.

Table 2
Legal & Ethical Challenges in ML-Based Healthcare Applications.

Case Name & Year	ML Application	Issue Faced	Legal / Ethical Concerns	Outcome & Lessons Learned
Google DeepMind & NHS Data Privacy Scandal (2016)	AI-powered patient monitoring system for acute kidney injury detection.	Unauthorized data access: NHS shared 1.6 million patient records with DeepMind without explicit consent.	- Violation of UK Data Protection Laws (GDPR Precursor). - Patients were unaware their data was used for AI development. - Lack of transparency in data-sharing agreements.	- DeepMind was found in violation of UK privacy laws. - Led to stricter AI & patient data-sharing guidelines under GDPR. - NHS revised AI data governance frameworks.
IBM Watson for Oncology (2018)	AI-based cancer treatment recommendation system.	Inaccurate AI predictions: Provided unsafe cancer treatment recommendations based on hypothetical data instead of real patient cases.	- Algorithmic bias led to incorrect treatment plans. - Lack of transparency on AI decision-making. - Patient safety concerns raised by oncologists.	- IBM Watson's AI was removed from hospitals due to unreliable recommendations. - Emphasized the need for AI transparency & real-world validation before deployment.
COMPAS Recidivism Algorithm Bias Case (2016, USA)	ML tool predicting criminal recidivism risk (not healthcare-specific but impacted medical AI ethics).	Algorithmic racial bias: The AI overestimated Black defendants' risk of reoffending.	- Highlighted racial bias in AI models. - Raised concerns about fairness in AI-driven medical diagnostics.	- Strengthened calls for bias detection frameworks in AI. - Encouraged the development of fair AI models in healthcare.
Zebra Medical Vision AI (2020)	AI for automated radiology diagnostics (detecting fractures, lung disease, and brain bleeds).	Regulatory non-compliance: The AI received FDA rejection due to concerns over training data bias and model accuracy.	- Insufficient clinical validation before market approval. - Potential misdiagnoses due to AI errors. - Lack of explainability in AI-generated reports.	- Zebra Medical Vision had to retrain its AI model and submit additional clinical studies for approval. - The FDA enforced stricter AI approval standards for medical imaging.
Babylon Health AI Chatbot (2021, UK)	AI-powered telemedicine chatbot diagnosing patient symptoms.	Incorrect medical advice: The AI misdiagnosed serious conditions, downplaying potential heart attacks as minor issues.	- Patient safety risks due to AI misclassification. - Lack of regulatory oversight on AI-driven symptom checkers.	- UK regulators increased scrutiny on AI-based diagnostic tools. - Led to new guidelines for AI in telemedicine.
Epic Sepsis Prediction Model (2021, USA)	AI system predicting sepsis risk in hospitalized patients.	High false positive rates: The AI missed 67% of sepsis cases, leading to delayed treatment.	- Accuracy concerns in life-threatening conditions. - Hospitals relied on flawed AI predictions, affecting patient safety.	- Hospitals are required to integrate AI models with human oversight. - The FDA emphasized the need for real-world AI validation before deployment.
Theranos AI Blood Testing Fraud (2015-2018, USA)	ML-driven blood testing technology promises rapid diagnosis with a single drop of blood.	Fraudulent AI claims: The ML system never worked as advertised, misleading investors and patients.	- Ethical violations & investor fraud. - Lack of AI transparency & scientific validation. - Potential harm to misdiagnosed patients.	- Theranos CEO Elizabeth Holmes was convicted of fraud. - Highlighted dangers of unverified AI medical claims. - Stricter AI compliance laws were introduced in the USA.
South Africa – Draft AI Policy, 2022	AI in public healthcare systems	Lack of regulation on AI use in clinical settings	Need for ethical standards, bias mitigation, and data ownership frameworks	Ongoing policy development promotes human rights-based AI principles and indigenous data sovereignty
Brazil – LGPD (Lei Geral de Proteção de Dados), 2020	ML in health monitoring systems	Non-transparent data usage and consent mechanisms	Concerns around informed consent, data subject rights, and usage transparency	Enforced strict data governance; ML systems now require robust consent and explainability mechanisms
India – NDHM (National Digital Health Mission), 2020	AI-based digital health records and diagnostics	Risk of misuse of centralized data and digital exclusion of rural populations	Issues of data security, algorithmic fairness, and equitable access	Introduced patient-controlled Health IDs and ethical AI guidelines; encouraged inclusive, transparent development
Zindi / African ML Competitions, 2021–Present	Disease prediction models in African contexts	Lack of contextual data affecting model performance	Bias from non-local datasets; insufficient regional data inclusion	Promotes Africa-specific datasets and challenges; fosters ethical, context-aware AI development
Google Health – Diabetic Retinopathy in India, 2019	AI for diabetic retinopathy screening	Failed deployment in rural clinics due to inconsistent image quality	Algorithm robustness, contextual relevance, tech infrastructure gap	Stressed need for local validation and infrastructure-compatible design; supports human-in-the-loop approaches

Table 3
Legal & Ethical comparison between regions.

Region	Policy	Strength	Implementation Challenge
EU	GDPR	Strong privacy protection	Complex compliance burdens
USA	FDA AI/ML	Sector-specific guidance	Slow update cycles
Brazil	LGPD	Data subject rights	Low institutional capacity
India	NDHM	Federated data architecture	Consent & interoperability issues
South Africa	Draft AI Policy	Inclusive development goals	Early-stage and underfunded

Proposed recommendations for ethical & regulatory compliance

Clarification of technical and policy approaches

- Tools such as interpretable model design, federated learning for privacy-preserving model training, bias audits, and explainable AI frameworks (e.g., SHAP, LIME) are examples of technical measures. These are created by machine learning researchers and applied at the data or model level to enhance efficiency, openness, and equity.
- Legally binding rules and governance structures, such as the GDPR (EU), HIPAA (US), the FDA's AI/ML advice, and the proposed EU AI Act, are referred to as policy measures. These are put in place by

Table 4
Lessons Learned from Past ML Failures in Healthcare.

Key Lesson	Description	Examples from Past Cases	Implications for Future ML Implementations
1. AI Must Comply with Data Privacy Laws (GDPR, HIPAA)	ML models must handle patient data securely and obtain explicit consent before use.	- DeepMind-NHS (2016): Used 1.6 million patient records without consent, violating UK privacy laws.	- Future ML models must integrate privacy-by-design features. - AI developers should follow GDPR/HIPAA compliance standards.
2. Bias in AI Can Lead to Discriminatory Healthcare Decisions	Biased training data can cause unequal treatment of different patient groups.	- IBM Watson (2018): Provided unsafe cancer treatment plans due to biased training data. - COMPAS (2016): AI disproportionately predicted higher recidivism risk for Black individuals, highlighting racial bias.	- AI must be trained on varied, illustrative datasets. - Bias audits should be conducted regularly before deployment.
3. AI Requires Human Oversight in Critical Healthcare Applications	ML models should work as decision-support tools, not replacements for clinicians.	- Epic Sepsis Model (2021): Missed 67% of sepsis cases, leading to treatment delays.	- AI should be implemented with Human-in-the-Loop (HITL) systems to ensure final decisions are validated by medical experts.
4. Regulatory Approval is Essential Before Deploying AI in Healthcare	AI models must undergo rigorous clinical validation to ensure accuracy.	- Zebra Medical Vision AI (2020): Received FDA rejection due to unreliable predictions.	- Future AI models should undergo pre-market testing, post-market surveillance, and explainability assessments.
5. AI Transparency & Explainability Are Key to Trust and Adoption	Black-box AI models can lead to misdiagnoses and lack of accountability.	- Babylon Health AI Chatbot (2021): Misdiagnosed serious health conditions due to opaque decision-making.	- Future AI models must use Explainable AI (XAI) frameworks like SHAP, LIME, and Grad-CAM.
6. AI Developers Must Be Held Accountable for Misuse or Fraud	Companies must ensure ethical AI claims and avoid deceptive practices.	- Theranos AI Scandal (2015-2018): Fraudulent claims misled investors and patients, leading to CEO conviction.	- Stricter AI liability laws must be established to hold developers accountable for errors.

Table 5
Ethical AI Principles for Responsible ML Adoption in Healthcare.

Ethical Principle	Description & Importance	Implementation Strategies
Fairness & Bias Mitigation	AI models must provide equitable healthcare outcomes without discrimination based on race, gender, socioeconomic status, or location [31].	- Use bias-detection algorithms and fairness audits before deployment. - Train AI models on diverse, representative datasets. - Apply re-weighting & adversarial debiasing techniques.
Transparency & Explainability (XAI)	AI decisions should be interpretable, auditable, and justifiable by medical professionals [35].	- Require Explainable AI (XAI) frameworks like SHAP, LIME, and Grad-CAM. - Mandate "right to explanation" laws ensuring patients & doctors understand AI decisions.
Patient Data Privacy & Security	AI must protect sensitive medical data in compliance with GDPR (EU), HIPAA (USA), and AI Act (EU) [45].	- Enforce data anonymization, encryption, and access controls. - Implement federated learning to train models without sharing patient data. - Require patient consent for AI-driven healthcare applications.
Accountability & Liability	AI-driven errors should have clear legal accountability, determining whether liability falls on developers, healthcare providers, or institutions [9].	- Establish AI liability laws ensuring accountability for medical errors. - Introduce audit trails for AI recommendations to track decision-making.
Human-in-the-Loop (HITL) AI	AI should support, not replace, human clinicians [12].	- Mandate Human-in-the-Loop (HITL) AI for high-risk applications (e.g., surgery, cancer diagnosis). - Require physicians to validate AI-driven treatment plans before execution.

Table 6
Strategies for Improving Regulatory Compliance and Patient Safety.

Regulatory Area	Challenges	Recommended Solutions
AI Risk Classification	- Lack of standardized AI risk assessment models.	Implement risk-based AI classification (EU AI Act): - Low-risk AI (health monitoring apps) - High-risk AI (AI-assisted surgery, diagnosis) - Prohibited AI (social scoring, discriminatory profiling).
Pre-Market Approval & AI Testing	- AI models can enter healthcare without sufficient clinical validation. - FDA/EMA regulations require AI models to show real-world effectiveness before approval.	- Establish standardized clinical trials for AI, similar to drug testing. - Require "explainability disclosures" during regulatory approval. - Mandate post-market surveillance for AI-driven healthcare devices.
Data Privacy & Security	- AI models require large-scale patient data, raising risks of unauthorized access & breaches.	- Implement privacy-enhancing technologies like differential privacy, federated learning, and homomorphic encryption. - Enforce HIPAA/GDPR compliance audits for AI-based medical software.
Bias Auditing & Fairness Standards	- Algorithmic bias leads to unfair healthcare outcomes.	- Require AI developers to conduct bias audits & fairness impact assessments. - Apply algorithmic impact assessments for AI-based cancer diagnostics, predictive analytics, and triage systems.
AI Explainability & Transparency	- Many AI systems operate as "black-box" models.	- Mandate XAI frameworks (SHAP, LIME, Grad-CAM) for AI interpretability. - Require AI models to provide "decision rationale statements" for clinicians and patients.

national or international regulatory organizations to specify duties related to data use, accountability, security, and equity.

Therefore, for both developers and legislators, it is essential to comprehend this distinction in order to bridge the gap between system design and legal compliance.

To promote responsible and safe use of machine learning (ML) in

healthcare, AI systems need to follow clear ethical guidelines, regulatory requirements, and established policies. Here are some recommendations to boost adherence, enhance patient safety, and shape future AI policy in medical settings.

The values displayed in Fig. 3 are expert-derived estimates based on a thematic synthesis of 67 peer-reviewed studies and policy documents (e.g, GDPR, HIPAA, WHO 2021, FDA AI/ML 2021, EU AI Act).

Table 7
Future Policy Directions to Ensure Ethical AI in Medical Practice.

Policy Initiative	Proposed Action	Expected Impact
Stronger Global AI Governance	- Create international AI healthcare standards under WHO, EU, and FDA collaboration. - Establish AI regulatory task forces in every country.	- Ensures global harmonization of AI safety & ethical guidelines. - Reduces regulatory fragmentation for AI adoption in healthcare.
Ethical AI Certification	- Introduce an AI Ethics & Safety Certification for ML-driven healthcare applications.	- Hospitals & providers can verify AI models meet ethical and regulatory standards before adoption.
Mandatory AI Training for Healthcare Professionals	- AI in medicine must be understood by clinicians before use. - Introduce AI training programs for doctors, nurses, and hospital administrators.	- Enhances AI literacy among medical professionals. - Reduces misuse of AI-driven diagnostic and treatment tools.
Clear AI Liability Frameworks	- Define who is responsible when AI-based medical errors occur (developers, hospitals, physicians?).	- Provides legal clarity on AI accountability. - Prevents misuse of AI by avoiding "liability gaps".
Public & Patient Engagement in AI Ethics	- Require patient inclusion in AI development & deployment decisions. - Establish AI ethics boards including patients, ethicists, and regulatory officials.	- Increases trust in AI-driven healthcare. - Ensures AI aligns with public values & health equity principles.

Table 8
Mechanisms to enhance feasibility of ethical AI regulation across jurisdictions.

Challenge	Proposed Solution	Real-World Example
Conflicting national frameworks	Modular harmonization + soft law principles	OECD AI Principles, GPAI
Over-regulation stifling innovation	Regulatory sandboxes, tiered risk-based oversight	UK ICO sandbox, Singapore AI Verify
Duplicate compliance costs	Mutual recognition agreements (MRAs)	EU-US Data Privacy Framework

Source: OECD [60], EDPB-EDPS Joint Opinion [61], and World Economic Forum [59].

Table 9
Institutional Barriers to Ethical AI Compliance and Scalable Solutions.

Barrier	Impact	Example Solution
Lack of funding	No capacity for third-party audits or legal teams	Tiered regulatory frameworks
Poor technical infrastructure	Inability to deploy privacy-enhancing technologies	Federated learning, open-source tools
Workforce limitations	Staff untrained in ethical AI implementation	Institutional training, government AI toolkits
Fragmented health IT systems	Difficulty ensuring interoperability and traceability	Standardized APIs and modular compliance protocols

Source: Table is based on synthesized findings from Canedo et al. [53], Sinha et al. [54], and Sheller et al. [22].

