

أحمد لطف

Early Prediction of Alzheimer%E2%80%99s Disease using Deep Learning and Data Fusion.docx

 Midocean University

Document Details

Submission ID

trn:oid::2945:404560988

Submission Date

Aug 3, 2026, 5:15 PM GMT+3

Download Date

Aug 3, 2026, 8:14 PM GMT+3

File Name

Early Prediction of Alzheimer's Disease using Deep Learning and Data Fusion.docx

File Size

3.6 MB

58 Pages

10,956 Words

62,944 Characters

16% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

-  **172 Not Cited or Quoted 15%**
Matches with neither in-text citation nor quotation marks
-  **7 Missing Quotations 1%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 9%  Internet sources
- 10%  Publications
- 12%  Submitted works (Student Papers)

Integrity Flags

1 Integrity Flag for Review

-  **Hidden Text**
440 suspect characters on 9 pages
Text is altered to blend into the white background of the document.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 172 Not Cited or Quoted 15%**
Matches with neither in-text citation nor quotation marks
- 7 Missing Quotations 1%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 9% Internet sources
- 10% Publications
- 12% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

- Submitted works**
University of Hertfordshire on 2024-09-02 <1%
- Submitted works**
Midocean University on 2025-09-05 <1%
- Internet**
mdpi-res.com <1%
- Internet**
arxiv.org <1%
- Internet**
www.mdpi.com <1%
- Internet**
pmc.ncbi.nlm.nih.gov <1%
- Internet**
www.frontiersin.org <1%
- Submitted works**
Liverpool John Moores University on 2025-12-14 <1%
- Submitted works**
University of Westminster on 2025-04-16 <1%
- Submitted works**
Liverpool John Moores University on 2026-04-20 <1%

11	Submitted works	British University in Egypt on 2026-04-16	<1%
12	Submitted works	Central Institute of Technology Kokrajhar on 2026-07-06	<1%
13	Internet	hdl.handle.net	<1%
14	Publication	Vijay Jeyakumar, N. Venkateswaran, Dinesh Bhatia. "Machine Learning for Biome...	<1%
15	Submitted works	Erasmus University Rotterdam on 2026-05-04	<1%
16	Submitted works	Midocean University on 2025-09-05	<1%
17	Internet	dspace.bracu.ac.bd	<1%
18	Publication	"Soft Computing: Theories and Applications", Springer Science and Business Medi...	<1%
19	Publication	Jaskaran Singh, Meenu Gupta, Rakesh Kumar. "Smart Technologies and Intelligen...	<1%
20	Internet	link.springer.com	<1%
21	Publication	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challeng...	<1%
22	Internet	dl.lib.uom.lk	<1%
23	Internet	www.researchgate.net	<1%
24	Submitted works	Colorado Technical University Online on 2026-06-21	<1%

25	Submitted works	Electrical & Computer Engineering on 2025-11-14	<1%
26	Submitted works	SRM University AP Amravati on 2026-08-02	<1%
27	Publication	Muhammad Irfan, Seyed Shahrestani, Mahmoud Elkhodr. "Enhancing Early Deme...	<1%
28	Submitted works	UNICAF on 2025-11-13	<1%
29	Internet	alzres.biomedcentral.com	<1%
30	Internet	webmail.thescipub.com	<1%
31	Publication	Mario Merone, Sebastian Luca D'Addario, Pierandrea Mirino, Francesca Bertino et...	<1%
32	Submitted works	Shaheed Benazir Bhutto University, Shaheed Benazirabad on 2026-06-03	<1%
33	Submitted works	University of East London on 2023-09-11	<1%
34	Internet	www.ijraset.com	<1%
35	Internet	www.spiedigitallibrary.org	<1%
36	Publication	Paolo Ferro, Harinadh Vemanaboina, Chander Prakash. "Computational Techniqu...	<1%
37	Publication	Saurav Mallik, Soumita Seth, Ben Othman Soufiene. "HealthTech Horizons - Chart...	<1%
38	Internet	constellation.uqac.ca	<1%

39	Internet	sutir.sut.ac.th	<1%
40	Internet	theses.lib.polyu.edu.hk	<1%
41	Internet	umpir.ump.edu.my	<1%
42	Internet	www.biospectrumasia.com	<1%
43	Internet	etd.aau.edu.et	<1%
44	Internet	mSPACE.lib.umanitoba.ca	<1%
45	Submitted works	Anna University on 2026-03-10	<1%
46	Publication	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dharendra Kumar Shukla. "Intelli...	<1%
47	Publication	Mandeep Kaur Arora, Jaspreet Singh, Anju Singh. "Development of intelligent sys...	<1%
48	Publication	Wusat Ullah, Qun Dai, Rana Muhammad Zulqarnain. "An Alzheimer's diagnosis te...	<1%
49	Internet	our.oakland.edu	<1%
50	Submitted works	GIFT University on 2026-07-13	<1%
51	Publication	Kokkula Lokesh, Nagendra Panini Challa, Abbaraju Sai Satwik, Jinka Chandra Kira...	<1%
52	Submitted works	Poornima University on 2026-07-30	<1%

53	Publication	V. Sekhar, P. Ajay Kumar Reddy, K. Logesh, T. Raghunadha Reddy, Dara Raju. "Ne...	<1%
54	Internet	api.mountainscholar.org	<1%
55	Internet	xmed.jmir.org	<1%
56	Submitted works	De Montfort University on 2025-05-30	<1%
57	Submitted works	KoE on 2026-06-12	<1%
58	Submitted works	University of Dundee on 2026-05-29	<1%
59	Internet	braininformatics.springeropen.com	<1%
60	Internet	theses.whiterose.ac.uk	<1%
61	Internet	repositorio-iul.iscte.pt	<1%
62	Publication	Asaf Raza, Huma Ayub, Javed Ali Khan, Ijaz Ahmad et al. "A Hybrid Deep Learning...	<1%
63	Publication	Javier Augusto Jurado Moncada. "Aplicação da técnica de contraste de fase da ord...	<1%
64	Submitted works	Siksha 'O' Anusandhan University on 2026-05-29	<1%
65	Submitted works	University of Hertfordshire on 2025-01-05	<1%
66	Submitted works	University of Sheffield on 2026-05-11	<1%

67	Internet	combinatorialpress.com	<1%
68	Publication	"Brain Informatics", Springer Science and Business Media LLC, 2025	<1%
69	Submitted works	Asian Institute of Technology on 2017-05-02	<1%
70	Publication	Besma Mabrouk, Nadia Bouattour, Noura Mabrouki, Lamia Sellami, Ahmed Ben H...	<1%
71	Submitted works	Centre for Technical Support on 2025-11-08	<1%
72	Submitted works	College of Engineering and IT on 2026-05-10	<1%
73	Submitted works	Department of Computer Science and Engineering (CSE) on 2026-06-10	<1%
74	Publication	Poonam Nandal, Tapas Kumar, Meeta Singh, Mamta Dahiya. "Progressive Compu...	<1%
75	Publication	Tariq Mahmood, Amjad Rehman, Tanzila Saba, Yu Wang, Faten S. Alamri. "Alzhei...	<1%
76	Submitted works	University of Bolton on 2026-01-13	<1%
77	Submitted works	University of East London on 2026-05-05	<1%
78	Submitted works	University of Wales Swansea on 2024-09-29	<1%
79	Publication	Yashvi Chandola, Vivek Uniyal, Yamini Bachheti. "Efficient Chest X-Ray Feature Ex...	<1%
80	Submitted works	cvrcollegehyderabad on 2024-03-07	<1%

81	Internet	download.bibis.ir	<1%
82	Internet	journals.savba.sk	<1%
83	Submitted works	universititeknologimara on 2026-07-09	<1%
84	Internet	worldwidescience.org	<1%
85	Internet	www.researchsquare.com	<1%
86	Publication	Arslan Abbas, Hsin-Chi Tsai, Yu-Ling Hsu, Mutebi John Kenneth, Bashir Hussain, Li...	<1%
87	Publication	Bal S. Virdee, Tanweer Ali, Jaume Anguera, Suman Lata Tripathy. "Connecting Int...	<1%
88	Publication	David Olatunde Olukanni, Oluseyi Olanrewaju Ajayi, Joshua Olusegun Okeniyi, Sh...	<1%
89	Submitted works	Dublin Business School on 2025-08-27	<1%
90	Submitted works	Liberty University on 2026-07-14	<1%
91	Publication	Linah M. Elnaghi, Yasmine M. Eltariny. "Evaluation of Deep Learning Models on AI...	<1%
92	Publication	Mohammed G. Alsubaie, Suhui Luo, Kamran Shaukat, Weijia Zhang. "A Novel Alz...	<1%
93	Submitted works	Queensland University of Technology on 2017-06-02	<1%
94	Publication	Samia Ishaque, Carlos Ordonez, Jinan A. Fiaidhi. "Exploration into Predicting Alzh...	<1%

95	Publication	Sarita Simaiya, Umesh Kumar Lilhore, Yogesh Kumar Sharma, Leela Prasad Gorre...	<1%
96	Submitted works	The University of the West of Scotland on 2025-04-21	<1%
97	Submitted works	University of Malakand, Chakdara on 2026-02-03	<1%
98	Submitted works	University of Portsmouth on 2024-05-22	<1%
99	Publication	Zhang, Haobai. "Exploring Mathematics Development Using Early Childhood Lon...	<1%
100	Internet	acikerisim.bahcesehir.edu.tr:8080	<1%
101	Internet	dspace.bracu.ac.bd:8080	<1%
102	Internet	dspace.daffodilvarsity.edu.bd:8080	<1%
103	Internet	eprints.gouni.edu.ng	<1%
104	Internet	ieomsociety.org	<1%
105	Internet	ir-library.ku.ac.ke	<1%
106	Internet	journal.50sea.com	<1%
107	Internet	repository.aaup.edu	<1%
108	Internet	sci-hub.red	<1%

109	Internet	www.it.iitb.ac.in	<1%
110	Publication	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dharendra Kumar Shukla. "Artific...	<1%
111	Submitted works	Otto-von-Guericke-Universität Magdeburg on 2026-01-16	<1%
112	Submitted works	Queen Mary and Westfield College on 2024-08-23	<1%
113	Publication	Udayakumar P, R. Subhashini. "Connectome Graph Attention Network (CGAT): A ...	<1%
114	Submitted works	Universiti Tun Hussein Onn Malaysia on 2026-05-25	<1%
115	Submitted works	University of Hertfordshire on 2025-01-04	<1%
116	Publication	Yuru Li, Xiaowei Chang, Jianlin Wu, Yuchen Liu, Hailu Wang, Yiyin Zhang. "Machin...	<1%
117	Publication	Adarsh Valoor, G.R. Gangadharan. "Unveiling the decision making process in Alzh...	<1%
118	Publication	Emrullah ŞAHİN, Naciye Nur Arslan, Durmuş Özdemir. "Unlocking the black box: ...	<1%
119	Publication	Kamlesh Kumar Pandey, Abhay Mishra, Ram Milan. "Predictive Modeling Approac...	<1%
120	Submitted works	Liverpool John Moores University on 2022-08-30	<1%
121	Publication	Rajaa Daami Resen, Laith Sabah Alzubaidi, Haider A. Alwzwazy, Manuel I. Capel. "...	<1%
122	Publication	Sakshi Gupta, Umesh Gupta, Moolchand Sharma, Kamal Malik. "Generative Intelli...	<1%

123	Publication	Trideva Sastri Koduru. "Nucleic Acids and Precision Therapies for Neuroprotectio...	<1%
124	Submitted works	UCSI KL Campus on 2025-08-29	<1%
125	Submitted works	Universiti Teknikal Malaysia Melaka - Library on 2026-06-22	<1%
126	Submitted works	University of Sydney on 2013-11-15	<1%
127	Publication	Bhabani Prasanna Pattanaik, Dharendra Kumar Shukla, Sambit Satpathy, Kamala...	<1%



Early Prediction of Alzheimer's Disease using Deep Learning and Data Fusion

Prepared By

Name: Ahmed Naji Lutf.

ID No.: 253000445

Name: Mohamed Amin Miqdad

ID No.: 252001179

Submitted in fulfillment of the requirements for the degree Master of Science in Artificial Intelligence

At

College of Informatics

Midocean University

Supervisor Name: Dr. Mohamed Essa

Collage of Informatics, Midocean University, Comoros

Email: 252001179@st.midocean.edu.km / 253000445@st.midocean.edu.km

Acknowledgements

We are truly grateful to our supervisor Dr. Mohamed Essa for his invaluable instructions and support. His insightful feedback helped us a lot to improve ourselves through this research.

We would like to thank the faculty members at the College of Informatics, Midocean University, for providing a consistence guidance and support which helped us.

Our gratitude to the Alzheimer's Disease Neuroimaging Initiative (ADNI) for providing us with required datasets that helped us with this research.

Finally, we would like to express our appreciation to our families for their patience and support.

Abstract

Detection of Alzheimer's disease (AD) is one of the toughest cognitive challenges in medical field (Hazarika et al., 2023). When disease is detected early this will often make a big difference in the treatment. The issue is that many current diagnostic models look at the problem from one angle which could lead to wrong decisions because it relies only to one single source of data, such as brain imaging without look to clinical data or vice versa, Because of this one-sided approach, doctors can miss the detection of disease symptoms ,especially during the first stages when shifting from a normal condition (NC) to mild cognitive impairment (MCI), before it developed to full Alzheimer's (AD). To deal with this limitation, this study presents approach that combines deep learning with data fusion. The framework examines 3D MRI scans with clinical data from the ADNI dataset. We will build a two-path model: one path uses a 3D-ResNet network to automatically pull-out visual features from the images, and the other relies on deep neural networks to process clinical data. Integration of MRI images and clinical data will give the model additional details to patient's condition. The approach was tested on Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (N = 296). The test show strong performance result with overall accuracy of 97.30%, a 97.20% F1-score, and a 0.9995 ROC-AUC and for better understanding we applied 3D Gradient-weighted Class Activation Mapping (Grad-CAM) to detect the brain regions that affected, especially the hippocampus, which is associated with Alzheimer's disease to improve early detection accuracy.

Keywords: Alzheimer's Disease, Deep Learning, Data Fusion, CNN, 3D-ResNet, Early Prediction, Medical Imaging, Explainable AI.

Table of Contents

i

Acknowledgements	ii
Abstract.....	iii
Table of Contents	iv
List of Abbreviations	vi
List of Figures.....	vii
Chapter 1: Introduction	1
Introduction	Error! Bookmark not defined.
1.1 Problem Definition.....	2
1.2 Thesis Objectives	3
1.3 Thesis Contribution.....	3
1.4 Thesis Organization	4
Chapter 2: Related Work.....	5
2.1 Introduction	5
2.2: Background	6
2.3 Related Work	10
2.3 Related Work Summary	15
Chapter 3: Proposed Methodology	18
3.1 Introduction	18
3.2 Dataset profile	19
3.2.1 ADNI Dataset.....	19
3.2.2 Dataset splits	19
3.3 Data Preprocessing Pipeline	20
3.4 Classification Models: Architectural Evolution	20
3.4.1 The Baseline 3D-CNN Model.....	21
3.4.2 The 3D VGG-Inspired Architecture	22

iv

3.3.4 Residual Network with 18 layers (3D ResNet-18)	23
3.5 Multimodal Fusion Framework	24
3.6 Classification Head	25
3.5 Training Protocol	27
3.6 Evaluation Metrics	28
3.4.1 Classification Accuracy	28
3.4.2 Precision (Positive Predictive Value)	28
3.4.3 Recall (Sensitivity).....	28
3.4.4 F1-Score	29
Chapter 4: Results and Discussion.....	30
4.1 Introduction	30
4.2 Preprocessing and Volumetric Standardization Analysis	30
4.2 Progressive Architecture Evaluation	32
4.3.1 Result of Baseline 3D CNN.....	33
4.3.2 Result of VGG-16	35
4.3.3 Result of Resnet-18.....	37
4.3.4 The Multimodal Cross-Modal Fusion (Proposed Solution):	39
4.3.5 Visual Interpretability via 3D Grad-CAM.....	42
4.3.6 Overall Comparative Analysis	43
Chapter 5: Conclusion And Future Work	46
5.1. Conclusion	46
5.4 Future Work	47
References.....	48

List of Abbreviations

- 3D CNN: Three-Dimensional Convolutional Neural Network
- AD: Alzheimer's Disease
- ADNI: Alzheimer's Disease Neuroimaging Initiative
- AI: Artificial Intelligence
- AUC-ROC: Area Under the Receiver Operating Characteristic Curve
- DL: Deep Learning
- FN: False Negative
- FP: False Positive
- Grad-CAM: Gradient-Weighted Class Activation Mapping
- GPU: Graphics Processing Unit
- MCI: Mild Cognitive Impairment
- MMSE: Mini-Mental State Examination
- MRI: Magnetic Resonance Imaging
- NC: Normal Control
- PET: Positron Emission Tomography
- ResNet: Residual Network
- sMRI: Structural Magnetic Resonance Imaging
- TN: True Negative
- TP: True Positive
- VGG: Visual Geometry Group
- ViT: Vision Transformer
- XAI: Explainable Artificial Intelligence

List of Figures

8	Figure 1.1. MRI images of Normal Cognitive (Row 1), Mild Cognitive Impairment (row 2) and last is Alzheimer's Disease2
14	Figure 2.1. ANN machine learning model.....7
	Figure 2.2. CNN model extracting data from MRI scan.....8
	Figure 2.3. VGG16 feature extractor and classifier to help for brain diagnoses.....9
	Figure 2.4. Unimodal 3D images into 3D CNN10
98	Figure 3.1. Overall framework of the proposed system integrating MRI classification and clinical data in fusion modal for AD detection.....18
120	Figure 3.2. Baseline CNN architecture for brain tumor classification.21
112	Figure 3.3. General Architecture of a VGG-16 Style CNN for Early Alzheimer's Disease Detection.....23
121	Figure 3.4. 3D ResNet-18 for Early Alzheimer's Disease Detection.]24
	Figure 4.1. show MRI sample images before and after preprocessing.....31
	Figure 4.2. show clinical data samples after preprocessing.....32
103	Figure 4.3. CNN baseline confusion matrix34
	Figure 4.4. CNN Loss curve (left) accuracy curve(right).....34
	Figure 4.5. VGG-16 confusion matrix36
	Figure 4.6. 3D VGG network curves during training.36
	Figure 4.7. 3D ResNet-18 confusion matrix38
	Figure 4.8. 3D ResNet-18 Loss (left) vs accuracy (right).....39
2	Figure 4.9. Multimodal confusion matrix40
	Figure 4.10. Multimodal loss (left) vs accuracy(right)41
4	Figure 4.11. Receiver Operating Characteristic (ROC) curves and corresponding Area Under the Curve (AUC) scores across the target diagnostic domains (NC, MCI, AD).....41
	Figure 4.12. Native Volumetric Cross Sections Across Diagnostic Classes42

List of Tables

Table 2.1. Summary and Comparison of Related Studies on Alzheimer’s Disease 15

Table 3.1. ADNI dataset splits and distribution19

Table 3.2. Hyperparameter configuration and training protocols across the evaluated networks.27

Table 4.1. Performance Comparison Across Architecture Evolution.....32

Table 4.2. Performance Metrics of the 3D CNN MRI-Only Baseline Model33

Table 4.3. Performance Metrics of VGG-16 Model35

Table 4.4. Performance Metrics of the Standalone Pretrained 3D ResNet-18 Model37

Table 4.5. Performance Metrics of multimodal fusion39

Table 4.6. Comprehensive Performance Benchmark, Structural Limitations, and Direct Architectural Comparisons43

Chapter 1 : Introduction

1.1 Overview

Alzheimer's disease is one of the most difficult brain conditions. While it is more common in older people, it can affect individuals at different stages of life, including those in their younger years. It slowly destroys memory and thinking skills, and in many cases, it is only discovered after the person shows clear symptoms. However, research has shown that changes in the brain start many years before these symptoms appear. This is why finding the disease early is the main goal of medical research today. For example, [1] showed that we can catch these early signs by identifying brain changes over time.

Within time, AI is getting more and more attention in every field and medical field as well for helping doctors, researchers and patients to give more accurate results about brain scans, especially MRIs, considered very helpful because they give a big and clear picture of the brain. According to that, recent papers by [2] and [3] take advantage of AI progress to scans and to put together some really accurate models to spot the disease. But worth noting that a significant portion of past research has focused on unimodal data, such as analyzing only MRIs. Although it is good to at certain degree, this approach fails to capture the full scope of Alzheimer's evolving condition. Recent research proves that combining different sources such as merging brain scans with a patient's clinical records and physical symptoms [4] leads to a much more precise and reliable diagnosis.

In this thesis, we propose an AI system for early diagnosis that looks at both brain changes and medical records over time. We use data from the ADNI, which is a famous collection of medical information used by many top researchers, including [5] and [6].

The main goal of this work is to see if combining brain images and clinical memory tests can help us find the disease in earlier stages which aims to give a clear and easy-to-understand tool to help patients as early as possible.

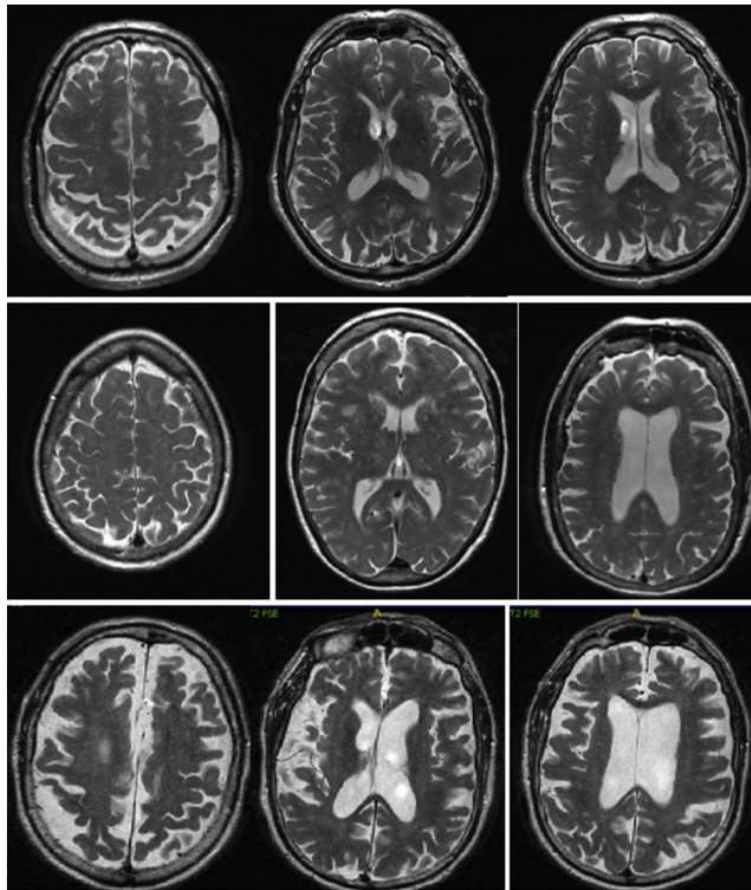


Figure 1.1. MRI images of Normal Cognitive (Row 1), Mild Cognitive Impairment (row 2) and last is Alzheimer's Disease

1.2 Problem Definition

Despite the large number of studies applying deep learning to Alzheimer's diagnosis, several important limitations remain.

First, many approaches focus exclusively on imaging data, particularly structural MRI. While MRI can reveal brain atrophy and structural abnormalities, it does not incorporate molecular or genetic information that may indicate susceptibility before structural damage becomes severe.

Second, both neuroimaging and genomic datasets are high-dimensional. Imaging volumes contain millions of voxels, and genomic profiles may include thousands of features. Directly training deep

learning models on such data can lead to overfitting if not handled properly. And for that we must have a good method for feature extraction and data size reducing.

The main issue with Alzheimer's is the changing withing the time, and most models just deal with data as one picture. They ignore the long-term changes of the disease and how it grows over time. Because of that, they will miss some clues that show at the very beginning of disease symptoms.

On top of all that, merging different types of data together which is known as (multimodal fusion) is challenging tasks. And There is no clear approach on how to mix brain scans with genetic data. Which lead us to find a solution for building the network that contain multi data type to do a give an excellent results one tested patients.

1.3 Thesis Objectives

The purpose of this research is to build and test an AI system that can read different data types to catch Alzheimer's in early stages. To do that we will apply the following steps:

- To create a deep learning model to import and MRI scans and extract important details.
- To develop multimodal fusion to mix the genetic data with brain pictures.
- To test and track the changes over time
- To check if our proposed framework give a better result by make a comparison with others that look only at one kind of data.

1.4 Thesis Contribution

The aim of this thesis is to contribute to the field of Alzheimer's disease by improving the early detection by implementing the following techniques:

- Deep learning model that handles different data type such as MRI images and clinical data for better diagnostic decisions
- A multi computer vision algorithms including baseline CNN and 3D ResNet-18 model for better comparing and classification.
- A framework to achieve better computing power performance with large scale of data sets.
- Using of Explainable AI (XAI) techniques to help to define and explain the exact brain areas as the computer do in order to increase the transpierce for doctors and their patients

39

1.5 Thesis Organization

The remainder of thesis organized as follows

Chapter 2: a discussion about history of Alzheimer's diagnosis over years

Chapter 3: present the proposed framework and its workflow

Chapter 4: analyze and cover testing phase and results.

109

Chapter 5: summarizes the study and the future direction of work.

Chapter 2 : Related Work

2.1 Introduction

With artificial intelligence (AI) things has changed in general not to mention the methods around how we diagnose brain diseases. In this chapter we will reviews past studies on detecting Alzheimer's Disease (AD). We will see how things shifted from simple practices to modern deep learning [6,7]. The goal here is to take a look and find out the problems in past researches with knowing why and try overcome with better ways, generally the major studies use **only one type of data at a time** which is not enough when you dealing with a human brain and with taking this into consideration a need of different data types (data fusion) is a real medical need, not just an extra . study [8] show that mixing 3D MRI scans with basic clinical records gave much better and increase the accuracy result. Combining this will lead to better tracking for a patient as they go from a healthy state (**Normal Control, NC**) to **Mild Cognitive Impairment (MCI)**, and finally to **AD**.

Diagnosing **AD** relied a lot on manual checks and the doctor's personal opinion this was situation before using the computers. Doctors usually checked memory loss using simple **tests like the Mini-Mental State Examination (MMSE)**. Even after using MRI scans, doctors had to look at many 2D image slices by eye to find changes one the brains, like brain shrinking [9,10]. Doing this manually cost a lot of time and resources, not only that but doctors can have a different opinion for same things they have looked for. And, as human cannot see the tiny brain changes with just our own eyes due to our limitations and for sure cannot see the tiny brain **changes that happen in the very early stages of the disease**. And the damage was already done withing the time and wrong diagnoses. Because of this delay it became more and more important to look for better and automated tools for medical field.

In order to understand what we mean to these tools, we need to look back at **traditional Machine Learning (ML)** and shifting to **Deep Learning (DL)**. Traditional ML is good in some case like basic clinical data. But it is not good when come to deal with medical images because it required humans interactions to pick the important features first [11]. To overcome with this limitations deep learning fixes this by automatically fetch an important details from the raw 3D data. And can deal with simple

shapes and move to large scale complex disease patterns. This automation means no humans interaction required to do the manual work, making DL a best choice for modern MRI scans [12].

2.2: Background

In the last two decades, the diagnosing for Alzheimer's disease changed a lot. At the beginning finding the small, early signs of the disease is still was a hard task for computers, but things getting easier withing time for doctors, researchers and patients due to evolving of tools that used machine learning and deep learning to help analyze patient data.

Firstly, machine learning methods required many manual work steps. Including collecting and data cleansing of raw data such as MRIs or genomic data after that researchers have to split the data. Then, with help human experts they need pick out important features, like tissue texture, brightness, or brain shape. Once these features extracted, they used standard models like Support Vector Machines or Decision Trees for learning and recognize the patterns.

These approaches have been used in the diagnosis of brain disorders, including Alzheimer's disease, where performance is often influenced by the quality of the extracted features [13,16].

Study [13] explored the use of machine learning to predict whether patients with mild cognitive impairment (MCI) may progress to Alzheimer's disease (AD). It employed XGBoost as a classification model and incorporated the Shapley value (SHAP) method to provide insights into feature contributions, including MRI and genetic data.

Studies [14–16] investigated different machine learning pipelines for Alzheimer's disease classification and prediction. Study [14] followed a supervised learning approach in which multiple classifiers, including Support Vector Machine, Random Forest, and Artificial Neural Network, were trained and compared using structured feature sets derived from non-medical and behavioral data. In the case of the Artificial Neural Network, the input layer received a fixed-length feature vector representing the selected variables, which were then processed through hidden layers to model non-linear relationships before producing the final output. The techniques used pre-processing, feature selection, and model evaluation. Study [15] improve this by using multimodal data and explainable AI to be able to know how the model makes its decisions. While study [16] focused on feature

extraction from images. They selected some details from brain scans and used them as inputs for models classification.

96

generally, these studies follow the same basic steps: pre-processing the data, extracting features, training the model, and checking the results. But things changed when researchers pay attention to mixing different data types (multimodal integration) to ensure that we can have better understand of how model thinking (interpretability).