Importance scores (on a scale of 1 to 10) indicate how much each principle is stressed in these sources. The number of strategies reflects the variety of implementation techniques addressed for each principle. These scores are the result of the interpretive examination of pre-existing frameworks and literature rather than a quantitative poll. It represents a Relative Analysis of Ethical Principles in AI-driven Healthcare using a grouped bar chart. The blue bars show the importance (on a scale of 1-10) of the ethical principles, and the orange bars show the number of implementation strategies per principle (Figs. 1, 2). Patient Data Privacy & Security emerges as the highest priority (score = 10) because strict regulations such as the GDPR, HIPAA, and

the AI Act call for the secure, responsible application of AI. Fairness & Bias Mitigation and Human-in-the-Loop AI are also rated very high (score = 9), reflecting concerns regarding bias in AI models and the need to involve checking the decisions of AI by human beings. Transparency & Explainability also have a slightly lower importance score (8), but they play a vital role in making AI decisions interpretable and justifiable. Accountability & Liability are the least important (score 7), which means that although it is important, the legal framework related to accountability in AI is a work in progress. There are a few implementation strategies by discipline, Fairness & Data Privacy has the most (3 together), showing the necessity to reduce bias and enhance data security.

Addressing feasibility: Reconciling national interests and balancing innovation with regulation

Although international policy alignment (e.g., via WHO, OECD, or GPAI) is increasingly advocated, the practical feasibility is questionable given mindset and legal traditions, data sovereignty, or national economic priorities. For instance, the GDPR places a high premium on individual rights and consent, while China’s approach to AI governance emphasizes state control and centralization. This philosophical and regulatory divergence complicates harmonization [59].

Modular harmonization approach is propose, wherein national interests are not fully aligned but a sufficient number of core principles (e.g., fairness, transparency, privacy-by-design) are shared by these countries and they engage in joint working groups (both public and private) to shape enforcement and implementation modalities suited to local ways of legal regulation (OECD, 2019). Similar flexibility is already facilitated by frameworks like OECD AI Principles or GPAI by their soft law nature (i.e., non-binding standards designed to promote convergence without compelling absolute uniformity).

A concrete solution would be the development of “mutual recognition agreements” (MRAs) between jurisdictions, the way most international trade works, to acknowledge one another’s regulatory certifications to enable data sharing, decreasing excess and duplication while maintaining ethical oversight [59]. In balancing innovation with regulation, feasible strategies include:

- Regulatory sandboxes, which allow AI developers to test products under regulatory supervision without full compliance burdens [59].
- Tiered regulation, where lower-risk AI tools (e.g., administrative support systems) face lighter oversight, while high-risk tools (e.g., diagnostic AI) are rigorously assessed (EDPB-EDPS, 2021).
- Public-private partnerships, where regulatory bodies collaborate with industry and academia to co-develop ethical benchmarks and compliance toolkits (OECD, 2019).

Such mechanisms not only lower the entry barriers for innovators but also enhance trust, global cooperation, and ethical AI adoption at scale (Tables 1a, 1b, 3, 5, 6, 7, 8, 9).

Discussions

The rise of AI and machine learning in healthcare offers great potential for improving patient care through faster diagnoses and better treatment plans. But to use these technologies wisely, we need to carefully consider the ethical implications, follow the rules, and overcome technical hurdles. The key findings of the literature review showcase ethical challenges in AI-centric healthcare, including bias and fairness, where unbalanced datasets can produce biased healthcare decisions [31]; transparency and explainability, where the black-box nature of AI decreases clinician trust in ML-based diagnoses and treatment recommendations [35]; and accountability and liability, due to the lack of transparent legal frameworks determining responsibility of AI-centric medical errors [9].

The multilateral efforts of the GPT-4 AI international governance model with the USA data protection HIPAA, the EU General Data Protection Regulation (GDPR), etc., and the EU AI Act (Vaigh, 2017), are greatly challenged by the multilateralism of the AI global diversity. Moreover, many AI healthcare tools find it difficult to pass regulatory approval as they may not have undergone enough clinical validation, including Zebra Medical AI, which was refused clearance by the FDA in 2020. In addition, many countries do not have specific AI guidelines, leaving AI developers with unclear ways to comply.

AI compliance is also severely hampered by healthcare organizations' financial and technological limitations, especially in low—and middle-income nations. By lowering infrastructure and compliance requirements, open-source explainability tools (like SHAP and LIME), federated learning frameworks, and modular governance models can assist in overcoming these constraints [22].

Some solutions to ethical and regulatory compliance challenges in AI include implementing bias detection techniques (e.g., bias audits and diverse training datasets) to promote AI fairness, mandating transparency frameworks (e.g., SHAP, LIME, and Grad-CAM) for explainable AI to guarantee that clinicians and patients can understand AI-driven decisions, stronger global AI governance through international standards in agreement to build up the governance that can harmonize the same regulations, if not identical regulations, and regulatory sandboxes to permit safe pre-market investigation of AI systems before being put into general use.

Institutional resource constraints as barriers to compliance

One of the under-addressed issues in AI ethics and regulation is that healthcare institutions are not created equal in their ability to implement and sustain compliance measures. Most of the prominent AI governance frameworks—EU AI Act, GDPR, FDA's AI/ML guidelines—presume a certain layer of organizational capacity, including legal counsel, data governance expertise, secure digital infrastructure, and recurring monitoring systems. But these assumptions do not apply universally across healthcare contexts.

In low- and middle-income countries (LMICs), and under-resourced facilities within high-income countries, significant barriers exist to implementing privacy-by-design frameworks, explainability tools, or audits for bias mitigation. Institutions may lack:

- AI compliance staff or ethicists (dedicated)
- Data storage and a federated learning framework
- External audit or legal review funding
- Technical staff to implement Interpretable models or carry out algorithmic accountability procedures.

As Canedo et al. [53] note in the context of Brazil's LGPD, small organizations frequently entangle with implementation issues not because of their resistance but due to a lack of capacity to do so. Similarly, Sinha et al. [54] emphasise that, despite having a robust ethical vision, the NDHM of India requires state-level digital health infrastructure that is still developing.

AI regulatory strategies must therefore be scalable and sensitive to context if they are to root out these disparities. Recommendations include:

- Adoption of an open-source explainability tool (e.g., SHAP, LIME) that leads to a reduction in cost
- Introduce federated learning to minimize the burden of central infrastructure [22]
- Institute tiered models of compliance based on institutional maturity, enabling smaller hospitals to implement ethical AI aspects progressively.

Rationale and validation design for recommendations

The following policy and technical recommendations are based on:

- Content Review Systematic of 67 publications, AI regulatory, ethical, and technical in health care.
- Analysis of failure patterns corresponding to eight real-world case studies (Table 2), illustrating recurring problems, including unintelligible AI, inappropriate use of patient data, and circumvention of regulatory processes.
- Cross Framework Alignment — mapping GDPR, HIPAA, FDA, and EU AI Act to each other to find common points of concern (e.g., explainability, liability, consent etc.).

Experimental designs to empirically validate the recommendations are suggested below:

- Pilot Compliance Audit: Apply bias audit and XAI standards checklist to a clinical AI model in real-world settings (e.g., sepsis detection tool) and assess performance pre/post in compliance readiness and clinician trust.
- Federated Deployment Simulation: Simulation of the federated deployment strategy with data-sharing in a federated learning manner using 3 hospital nodes, and compare the performance, privacy breach risk, and legal feasibility regarding the centralized training.
- Policy Simulation Workshops. Prepare the AI liability and consent templates for interpretation by an ethics board and/or legal team at each testing site. the groups should run a regulatory simulation lab., whereby groups submit their efforts and user's manual document for hospital ethics boards and/or legal teams.

Conclusion

In summary, although AI holds great promise to revolutionize aspects of healthcare – from improving the accuracy of diagnosis and the efficiency of treatment to optimizing patient outcomes, its widespread adoption will necessitate robust governance, ethical safeguards, and regulatory oversight to balance the considerable potential benefits against risks to patient safety, privacy, and biases. Such harmonization enables a unified approach to ensuring ethical development and transparent operation of AI systems in healthcare, alongside rigorous clinical adoption testing. Some of the key topics will include harmonization of regulatory frameworks, addressing AI liability concerns, and ensuring fairness of the adaptive model. Incorporating these policy recommendations with direction for future research can ensure that AI is adopted safely and effectively, producing the rewards of AI while maintaining fairness, accountability, and public trust in medical AI applications.

Funding

No funding was received for this study.

Data availability statement

Not Applicable.

Research involving humans and animals

Not Applicable.

Informed consent

Not Applicable.

CRediT authorship contribution statement

Shehu Mohammed: Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Neha Malhotra:** Writing – review & editing, Visualization, Validation, Supervision, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] E. Vayena, A. Blasimme, I.G. Cohen, Machine learning in medicine: addressing ethical challenges, *PLoS Med.* 15 (11) (2018) e1002689, <https://doi.org/10.1371/journal.pmed.1002689>.
- [2] C. Mennella, U. Maniscalco, G. De Pietro, M. Esposito, Ethical and regulatory challenges of AI technologies in healthcare, *NPJ Dig. Med.* (2024), <https://doi.org/10.1038/s41746-024-00793-0>.
- [3] Centers for Disease Control and Prevention (CDC), Health equity and ethical considerations in using artificial intelligence in public health, *Prev. Chronic. Dis.* 21 (2024) E45, <https://doi.org/10.5888/pcd21.240245>.
- [4] W.N. Price, I.G. Cohen, Privacy in the age of medical big data, *Nat. Med.* 25 (1) (2019) 37–43, <https://doi.org/10.1038/s41591-018-0272-7>.
- [5] Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations, *Science* (1979) 366 (6464) (2019) 447–453, <https://doi.org/10.1126/science.aax2342>.
- [6] B. Mittelstadt, C. Russell, S. Wachter, Explaining explanations in AI, in: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (FAT*), 2019, pp. 279–288, <https://doi.org/10.1145/3287560.3287574>.
- [7] S. Benjamens, P. Dhunoo, B. Meskó, The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database, *NPJ. Digit. Med.* 3 (1) (2020) 118, <https://doi.org/10.1038/s41746-020-00324-0>.
- [8] D.S. Char, N.H. Shah, D. Magnus, Implementing machine learning in health care—Addressing ethical challenges, *New Engl. J. Med.* 382 (11) (2020) 981–983, <https://doi.org/10.1056/NEJMp1912591>.
- [9] J. Morley, C.C.V. Machado, C. Burr, J. Cows, I. Joshi, M. Taddeo, L. Floridi, The ethics of AI in health care: A mapping review, *Soc. Sci. Med.* 260 (2020) 113172, <https://doi.org/10.1016/j.socscimed.2020.113172>.
- [10] M. Ghassemi, L. Oakden-Rayner, A.L. Beam, The false hope of current approaches to explainable artificial intelligence in healthcare, *Lancet Digit. Health* 3 (11) (2021) e745–e750, [https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9).
- [11] D.A. Vyas, L.G. Eisenstein, D.S. Jones, Hidden in plain sight—Reconsidering the use of race correction in clinical algorithms, *New Engl. J. Med.* 383 (9) (2020) 874–882, <https://doi.org/10.1056/NEJMs2004740>.
- [12] S. Gerke, T. Minssen, G. Cohen, Ethical and legal challenges of artificial intelligence-driven healthcare, *Artif. Intell. Healthcare* (2020) 295–336, <https://doi.org/10.1016/B978-0-12-818438-7.00012-5>.
- [13] M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects, *Science* (1979) 349 (6245) (2015) 255–260, <https://doi.org/10.1126/science.aaa8415>.
- [14] A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, E. Topol, Deep learning-enabled multi-modal fusion of medical imaging and electronic health records for improved diagnostics and prognostics, *Nat. Commun.* 12 (1) (2021) 6675, <https://doi.org/10.1038/s41467-021-26946-4>.
- [15] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.W. M van der Laak, A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88, <https://doi.org/10.1016/j.media.2017.07.005>.
- [16] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, A.Y. Ng, CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning, 2017 arXiv preprint arXiv:1711.05225, <https://arxiv.org/abs/1711.05225>.
- [17] A. Ribas, J.D. Wolchok, J. Schlom, E.M. Jaffee, Cancer immunotherapy comes of age, *Nat. Commun.* 9 (1) (2018) 1–14, <https://doi.org/10.1038/s41467-018-04388-3>.
- [18] J. He, S.L. Baxter, J. Xu, J. Xu, X. Zhou, K. Zhang, The practical implementation of artificial intelligence technologies in medicine, *Nat. Med.* 25 (1) (2019) 30–36, <https://doi.org/10.1038/s41591-018-0307-0>.
- [19] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, N. Elhadad, Intelligible models for healthcare: predicting pneumonia risk and hospital 30-day readmission, in: Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015, pp. 1721–1730, <https://doi.org/10.1145/2783258.2788613>.
- [20] E. Tjoa, C. Guan, A survey on explainable artificial intelligence (XAI): towards medical AI transparency, *Nat. Mach. Intell.* 2 (1) (2020) 56–67, <https://doi.org/10.1038/s42256-020-00273-3>.
- [21] J. Shen, C.J.P. Zhang, B. Jiang, J. Chen, J. Song, Z. Liu, Z. He, Artificial intelligence versus clinicians in disease diagnosis: systematic review, *JMIR. Med. Inform.* 10 (4) (2022) e32912, <https://doi.org/10.2196/32912>.
- [22] M.J. Sheller, B. Edwards, G.A. Reina, J. Martin, S. Pati, A. Kotrotsou, Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data, *Sci. Rep.* 10 (1) (2020) 12598, <https://doi.org/10.1038/s41598-020-69250-1>.
- [23] N. Rieke, J. Hancox, W. Li, F. Milletari, H.R. Roth, S. Albarqouni, G. Kaissis, The future of digital health with federated learning, *NPJ. Digit. Med.* 3 (2020) 119, <https://doi.org/10.1038/s41746-020-00323-1>.
- [24] J.M. Stokes, K. Yang, K. Swanson, W. Jin, A. Cubillos-Ruiz, N.M. Donghia, J. J. Collins, A deep learning approach to antibiotic discovery, *Cell* 180 (4) (2020) 688–702, <https://doi.org/10.1016/j.cell.2020.01.021>, e13.
- [25] A. Zhavoronkov, Y.A. Ivanenkov, A. Aliper, M.S. Veselov, V.A. Aladinskiy, A. V. Aladinskaya, A. Aspuru-Guzik, Deep learning enables rapid identification of potent DDR1 kinase inhibitors, *Nat. Biotechnol.* 37 (9) (2019) 1038–1040, <https://doi.org/10.1038/s41587-019-0224-x>.
- [26] Y. Yang, M.S. Islam, J. Wang, Y. Li, A comprehensive survey on machine learning techniques in medical diagnosis, *Comput. Biol. Med.* 101 (2017) 107–128, <https://doi.org/10.1016/j.combiomed.2018.06.016>.
- [27] J. De Fauw, J.R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, P.A. Keane, Clinically applicable deep learning for diagnosis and referral in retinal disease, *Nat. Med.* 24 (9) (2018) 1342–1350, <https://doi.org/10.1038/s41591-018-0174-4>.
- [28] W. Streeter, Can better data save the NHS? *Financial Times* (2024). <https://www.ft.com/content/b4c57347-d64d-436a-a2f1-33b7049a74b7>.
- [29] M. Roy, S.J. Minar, P. Dhar, A.T.M.O. Faruq, Machine Learning Applications In Healthcare: The State Of Knowledge and Future Directions, 2023 arXiv preprint arXiv:2307.14067, <https://arxiv.org/abs/2307.14067>.
- [30] Vatican News, New Vatican document offers AI guidelines from warfare to health care, Associated Press, 2025. <https://apnews.com/article/231b4b7b8ed6a195ec920f1362c15e2>.
- [31] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, A. Galstyan, A survey on bias and fairness in machine learning, *ACM Comput. Surv. (CSUR)* 54 (6) (2021) 1–35, <https://doi.org/10.1145/3457607>.
- [32] M.A. Gianfrancesco, S. Tamang, J. Yazdany, G. Schmajuk, Potential biases in machine learning algorithms using electronic health record data, *JAMA Intern. Med.* 178 (11) (2018) 1544–1547, <https://doi.org/10.1001/jamainternmed.2018.3763>.
- [33] J. Buolamwini, T. Gebru, Gender shades: intersectional accuracy disparities in commercial gender classification, in: Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT*), 2018, pp. 77–91, <https://doi.org/10.1145/3287560.3287583>.
- [34] S.M. Lundberg, G.G. Erion, S.I. Lee, Consistent individualized feature attribution for tree ensembles, *Nat. Mach. Intell.* 2 (1) (2020) 56–67, <https://doi.org/10.1038/s42256-019-0138-9>.
- [35] Z.C. Lipton, The myths of model interpretability, *Queue*, 16 (3) (2018) 31–57, <https://doi.org/10.1145/3236386.3241340>.
- [36] F. Doshi-Velez, B. Kim, Towards a rigorous science of interpretable machine learning, 2017 arXiv preprint arXiv:1702.08608.
- [37] W. Samek, T. Wiegand, K.R. Müller, arXiv preprint, 2019, <https://doi.org/10.48550/arXiv.1904.00026>.
- [38] S. Wachter, B. Mittelstadt, C. Russell, Counterfactual explanations without opening the black box: automated decisions and the GDPR, *Harv. J. Law Technol.* 31 (2) (2017) 841–887, <https://doi.org/10.2139/ssrn.3063289>.
- [39] J. Amann, A. Blasimme, E. Vayena, D. Frey, V.I. Madai, Explainability for artificial intelligence in healthcare: A multidisciplinary perspective, *BMC. Med. Inform. Decis. Mak.* 20 (1) (2020) 1–9, <https://doi.org/10.1186/s12911-020-01332-6>.
- [40] W. Wang, K. Siau, Artificial intelligence governance in healthcare: challenges, opportunities, and future research directions, *Int. J. Inf. Manage* 64 (2022) 102466, <https://doi.org/10.1016/j.ijinfomgt.2022.102466>.
- [41] P. Voigt, dem von, A. Bussche, The EU General Data Protection Regulation (GDPR): A practical guide, Springer International Publishing, 2017, <https://doi.org/10.1007/978-3-319-57959-7>.
- [42] B. Goodman, S. Flaxman, European Union regulations on algorithmic decision-making and a "right to explanation", *AI. Mag.* 38 (3) (2017) 50–57, <https://doi.org/10.1609/aimag.v38i3.2741>.
- [43] J. Powles, H. Hodson, Google DeepMind and healthcare in an age of algorithms, *Health Technol. (Berl)* 7 (4) (2017) 351–367, <https://doi.org/10.1007/s12553-017-0179-1>.
- [44] S. Hoffman, A. Podgurski, Artificial intelligence and the law: regulation and accountability for AI in global health, *J. Law Biosci.* 8 (1) (2021), <https://doi.org/10.1093/jlb/lsab014>.
- [45] D. McGraw, Building public trust in uses of Health Insurance Portability and Accountability Act de-identified data, *JAMA Intern. Med.* 173 (17) (2013) 1581–1582, <https://doi.org/10.1001/jamainternmed.2013.7111>.
- [46] Food and Drug Administration, Artificial intelligence/machine learning-based software as a medical device (SaMD), FDA Regulatory Guidelines, 2021. <https://www.fda.gov/media/145022/download>. Accessed 17 April 2025.
- [47] Food and Drug Administration, Predetermined change control plans for machine learning-enabled medical devices: guiding principles. U.S. Food and Drug Administration, Retrieved from, <https://www.fda.gov/media/145022/download>, 2023.
- [48] European Medicines Agency, Guideline on AI applications in medicine, EMA AI Regulat. Framework (2021). https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-use-ai-medicines_en.pdf. Accessed 18 April 2025.
- [49] M.D. Abramoff, Y. Lou, A. Erginay, W. Clarida, R. Amelon, J.C. Folk, M. Niemeijer, Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning, *Invest. Ophthalmol. Vis. Sci.* 59 (10) (2018) 4167–4175, <https://doi.org/10.1167/iov.18-24673>.

- [50] European Commission, Amended proposal for the AI Act (updated draft). <https://digital-strategy.ec.europa.eu/en/library>, 2023. Accessed on 18 April 2025.
- [51] M. Veale, F.Z. Borgesius, Demystifying the Draft EU Artificial Intelligence Act, *Comput. Law Rev. Int.* 22 (4) (2021) 97–112, <https://doi.org/10.9785/crl-2021-220402>.
- [52] M. Brkan, Do algorithms rule the world? Algorithmic decision-making and the EU General Data Protection Regulation (GDPR), *Int. J. Law Inf. Technol.* 29 (2) (2021) 91–121, <https://doi.org/10.1093/ijlit/eaad023>.
- [53] E.D. Canedo, A.D. Silva, F.B. Araujo, V.S. Carvalho, J.D. Costa, et al., Challenges regarding the compliance with the General Data Protection Law by Brazilian organizations: A survey, in: O. Gervasi, et al. (Eds.), *Computational science and its applications – ICCSA 2021: Vol. 12951. Lecture Notes in Computer Science*, Springer, 2021, pp. 460–475, https://doi.org/10.1007/978-3-030-86970-0_31.
- [54] R. Sinha, A. Agarwal, N. Jain, Ethical implications of India's National Digital Health Mission (NDHM): A policy analysis, *Health Policy. Technol.* 11 (2) (2022) 100624.
- [55] R. Garg, Analysing the NDHM's health data management policy: part 1. Internet Freedom Foundation. <https://internetfreedom.in/analysing-the-ndhms-health-data-management-policy-part-1/accessed>, 2021 on 17April, 2025.
- [56] South African Department of Communications and Digital Technologies, Draft National Artificial Intelligence Policy Framework, Retrieved from, <https://www.dcdt.gov.za/on>, 2021, 17 April, 2025.
- [57] European Commission, Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), COM (2021) (2021) 206 final, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>. Accessed on 18 April 2025.
- [58] European Commission, The Artificial Intelligence Act: AI Regulation in the European Union, EU Digital Strategy Report, 2023. https://ec.europa.eu/digital-strategy/ai-act_en. Accessed on 18 April 2025.
- [59] World Economic Forum, Global technology governance report 2021: harnessing Fourth Industrial Revolution technologies in a COVID-19 world. <https://www.weforum.org/reports/global-technology-governance-report-2021>, 2020. Accessed on 18 April 2025.
- [60] Organisation for Economic Co-operation and Development (OECD), OECD principles on artificial intelligence. <https://oecd.ai/en/ai-principles>, 2019. Accessed on 18 April 2025.
- [61] European Data Protection Board (EDPB) & European Data Protection Supervisor (EDPS), Joint opinion 5/2021 on the proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). <https://edpb.europa.eu>, 2021. Accessed on 18 April 2025.

CHAPTER 7: FINAL IMPLEMENTATION AND SYSTEM INTEGRATION

7.1 Methodology and Experimental Setup

7.1.1 Overview of the proposed framework

The final implementation of this research presents a multimodal machine learning framework designed for the early detection of Alzheimer’s disease (AD) by integrating neuroimaging (MR) and structured clinical data. The system follows a dual-pathway architecture consisting of:

- a. A **Convolutional Neural Network (CNN)** specifically **EfficientNetV2-S** for MRI-based feature extraction: This component analyzes MRI brain images to detect structural abnormalities associated with AD. EfficientNetV2-S leverages high-capacity convolutional feature extraction and transfer-learning optimization to enhance early-detection sensitivity. The model is capable of identifying subtle neurodegenerative patterns that may not be visibly apparent to clinicians.
- b. A **Feedforward Neural Network (FNN)** for Clinical data modeling: The FNN processes structured clinical attributes, including demographic data, cognitive assessments, and basic laboratory indicators, to complement image-based predictions. This dual-modality approach increases diagnostic robustness by capturing both anatomical and clinical evidence.

These independent pathways capture complementary aspects of Alzheimer’s pathology and contribute to improved diagnostic performance.

7.1.2 Dataset Description and Data Splitting Strategy

The dataset used in this study consists of:

- 2000 MRI brain scans (balanced: Demented/Non-Demented)

- 2000 Structured clinical records

To ensure robust model training and unbiased evaluation, the dataset was divided into:

- 70% Training Set
- 15% Validation Set
- 15% Test Set

This three-way split was adopted instead of a conventional 80-20 split to ensure a clear separation between model tuning and final evaluation. The validation set was used for hyperparameter tuning and early stopping, while the test set remained completely unseen during training, ensuring unbiased performance estimation and improved generalization.

Furthermore, the dataset splitting was performed before preprocessing to prevent data leakage. Techniques such as SMOTE were applied only to the training set, ensuring realistic evaluation conditions.

7.1.3 Dataset Preprocessing and Feature Engineering

MRI Preprocessing

- Resizing images to 224×224 pixels
- Intensity normalization
- Data augmentation (rotation, flipping, zooming)

Clinical Data Preprocessing

- Missing value imputation using the **median strategy**
- One-hot encoding for categorical variables
- Feature scaling using **StandardScaler**

Feature Subset Evaluation

To enhance model efficiency and interpretability, multiple feature subsets ranging from 8 to 20 key clinical variables were experimentally evaluated. This allowed identification of the most informative predictors, including:

- MMSE (cognitive score)
- ADL (functional ability)
- Cholesterol levels (metabolic indicator)

This step ensures optimal feature representation while reducing redundancy.

7.1.4 Model Architecture and Design

(a) MRI-Based Model (CNN- EfficientNetV2-S)

The MRI pathway utilizes the **EfficientNetV2_S architecture**, selected due to its superior balance between accuracy and computational efficiency.

- Transfer learning with pretrained weights
- Early layers frozen, final layers fine-tuned
- Input size: 224×224
- Loss function: Binary Cross-Entropy

Optimizer Selection (Justification)

The **Adam optimizer** was employed due to its adaptive learning rate mechanism and efficient convergence properties. Compared to traditional optimizers such as Stochastic Gradient Descent (SGD), Adam provides:

- Faster convergence
- Improved stability
- Better performance on complex multimodal datasets

(a) Clinical Model (Feedforward Neural Network - FNN)

The clinical pathway is implemented using a **Feedforward Neural Network**, designed to capture complex relationships within structured biomedical data.

- Hidden layers: 57 and 129 neurons
- Activation: ReLU
- Output: Sigmoid neuron

Hyperparameter Optimization

Hyperparameters were optimized using **Optuna (50 trials)**, ensuring optimal model configuration.

7.1.5 Justification for Non-Linear Modeling

Alzheimer's disease progression involves complex, non-linear interactions among biomarkers, cognitive scores, and imaging features. Traditional linear models, such as logistic regression, are insufficient to capture these relationships.

Therefore, this study adopts non-linear modeling approaches, including:

- Deep Neural Networks
- Convolutional Neural Networks
- Gradient-based optimization

These models enable the learning of hierarchical and non-linear feature representations, resulting in significantly improved diagnostic performance.

7.1.6 Training Strategy

The training process incorporated several strategies to improve generalization and prevent overfitting:

- Data augmentation (MRI images)
- Early stopping (patience = 10 epochs)
- Validation-based performance monitoring
- Regularization using dropout

Learning Rate

An initial learning rate of $1e-4$ was used to ensure stable and gradual convergence during training.

7.1.7 Evaluation Metrics and Performance Analysis

To ensure a comprehensive and clinically relevant evaluation, this study employed an extended set of performance metrics, going beyond conventional measures.

Core Metrics

- Accuracy
- Precision
- Recall (Sensitivity)
- F1 –Score

Clinical Metrics

- Sensitivity: Ability to detect AD patients
- Specificity: Ability to avoid false positives

Advanced Metrics

- ROC-AUC
- Precision-Recall AUC
- Confusion Matrix
- Log Loss
- Mathews Correlation coefficient (MCC)
- Cohen's kappa

This multi-metric evaluation framework ensures both statistical robustness and clinical reliability, addressing limitations in prior studies that rely on accuracy.