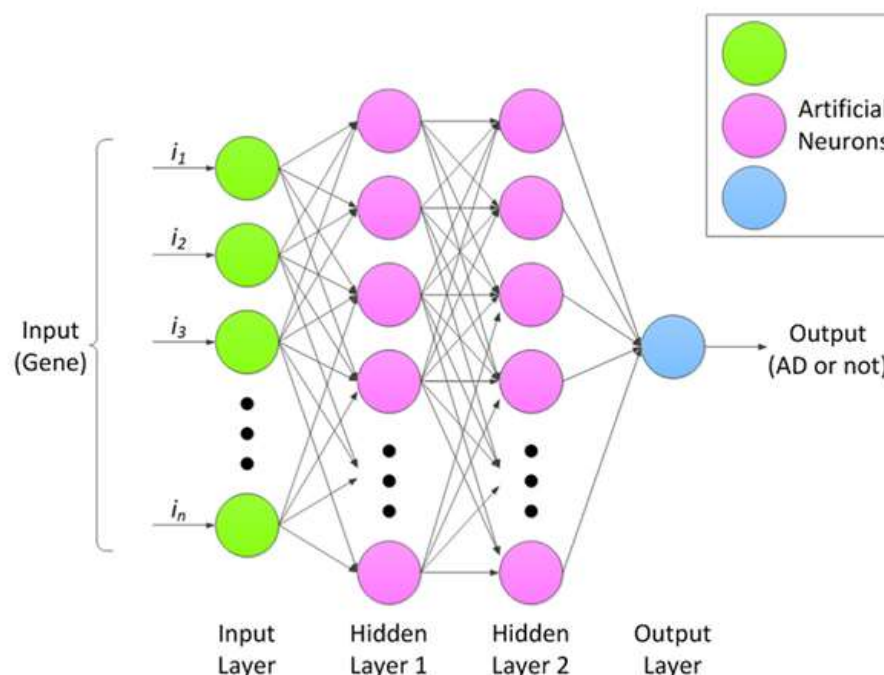


Figure 2.1.ANN machine learning model. Adopted by [14]

Although machine learning provided a big leap in detecting Alzheimer's, working directly with raw 3D brain scans has a limitation problem. Extracting good features required a lot human interactions to deal with brain areas not only this but flatten a complex 3D brain into a simple data table caused an important spatial detail to be lost.

To solve these extraction issues, the researchers used a deep learning computer vision, especially Convolutional Neural Networks (CNNs). Without relying on human input features, CNNs find a way to visualize patterns directly from the MRI data. By passing the adjusted scans through convolutional and pooling layers, these networks handle everything from basic lines to complex brain damage.

And for getting better accuracy research seeks to using deeper networks for better result, like AlexNet which changed how we classify medical images. They proved that having deeper layer such as ReLU activations and dropout makes it much easier for the model to recognize important features.

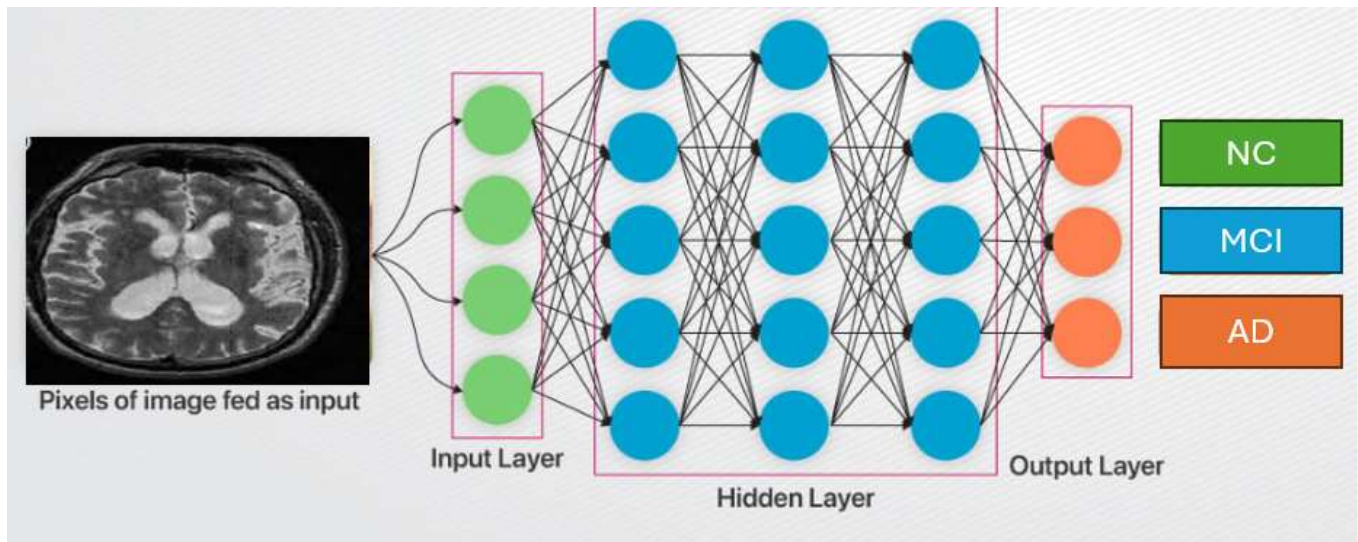


Figure 2.2. CNN model extracting data from MRI scan

Following these early models, VGG-16 became a very popular network due to its effective algorithms. It uses small filters to pull a complex visual detail. It can learn features step by step by moving from simple to complex brain tissues. In Alzheimer's research, VGG-16 is often used with a technique called transfer learning. Researchers take models that are already trained on huge datasets like ImageNet and fine-tune them to help for MRI scans.

Nowadays, VGG-16 is mainly used as a simple feature extractor. It finds a features and passed it to fully connected layers then to classifiers to group patients into different stages. Recent studies show that using this network with data augmentation and normalization will greatly improves the model's accuracy with better data handling [17].

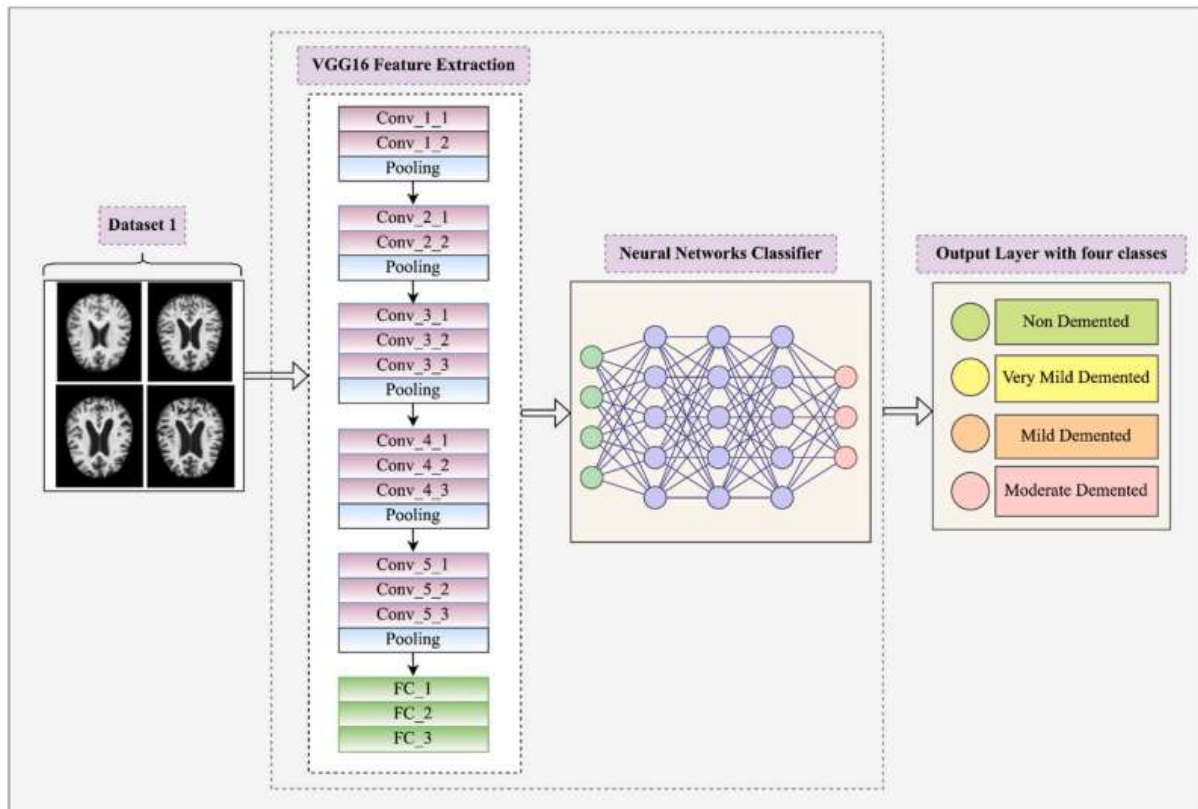


Figure 2.3. VGG16 feature extractor and classifier to help for brain diagnoses. Adopted by [17]

While 2D CNNs was helpful but using them for 3D MRI scans lead use to real challenges. In order to make these models work, the 3D brain scan has to be cut into flat 2D slices. But slicing images will breaks the natural connections between different brain regions. As a result, the model loses a lot of important spatial information, which at the end lead to make it difficult to understand the brain full structure.

To solve this, researchers moved toward 3D-CNNs. Instead of using flat pixels, these networks used and process voxels (3D pixels). This allows the model to take in the entire MRI scan at once, preserving the actual shape and structural details of the brain [18].

However, even with these improved 3D models, a major clinical issue sill exist. Almost all of these systems were unimodal, meaning they only looked at one type of information. For example, a deep learning model might be perfect at analyzing an MRI, but it completely ignores the patient's clinical history or the model might predict the disease using clinical history but never actually look at the physical brain. This is not how doctors work in real life , they did not diagnose a patient based on just one test. Instead, they look at the brain scan alongside the patient's age, education level, and cognitive tests like the MMSE [19,21].

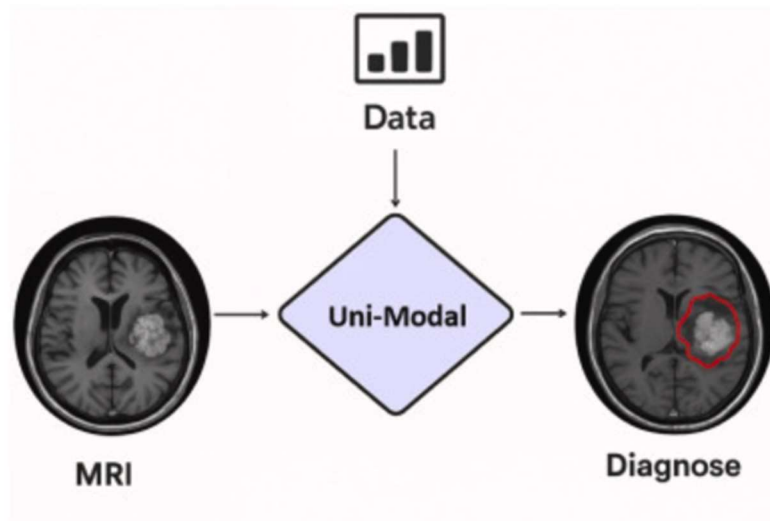


Figure 2.4. Unimodal 3D images into 3D CNN

Put all this history together the hard manual work of early machine learning, the challenges of cutting 3D images into 2D, and how 3D models can only use one type of data , we will clearly see a big missing part in the current studies. There is a need for an AI system that actually supposed to work like doctors in clinics. We want a model that take a look at full 3D brain scans properly, and at the very same time, read the patient's medical records over time. Combining these different kinds of information together into one single system is simply the best way to make early diagnosis much better.

2.3 Related Work

Muksimova et al (2025) proposed a multi-modal deep learning framework called FusionNet for the early detection and classification of Alzheimer's disease. The model integrates deep learning techniques: a 3D CNN with an LSTM network to track brain changes over time. Using the ADNI dataset, the study combined MRI, PET, and CT scans to train the system effectively. The proposed framework achieved a high diagnostic accuracy of 94%. However, its main limitations are the high computational power required for processing 3D multi-modal scans, and the practical difficulty and high cost of obtaining all three types of brain scans for a single patient in standard clinical settings.

Li et al (2025). presented a multi-modal deep learning framework to improve the early detection of Alzheimer's disease. The researchers focus on integration on two main deep learning techniques: a Convolutional Neural Network (CNN) to analyze brain images, and a Transformer model to process

genetic data, combined via a cross-transformer fusion. The study used the ADNI dataset, including MRI scans and genetic records from 1,541 patients across different stages: healthy (NC), stable memory loss (sMCI), progressive memory loss (pMCI), and Alzheimer's disease (AD). This system achieved a high accuracy of 93.04%. However, its main limitations are the use of a basic CNN model that restricts performance and the difficulty of mixing different genetic records.

Bricaud and Masala (2026) presented a comprehensive framework to evaluate various deep learning models for detecting early Alzheimer's disease. Instead of relying on a single network, the researchers compared several different methods, including traditional CNNs and modern Vision Transformers. They used the ADNI data set for MRI scans and grouping the patients into normal and Alzheimer's disease they used individual 2D slices and entire 3D, this hybrid model (ConvNeXtV2-L) achieved 91.07% as accuracy, however the main limitation with the study the is relies fully on MRI images without using of patient's clinical records.

Bekhet et al. (2026) introduced a deep learning framework named (DeepALZNET) for the early prediction and detection of Alzheimer's disease. The researchers mainly used two independent techniques: the first one is 1D Convolutional Neural Network (CNN) and the second one is a Random Forest model, they used both models to analyze patient clinical records, and to detect changes in brain MRI scans by using VGG19 network. The study used large public datasets, specifically ADNI and OASIS to categorize the patients into three stages: normal cognitive (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD). The framework achieved over 90% as accuracy. However, its main limitations are the combine images and clinical data to unified decision, and it relies on 2D images which might miss some 3D details of the brain.

Zhang et al. (2023) Proposed a framework to improve the early detection of Alzheimer's disease by Neural Network (GNN) as multi model. Instead of analyzing flat images, they used an advanced GNN deep learning techniques to understand the network connections and relationships between different brain regions. The study used the ADNI dataset and combine brain scans (MRI and PET) for a more advanced analysis. Patients categorized into four main developmental stages: healthy (NC), early memory loss (EMCI), late memory loss (LMCI), and Alzheimer's disease (AD). The proposed framework achieved a 92% accuracy in detecting the disease. Nevertheless, the study main

9

limitations related to high clinical cost for patients to go both for MRI and PET scans and the heavy computational power needed for each individual patient.

6

108

1

Helaly et al. (2021) used a deep learning framework called (E2AD2C) to detect and classify of Alzheimer's disease in early stages. The methods used in the study focused mainly on using simple computer networks (CNN) and a more advanced, pre-trained system called VGG19. The study used ADNI dataset to brain scan images, covering four different stages: healthy (NC), early memory loss (EMCI), late memory loss (LMCI), and Alzheimer's disease (AD). And for accuracy the study reached to 95.17% for using the simple computer, and accuracy of 97%. for advanced VGG19 model. Yet the study drawbacks are the high computer power needed for the 3D brain scans without considering to use other patient data like memory tests or age.