7.1.8 Data Leakage Prevention and Validation Strategy

To ensure unbiased model evaluation:

- Dataset splitting was performed before preprocessing
- SMOTE is applied only on training data
- Preprocessing pipelines fitted only on the training set
- Validation and test sets remained unseen

This approach ensures true generalization performance.

7.2 Summary of Contributions (Final Implementation)

The final implementation of this research provides:

- A multimodal AI framework integrating MRI and Clinical data
- A dual-pathway architecture (CNN and FNN)
- Use of non-linear deep learning models for improved prediction
- Robust evaluation using multiple performance metrics
- A validated data splitting strategy ensuring unbiased results
- A clinically interpretable and scalable diagnostic system
- Implementation of NeuroBot, an interactive AI assistant designed to support clinical workflows and enhance patient engagement.
- Integration of explainable AI (XAI) techniques, specifically SHAP-based feature attribution, to improve decision transparency and support clinician trust.
- Demonstration of strong diagnostic performance, particularly in early-stage AD detection.
- Advancement of patient-centered diagnostic design through the development of an intuitive GUI and interactive user experience.

7.3 Connection to Thesis Objectives

The work plays a foundational role in the overall thesis trajectory. It illustrates how deep learning architectures can be translated from theoretical models into practical, interpretable, and clinically usable diagnostic tools. Specifically, the work advances two central thesis objectives:

- a. **Development of Multimodal Machine Learning Frameworks:** This work establishes the initial hybrid integration of imaging and clinical biomarkers, an approach expanded into more complex multimodal, temporal, and graph-based models in Papers 1 and 2.
- b. **Promotion of Transparency and Clinical Trust:** Through Grad-CAM explainability, GUI-based interaction, and the NeuroBot assistant, the study directly contributes to the thesis goal of creating transparent, user-centered, and ethically deployable AD diagnostic systems.

Overall, the final work serves as the applied entry point in the thesis pipeline, providing the practical and conceptual foundation upon which the subsequent multimodal and governance-focused studies are built.

CHAPTER 8: CONCLUSIONS AND FUTURE WORK

8.1 Summary of Contributions

This thesis offers several substantial contributions to the field of Alzheimer’s Disease (AD) detection, combining methodological innovation, multimodal data integration, clinical interpretability, and responsible AI governance. By addressing longstanding challenges related to early diagnosis, model robustness, and real-world deployment, the work advances the current state of biomedical machine learning. The key contributions are summarized below.

a. **Development of Novel Hybrid Deep Learning and Machine Learning Architectures:**

The research introduces multiple hybrid AI frameworks that integrate modern deep learning architectures such as EfficientNetV2-S, CNNs, LSTMs, and GNNs with powerful machine learning models, including Gradient Boosting Machines and Deep Neural Networks. These systems jointly analyze imaging biomarkers (MRI/CT) and structured clinical data to classify AD, MCI, and cognitively normal states with high accuracy.

By outperforming unimodal baselines, these hybrid architectures establish a technically strong and clinically relevant foundation for deployable AD diagnostic tools.

b. **Multimodal Fusion of Spatial, Temporal, and Relational Biomarkers:** A central scientific contribution of the thesis is the design of a unified multimodal fusion framework that integrates:

- i. ***Spatial neuroimaging information*** captured through CNNs
- ii. ***Temporal biomarker and cognitive trajectories*** modeled using LSTM networks
- iii. ***Relational patient similarity patterns*** learned through Graph Neural Networks

This holistic approach enables more robust early-stage AD detection and significantly improves the prediction of MCI-to-AD conversion. The findings provide empirical evidence that AD pathology is inherently multimodal and benefits from multi-perspective computational modeling.

c. **Explainable and Patient-Centric Diagnostic Ecosystems:** Beyond predictive performance, the thesis emphasizes transparency and usability. Key contributions include:

- i. SHAP-based feature attribution for biomarker-level interpretability

- ii. Model explanation dashboards that highlight influential clinical and cognitive factors
- iii. A clinician-friendly Graphical User Interface (GUI) to streamline diagnostic workflows
- iv. An interactive conversational assistant (NeuroBot) for guided data entry, explanation delivery, and patient support

These features ensure that the developed systems are interpretable, trustworthy, and accessible, supporting clinical decision-making and facilitating broader adoption in healthcare settings.

d. **Federated Learning and GAN-Based Domain Adaptation for Multicenter Robustness:** To address cohort heterogeneity, imaging variability, and strict data governance constraints, the thesis integrates:

- i. Federated learning, enabling distributed training without transferring sensitive patient data
- ii. GAN-based domain adaptation, harmonizing imaging and clinical data from different institutions and scanners

These techniques substantially enhance cross-site generalizability, reduce distributional bias, and demonstrate the feasibility of privacy-preserving AI for large-scale clinical applications.

e. **Ethical and Regulatory Frameworks for Responsible AI Deployment:** Recognizing that clinical impact requires more than technical innovation, the thesis provides a thorough analysis of global ethical and regulatory frameworks, including GDPR, HIPAA, FDA AI/ML guidance, and the EU AI Act (2024–2025). The work identifies challenges related to bias, accountability, explainability, transparency, and international data governance.

The thesis proposes guidelines and best practices to ensure that AI-driven diagnostic systems comply with modern expectations for fairness, privacy, safety, and clinician oversight.

8.2 Limitations

Although this thesis presents substantial methodological and practical contributions to early Alzheimer’s Disease (AD) detection, several limitations must be acknowledged. These

reflect challenges inherent to multimodal AD research, constraints in data availability, and broader complexities associated with deploying AI systems in real clinical environments.

- a. **Limited External Validation Across Diverse Cohorts:** The proposed models demonstrated strong performance on institutional datasets and ADNI-derived imaging and biomarker data. However, access to large, demographically diverse, multi-cohort datasets remained limited. Due to strict privacy regulations, data-sharing restrictions, and institutional constraints, extensive external validation across international or multi-hospital cohorts could not be conducted. This limits the ability to comprehensively evaluate the models' generalizability across different populations, scanners, and clinical settings, which is essential for real-world deployment.
- b. **Variability in Imaging Protocols and Acquisition Devices:** MRI and CT scans vary considerably across scanner manufacturers, acquisition parameters, and clinical environments. These differences introduce domain shifts in signal intensity, resolution, and noise characteristics factors that can degrade deep learning performance. Although GAN-based domain adaptation was employed to harmonize cross-site imaging differences, residual inconsistencies remain. Such variability continues to challenge multimodal pipelines reliant on neuroimaging data and may limit consistent performance in large-scale or multicenter applications.
- c. **Multimodal Data Are Expensive, Sparse, and Often Incomplete:** A key strength of this thesis is its integration of spatial (MRI/CT), temporal (longitudinal scores), biomarker (CSF/blood), and relational (graph-based) information. However, multimodal datasets that include all these modalities are rare, costly to acquire, and often incomplete. Many clinical records contain missing CSF measurements, absent genetic profiles, or irregular longitudinal follow-up. This data sparsity limits the full potential of multimodal fusion and may unintentionally skew the dataset toward individuals receiving more comprehensive evaluations typically those in specialized research centers rather than general clinical populations.
- d. **Ethical Challenges: Underrepresentation and Algorithmic Bias:** AI models are highly sensitive to the demographic composition of their training data. Underrepresentation of ethnic minorities, rural communities, or individuals with low educational attainment can introduce systematic bias. Such disparities may result in unequal diagnostic accuracy or

misclassification rates across subgroups. Although fairness considerations and regulatory frameworks were examined in Paper 4, achieving true representativeness remains a major challenge due to the global scarcity of diverse AD cohorts.

- e. **Limitations of GAN-Based Harmonization:** GAN-based harmonization techniques were applied to reduce domain variability and support cross-cohort training. While effective, these methods may inadvertently introduce synthetic artifacts or subtly distort biologically meaningful imaging features. These modifications, though often visually imperceptible, may influence feature extraction and bias downstream predictions. Ensuring that harmonization preserves clinically relevant information remains an unresolved research challenge and may require more robust validation, hybrid correction strategies, or direct clinician oversight.

8.3 Recommendations for Future Research

Building upon the methodological innovations and empirical findings of this thesis, several promising directions for future research emerge. These opportunities address current limitations related to dataset diversity, model generalizability, clinical integration, and regulatory preparedness. Advancing research along these lines will support the development of more robust, equitable, and clinically deployable AI systems for Alzheimer’s Disease (AD) detection and progression modeling.

- a. **Expansion of Multimodal Datasets with Greater Ethnic and Demographic Diversity:** Most existing AD datasets are dominated by Western, urban, and highly educated populations, which can constrain model generalizability and exacerbate demographic biases. Future work should focus on assembling large-scale, diverse multimodal datasets that include participants from underrepresented populations across Asia, Africa, South America, and rural regions globally. Such efforts would:
 - i. Improve fairness and reduce algorithmic bias
 - ii. Enhance the clinical validity of models across heterogeneous populations
 - iii. Support equitable deployment in varied healthcare environments

Broadening representation is essential for achieving diagnostic systems that perform reliably for all demographic groups.

b. **Integration of Large Language Models (LLMs) for Advanced Clinical Decision**

Support: Recent advances in generative AI, such as GPT, BioGPT, Med-PaLM, and LLaMA, offer new opportunities for enhancing clinical reasoning and patient communication. Future research could integrate LLMs with multimodal diagnostic pipelines to:

- i. Generate clinician-friendly explanations and reports
- ii. Summarize complex multimodal findings
- iii. Support triage and differential diagnosis
- iv. Provide personalized patient and caregiver guidance

Coupling LLM reasoning with imaging, biomarker, and longitudinal models may lead to more holistic and context-aware decision-support systems that extend beyond simple classification tasks.

c. **Development of Standardized AI Governance and Oversight Frameworks:** Although Paper 3 examined regulatory and ethical considerations, there remains a pressing need for consistent, reproducible AI governance frameworks that span the full lifecycle of clinical AI systems. Future work should explore the development of:

- i. Risk-classification schemes for adaptive AI algorithms
- ii. Standardized documentation tools (model cards, dataset datasheets)
- iii. Fairness auditing and demographic bias assessment protocols
- iv. Privacy-preserving data-sharing methods (federated analytics, secure multiparty computation)

Creating unified governance frameworks would help clinicians, regulators, and developers ensure the safety, transparency, and accountability of emerging AI tools.

d. **Advanced Longitudinal Modeling with Transformer Architectures:** While LSTMs were effective for modeling biomarker and cognitive trajectories, transformer-based architectures, including Time-Series Transformers, Vision Transformers, and Multimodal Transformers, offer enhanced capabilities for capturing long-range temporal dependencies and complex multimodal interactions. Future work could leverage transformers for:

- i. Predicting MCI-to-AD conversion

- ii. Modeling long-term disease progression
- iii. Integrating multimodal temporal sequences (imaging, biomarkers, EHR data)

Transformers may also improve interpretability and sensitivity to subtle early pathological changes, further enhancing clinical relevance.

- e. **Deployment as FDA/EMA-Compliant Medical Devices (SaMD):** To transition from research prototypes to clinically approved tools, future systems should aim for compliance with international regulatory frameworks, including:
 - i. FDA Good Machine Learning Practices (GMLP)
 - ii. FDA/EMA guidelines for Software as a Medical Device (SaMD)
 - iii. EU AI Act requirements for high-risk medical AI

Achieving regulatory readiness will require:

- i. Detailed model traceability and documentation
- ii. Rigorous validation and verification pipelines
- iii. Continuous performance monitoring and post-market surveillance
- iv. Robust safety and risk-mitigation strategies

Regulatory certification represents a critical step toward real-world clinical deployment.

- f. **Real-World Clinical Trials and Prospective Validation:** Prospective, real-world evaluations remain the gold standard for assessing clinical effectiveness and safety. Future research should conduct:
 - i. Multicenter clinical trials
 - ii. Real-time clinical workflow integration studies
 - iii. Physician-in-the-loop validation
 - iv. Longitudinal observational studies
 - v. Cross-site robustness assessments

These studies will provide insights into patient outcomes, clinician trust, usability, and real-world diagnostic performance, essential factors for successful adoption.

References

- Abdelaziz, M., Wang, T., Anwaar, W., & Elazab, A. (2024). Multi-scale multimodal deep learning framework for Alzheimer's disease diagnosis. *Computer Methods and Programs in Biomedicine*, 242, 109438. <https://doi.org/10.1016/j.compbimed.2024.109438>
- Abuhantash, M., Al-Tamimi, H., Abu Khalaf, J., Abu-Naser, S. S., & Naser, S. S. A. (2024). Comorbidity-based framework for Alzheimer's disease classification using graph neural networks. *Scientific Reports*, 14(1), 19924. <https://doi.org/10.1038/s41598-024-72321-2>
- Alatrany, A.S., Khan, W., Hussain, A. et al. (2024). An explainable machine learning approach for Alzheimer's disease classification. *Sci Rep* 14, 2637. <https://doi.org/10.1038/s41598-024-51985-w>[56]
- Zhang, F., Pan, B., Shao, P., Liu, P., Shen, S., Yao, P., & Xu, R. X. (2022). A single model deep learning approach for Alzheimer's disease diagnosis. *Neuroscience*, 491, 200-214. <https://doi.org/10.1016/j.neuroscience.2022.03.026>
- Alayba, A. M., Senan, E. M., & Alshudukhi, J. S. (2024). Enhancing early detection of Alzheimer's disease through hybrid models based on feature fusion of multi-CNN and handcrafted features. <https://doi.org/10.1038/s41598-024-82544-y>
- Al-Hammadi, N., Abouelyazid, M., Brown, D. C., Lalwani, P., Devos, H., Carr, D. B., & Babulal, G. M. (2025). Integrating machine learning and environmental and genetic risk factors for the early detection of preclinical Alzheimer's disease. *The Journals of Gerontology: Series B*. <https://doi.org/10.1093/geronb/gbaf023>
- Alsubaie, M. G., Luo, S., & Shaukat, K. (2024). Alzheimer's disease detection using deep learning on neuroimaging: A systematic review. *Machine Learning and Knowledge Extraction*, 6(1), 464–505. <https://doi.org/10.3390/make6010024>
- Amann, J., Blasimme, A., Vayena, E. et al. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak* 20, 310 (2020). <https://doi.org/10.1186/s12911-020-01332-6>
- Andersson, E., Schultz, N., Saito, T., Saido, T. C., Blennow, K., Gouras, G. K., Zetterberg, H., & Hansson, O. (2023). Cerebral A β deposition precedes reduced cerebrospinal fluid and serum A β 42/A β 40 ratios in the AppNL-F/NL-F knock-in mouse model of Alzheimer's disease. *Alzheimer's Research & Therapy*, 15(1), 64. <https://doi.org/10.1186/s13195-023-01196-8>

- Asaduzzaman, M., Alom, M. K., & Karim, M. E. (2025). ALZENET: Deep learning-based early prediction of Alzheimer's disease through magnetic resonance imaging analysis. *Telematics and Informatics Reports*, 17, 100189. <https://doi.org/10.1016/j.teler.2025.100189>
- Balbal, K. F. (2024). Machine learning approaches for prediction of Alzheimer's disease. *Journal of Artificial Intelligence and Data Science*, 4(2), 110–118. Retrieved from <https://dergipark.org.tr/en/pub/jaida/issue/89048/1599026>
- Dubois, B., Villain, N., Schneider, L., Fox, N., Campbell, N., Galasko, D., Kivipelto, M., Jessen, F., Hanseeuw, B., Boada, M., Barkhof, F., Nordberg, A., Froelich, L., Waldemar, G., Frederiksen, K. S., Padovani, A., Planche, V., Rowe, C., Bejanin, A., ... Frisoni, G. B. (2024). Alzheimer disease as a clinical-biological construct—An International Working Group recommendation. *JAMA Neurology*, 81(12), 1304–1311. <https://doi.org/10.1001/jamaneurol.2024.3770>
- Erdoğan, O., İyigün, C., & Aydın Son, Y. (2024). EnSCAN: ENsemble Scoring for prioritizing CAusative variaNts across multi-platform GWAS for Late-Onset Alzheimer's Disease. Research Square. <https://doi.org/10.21203/rs.3.rs-4031105/v1>
- Farhatullah, Chen, X., Zeng, D., Ullah, R., Nawaz, R., Xu, J., & Arslan, T. (2025). A deep learning approach for non-invasive Alzheimer's monitoring using microwave radar data. *Neural Networks*, 181, 106778. <https://doi.org/10.1016/j.neunet.2024.106778>
- Fathi, S., Ahmadi, A., Dehnad, A., Almasi, M., & Melika, D. (2024). A Deep Learning-Based Ensemble Method for Early Diagnosis of Alzheimer's Disease using MRI Images. *Neuroinformatics*, 22(1), 89–105. <https://doi.org/10.1007/s12021-023-09646-2>
- Feng, Y., Gao, B., Deng, S., Qiu, A., & Qin, J. (2024). Unified Multi-modal Learning for Any Modality Combinations in Alzheimer's Disease Diagnosis. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024* (pp. 487–497). Springer. https://doi.org/10.1007/978-3-031-72384-1_46
- Gasmi, K., Alyami, A., Hamid, O., Altaieb, M. O., Shahin, O. R., Ben Ammar, L., Chouaib, H., & Shehab, A. (2024). Optimized hybrid deep learning framework for early detection of Alzheimer's disease using adaptive weight selection. *Diagnostics*, 14(24), 2779. <https://doi.org/10.3390/diagnostics14242779>
- Gaeta, A. M., Quijada-López, M., Barbé, F., Vaca, R., Pujol, M., Minguez, O., Sánchez-de-la-Torre, M., Muñoz-Barrutia, A., & Piñol-Ripoll, G. (2024). Predicting Alzheimer's disease

- CSF core biomarkers: A multimodal machine learning approach. *Frontiers in Aging Neuroscience*, 16. <https://doi.org/10.3389/fnagi.2024.1369545>
- Golovanevsky, M., Eickhoff, C., & Singh, R. (2022). Multimodal attention-based deep learning for Alzheimer's disease diagnosis. *Journal of the American Medical Informatics Association*, 29(12), 2014–2022. <https://doi.org/10.1093/jamia/ocac168>
- Hampel, H., Hu, Y., Cummings, J., Mattke, S., Iwatsubo, T., Nakamura, A., Vellas, B., O'Bryant, S., Shaw, L. M., Cho, M., Batrla, R., Vergallo, A., Blennow, K., Dage, J., & Schindler, S. E. (2023). Blood-based biomarkers for Alzheimer's disease: Current state and future use in a transformed global healthcare landscape. *Neuron*, 111(18), 2781–2799. <https://doi.org/10.1016/j.neuron.2023.05.017>
- Hansson, O., Edelmayer, R. M., Boxer, A. L., Carrillo, M. C., Mielke, M. M., Rabinovici, G. D., Salloway, S., Sperling, R., Zetterberg, H., & Teunissen, C. E. (2022). The Alzheimer's Association appropriate use recommendations for blood biomarkers in Alzheimer's disease. *Alzheimer's & Dementia*, 18(12), 2669–2686. <https://doi.org/10.1002/alz.12756>
- Hassan, N., Saleh, A., & Miah, M. (2024). Residual-Based Multi-Stage Deep Learning Framework for Computer-Aided Alzheimer's Disease Detection. <https://doi.org/10.3390/jimaging10060141>
- Hazarika, R. A., Kandar, D., & Kumar, A. (2024). A novel machine learning-based technique for classification of early-stage Alzheimer's disease using brain images. <https://doi.org/10.1007/s11042-023-16379-6>
- Hoang, G. M., Kim, U.-H., & Kim, J. G. (2023). Vision transformers for the prediction of mild cognitive impairment to Alzheimer's disease progression using mid-sagittal sMRI. *Frontiers in Aging Neuroscience*, 15, 1102869. <https://doi.org/10.3389/fnagi.2023.1102869>
- Hk, R., Deepak, D. R., & Kumar, H. A. P. (2024). A robust framework for Alzheimer's disease detection and staging: Incorporating multi-feature integration, MRMR feature selection, and Random Forest classification. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-19875-5>
- Jiang, H., Qian, Y., Zhang, L., Jiang, T., & Tai, Y. (2025). ReIU: An efficient preliminary framework for Alzheimer patients based on multi-model data. *Frontiers in Public Health*, 12, 1449798. <https://doi.org/10.3389/fpubh.2024.1449798>

- Kadem, M. (2023). Interpretable machine learning in Alzheimer's disease dementia (Master's thesis, McMaster University). <http://hdl.handle.net/11375/28503>
- Kang, W., Lin, L., Zhang, B., Shen, X., Wu, S., & Alzheimer's Disease Neuroimaging Initiative. (2021). Multi-model and multi-slice ensemble learning architecture based on 2D convolutional neural networks for Alzheimer's disease diagnosis. *Computers in Biology and Medicine*, 136, 104678. <https://doi.org/10.1016/j.combiomed.2021.104678>
- Kavitha, C., Mani, V., Srividhya, S. R., Khalaf, O. I., & Tavera Romero, C. A. (2022). Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models. *Frontiers in Public Health*, 10, 853294. <https://doi.org/10.3389/fpubh.2022.853294>
- Kim YJ, Kim SM, Jeong DH et al (2021) Associations between metabolic syndrome and type of dementia: analysis based on the National Health Insurance Service database of Gangwon province in South Korea. *Diabetol Metab Syndr* 13:4. <https://doi.org/10.1186/s13098-020-00620-5>
- Kim, Y., Lim, J., & Oh, J. (2024). Taming neuroinflammation in Alzheimer's disease: The protective role of phytochemicals through the gut–brain axis. *Biomedicine & Pharmacotherapy*, 178, 117277. <https://doi.org/10.1016/j.biopha.2024.117277>
- Khasanah, I. (2024). Enhancing Alzheimer's disease diagnosis with K-NN: A study on preprocessed MRI data. *International Journal of Artificial Intelligence in Medical Issues*, 2(1). <https://doi.org/10.56705/ijaimi.v2i1.150>
- Khatun, M., Islam, M. M., Rifat, H. R., Shahid, M. S. B., Talukder, M. A., & Uddin, M. A. (2024). Hybridized Convolutional Neural Networks and Long Short-Term Memory for Improved Alzheimer's Disease Diagnosis from MRI Scans. *arXiv preprint arXiv:2403.05353*.
- Lalitha, & Poovaraghan. (2024). A deep learning-based early Alzheimer's disease detection using magnetic resonance images. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-19677-9>
- Lantero Rodríguez, J., Karikari, T. K., Suárez-Calvet, M., Troakes, C., King, A., Emersic, A., Aarsland, D., Hye, A., Zetterberg, H., Blennow, K., & Ashton, N. J. (2020). Plasma ptau181 accurately predicts Alzheimer's disease pathology at least 8 years prior to postmortem and improves the clinical characterisation of cognitive decline. *Acta Neuropathologica*, 140(3), 267–278. <https://doi.org/10.1007/s00401-020-02195-x>

- Leandrou, S., Lamniso, D., Bougias, H., Stogiannos, N., Georgiadou, E., & Pattichis, C. S. (2023). A cross-sectional study of explainable machine learning in Alzheimer's disease: Diagnostic classification using MR radiomic features. *Frontiers in Aging Neuroscience*, 15, 1149871. <https://doi.org/10.3389/fnagi.2023.1149871>
- Lee, G.-B., Jeong, Y.-J., Kang, D.-Y., Yun, H.-J., & Yoon, M. (2024). Multimodal feature fusion-based graph convolutional networks for Alzheimer's disease stage classification using F-18 florbetaben brain PET images and clinical indicators. *PLOS ONE*, 19(12), e0315809. <https://doi.org/10.1371/journal.pone.0315809>
- Lee, J. C., Kim, S. J., Hong, S., et al. (2019). Diagnosis of Alzheimer's disease utilizing amyloid and tau as fluid biomarkers. *Experimental & Molecular Medicine*, 51(1), 1–10. <https://doi.org/10.1038/s12276-019-0250-2>
- Liu, J., Wu, S., Fu, Q., Luo, X., Luo, Y., Qin, S., Huang, Y., & Chen, Z. (2025). Multimodal diagnosis of Alzheimer's disease based on resting-state electroencephalography and structural magnetic resonance imaging. *Frontiers in Physiology*, 16, 1515881. <https://doi.org/10.3389/fphys.2025.1515881>
- Lombardi, G., Crescioli, G., Cavedo, E., Lucenteforte, E., Casazza, G., Bellatorre, A. G., Lista, C., Costantino, G., Frisoni, G., Virgili, G., & Filippini, G. (2020). Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment. *Cochrane Database of Systematic Reviews*, 2020(3), CD009628. <https://doi.org/10.1002/14651858.CD009628.pub2>
- Ma, K., Zhang, J., Huang, X., Wang, M., & Zhang, J. (2025). Leveraging transformer models to predict cognitive impairment: Accuracy, efficiency, and interpretability. *BMC Public Health*, 25, 504. <https://doi.org/10.1186/s12889-025-21762-z>
- McGraw, D., & Mandl, K. D. (2021). Privacy protections to encourage use of health-relevant digital data in a learning health system. *NPJ Digital Medicine*, 4(1), 2. <https://doi.org/10.1038/s41746-020-00362-8>
- Mohammed, Shehu; Malhotra, Neha; Singh, Arun; Awadelkarim, Awad M.; Ahmed, Shakeel; Potharaju, Saiprasad. (2025). Novel hybrid intelligence model for early Alzheimer's diagnosis utilizing multimodal biomarker fusion. *Informatics in Medicine Unlocked*, 57, 101668. <https://doi.org/10.1016/j.imu.2025.101668>

- Mohammed, S., Malhotra, N. (2025). Predicting Alzheimer's Disease onset: A machine learning framework for early diagnosis using biomarker data. *Computer Methods and Programs in Biomedicine Update*, 100209. <https://doi.org/10.1016/j.cmpbup.2025.100209>
- Mohammed, S., & Malhotra, N. (2025). Ethical and regulatory challenges in machine learning-based healthcare systems: A review of implementation barriers and future directions. *Technology and Health Care Benchmarks*, 100, 215. <https://doi.org/10.1016/j.tbench.2025.100215>
- Moscoso, A., Karikari, T. K., Grothe, M. J., Ashton, N. J., Lantero-Rodriguez, J., Snellman, A., Zetterberg, H., Blennow, K., & Schöll, M. (2022). CSF biomarkers and plasma p-tau181 as predictors of longitudinal tau accumulation: Implications for clinical trial design. *Alzheimer's & Dementia*, 18(7), 1312–1324. <https://doi.org/10.1002/alz.12570>
- Oh, W., & Nadkarni, G. N. (2023). Federated learning in healthcare using structured medical data. *Advances in Kidney Disease and Health*, 30(1), 4–16. <https://doi.org/10.1053/j.akdh.2022.11.007>
- Pan, F. F., Huang, Q., Huang, C. C., Lu, Y., Cui, L., Huang, L., Guan, Y., Xie, F., & Guo, Q. H. (2025). Associations of hippocampal volumes, brain hypometabolism, and plasma NfL with amyloid, tau, and cognitive decline. *Alzheimer's & Dementia*. <https://doi.org/10.1002/alz.70005>
- Pang, Z., Wang, X., Wang, X., & others. (2021). A multi-modal data platform for diagnosis and prediction of Alzheimer's disease using machine learning methods. *Mobile Networks and Applications*, 26, 2341–2352. <https://doi.org/10.1007/s11036-021-01834-1>
- Park, J. H., Cho, H. E., Kim, J. H., Wall, M. M., Stern, Y., Lim, H., Yoo, S., Kim, H. S., & Cha, J. (2020). Machine learning prediction of incidence of Alzheimer's disease using large-scale administrative health data. *npj Digital Medicine*, 3, 46. <https://doi.org/10.1038/s41746-020-0256-0>
- Pradhan, N., Sagar, S., & Singh, A. S. (2024). Analysis of MRI image data for Alzheimer disease detection using deep learning techniques. *Multimedia Tools and Applications*, 83,17729–17752. <https://doi.org/10.1007/s11042-023-16256-2>
- Rafsan, M., Oraby, T., Roy, U., Kumar, S., & Rodrigo, H. (2025). Deep learning for early Alzheimer disease detection with MRI scans. *arXiv*. <https://doi.org/10.48550/arXiv.2501.09999>