75

21

Warren and Moustafa (2026) in their study they used a multimodal scan system that combines functional brain imaging (fMRI) with clinical tests to identify Alzheimer's disease. The proposed method uses a 3D Convolutional Neural Network (3D CNN) that analyzes a brain region that could be affected early in the disease. To make the model more trustworthy, the researchers enhanced the fMRI scans with five clinical memory tests to prevent any data errors. A key part of their method is the use of "explainable AI" (XAI) through a technique called perturbation ranking, which helps in the examination. The study selected a small group of 52 participants from the ADNI database. The accuracy was 90%, which is much better than the 58% accuracy reached when using brain scans, even so mixing the brain data with clinical tests give a better accuracy but research dealt with very small sample size and need to be tested on large patients' data set.

73

122

Wang et al. (2024) In this paper, the authors developed a deep learning model to catch Alzheimer's early by mixing data from different data sources. For brain changes, they used a ResNet to pull features from specific brain areas, and with help of an attention convolution module they have a close look to a sick spot. For achieving this the ADNI data set has been using to mix MRI brain scans with genetic data to create special 3D images. The patients have been split into three groups: NC, MCI, and AD. By using this approach, they reached to 93.44% accuracy. Still the main issues are building 3D image fusions required a lot of computing power with noting the difficulty of getting the clinic data and ask every single patient to do both an advanced MRI and a full genetic test.

118

Alsubaie et al. (2026) built a new deep learning model called BiFPN3DViT to find Alzheimer's using brain pictures. The framework basically mixed a 3D CNN to grab the basic brain shapes, a Bi-directional Feature Pyramid Network (BiFPN) network to link details of different sizes, and Vision Transformers to get the whole brain picture. They tested on large scale of data set from ADNI, nearly 4,706 MRI scans, which grouped by healthy people, patients with mild memory loss, and Alzheimer's cases. By putting these together, their model hit a 92.1% accuracy in sorting the three disease stages. But the main downside of this model is that depend only on MRI scans without taking into consideration the patient's medical files.

Ahmed et al (2023) tried to spot Alzheimer's by looking at brain pictures. They built a Deep CNN model to put people into three groups, completely healthy, having mild memory problems, and actual Alzheimer's patients. They used a regular MRI scan from ADNI data set to train their network. The model worked very well and getting a 93.86% accuracy but it has a limitation because it only takes a look to brain scans and ignores the patient's medical files and without Explainable AI, doctors facing a challenge of figuring out why the computer made its choice.

71

Lu et al. (2022) used a deep learning system to check for Alzheimer's by depend only on brain pictures. The system collected a huge number of MRI scan, approximately over 85,000 from a lot of different hospitals, including the ADNI dataset. They used a transfer learning, which means they took an Inception-ResNet-V2 network that was already trained on simple stuff, and just did a little modification on it to find Alzheimer's. Because they trained it on so many images, the model got an accuracy between 90.9% and 94.5% across different test groups. Nonetheless the big downsides related to the dependency of scans without looking to the patients' medical files.

Khalid et al (2026) found out that depending on MRI only is not enough to detect Alzheimer's , so they came up with dual-modal approach. With using data set pulled from OASIS-3. And then using a CNN to work with MRI images and for clinical data they decided to go with Random Forest model. This helped them to combine brain scans with clinical symptoms to give a full profile insight. Because of dual-model approach they achieved 94%. Above all of that they used also Explainable AI to map out the exact damaged areas, like the hippocampus. Yet the study downside is not using a single

unified model with fused clinical and MRI data instead they rely totally on two separate models for processing the MRI scan and clinical data.

Sheikh et al. (2025) focused on detecting early Alzheimer's in fast and low-cost using brain pictures. While many deep learning models are heavy and require a lot of computing powers, the researchers decided to focus on lightweight networks, specifically EfficientNetV2B0 and MobileNetV2 which allow clinic data to be processed efficiently. They took flat 2D slices from ADNI MRI scans to sort patients into three stages: totally normal, early mild memory loss, and late mild memory loss. The best their model get is 88.0% accuracy. They also added Explainable AI to highlight the same brain areas the computer was looking at and to increase the result trust for doctors. Although the study shows that small models can be useful but the main issues here are that they only used flat 2D slices instead of the whole 3D brain shape with ignoring of the patients' clinical history.

Ghosh et al. (2023) came up with a safe way for hospitals to work together to find Alzheimer's using brain scans. To overcome the issue of patient's data distributed in different places and the privacy concerns, they used something called Federated Learning. Using this they allow multiple clinics to train a small network like MobileNet together without sharing any private pictures. They used flat 2D MRI scans from ADNI and OASIS data sets. This helped them reach 95.24% accuracy on the OASIS data, and 84% accuracy when mixed both datasets. Although the approach gives a good accuracy, the main drawback is that study using only flat 2D slices instead of the whole 3D brain and ignored the patients' medical files totally in the study.

Abbasian and Hammond (2024) presented a 3D deep learning model to detect Alzheimer's in early stages. By using brain pictures, they categories patients' brains between healthy or early mild cognitive impairment (EMCI) only. And with help of 3D-CNN they look at the whole 3D brain all at once instead of slicing it into flat 2D slices. They used ADNI dataset for 3D MRI scans. The research succeeds in reaching 81.80% accuracy in separating the early-stage patients from the healthy ones. This gives an indication that using a full 3D brain gives more details than 2D methods. However the study main limitations are, it only looks at two groups (early sick versus normal) and it only uses MRI scans while ignoring the other clinical data.

2.3 Related Work Summary

The above fifteen reviewed studies used variant Deep learning methods for Alzheimer's daises which show promising indication for Deep learning usage in medical field generally and brain daises specifically like AD detection and classifications. And in terms of high classification accuracy some studies reach to 94% in accuracy and others up to 97%. However, despite the high performance of DL techniques in the field as shown in the studies but there are limitations like lack of clinical data in some studies or small dataset in others not to mention the high computer power required in some studies as well. Table 2.1 will summarize each proposed method, accuracy, data set and the main limitations.

Table 2.1. Summary and Comparison of Related Studies on Alzheimer's Disease

Study (Year)	Proposed Method	Dataset	Key Accuracy	Main Limitations
Muksimova et al. (2025)	CNN + LTMS	OASIS	94.00%	Fuses images only; completely lacks integration of clinical medical records.
Li et al. (2025)	CNN + Transformer	ADNI	93.04%	Struggles with MCI-to-AD prediction.
Bricaud & Masala (2026)	Hybrid CNN (ConvNeXtV2-L)	ADNI	91.07%	MRI-only; lacks integration of clinical medical records.

Bekhet et al. (2025)	1D CNN + Random Forest	OASIS	> 90.0%	No simultaneous fusion; relies on 2D images.
Zhang, Y., et al. (2023)	GNN framework	ADNI	92%	Highly computational cost with required of both MRI + PET scans
Helaly et al. (2021)	E2AD2C (CNN & VGG19)	ADNI	97.0%	High power for 3D scans; images-only.
Warren & Moustafa (2026)	3D CNN with XAI	ADNI	90.0%	Very small sample size; lacks validation.
Wang, J., et al. (2024)	ResNet + Attentation convolution	ADNI	93.44%	Highly computational cost with required of both MRI + PET scans
Alsubaie et al. (2026)	BiFPN3DViT (3D CNN + BiFPN + ViT)	ADNI	92.10%	MRI-only; lacks integration of clinical medical records.

Ahmed et al. (2023)	Deep CNN	OASIS	93.86%	MRI-only; lacks Explainable AI tools.
Lu et al. (2022)	Transfer Learning (Inception-ResNet-V2)	ADNI	94.50%	MRI-only; data mostly from Western populations.
Khalid et al. (2026)	CNN & Random Forest	ADNI	94.00%	Processes images and clinical data separately; lacks a fully unified fusion network.
Sheikh et al. (2025)	Lightweight CNN (EfficientNetV2B0) & Explainable AI	Kaggle Alzheimer's MRI Dataset	88.00%	2D images only; lacks integration of clinical medical records.
Ghosh et al. (2023)	Federated Learning & MobileNet	Kaggle Alzheimer's MRI Dataset	95.24%	2D images only; lacks integration of clinical medical records.
Abbasian & Hammond (2024)	3D CNN	ADNI	81.80%	MRI-only; lacks clinical data integration.

41

Chapter 3 : Proposed Methodology

3.1 Introduction

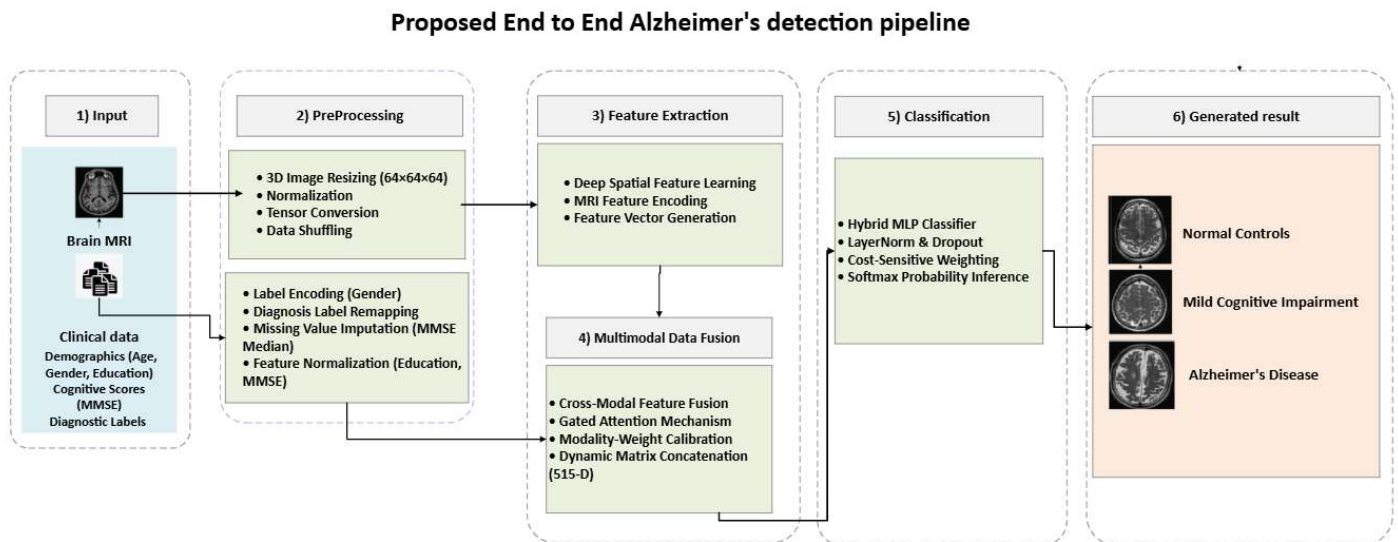
70

125

1

This chapter describes the research methodology used in this study to develop a multi-class classification framework for Alzheimer's disease. The proposed approach focuses on identifying diagnosis based on three categories: Normal Control (NC), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). The accuracy of system diagnosis depends on using an effective learning model together with properly prepared data. This chapter explains how our proposed framework pipelines work side by side. We built a dual stream model that combines brain images with simple clinical records. The system has four main parts. First, we collect and organize the dataset brain images and clinical data. Second, we preprocess both types of data. Third, we design the model to extract a feature from MRI and classify them using baseline CNN, vgg-16 and resnet-18 after that we used resnet-18 results and fuse it with clinical data then Finally, we use a gated attention classifier to get the results. Figure 3.1 shows the full structure of this multimodal system

69



9

Figure 3.1. Overall framework of the proposed system integrating MRI classification and clinical data in fusion modal for AD detection.

3.2 Dataset profile

3.2.1 ADNI Dataset

The data set in this research was obtained from **Alzheimer's Disease Neuroimaging Initiative (ADNI)**, ADNI is globally recognized and widely used as a benchmark dataset for **Alzheimer's Disease research**. It provide high-quality structural Magnetic Resonance Imaging (sMRI) scans besides that it also provides a rich clinical profile. This dataset selected because of the quality of the data which helps the proposed model to achieve better and more reliable results.

3.2.2 Dataset splits

The framework used dataset from the **Alzheimer's Disease Neuroimaging Initiative (ADNI)**. In total, we used 1,971 clinical records and their matching MRI files. We split this data into three parts. We used 70% of the data for training the model. Then, we used 15% for validation. We kept the final 15% for independent testing. **Table 3.1** summarizes the dataset distribution and splits.

Table 3.1. ADNI dataset splits and distribution

Medical Group (Diagnosis)	Data Split	Number of Patients (N)
Normal Control (NC)	Training (70%)	335
Normal Control (NC)	Validation (15%)	72
Normal Control (NC)	Testing (15%)	72
Early MCI (EMCI)	Training (70%)	536
Early MCI (EMCI)	Validation (15%)	115
Early MCI (EMCI)	Testing (15%)	115
Alzheimer's (AD)	Training (70%)	508
Alzheimer's (AD)	Validation (15%)	109
Alzheimer's (AD)	Testing (15%)	109
Total Cohort	Cumulative	1971

3.3 Data Preprocessing Pipeline

ADNI has two raw data sets the first is MRI images and the second is clinical data the main format of MRI come with uncompressed format (.nii) with $196 \times 256 \times 256$ dimensions. Processing these massive 3D matrices directly during the training loop creates will cause a computational bottleneck and GPU memory exhaustion. To overcome this constraint without losing any important feature we have to Resizing the native $196 \times 256 \times 256$ to $64 \times 64 \times 64$ voxels. This resolution compression is implemented via 3D Cubic Spline Interpolation algorithm ($order = 3$) and for Nan Value Eradication (NaN or $\pm\infty$). The pipeline converts any invalid value to absolute zero (0.0), protecting the model from generating breaking errors during backpropagation, and since MRI scans in the ADNI dataset were acquired using different scanners, intensity values not equal across images. To reduce these variations, Min–Max normalization was applied independently to each 3D brain volume. In addition to the MRI preprocessing pipeline, the patients' clinical data were prepared before merging it with images feature. The gender variable was converted into numerical values suitable for model training. Also Missing values in the Mini-Mental State Examination scores replaced using the median value calculated from the diagnostic group in the training set to avoid data leakage. In addition, the numerical features like education normalized to a common range of 0–1 to ensure consistency and balanced learning across all input features. Finally, the training data were randomly shuffled at the beginning of each epoch to prevent the model from learning patterns.

3.4 Classification Models: Architectural Evolution

To identify the most suitable architecture for classifying three-dimensional structural brain MRI volumes, three deep learning models were evaluated. A baseline 3D-CNN was first implemented to extract volumetric spatial features, but its limited depth resulted in poor feature representation and a bias toward the majority class. A 3D VGG network was then examined using stacked $3 \times 3 \times 3$ convolutional layers; however, the absence of skip connections led to the vanishing gradient problem during training. These limitations lead to adopt of a residual learning architecture Resnet-18 where residual connections improve gradient propagation, stabilize optimization, and enable more effective feature learning in deep networks.

3.4.1 The Baseline 3D-CNN Model

Baseline 3D-CNN architecture presented in Figure 3.2 extends conventional two-dimensional convolution to 3D volumetric brain MRI analysis. The input scan is represented as a tensor ($X \in \mathbb{R}^{H \times W \times D \times C}$), where (H), (W), and (D) denote the spatial dimensions and (C) is the number of input channels. The forward pass of the 3D convolutional layer is computed as:

$$\left[v_{l,v}^{x,y,z} = f \left(b_{l,v} + \sum_m \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} \sum_{r=0}^{R-1} w_{l,v,m}^{p,q,r} \cdot v_{l-1,m}^{x+p,y+q,z+r} \right) \right] \quad (1)$$

In Equation (1) where ($f(\cdot)$) is the ReLU activation function, ($b_{l,v}$) is the bias term, and ($w_{l,v,m}^{p,q,r}$) represents the trainable kernel weights. Spatial resolution is then reduced using the 3D Max-Pooling operation:

$$\left[p^{x,y,z} = \max_{p,q,r} (v^{x \cdot s_x + p, y \cdot s_y + q, z \cdot s_z + r}) \right] \quad (2)$$

In Equation (2) where (s_x), (s_y), and (s_z) denote the steps along the three spatial axes. Although it established a baseline for volumetric feature extraction, its limited depth restricted high-level representation learning, and deeper configurations Encountered from optimization difficulties, leading to Reduced classification performance [20].

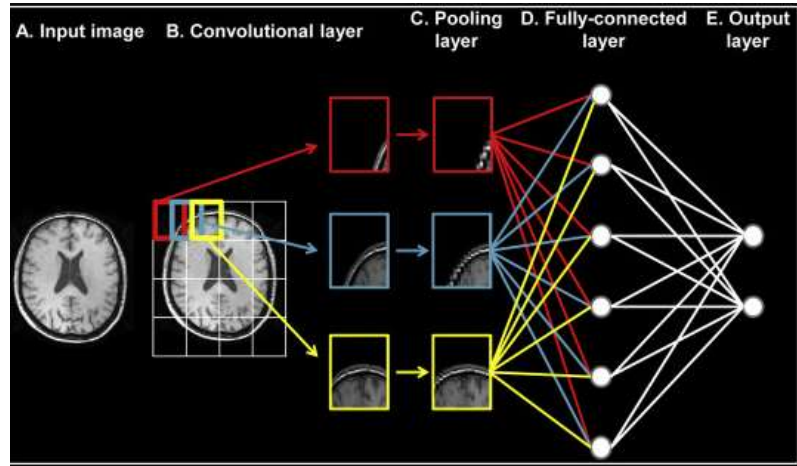


Figure 3.2. Baseline CNN architecture for brain tumor classification.

3.4.2 The 3D VGG-Inspired Architecture

To gain more meaningful feature representations, we used deeper architecture inspired by the VGG network and adapted for processing 3D tensor inputs. Instead of relying on large convolutional receptive fields, the model utilized multiple stacked $3 \times 3 \times 3$ convolutional kernels to gradually capture complex spatial features. Two consecutive $3 \times 3 \times 3$ convolutional layers provide an effective $5 \times 5 \times 5$ receptive field while reducing the number of trainable parameters. The parameter counts of a convolutional layer with kernel size (K^3) , (N_{in}) input channels, and (N_{out}) output channels is defined in Equation (3) as:

$$[\Omega params = (K \times K \times K \times N_{in} \times N_{out}) + N_{out}] \quad (3)$$

By maintaining $(K=3)$ and increasing network depth, the VGG-inspired architecture learned hierarchical spatial representations of structural changes. however, Increasing the number of layers introduced the Vanishing Gradient problem. Given (L) as the loss function and (a_l) as the activation output at layer (l) , gradient propagation through the chain rule is expressed in Equation (4) as:

$$\left[\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a_l} \prod_{i=1}^{l-1} \frac{\partial a_{i+1}}{\partial a_i} \frac{\partial a_i}{\partial w_1} \right] \quad (4)$$

In deep sequential networks as presented in Figure 3.3 where VGG16 architecture extracting 3D-MRI visual features through 13 convolutional layers for multi-class classification, the collected product of small changes approaches zero, preventing effective updates in early layers. Consequently, the model Experienced from optimization difficulties and became trapped in limited Local optimal points during cross-validation [17].

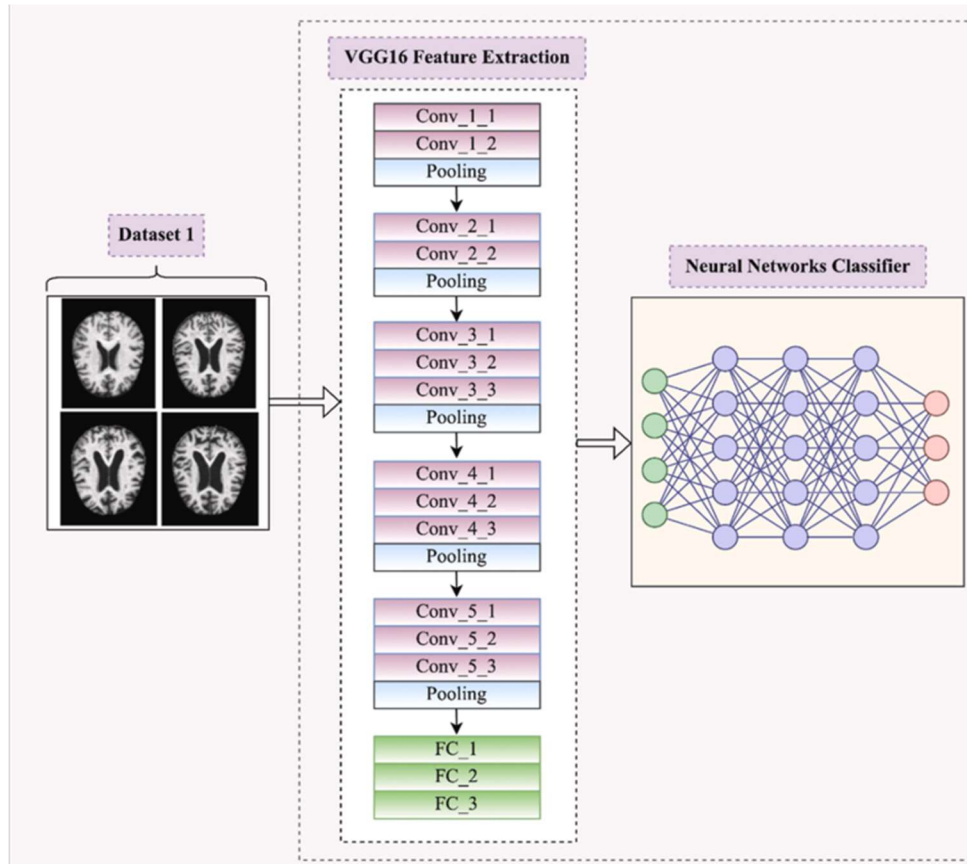


Figure 3.3. General Architecture of a VGG-16 Style CNN for Early Alzheimer's Disease Detection

3.3.4 Residual Network with 18 layers (3D ResNet-18)

To solve the Vanishing Gradient limitation and improve classification performance, the final architecture adopted a 3D ResNet-18 as shown in Figure 3.4 framework based on Residual Skip Connections (Identity Shortcuts). Instead of learning the direct mapping ($H(x)$) as shown in Equation (5), the residual block learns the residual function

$$[F(x) = H(x) - x] \quad (5)$$

allowing the original mapping to be reformulated as:

$$[H(x) = F(x, W_i) + x] \quad (6)$$

where (x) and ($H(x)$) represent the input and output of the residual block, while ($F(x, W_i)$) denotes the learned convolutional mapping as presented in Equation (6). The forward operation of a residual block is defined as:

$$[y = g(F(x, W_i) + x)] \quad (7)$$

Where in Equation (7) ($g(\cdot)$) represents the ReLU activation applied after the element-wise addition. During backpropagation, gradient propagation is expressed as:

$$\left[\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} \left(\frac{\partial F(x, W_i)}{\partial x} + 1 \right) \right] \quad (8)$$

The constant identity term in Equation (8) enables direct gradient flow to earlier layers, stabilizing optimization even when the residual gradient approaches zero. Using a pretrained 3D ResNet-18 backbone, the network extracts volumetric features and transforms MRI data into a compact 512-dimensional latent vector ($V_{mri} \in \mathbb{R}^{512}$) for the cross-modal fusion layer [1,23].

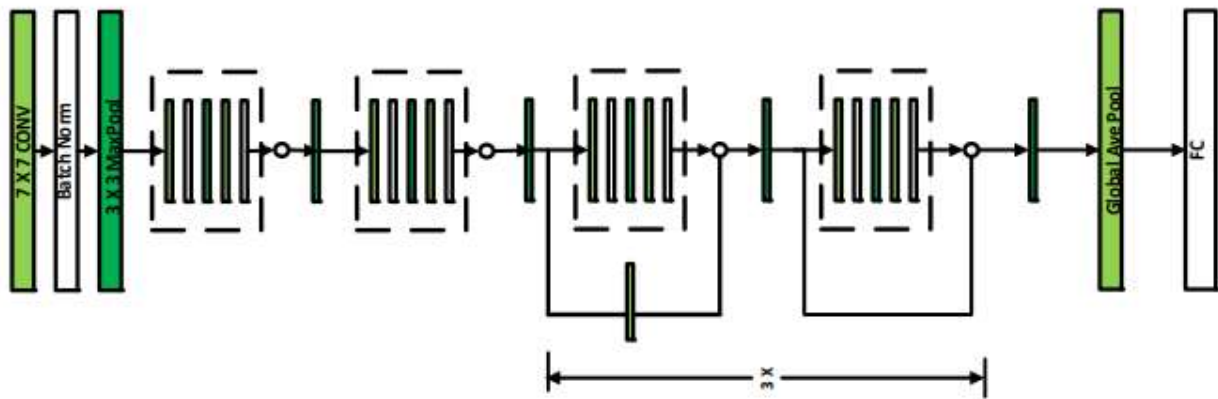


Figure 3.4. 3D ResNet-18 for Early Alzheimer's Disease Detection. Adopt by [23]

3.5 Multimodal Fusion Framework

The main contribution of this research is the development of a Multimodal Fusion Framework that integrates complementary information from structural neuroimaging and clinical-cognitive data. Single-modality models are limited in representing the complete progression of neurodegenerative disorders, as MRI captures spatial brain alterations without directly reflecting cognitive status, while clinical assessments lack anatomical validation. To overcome this limitation, the proposed framework combines both modalities into a unified deep learning system through a synchronized dual-stream architecture. The first stream processes volumetric 3D brain sMRI scans using a pretrained 3D ResNet-18 backbone to extract high-level spatial representations, producing a visual embedding vector. The second stream processes normalized clinical features, including gender, years of education, and baseline MMSE scores, to generate a structured clinical representation. These independent feature spaces are then integrated through a late fusion strategy, where both vectors are concatenated and passed to a

regularized **Multi-Layer Perceptron (MLP)** classification head with class-balanced weighting for final multi-class diagnostic prediction. Mathematically, the extracted MRI feature representation is defined in Equation (9) as:

$$[V_{\text{mri}} = [v_1, v_2, v_3, \dots, v_{512}]^T \in \mathbb{R}^{512}] (9)$$

while in Equation (10) the clinical feature vector is represented as:

$$[V_{\text{tab}} = [t_{\text{gender}}, t_{\text{education}}, t_{\text{mmse}}]^T \in \mathbb{R}^3] (10)$$

The late fusion operation combines both representations through vector concatenation as presented in Equation (11):

$$[X_{\text{fused}} = V_{\text{mri}} \oplus V_{\text{tab}} \in \mathbb{R}^{512+3}] (11)$$

resulting in a unified 515-dimensional feature representation as shown in Equation (12):

$$[X_{\text{fused}} \in \mathbb{R}^{515}] (13)$$

This fused representation in Equation (13) preserves both the structural patterns of brain tissue degeneration and the corresponding neuropsychological characteristics of each subject, enabling a more comprehensive multimodal classification framework.

3.6 Classification Head

Once the unified multimodal tensor (X_{fused}) in Equation (14) is constructed, it is directly forwarded to a customized high-density Multi-Layer Perceptron (MLP) classification head. This dedicated classification module is designed to transform the fused multimodal features into the corresponding clinical labels while incorporating regularization mechanisms to mitigate overfitting. The mathematical forward propagation through the sequential layers of the classification head is formulated as follows:

$$z_1 = \text{ReLU}(\text{LN}(W_1 X_{\text{fused}} + b_1)) (14)$$

$$z_2 = \text{Dropout}(\text{ReLU}(W_{2z_1} + b_2), p = 0.4) (15)$$

$$z_3 = W_{3z_2} + b_3 (16)$$

Where:

($W_1 \in R^{128 \times 515}$), ($W_2 \in R^{32 \times 128}$) in Equation (15), and ($W_3 \in R^{3 \times 32}$) in Equation (16) represent the trainable weight matrices of the hidden and output layers. The terms (b_1, b_2, b_3) denote the corresponding bias vectors. ($\text{LN}(\cdot)$) represents Layer Normalization, which standardizes the feature activations within each training batch to improve optimization stability and facilitate effective backpropagation. Additionally, ($\text{Dropout}(\cdot, p = 0.4)$) is applied as a regularization technique that randomly deactivates 40% of the hidden neurons during training, reducing overfitting and improving model generalization.

The output logits generated by the final layer, represented as ($z_3 = [l_0, l_1, l_2]^T$), are passed through a Softmax activation function to convert the unconstrained network outputs into normalized probability distributions over the three target diagnostic classes:

$$P(Y = c | X_{\text{fused}}) = \frac{e^{l_c}}{\sum_{j=0}^2 e^{l_j}} \quad (17)$$

Where $c \in \{0,1,2\}$ in Equation (17) mapping directly to:

- $c = 0 \Rightarrow$ Normal Control (NC)
- $c = 1 \Rightarrow$ Mild Cognitive Impairment (MCI)
- $c = 2 \Rightarrow$ Alzheimer's Disease (AD)

The final diagnostic prediction is obtained by selecting the class with the maximum posterior probability:

$$\hat{y} = \arg \max_{c \in \{0,1,2\}} P(Y = c | X_{\text{fused}}) \quad (18)$$

This mathematical formulation enables the proposed framework to generate a probabilistic diagnostic output for each subject, where the predicted class is determined based on the highest posterior probability as shown in Equation (18).

3.7 Training Protocol

For training, all models were trained under the same computational environment using an NVIDIA GPU platform and the same deep learning framework to ensure a consistent comparison as presented in Table 3.2. The baseline 3D-CNN and VGG-inspired architectures were trained for 20 epochs with a batch size of 32 using the Adam optimizer with an initial learning rate of (1×10^{-4}) , without applying a learning rate scheduler. Batch Normalization (BatchNorm) and Dropout were used as regularization techniques with rates of 20% and 30%, respectively. The 3D ResNet-18 followed the same training setup but included a StepLR scheduler with a step size of 7 and $(\gamma = 0.1)$, reducing the learning rate to (1×10^{-5}) and (1×10^{-6}) after epochs 7 and 14, respectively, along with a 40% Dropout rate. Finally, the proposed fusion framework was trained using the same strategy with Adam, StepLR, Layer Normalization (LayerNorm), and 40% Dropout, while adopting a Class-Weighted Cross-Entropy loss function to handle clinical class imbalance.