- Reitz, C., Brayne, C., & Mayeux, R. (2011). Epidemiology of Alzheimer disease. *Nature Reviews Neurology*, 7(3), 137–152. <https://doi.org/10.1038/nrneurol.2011.2>
- Sharma, M., Pal, P., & Gupta, S. K. (2024). Advances in Alzheimer's disease: A multifaceted review of potential therapies and diagnostic techniques for early detection. *Neurochemistry International*, 177, 105761. <https://doi.org/10.1016/j.neuint.2024.105761>
- Shen, X. N., Huang, Y. Y., Chen, S. D., Guo, Y., Tan, L., Dong, Q., Yu, J. T., & Alzheimer's Disease Neuroimaging Initiative. (2021). Plasma phosphorylated-tau181 as a predictive biomarker for Alzheimer's amyloid, tau and FDG PET status. *Translational Psychiatry*, 11(1), 585. <https://doi.org/10.1038/s41398-021-01709-9>
- Sheng, J., Zhang, Q., Zhang, Q., Wang, L., Yang, Z., Xin, Y., & Wang, B. (2024). A hybrid multimodal machine learning model for detecting Alzheimer's disease. *Computers in Biology and Medicine*, 170, 108035. <https://doi.org/10.1016/j.compbiomed.2024.108035>
- Singh, M. K., Shin, Y., Ju, S., Han, S., Kim, S. S., & Kang, I. (2024). Comprehensive overview of Alzheimer's disease: Etiological insights and degradation strategies. *International Journal of Molecular Sciences*, 25(13), 6901. <https://doi.org/10.3390/ijms25136901>
- Suchitra, S., Lalitha, K., & Poovaraghan, R. J. (2024). A deep learning-based early Alzheimer's
 [zhttps://doi.org/10.1007/s11042-024-19677-9](https://doi.org/10.1007/s11042-024-19677-9)
- Tagai, K., Tatebe, H., Matsuura, S., Kikuchi, T., Takeuchi, T., & Nakayama, S. (2024). A novel plasma p-tau181 assay as a specific biomarker of tau pathology in Alzheimer's disease. *Translational Neurodegeneration*, 13(1), 44. <https://doi.org/10.1186/s40035-024-00439-4>
- Taiyeb Khosroshahi, M., Morsali, S., Gharakhanlou, S., Motamedi, A., Hassanbaghlou, S., Vahedi, H., Pedrammehr, S., Kabir, H. M. D., & Jafarizadeh, A. (2025). Explainable artificial intelligence in neuroimaging of Alzheimer's disease. *Diagnostics*, 15(5), 612. <https://doi.org/10.3390/diagnostics15050612>
- Vanaja, T., Shanmugavadivel, K., Subramanian, M., & Others. (2025). Advancing Alzheimer's detection: Integrative approaches in MRI analysis with traditional and deep learning models. *Neural Computing & Applications*. <https://doi.org/10.1007/s00521-025-10993-1>
- Venkatesan, D., Muthukumar, S., Iyer, M., Suresh Babu, H. W., Gopalakrishnan, A. V., Yadav, M. K., & Vellingiri, B. (2024). Heavy metals toxicity on epigenetic modifications in the pathogenesis of Alzheimer's disease (AD). *Journal of Biochemical and Molecular Toxicology*. <https://doi.org/10.1002/jbt.23741>

- Venugopalan, J., Tong, L., Hassanzadeh, H. R., & Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific Reports*, 11(1), 3254. <https://doi.org/10.1038/s41598-020-74399-w>
- Venugopalan, J., Tong, L., Hassanzadeh, H. R., & Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific Reports*, 11(1), 3254. <https://doi.org/10.1038/s41598-020-74399-w>
- Vo, T., Ibrahim, A. K., & Zhuang, H. (2025). A multimodal multi-stage deep learning model for the diagnosis of Alzheimer's disease using EEG measurements. *Neurology International*, 17(6), 91. <https://doi.org/10.3390/neurolint17060091>
- Wang, F., Liang, Y., & Wang, Q.-W. (2025). Uncovering the epigenetic regulatory clues of PRRT1 in Alzheimer's disease: a strategy integrating multi-omics analysis with explainable machine learning. *Alzheimer's Research & Therapy*, 17, Article 12. <https://doi.org/10.1186/s13195-024-01646-x>
- Wang, R., He, Q., Han, C., Wang, H., Shi, L., & Che, Y. (2023). A deep learning framework for identifying Alzheimer's disease using fMRI-based brain networks. *Frontiers in Neuroscience*, 17, 1177424. <https://doi.org/10.3389/fnins.2023.1177424>
- Wang, J., Wen, S., Liu, W., & others. (2024). Deep joint learning diagnosis of Alzheimer's disease based on multimodal feature fusion. *BioData Mining*, 17, 48. <https://doi.org/10.1186/s13040-024-00395-9>
- World Health Organization. (2023). Dementia: Key facts. <https://www.who.int/news-room/fact-sheets/detail/dementia> access on 28/09/2025
- Yakemow, G., Kolesar, T. A., Beheshti, I., Wright, N., Ryner, L., Chaulk, S., Patel, R., & Ko, J. H. (2024). Investigating neural markers of Alzheimer's disease in posttraumatic stress disorder using machine learning algorithms and magnetic resonance imaging. *Frontiers in Neurology*. <https://doi.org/10.3390/fneur.2024.1470727>
- Yang, X., Wang, Q., Wu, C., Zhang, L., Chen, Y., & Lan, Z. (2025). Predicting conversion of Alzheimer's disease based on multi-modal fusion of neuroimaging and genetic data. *Complex & Intelligent Systems*, 11, 58. <https://doi.org/10.1007/s40747-024-01680-0>
- Yuan, S., Liu, Q., Huang, X., Tan, S., Bai, Z., Yu, J., Lei, F., & Le, H. (2024). Development of an individualized dementia risk prediction model using deep learning survival analysis

incorporating genetic and environmental factors. *Alzheimer's Research & Therapy*.
<https://doi.org/10.1186/s13195-024-01663-w>

Zhang, F., Pan, B., Shao, P., Liu, P., Shen, S., Yao, P., & Xu, R. X. (2022). A single model deep learning approach for Alzheimer's disease diagnosis. *Neuroscience*, 491, 200-214.
<https://doi.org/10.1016/j.neuroscience.2022.03.026>

Zhang, Minyue; Tang, Enze; Ding, Hongwei; Zhang, Yang (2024). AI in communication sciences and disorders (Zhang et al., 2024). *ASHA journals. Dataset*.
<https://doi.org/10.23641/asha.27162564.v1>

LIST OF PUBLICATIONS

1. A research article entitled “**Novel hybrid intelligence model for early Alzheimer's diagnosis utilizing multimodal biomarker fusion**” has been published in the Journal of *Informatics in Medicine Unlocked*. Indexing: Scopus [Q2].
2. A research article entitled “**Predicting Alzheimer's Disease onset: A machine learning framework for early diagnosis using biomarker data**” has been published in the *Computer Methods and Programs in Biomedicine Update*. Indexing: Scopus [Q1].
3. A research article entitled: “**Ethical and regulatory challenges in machine learning-based healthcare systems: A review of implementation barriers and future directions**” has been published in the Journal of *BenchCouncil Transactions on Benchmarks, Standards and Evaluations* [Q1].
4. A research article entitled: **A Hybrid Deep Learning Framework with an Interactive AI Assistant for Early and Interpretable Alzheimer's Disease Detection** has been **communicated** for publication to the journal of “*Informatics in Medicine Unlocked*”. Indexing: Scopus [Q2]. **Status: with editor**

LIST OF CONFERENCES

1. Certificate of presentation on titled: **“Machine Learning and Artificial Intelligence for Enhanced Alzheimer's Disease Detection: A State-of-the-Art Review”** In Proceedings of the IEEE International Conference on Ubiquitous Intelligent Systems (ICUIS 2024), India. Paper published in IEEE Xplore: 13 February 2025
2. Certificate of presentation on titled: **“Dual-Model Framework for Alzheimer's Detection in CT Imaging: A Comparative Study of EfficientNetV2-B0 and ResNet50V2”**. In Proceedings of the International Conference on Emerging Trends in Artificial Intelligence (ASIANCON 2025). Sponsored by IEEE Bombay Section and Hosted by Pimpri Chinchwad College of Engineering and Research. Pune, India
3. Certificate of presentation on titled: **“Bridging the Digital Divide: The Role of Digital Health Literacy in Enhancing Hybrid Machine Learning Models for Early Alzheimer’s Disease Detection”**. International Conference on “Digital Health Literacy among Elderly Populations: Levels, Influencing Factors, Barriers and Health Outcomes” Mittal School of Business, Lovely Professional University, Punjab(February 2025).
4. Certificate of presentation on the topic: **“Transforming Agricultural Supply Chains: The Power of Blockchain Technology”**. In the Proceedings of the International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE 2025), India. Paper published in IEEE Xplore: 14 March 2025
5. Certificate of presentation on the topic: **“Transforming Agriculture with Cloud Computing: Data Processing and Analysis Innovations”**. In the Proceedings of the International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE 2025), India. Paper published in IEEE Xplore: 14 March 2025
6. Certificate of presentation on the topic: **“Integrating Big Data Analytics in Soil Quality Assessment: Techniques, Challenges, and Applications”**. In the Proceedings of the International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE 2025), India. Paper published in IEEE Xplore: 14 March 2025

7. Certificate of presentation on the topic: **“Enhancing Crop Growth Monitoring through Satellite Imagery Analysis: A Precision Agriculture Approach”**. In the Proceedings of the 2nd International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS 2025), India. Paper published in IEEE Xplore: 24 November 2025
8. Certificate of presentation on the topic: **“Artificial Intelligence and the Future of Data Professional: Adapting, Thriving and Leading”** In the Proceedings of the 1st International Conference on Artificial Intelligence and the Future of Data Professional: Adapting, Thriving and Leading. Organized by the Nigerian Society of Data Analysts and Business Intelligence experts (NSDABIE 2025).

LIST OF WORKSHOPS

1. Certificate of participation in the Five – days physical workshop on **“Workshop cum Hands-on Training on Ethics in Biomedical Research”** held between January 13th, 2025, and January 18th, 2025, organized by the School of Pharmaceutical Sciences in collaboration with Human Resource Development Center, Lovely Professional University, India
2. Certificate of participation in the one-day online workshop on **“ChatGPT & AI in Microsoft Office”** held virtually on Thursday, January 9th, 2025, by Skill Nation, India.

APPENDICES

Appendix 1: Data Collection Form

Sample of the data collection form used for clinical data at Shrimann Super Speciality Hospital Jalandhar.

Alzheimer's Research Data Collection Form

Development of a multi-model machine learning-based approach for early detection of Alzheimer's disease

This form is designed to be filled out manually by the patient or the caregiver. Each section captures a different dimension of the dataset, including demographic information, lifestyle factors, medical history, and cognitive assessments, which are crucial for diagnosing Alzheimer's. Each field aligns with the dataset attributes to ensure consistent and accurate data entry.

Patient Information

1. **Patient ID:** -----
2. **Age(yrs):** -----
3. **Gender:**
 - ☐ Male
 - ☐ Female
4. **Ethnicity:**
 - ☐ African
 - ☐ African American
 - ☐ Asian
 - ☐ Other (Specify): _____
5. **Education Level:**
 - ☐ No formal education
 - ☐ Primary school
 - ☐ Secondary school
 - ☐ College or higher
6. **BMI (Body Mass Index):** [15 to 40] []

Lifestyle Factors

7. **Smoking:**
 - ☐ Yes

- ☐ No
- 8. **Alcohol Consumption (Units per week):** [0 to 20] []
- 9. **Physical Activity (Hours per week):** [0 to 10] []
- 10. **Diet Quality (Scale 1-10):** [0 to 10] []
- 11. **Sleep Quality (Scale 1-10):** [4 to 10] []

Medical History

12. Family History of Alzheimer's Disease:

- ☐ Yes
- ☐ No

13. Cardiovascular Disease:

- ☐ Yes
- ☐ No

14. Diabetes:

- ☐ Yes
- ☐ No

15. Depression:

- ☐ Yes
- ☐ No

16. History of Head Injury:

- ☐ Yes
- ☐ No

17. Hypertension:

- ☐ Yes
- ☐ No

Vital Signs and Lab Results/Clinical Measurements

- 18. **Systolic Blood Pressure (mmHg):** [90 to 180] []
- 19. **Diastolic Blood Pressure (mmHg):** [60 to 120] []
- 20. **Total Cholesterol (mg/dL):** [150 to 300] []
- 21. **Low-Density Lipoprotein Cholesterol (mg/dL):** [50 to 200] []
- 22. **HDL Cholesterol (mg/dL):** [20 to 100] []
- 23. **Triglycerides (mg/dL):** [50 to 400] []

Cognitive and Functional Assessments

24. **Mini-Mental State Exam (MMSE) Score:** [0 to 30] []

25. **Functional Assessment Score:** [0 to 10] []

26. **Memory Complaints:**

- ☐ Yes
- ☐ No

27. **Behavioral Problems:**

- ☐ Yes
- ☐ No

28. **Activities of Daily Living (ADL) Score:** [0 to 10] []

Symptoms

29. **Confusion:**

- ☐ Yes
- ☐ No

30. **Disorientation:**

- ☐ Yes
- ☐ No

31. **Personality Changes:**

- ☐ Yes
- ☐ No

32. **Difficulty Completing Tasks:**

- ☐ Yes
- ☐ No

33. **Forgetfulness:**

- ☐ Yes
- ☐ No

Diagnosis and Care

34. **Diagnosis:**

- ☐ No Alzheimer's
- ☐ Alzheimer's Disease

35. **Doctor in Charge:**-----

Appendix 2: Dataset Description

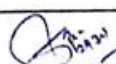
Detailed Meaning of Features in Alzheimer's Disease Datasets Template