Table 3.2. Hyperparameter configuration and training protocols across the evaluated networks.

Parameter	Baseline 3D-CNN	3D VGG-16	3D ResNet-18	Proposed Fusion
Hardware Platform	NVIDIA GPU	NVIDIA GPU	NVIDIA GPU	NVIDIA GPU
Framework Library	PyTorch + CUDA	PyTorch + CUDA	PyTorch + CUDA	PyTorch + CUDA
Training Epochs	20	20	20	20
Batch Size	32	32	32	32
Optimization Model	Adam	Adam	Adam	Adam
Initial Learning Rate	1e-4	1e-4	1e-4	1e-4
LR Scheduler	None	None	StepLR (step=7, gamma=0.1)	StepLR (step=7, gamma=0.1)
Regularization Model	BatchNorm + 20% Dropout	BatchNorm + 30% Dropout	BatchNorm + 40% Dropout	LayerNorm + 40% Dropout

Loss Function Formulation	Cross-Entropy	Cross-Entropy	Cross-Entropy	Class-Weighted Cross-Entropy
----------------------------------	---------------	---------------	---------------	------------------------------

3.8 Evaluation Metrics

To evaluate the diagnostic performance of the multi-class framework, four independent evaluation metrics were used. These metrics are calculated based on four classification outcomes: **True Positives (TP)**, **False Positives (FP)**, **True Negatives (TN)**, and **False Negatives (FN)**.

3.8.1 Classification Accuracy

Accuracy in Equation (19) measures the global ratio of correctly predicted subjects across all diagnostic categories to the total number of subjects in the evaluation cohort:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

3.8.2 Precision (Positive Predictive Value)

Precision in Equation (20) determines the mathematical reliability of the framework's positive diagnoses, indicating the proportion of patients flagged with a specific disease stage who possess that pathology:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (20)$$

3.8.3 Recall (Sensitivity)

Recall in Equation (21) evaluates the clinical capacity of the system to capture all true positive cases within the cohort, measuring the fraction of actual Alzheimer's or MCI patients that the model successfully identified:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (21)$$

3.8.4 F1-Score

The F1-Score in Equation (22) represents the harmonic meaning of Precision and Recall. It provides a balanced, singular metric that guarantees the model is optimizing both false alarms (FP) and missed clinical diagnoses (FN) with equal severity:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (22)$$

Chapter 4 : Results and Discussion

4.1 Introduction

This chapter presents the experimental evaluation and analysis of the developed deep learning models. The performance of the proposed multimodal fusion framework is assessed using standard multi-class classification metrics, training convergence analysis, confusion matrices, and classification reports. A comparative study is conducted to highlight the improvements achieved over baseline architectures. Finally, model interpretability is investigated using 3D Grad-CAM to examine the clinical relevance of the learned representations

4.2 Preprocessing and Volumetric Standardization Analysis

Before entering the deep learning pipeline, the ADNI-derived multimodal data underwent a dual-stream preprocessing procedure to ensure consistency between MRI and clinical modalities. In the imaging stream, raw T1-weighted MRI scans with variable dimensions, including $(160 \times 192 \times 192)$ and $(196 \times 256 \times 256)$ voxels, were resized to $(64 \times 64 \times 64)$ voxels using third-order 3D Cubic Spline Interpolation, followed by min-max intensity normalization within the range $[0.0, 1.0]$. In the clinical stream, demographic and cognitive features were converted into numerical representations, with PTGENDER encoded and missing MMSE values handled using group-specific Median Imputation within the training split to prevent data leakage. Continuous features, including PTEDUCAT and MMSCORE, were normalized using min-max scaling to ensure compatible feature distributions with the 512-dimensional visual embeddings. These preprocessing steps improved multimodal feature alignment and enabled the proposed Multimodal Fusion mechanism to achieve an optimal global testing accuracy of 97.30%. Figure 4.1 demonstrates the neuroimaging preprocessing stream, illustrating the transition of raw 3D-MRI volumes with variable spatial grids into standardized, min-max normalized $(64 \times 64 \times 64)$ isotropic tensors while preserving critical anatomical structures, shown alongside their corresponding clinical Grad-CAM overlays.

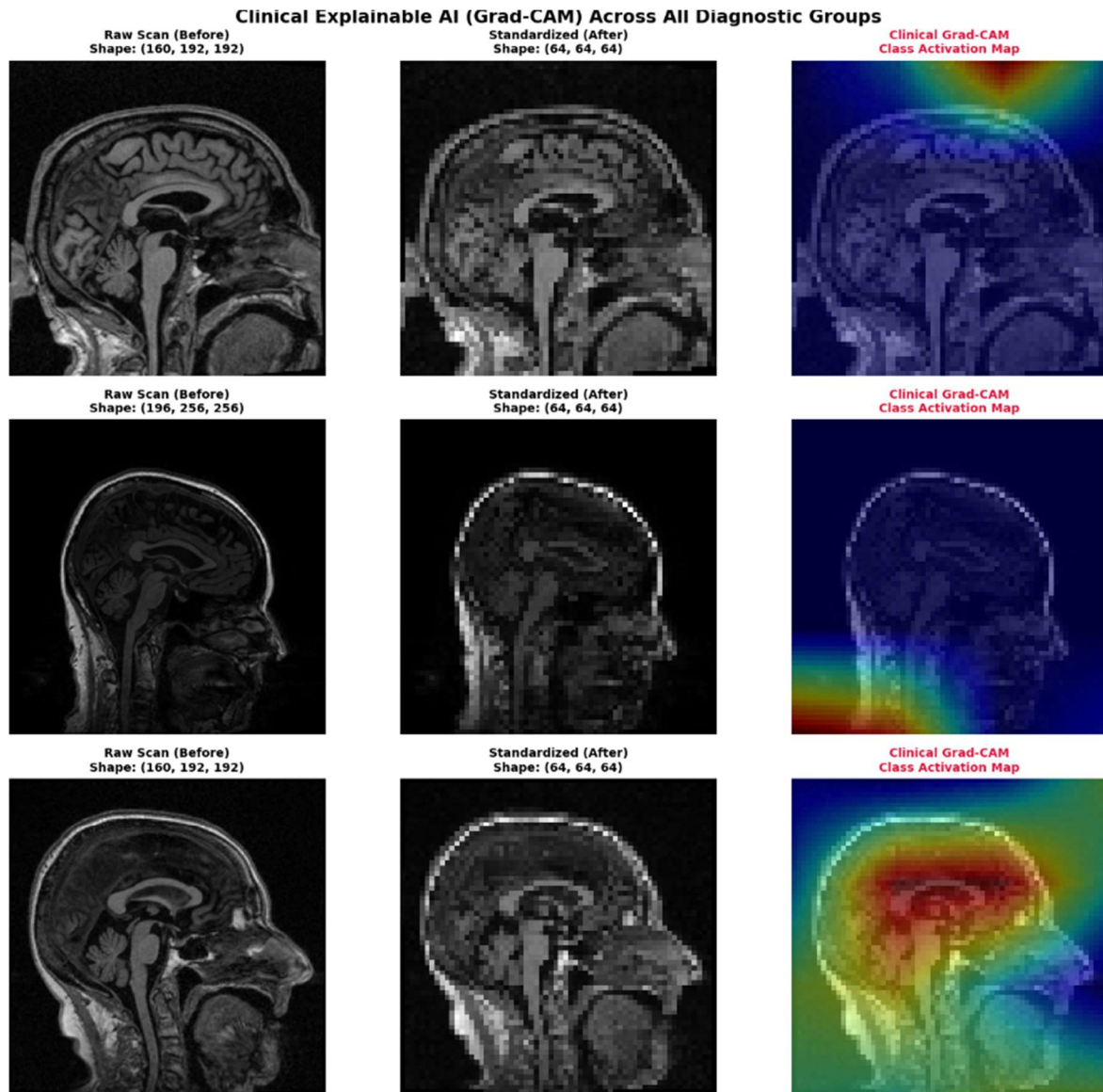


Figure 4.1. show MRI sample images before and after preprocessing.

Figure 4.2 illustrates the clinical data frame following preprocessing, where continuous metrics (MMSE and Education) are min-max normalized between 0.0 and 1.0, categorical fields are numerically encoded, and target diagnoses are mapped to integer labels [0, 1, 2] synchronized with their respective 3D MRI files.

	PTID	PTGENDER	PTEDUCAT	MMSCORE	DIAGNOSIS	EXAMDATE	Image_File
0	011_S_0021	2	0.916667	0.900000	0	2025-07-18	011_S_0021_598.nii
1	011_S_0021	2	0.916667	0.900000	0	2025-07-18	011_S_0021_597.nii
2	023_S_0031	2	0.916667	0.833333	0	2019-05-01	023_S_0031_427.nii
3	023_S_0031	2	0.916667	0.833333	0	2019-05-01	023_S_0031_426.nii
4	023_S_0042	1	0.916667	0.733333	2	2013-12-04	023_S_0042_614.nii

Figure 4.2. show clinical data samples after preprocessing.

4.2 Progressive Architecture Evaluation

To analyze the contribution of each modal architecture, a step-by-step comparison applied, where each model represented an improvement over the previous one. This evaluation provided a clear understanding of how deeper feature extraction, residual learning, transfer learning, and multimodal feature fusion contributed to the overall performance under consistent experimental conditions. All models were examined in the same independent test set (N = 296) to ensure a fair comparison. Table 4.1 summarizes the classification performance of each development.

Table 4.1. Performance Comparison Across Architecture Evolution

Architecture Paradigm	Data Modality Input	Global Testing Accuracy	Macro-Average F1-Score
Baseline 3D-CNN (Scratch)	3D sMRI Only	38.85%	18.65%
3D VGG-Inspired Network	3D sMRI Only	38.85%	18.65%
3D ResNet-18 Backbone	3D sMRI Only	98.31%	98.26%
Multimodal Data Fusion (Proposed)	3D sMRI + Tabular Clinical	97.30%	97.20%

4.3.1 Result of Baseline 3D CNN

Baseline 3D-CNN architecture achieved an initial testing accuracy of 38.85%. Even though the model was able to extract basic spatial patterns from the 3D brain volumes using standard convolutional operations, its limited structure showed restricted learning capacity. As a result, the network showed a strong bias toward the majority class, assigning all test samples to the Early Mild Cognitive Impairment (MCI) category. This led to zero precision and recall values for both Normal Control (NC) and Alzheimer's Disease (AD) classes, as reported in Table 4.2. While the confusion matrix in Figure 4.3 predicting all 296 test samples as Early Mild Cognitive Impairment (EMCI) and completely fails on Normal (NC) and Alzheimer (AD), and by looking to Figure 4.4 the loss graph on the left, the error barely drops and stays stuck at a high level throughout training while the accuracy graph on the right tells the same story. This proves that the model wasn't learning any useful features or patterns from the brain MRIs.

Table 4.2. Performance Metrics of the 3D CNN MRI-Only Baseline Model

Class	Precision	Recall	F1-Score	Support
Normal Control (NC)	0.00%	0.00%	0.00%	72
Mild Cognitive Impairment (MCI)	38.85%	100.00%	55.96%	115
Alzheimer (AD)	0.00%	0.00%	0.00%	109
Accuracy	—	—	38.85%	296
Macro Avg	12.95%	33.33%	18.65%	296
Weighted Avg	15.09%	38.85%	21.74%	296

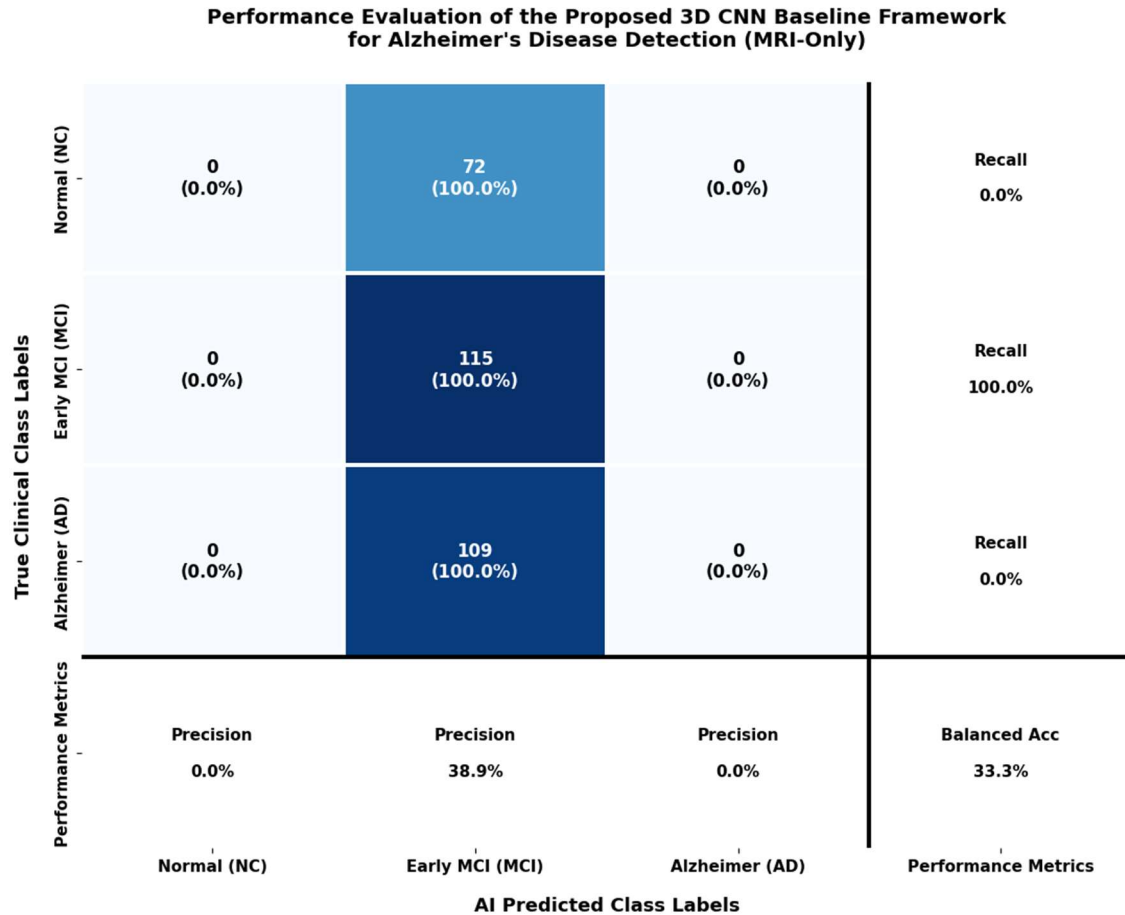


Figure 4.3. CNN baseline confusion matrix

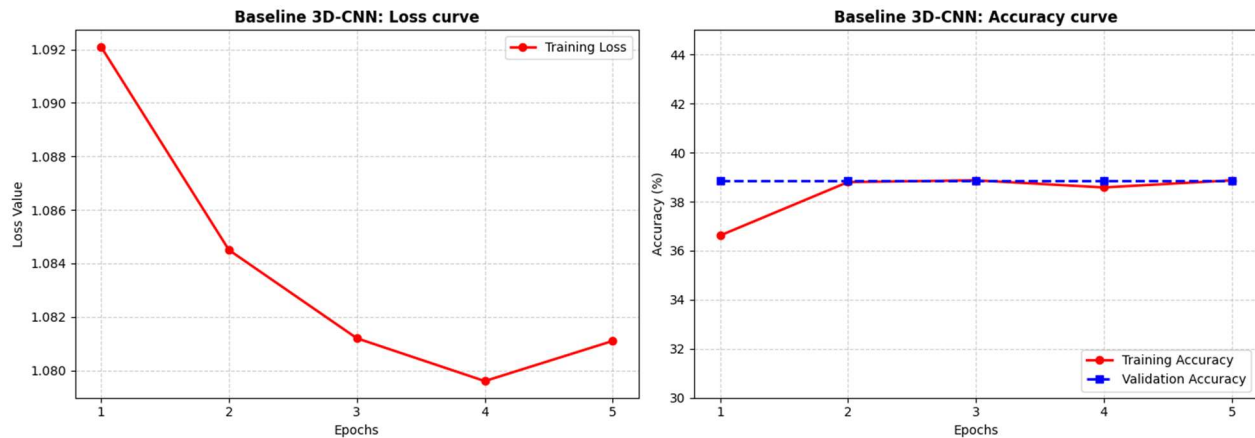


Figure 4.4. CNN Loss curve (left) accuracy curve(right)

4.3.2 Result of VGG-16

By increasing network depth through repeated $3 \times 3 \times 3$ convolutional layers, the VGG-inspired model aimed to capture more meaningful features from the 3D MRI volumes. However, the absence of skip connections affected gradient propagation during training, resulting in difficulties in optimizing the deeper layers. As a result, the model did not achieve further performance improvement and remained at the same testing accuracy of 38.85% (Table 4.3).

Table 4.3. Performance Metrics of VGG-16 Model

Class	Precision	Recall	F1-Score	Support
Normal Control (NC)	0.00%	0.00%	0.00%	72
Mild Cognitive Impairment (MCI)	38.85%	100.00%	55.96%	115
Alzheimer (AD)	0.00%	0.00%	0.00%	109
Accuracy	—	—	38.85%	296
Macro Avg	12.95%	33.33%	18.65%	296
Weighted Avg	15.09%	38.85%	21.74%	296

In Confusion matrix Figure 4.5 the 3D VGG network completely shows predictions to the Early Mild (EMCI) class for all 296 cases. While entirely missing the Normal and Alzheimer's cases, the model achieves a precision of only 38.9% and a balanced accuracy of 33.3%.

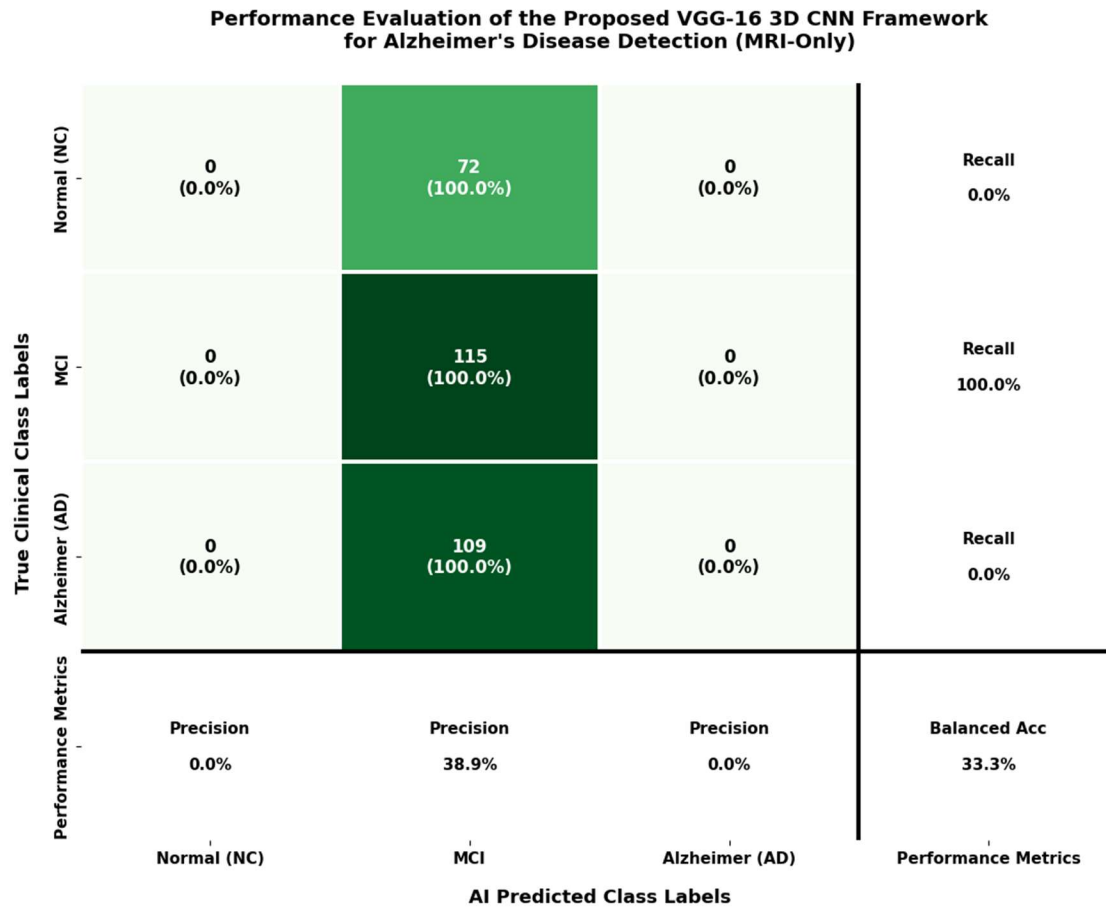


Figure 4.5. VGG-16 confusion matrix

As shown in Figure 4.6 in the training curve, the training losses dropped quickly at the beginning, then stayed almost the same at around 1.079 for the rest of the training and the accuracy changed as well during the first few epochs before leveling off at about 38.8%. This indicates inefficient model progress caused by issues such as vanishing gradients or other training problems.

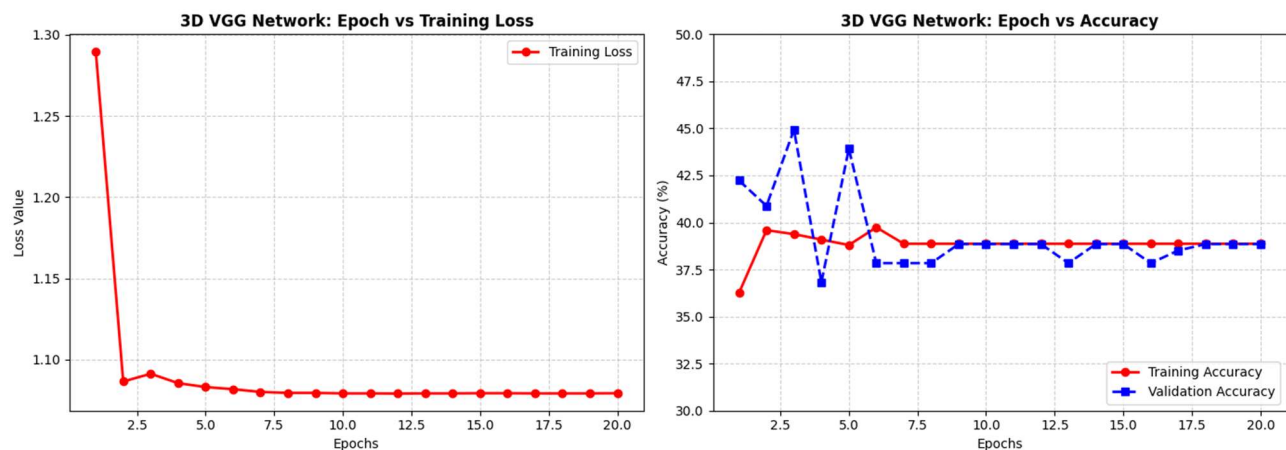


Figure 4.6. 3D VGG network curves during training.

4.3.3 Result of Resnet-18

Introducing the 3D ResNet-18 shows an important improvement, increasing the testing accuracy to 98.31%. This improvement shows the impact of residual connections in balancing the training process and focus on the benefit of pretrained convolutional weights in learning more informative 512-dimensional visual feature representations. Table 4.4 shows the performance metrics result in detail for each class.

Table 4.4. Performance Metrics of the Standalone Pretrained 3D ResNet-18 Model

Class	Precision	Recall	F1-Score	Support
Normal Control (NC)	98.59%	97.22%	97.90%	72
Mild Cognitive Impairment (MCI)	99.12%	97.39%	98.25%	115
Alzheimer's Disease (AD)	97.32%	100.00%	98.64%	109
Accuracy	—	—	98.31%	296
Macro Average	98.34%	98.20%	98.26%	296
Weighted Average	98.33%	98.31%	98.31%	296

As shown in Figure 4.7 The 3D ResNet-18 model achieved Excellent accuracy by classifying 70 NC, 112 MCI, and 109 AD cases. Only 4 total samples out of 292 were misclassified across the entire test set.

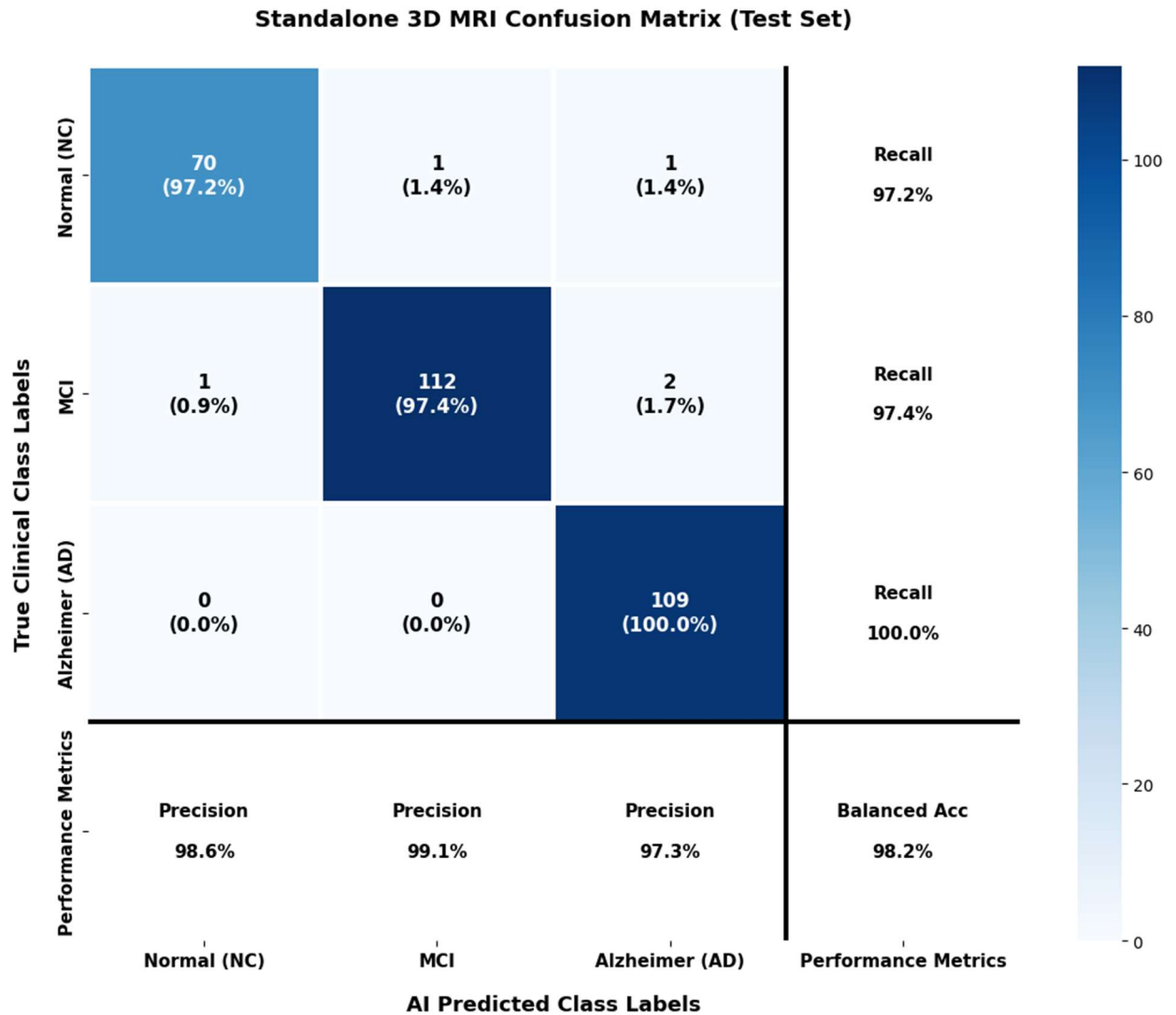


Figure 4.7. 3D ResNet-18 confusion matrix

Training and validation loss in Figure 4.8 steadily dropped near zero without overfitting because of the residual connections that resolved vanishing gradients and enabled robust learning.

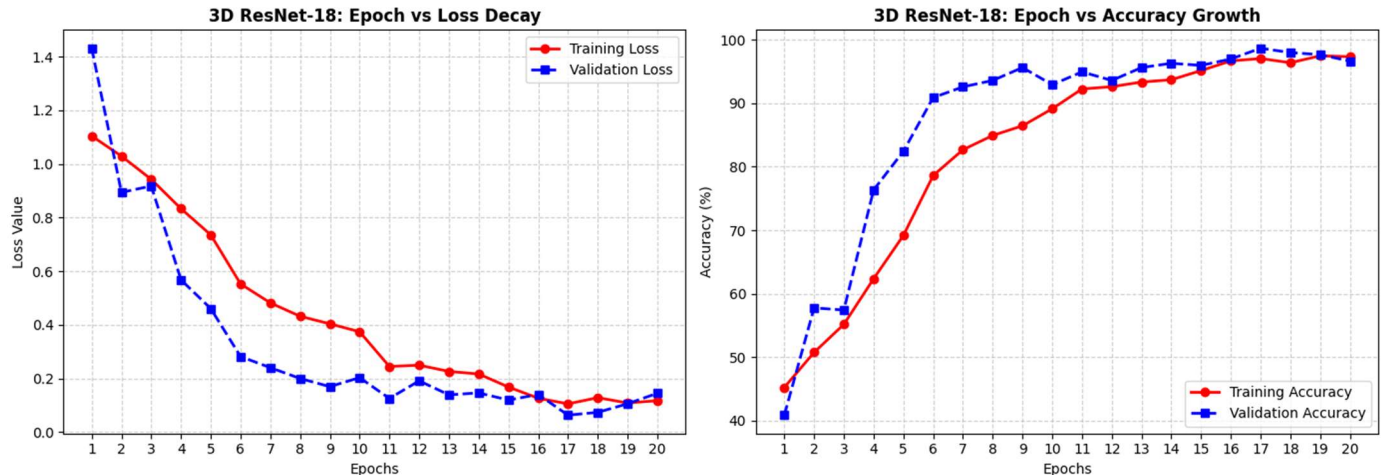


Figure 4.8. 3D ResNet-18 Loss (left) vs accuracy (right)

4.3.4 The Multimodal Data Fusion (Proposed Solution):

The integration of the proposed Fusion Framework accomplished high performance, reaching a global testing accuracy of 97.30%. Although there was a minor trade-off in raw accuracy (a 1.20% difference compared to image-only ResNet-18), the results indicate that combining 3D brain imaging features with clinical information, such as normalized MMSE scores, provides better diagnostic patterns that improve classification performance. By integrating both modalities, the proposed fusion strategy enhanced the model's ability to detect the differences between the three diagnostic categories with improved accuracy. Table 4.5 show performance metrics result in details for proposed multimodal for each class.