Feature	Full Meaning	Relevance to Alzheimer's Disease
PatientID	Unique Patient Identifier	A unique code is assigned to each patient in the dataset for tracking medical records.
Age	Patient's Age (in years)	Advanced age is the strongest risk factor for Alzheimer's disease.
Gender	Biological Sex (Male/Female/Other)	Some studies indicate higher AD prevalence in females.
Ethnicity	Patient's Ethnic Background	Certain ethnic groups have a higher or lower risk of developing AD due to genetic and lifestyle factors.
EducationLevel	Highest Level of Education Attained	Higher education is linked to cognitive resilience and a lower risk of AD.
BMI	Body Mass Index	Obesity in midlife is associated with a higher risk of cognitive decline and AD.
Smoking	Tobacco Use Status (Yes/No/Former)	Smoking increases oxidative stress and vascular damage, contributing to dementia risk.
AlcoholConsumption	Alcohol Intake Frequency (None/Moderate/Heavy)	Excessive alcohol use is linked to brain atrophy and increased AD risk.
PhysicalActivity	Exercise Frequency (Sedentary/Moderate/Active)	Regular exercise helps reduce AD risk by improving brain vascular health.
DietQuality	Nutritional Intake Assessment	Diets rich in antioxidants (e.g., Mediterranean diet) may lower AD risk, while high-fat diets increase risk.
SleepQuality	Restorative Sleep Measurement	Poor sleep quality and sleep disorders (e.g., sleep apnea) are linked to cognitive decline and increased amyloid plaque accumulation.
FamilyHistoryAlzheimers	Family History of AD (Yes/No)	A family history of AD increases the genetic risk, especially with the APOE ε4 allele.
CardiovascularDisease	Presence of Heart Disease	Cardiovascular issues (e.g., atherosclerosis) impair blood flow to the brain, increasing AD risk.
Diabetes	Diabetes Mellitus Diagnosis	Type 2 diabetes is linked to insulin resistance in the brain, contributing to neurodegeneration.
Depression	History of Depression (Yes/No)	Depression is a known risk factor for cognitive decline and dementia.
HeadInjury	History of Traumatic Brain Injury (TBI)	Past head injuries increase the risk of developing AD later in life.
Hypertension	High Blood Pressure Diagnosis	Chronic hypertension can cause brain microvascular damage, leading to cognitive impairment.
SystolicBP	Systolic Blood Pressure (mmHg)	High systolic BP is a known contributor to vascular dementia and AD.
DiastolicBP	Diastolic Blood Pressure (mmHg)	Low diastolic BP may lead to reduced brain perfusion and cognitive decline.
CholesterolTotal	Total Cholesterol Levels (mg/dL)	High cholesterol is linked to AD through vascular and inflammatory mechanisms.

CholesterolLDL	Low-Density Lipoprotein (LDL) "Bad" Cholesterol	High LDL is associated with cerebrovascular issues and increased dementia risk.
CholesterolHDL	High-density lipoprotein (HDL) "Good" Cholesterol	Low HDL may contribute to cognitive decline.
CholesterolTriglycerides	Triglyceride Levels (mg/dL)	High triglycerides are linked to metabolic syndrome and neuroinflammation, increasing AD risk.
MMSE	Mini-Mental State Examination Score	A cognitive screening test for AD; scores below 24 suggest cognitive impairment.
FunctionalAssessment	Cognitive & Daily Function Evaluation	Measures a patient's ability to perform everyday activities, declining as AD progresses.
MemoryComplaints	Self-Reported Memory Issues (Yes/No)	Early signs of cognitive impairment, are often reported before formal AD diagnosis.
BehavioralProblems	Presence of Behavioral Symptoms	Includes aggression, anxiety, hallucinations, and depression, common in AD patients.
ADL	Activities of Daily Living Assessment	Assesses a patient's ability to independently perform daily tasks (e.g., dressing, eating).
Confusion	Episodes of Disorientation & Confusion	A hallmark symptom of AD, worsening as the disease progresses.
Disorientation	Spatial & Temporal Disorientation	Patients struggle with recognizing places, time, and personal identity.
PersonalityChanges	Altered Behavior & Mood Swings	AD patients often exhibit uncharacteristic irritability, aggression, or withdrawal.
DifficultyCompleting Tasks	Challenges in Routine Activities	AD patients struggle with problem-solving and following step-by-step instructions.
Forgetfulness	Frequent Memory Lapses	One of the earliest symptoms of AD, particularly affecting recent memories.
Diagnosis	Final AD Classification	Indicates whether the patient has Alzheimer's Disease (AD) or Normal Cognition (NC).
DoctorInCharge	Assigned Neurologist or Physician	Tracks medical decisions, treatments, and interventions.

Appendix 3: Ethics Approvals

Certificate of Data collection issue by the Center for Research Degree Program, Lovely Professional University, Phagwara, Punjab, India.

 LOVELY PROFESSIONAL UNIVERSITY <i>Transforming Education Transforming India</i>	NAAC A++ Division of Examination Recognized by UGC under Section 2(f) of the University Grants Commission Act, 1956 by the Notification No. F.9-10/ (CPP-I) Member of Association of Indian Universities (AIU), New Delhi									
Sr. No. 105 Code No. CRT206	LPU/10.2/SC/2024/029/250116/0001 Date: 16-Jan-25									
<u>TO WHOM SO EVER IT MAY CONCERN</u>										
Name of the Student Mr.	Shehu Mohammed									
Son of Mr.	Mohammed									
Registration No.	12325250									
Name of the Programme	Doctor of Philosophy (Computer Applications)									
Batch	2023									
Faculty	Lovely Faculty of Technology & Sciences									
This is to certify that Mr. Shehu Mohammed registration no. 12325250 is pursuing Ph.D. in Computer Applications under the supervision of Dr. Neha Malhotra, Associate Professor, Lovely Professional University. His research topic is "Development of A Multi-Model Machine Learning Based Approach for Early Detection of Alzheimer Disease". To carry out his research work he intends to visit following institution/organization for the purpose of data collection which may be allowed as per the rules and regulations established by your institution/organization.										
<table border="1"><thead><tr><th>Sr.No.</th><th>Name of Institution/Organization</th><th>Location</th></tr></thead><tbody><tr><td>1</td><td>Shrimann Superspeciality Hospital, Jalandhar</td><td>Pathankot Road Near Reru, Chowk, Jalandhar, Punjab</td></tr><tr><td>2</td><td>Federal Medical Centre Gusau</td><td>Zamfara, Nigeria</td></tr></tbody></table>		Sr.No.	Name of Institution/Organization	Location	1	Shrimann Superspeciality Hospital, Jalandhar	Pathankot Road Near Reru, Chowk, Jalandhar, Punjab	2	Federal Medical Centre Gusau	Zamfara, Nigeria
Sr.No.	Name of Institution/Organization	Location								
1	Shrimann Superspeciality Hospital, Jalandhar	Pathankot Road Near Reru, Chowk, Jalandhar, Punjab								
2	Federal Medical Centre Gusau	Zamfara, Nigeria								
 Deputy Registrar (Records) Deputy Registrar (Records) Lovely Professional University										
 Prepared by	 Verified by									
 Checked by										
Disclaimer: This certificate is issued on the basis of information available in the office of records on the date of its issue and the University reserves the right to										

Certificate of participation in workshop cum Hands-on Training on Ethics in Biomedical research, organized by the School of Pharmaceutical Sciences in collaboration with the Human Resource Development Center, Lovely Professional University

	  HUMAN RESOURCE DEVELOPMENT CENTER		Certificate No. 368966
Certificate of Participation			
This is to certify that Mr. Shehu Mohammed S/o Mohammed participated in Workshop cum Hands-on Training on Ethics in Biomedical Research organized by School of Pharmaceutical Sciences in collaboration with Human Resource Development Center, Lovely Professional University w.e.f. January 13, 2025 to January 18, 2025 and obtained "A" Grade.			
Date of Issue : 06-03-2025 Place : Phagwara (Punjab), India	 Prepared by (Administrative Officer-Records)	 Head of School School of Pharmaceutical Sciences	 Head of Division Human Resource Development Center

Institutional Ethics Committee, Lovely Professional University, approval certificate



**INSTITUTIONAL ETHICS
COMMITTEE (IEC)**

(Registration No: EC/NEW/INST/2022/3110)

Ref: LPU/IEC-LPU/2025/1/2

Date: 15.02.25

Dear Shelu Mohammed

The Institutional Ethics Committee, Lovely Professional University (IEC-LPU) reviewed and discussed your application to conduct the study entitled "Development of a Multi-Model Machine Learning Based Approach for Early detection of Alzheimer's Disease"

The following documents were reviewed:

- (a) Application form
- (b) Research Protocol
- (c) Principal investigator's current Curriculum Vitae.

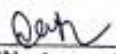
The following members of the ethics committee were present at the meeting held on 29.01.25

Sr. No	Name	Designation as per IEC
1	Dr. Shivani Tandon	Chairperson
2	Dr. Navneet Khurana	Member Secretary
3	Dr. Kanav Khara	Alternate Member Secretary
4	Dr. Anil Malhotra	Clinician-I
5	Dr. Naresh Kundra	Clinician-II
6	Mr. Sumesh Mathew	Social scientist/Philosopher/Ethicist/Theologian
7	Dr. Rajan Kumar Sharma	Lay person

The Ethics committee **approved** the study to be conducted in its presented form subject to the condition of

- Approval of the ethics committee where the study is being conducted
- The principal investigator and all members of the project shall ensure compliance with current regulatory provisions (as per New Drug Clinical Trial Rules 2019 and ICH-GCP), Ethical Guidelines for Biomedical Research on Human Participants by ICMR, and the standard operating procedure of IEC including timely submission of Interim Annual Report and Final Closure Report.
- Participant Information Sheet and a copy of signed Informed Consent shall be given to every research participant.
- Inform IEC in case of any proposed amendments (change in protocol/procedure, site / Investigator, etc).
- Inform IEC immediately in case of any adverse events and serious adverse events.
- Members of the IEC have the right to monitor any project with prior intimation.

Yours sincerely,


Chairperson/Member secretary/Alternate member secretary, IEC-LPU

15/2/25
dd/mm/yyyy

Member Secretary
Institutional Ethics Committee
Lovely Professional University
Phagwara-Punjab-144411



Jalandhar-Delhi G.T. Road, Phagwara, Punjab (India) - 144411
Ph: +91-1824-444039 E-mail: ms_ieclpu@lpu.co.in website: www.lpu.in

Permission Letter for Conducting Research at Shrimann Superspeciality Hospital,
Jalandhar

 Shrimann SUPERSPECIALITY HOSPITAL —Reliable & Transparent—	
<u>Dated 26th March 2025</u>	
To The Deputy Registrar(Records) Lovely Professional University Phagwara-144411	
<u>Subject: Permission Letter for conducting Research Mr. Shehu Mohammed</u>	
This is to certify that Mr. Shehu Mohammed S/o Mr. Mohammed, a student of Lovely Professional University with Regd. No. 12325250, undergoing a research project for Ph.D. (Computer Applications) with the title "Development of A Multi-Model Machine Learning based approach for Early Detection of Alzheimer Disease" He is hereby permitted to get data from the hospital to carry out his research work under the guidance of Dr. Anmol Singh Rai, Consultant, Neurologist provided an approval from the Ethics Committee and MoU signed from Lovely Professional University, Phagwara has to be submitted within 15 days of receiving this document.	
 Aman Madhok Assistant General Manager Department of Human Resources	
Shrimann Superspeciality Hospital (A Unit of Shriman Enterprises) Jalandhar Pathankot Road, Near Reru Chowk, Jalandhar 144012, Punjab. Tel: 0181-5017777, 81451-81451 www.shrimannhospitals.com	

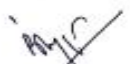
Recommendation/certification letter for conducting the research and developing an AI assisted tool in medical field by the Shrimann Superspeciality hospital

	 Shrimann SUPERSPECIALITY HOSPITAL (A Fortis Network Hospital) NHS-44, Jalandhar Pathankot Road, Near Reru Chowk, Jalandhar, Punjab - 144012 (India) Tel : 0181 501 7777 Emergency: 81451 81451 Email : contactus@fortishealthcare.com Website : www.fortishealthcare.com
---	--

Dated : 25th November 2025

TO WHOM IT MAY CONCERN

This is to certify that **Mr. Shehu Mohammed, S/o Mr. Mohammed**, a research scholar of **Lovely Professional University, Punjab** with Regd. No. 12325250 has conducted significant research during his Ph. D (Computer Applications) in regards to his research project with title **"Development of A Multi-Model Machine Learning based approach for Early Detection of Alzheimer Disease"**. His work has practical implication in the medical field. The Artificial Intelligence based system he developed, assists doctors in diagnosing **Alzheimer's Disease**.

 Dr. Anmol Singh Rai Consultant-Neurology	 Dr. Radha Raman Sagar Medical Superintendent
--	--

FORTIS HOSPOTEL LIMITED

Regd. Office: Fortis Memorial Research Institute, Sector-44, Near Metro Station, Gurgaon, Haryana - 122002 (India)
Tel: +91 124 438 6666, Fax: +91 124 496 2222, E-mail: secretary@fortishealthcare.com, Web: www.fortishealthcare.com, CIN: U74899HR130021C054270

Appendix 4: Sites Visitation for Data Collection

Evidence of Images during data collection at Radiology Department of Shrimaan Superspeciality Hospital in Jalandhar





Appendix 5: Proof of All Journal Publications

Images of all the three published papers in Scopus Indexed journals.

Informatics in Medicine Unlocked 57 (2025) 101668



ELSEVIER

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu





Novel hybrid intelligence model for early Alzheimer's diagnosis utilizing multimodal biomarker fusion

Shehu Mohammed^a, Neha Malhotra^a, Arun Singh^a, Awad M. Awadelkarim^b, Shakeel Ahmed^c, Saiprasad Potharaju^{d,*}

^a School of Computer Applications, Lovely Professional University, Phagwara, 144411, India
^b Department of Information Technology, Faculty of Computers and Information Technology, University of Tabuk, Tabuk, Saudi Arabia
^c School of Computer Science (SCS), Taylor's University, Subang Jaya, 47500, Malaysia
^d Department of CSE, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India

ARTICLE INFO

Keywords:
Alzheimer's disease diagnosis
Hybrid machine learning model
Gradient boosting machines (GBM)
Deep neural networks (DNN)
Feature selection
Biomarker identification

ABSTRACT

One of the significant causes of dementia and a leading peril to global public health is Alzheimer's disease (AD), which calls for early and accurate diagnosis. The paper proposes a novel hybrid machine learning framework that integrates Gradient Boosting Machine (GBM) and Deep Neural Networks (DNN) for predicting Alzheimer's disease from multimodal biomarkers. The database comprises 35 demographic, behavioral, and clinical features. Feature selection procedures produced key predicting variables (i.e., MMSE scores, performance in Activities of Daily Living (ADL), cholesterol level, and behavior problems). A hybrid model was created by combining individual models, and it proved to be the most effective compared to particular models, achieving 92.6 % accuracy and a 0.94 AUC score on the database. The synergy between the capability of GBM for tabular data and the ability of DNN for complex interaction gives a good outcome. The research demonstrates the efficacy of blending machine learning techniques for supporting Alzheimer's disease (AD) identification and provides a method for early identification at a broader level. It is hoped that more biomarkers will be incorporated, and the model will be validated on larger and more phenotypically diverse databases to achieve clinical usability and generalizability.



Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Computer Methods and Programs in Biomedicine Update

journal homepage: www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine-update



Predicting Alzheimer's Disease onset: A machine learning framework for early diagnosis using biomarker data

Shehu Mohammed^{*}, Neha Malhotra

Department of Computer Applications, Lovely Professional University Phagwara, India

ARTICLE INFO

Keywords:

Alzheimer's Disease (AD)
Early diagnosis
Multimodal biomarkers
Machine learning
Deep learning
Neuroimaging

ABSTRACT

Alzheimer's Disease (AD) is a significant global health issue, and the current diagnostic techniques cannot diagnose the disease at its early stages, hence the difficulty of early therapeutic management. In response to the formulated research problem, this study articulates a new multimodal machine-learning framework for early AD diagnosis. The main goal is to combine multiple biomarkers: neuroimaging, CSF, genetic, and longitudinal cognitive data and develop a robust model for accurate early AD diagnosis. The importance of this work is in the opportunity to shift diagnostic paradigms by employing deep learning algorithms, including CNNs, LSTM networks, and GNNs to analyze spatial, temporal, and relational patterns across multi-modal data. The methodology involves federated learning and domain adaptation with GANs to integrate data from multiple centers with the patient's privacy intact. It shows that the proposed multimodal model is superior to single-modality models with an AUC-ROC of 0.94 and reveals that hippocampal volume and plasma p-tau are the most informative biomarkers in the early diagnosis of AD. The study's implications indicate that combining multimodal data improves diagnostic accuracy and clinical relevance by providing a roadmap to developing personalized medicine and better patient care. Future work will be aimed at increasing the variability of the dataset and the clinical trials to test the model to improve its applicability and performance in actual practice.



Contents lists available at ScienceDirect

BenchCouncil Transactions on Benchmarks, Standards and Evaluations

journal homepage: www.keaipublishing.com/en/journals/benchcouncil-transactions-onbenchmarks-standards-and-evaluations/



Review Article

Ethical and regulatory challenges in machine learning-based healthcare systems: A review of implementation barriers and future directions[☆]

Shehu Mohammed^{*,} Neha Malhotra

School of Computer Applications, Lovely Professional University, Jalandhar, 144111, India



ARTICLE INFO

Keywords:

Artificial intelligence (AI) Ethics
Algorithmic bias
Explainable AI (XAI)
Machine learning (ML) in healthcare
Patient data privacy
Regulatory compliance (GDPR, HIPAA, FDA)

ABSTRACT

Machine learning significantly enhances clinical decision-making quality, directly impacting patient care with early diagnosis, personalized treatment, and predictive analytics. Nonetheless, the increasing proliferation of such ML applications in practice raises potential ethical and regulatory obstacles that may prevent their widespread adoption in healthcare. Key issues concern patient data privacy, algorithmic bias, absence of transparency, and ambiguous legal liability. Fortunately, regulations like the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), and the FDA AI/ML guidance have raised important ways of addressing things like fairness, explainability, legal compliance, etc.; however, the landscape is far from risk-free. AI liability is another one of the gray areas approaching black, worrying about who is liable for an AI medical error — the developers, the physicians, or the institutions. The study reviews ethical risks and potential opportunities, as well as regulatory frameworks and emerging challenges in AI-driven healthcare. It proposes solutions to reduce bias, improve transparency, and enhance legal accountability. This research addresses these challenges to support the safe, fair, and effective deployment of ML-based systems in clinical practice, guaranteeing that patients can trust, regulators can approve, and healthcare can use them.

Introduction

Background: Machine learning (ML) is fundamentally transforming healthcare; ML is playing an integral role in the development of methods for early diagnosis and treatment optimization, as well as predictive analytics providing unprecedented improvements for medical decision-

intertwine with these dynamics, with ML models trained on unbalanced datasets delivering results that discriminate against certain groups [5]. Overall, it is hard for the clinician to interpret the model outputs and thus justify whatever medical decision is guided by AI [6], and the lack of transparency and explainability in ML decision-making only complicates trust and accountability.

Appendix 6: Proof of Conference Papers Publications

Images of two conference papers published in IEEE Xplore indexed Scopus journal.

The screenshot shows the IEEE Xplore digital library interface. At the top, there is a navigation bar with links to IEEE.org, IEEE Xplore, IEEE SA, IEEE Spectrum, and More Sites. On the right, there are links for Subscribe, Donate, Cart, Create Account, and Personal Sign In. Below this, the IEEE Xplore logo is on the left, and a search bar with a dropdown menu set to 'All' is in the center. To the right of the search bar is an 'Institutional Sign In' button. Below the search bar, there is a link to 'ADVANCED SEARCH'.

The main content area displays the article title: "Machine Learning and Artificial Intelligence for Enhanced Alzheimer's Disease Detection: A State-of-the-Art Review". Below the title, it indicates the publisher is IEEE and provides links to 'Cite This' and 'PDF'. The authors listed are Shehu Mohammed and Neha Malhotra, with a link to 'All Authors'. On the left side of the article, there is a sidebar with a '47 Full Text Views' counter and a list of document sections: I. Introduction, II. Scope of Machine Learning, III. Literature Review, IV. Related Work Comparison, and V. Our Findings and Discussion. Below the sections, there are links for 'Show Full Outline', 'Authors', 'Figures', 'References', 'Keywords', and 'Metrics'.

The abstract section is titled 'Abstract:' and contains the following text: "Alzheimer's disease (AD) is a substantial public health challenge due to its rising prevalence on a global scale. Timely identification and tailored treatment are essential for enhancing patient results and maximizing healthcare resources. This paper offers a thorough examination of the latest progress in applying machine learning (ML) and artificial intelligence (AI) methods to improve Alzheimer's disease (AD) detection and customize care tactics. We will first discuss the etiology of Alzheimer's disease and emphasize the need of early identification. We will explore different machine learning and artificial intelligence techniques such as deep learning, support vector machines, and ensemble approaches used for Alzheimer's disease detection and prognosis. We investigate combining many types of data sources, including neuroimaging, genetic markers, and clinical assessments, to improve predictive models. We also examine how AI-driven decision support systems help create tailored care plans and optimize treatment for people with Alzheimer's disease. We discuss present obstacles and prospective paths in the area, such as data diversity, model explainability, and ethical concerns. This study seeks to offer insights into the current state of machine learning (ML) and artificial intelligence (AI) applications in Alzheimer's disease (AD) research and clinical practice, with the goal of advancing early diagnosis and individualized management strategies for this severe neurological ailment."

Below the abstract, it states 'Published in: 2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)'. It also provides the 'Date of Conference: 12-13 December 2024', 'DOI: 10.1109/ICUIS64676.2024.10867088', 'Date Added to IEEE Xplore: 13 February 2025', 'Publisher: IEEE', and 'Conference Location: Gobichettipalayam, India'.

On the right side of the article, there is a promotional banner for 'Need Full-Text' access to IEEE Xplore for organizations, with a 'CONTACT IEEE TO SUBSCRIBE' button. Below this, there is a 'More Like This' section with two recommendations: 'Detecting Alzheimer's Disease (AD) Using Machine Learning Algorithms' (published 2024) and 'Predicting Alzheimer's Disease Progression through Machine Learning Algorithms' (published 2023). At the bottom right, there is a small image with the text 'A Step-by'.

Q

[ADVANCED SEARCH](#)

[Conferences > 2025 5th Asian Conference on ...](#)[Back to Results](#) | [Next >](#) PDF

Shehu Mohammed : Neha Malhotra : Sourabh Kumar All Authors

Text Views

CONTACT IEEE TO SUBSCRIBE >

Abstract:

Document Sections

- I. Introduction
- II. Related Work
- III. Proposed Methodology and Model Architecture
- IV. Experimental Results and Model Evaluation
- V. Conclusion

Authors

Figures

References

Keywords

Metrics

AD is a neurodegenerative disease that gradually impairs memory and cognitive function, requiring prompt and accurate recognition for effective management. Traditional radiological methods are effective but limited by subjective interpretation and time-consuming analysis, especially when examining large quantities of brain CT scans. This paper introduces a two-model deep learning framework that includes EfficientNetV2-B0 and ResNet50V2 to automate Alzheimer's Disease classification based on CT images, outperforming previous methods. Both models are trained on a carefully curated dataset of labeled brain CT images and utilize advanced transfer learning with preprocessing and data augmentation to ensure robust learning. The EfficientNetV2-B0 model is adapted with compound scaling to optimize accuracy and computational efficiency, while ResNet50V2 uses deep residual learning to detect subtle pathology patterns in the CT data. Visual interpretability is enhanced through Grad-CAM heatmaps, allowing medical practitioners to identify and verify decision regions and gain insights into each classification. Experimental results demonstrate that these models outperform other state-of-the-art approaches such as MobileNetV2, VGG16, and DenseNet121 across metrics like classification accuracy, precision, recall, F1-score, and AUC. Furthermore, these models are suitable for clinical and resource-limited settings due to their fast and stable inference speeds. The findings confirm that the hybrid deep learning approach utilizing EfficientNetV2-B0 and ResNet50V2 can serve as an effective and interpretable method for detecting Alzheimer's disease via CT imaging.

Published in: 2025 5th Asian Conference on Innovation in Technology (ASIANCON)

Date of Conference: 22-23 August 2025

DOI: 10.1109/ASIANCON66527.2025.11281242

Date Added to IEEE Xplore: 16 December 2025

Publisher: IFFF

▼ ISBN Information:

Conference Location: PIMPRI, India

I. Introduction

One of the most common types of neurodegenerative disease is Alzheimer's Disease (AD), which causes around 60 to 70 percent of all dementia in the world. Marked by gradual memory loss, poor decision making, and decline in cognitive ability, it inflicts a huge emotional and financial burden on families and health care systems. Timely and effective diagnosis is very important because it enables patients and caregivers to discuss treatment

More Like This

Early Detection of Alzheimer's Disease Using Deep Learning and MRI: A CNN-Based Approach with Enhanced Data Augmentation

2025 3rd International Conference on
Foundation and Large Language Models
(FLLM)

Published: 2025

CNN and Transfer Learning-based Deep Learning

Architectures for Alzheimer's Disease Detection from MRI

Scan: A Comparative Analysis

2024 IEEE International Conference on
Biomedical Engineering, Computer and
Information Technology for Health
(BECITHCON)

Published: 2024

[Show More](#)

IEEE Authors:

Increase

Appendix 7: Conference Presentation Certificates

Images of the three conference Papers presented at different International conferences.





IEEE BOMBAY
SECTION

**2025 5th Asian Conference on Innovation in Technology
(ASIANCON)**

22nd – 23rd August 2025

Certificate

*This is to certify that Dr./Prof./Mr./Ms. **Shehu Mohammed** has presented paper entitled **Dual-Model Framework for Alzheimer's Detection in CT Imaging: A Comparative Study of EfficientNetV2-B0 and ResNet50V2** in 2025 5th Asian Conference on Innovation in Technology (ASIANCON) during 22nd & 23rd August 2025.*

Mrs. V.S. Kumbhar
Publication Chair

Prof. Dr. R.G. Mapari
Convener

Prof. Dr. H.U. Tiwari
General Chair



MITTAL
SCHOOL OF BUSINESS



Certificate No. 371478

Certificate of Presentation

This is to certify that Dr./Mr./Ms. **Shehu Mohammed and Neha Malhotra** has presented a paper on **"Bridging the Digital Divide: The Role of Digital Health Literacy in Enhancing Hybrid Machine Learning Models for Early Alzheimer's Disease Detection"** in International Conference on **"Digital Health Literacy among Elderly Populations: Levels, Influencing Factors, Barriers and Health Outcomes,"** held on **February 15, 2025**, organized by Mittal School of Business at Lovely Professional University, Punjab.

Date of Issue : 22-03-2025
Place : Phagwara (Punjab), India

Prepared by
(Administrative Officer-Records)

Convener
Mittal School of Business, LPU

Head of school
Mittal School of Business, LPU

Appendix 8: Certificates of Training Workshops

Images of the two training workshops attended.





Scan To Verify

CERTIFICATE OF COMPLETION

PROUDLY PRESENTED TO

Shehu Mohammed

FOR COMPLETING THE **CERTIFIED WORKSHOP ON**



**ChatGPT & AI in
Microsoft Office**

A handwritten signature in black ink, appearing to read "Jatan Shah".

Conducted By

Mr. Jatan Shah

Microsoft Certified Trainer
Co-Founder & CEO, Skill Nation

Date of workshop:

**9 January,
2025**

www.jatanshah.com