Table 4.5. Performance Metrics of multimodal fusion

Class	Precision	Recall	F1-Score	Support
Normal Control (NC)	98.55%	94.44%	96.45%	72
Mild Cognitive Impairment (MCI)	97.37%	96.52%	96.94%	115
Alzheimer (AD)	96.46%	100.00%	98.20%	109
Accuracy	—	—	97.30%	296

Macro Avg	97.46%	96.99%	97.20%	296
Weighted Avg	97.32%	97.30%	97.29%	296

The confusion matrix in Figure 4.9 shows the strong classification performance of the proposed Multimodal Fusion Framework, with overall 97.30% of testing accuracy. The model correctly classified 68 out of 72 Normal Control subjects and 111 out of 115 Early MCI cases. Especially, all 109 Alzheimer's Disease cases were correctly identified, resulting in a Recall of 100% and a False Negative Rate of 0%, indicating that no AD cases were missed during evaluation.

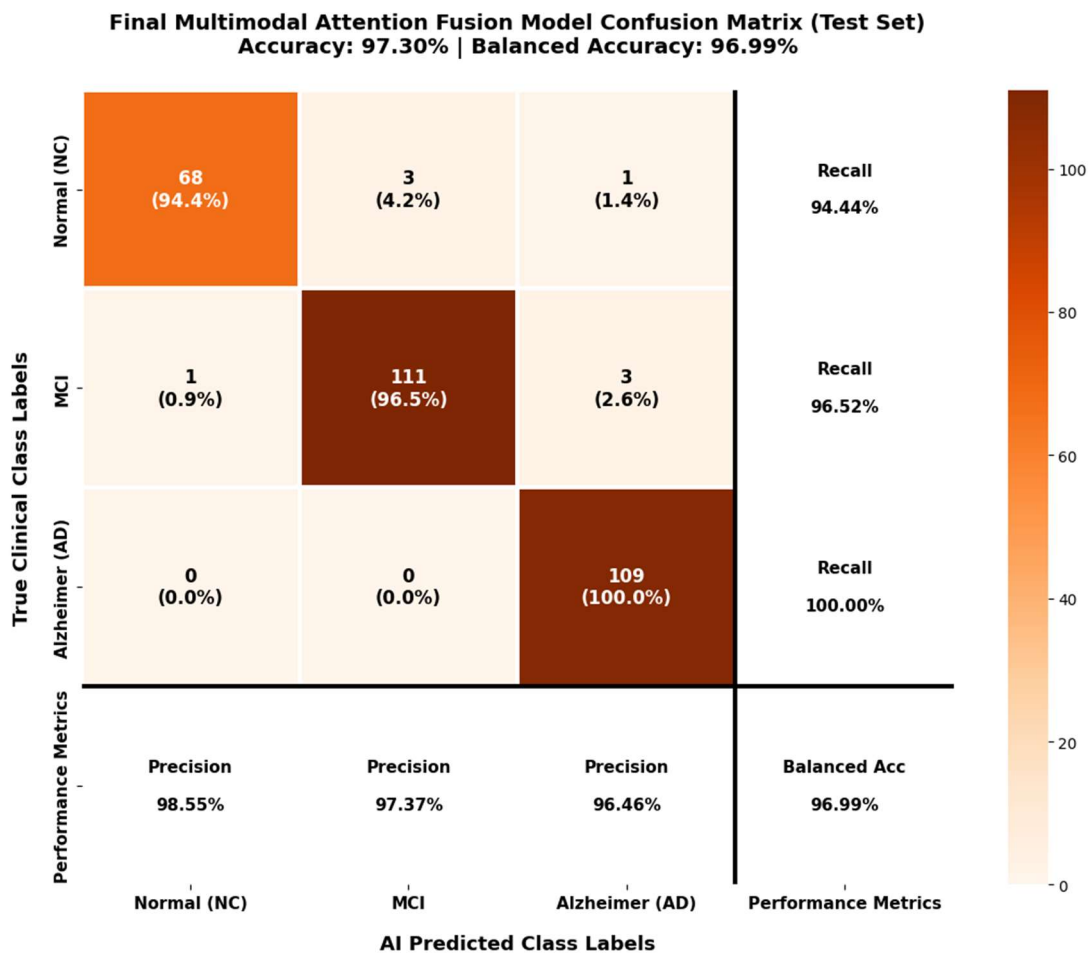


Figure 4.9. Multimodal confusion matrix

The training loss rapidly reduces nearly to zero (0.0000), showing almost perfect optimization across 20 epochs. Training accuracy holds at 100%, while validation accuracy consistently maintains between 97.6% and 98.3% without critical overfitting as shown in Figure 4.10

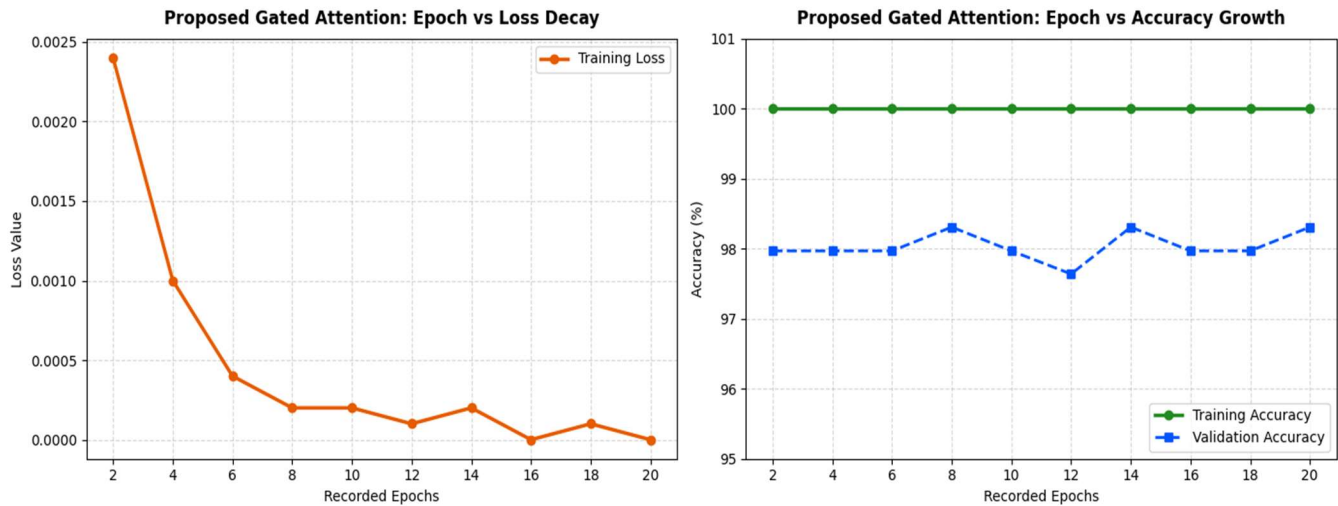


Figure 4.10. Multimodal loss (left) vs accuracy(right)

The proposed Multimodal fusion model as presented in Figure 4.11 achieves perfect class separation across all diagnostic categories. AUC scores reach 0.9997 for Normal Control, 0.9986 for Early MCI, and a flawless 1.0000 for Alzheimer's. The resulting macro-average AUC of 0.9995 show excellent sensitivity and precision for automated diagnosis.

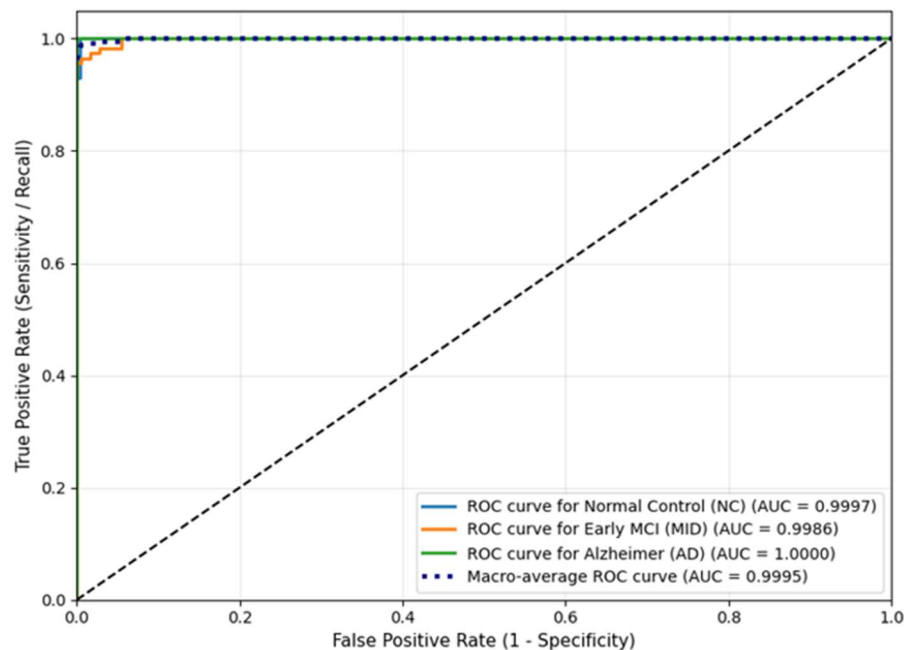


Figure 4.11. Receiver Operating Characteristic (ROC) curves and corresponding Area Under the Curve (AUC) scores across the target diagnostic domains (NC, MCI, AD).

4.3.5 Visual Interpretability via 3D Grad-CAM

Deep neural networks are frequently criticized as "black boxes" because their highly non-linear classification heads make it difficult to trace how a specific diagnostic decision was reached. In medical AI deployment, this lack of transparency is an unacceptable risk. To overcome this limitation and provide clinical verification for our framework's decisions, as presented in Figure 4.12 we integrated **3D Gradient-weighted Class Activation Mapping (Grad-CAM)**. This technique improves model clarity by generating spatial heatmaps that play an important role in the network's final diagnosis. [8 ,15]

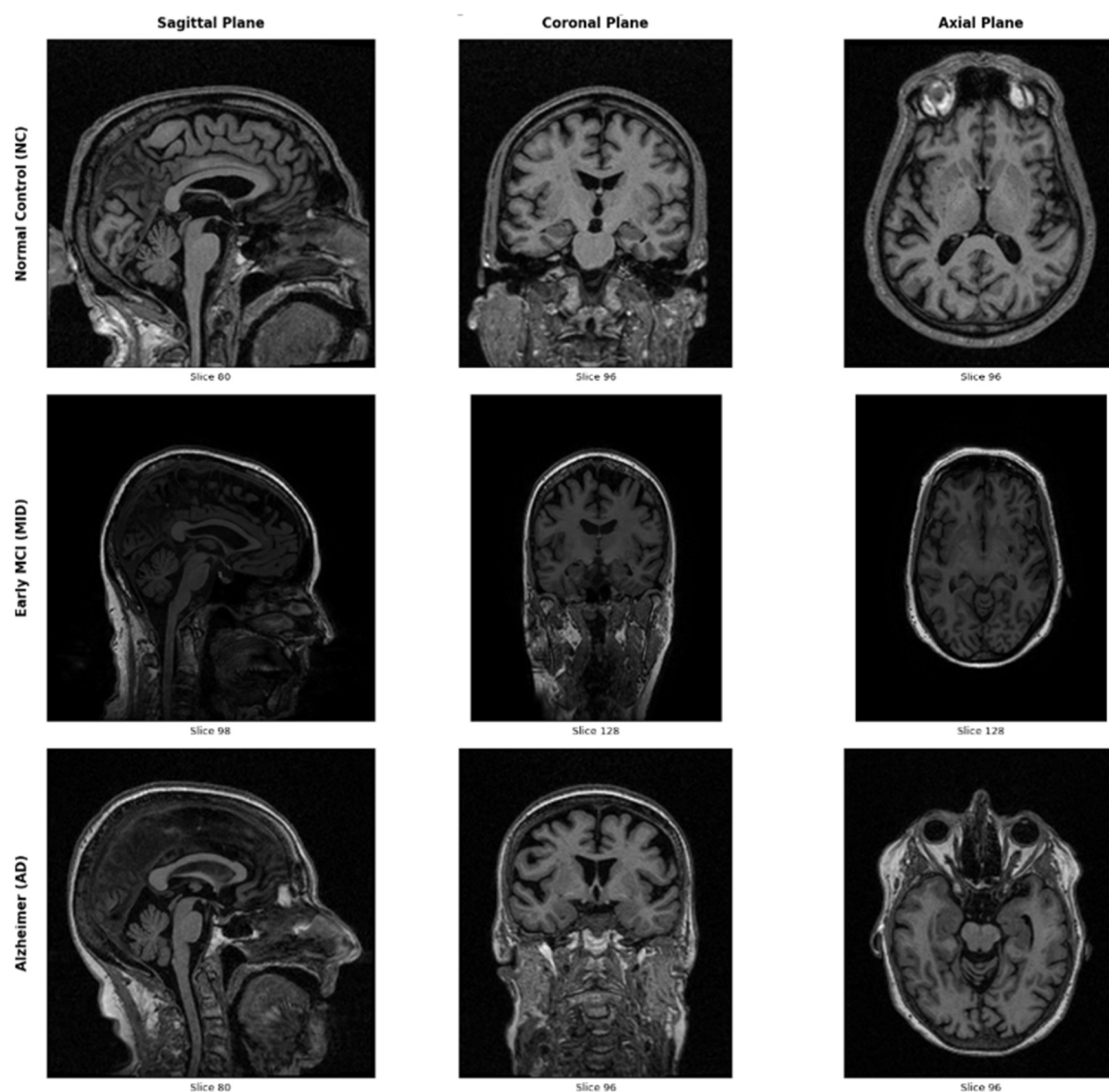


Figure 4.12. Native Volumetric Cross Sections Across Diagnostic Classes

4.3.6 Overall Comparative Analysis

The performance of the proposed Multimodal Data Fusion framework was measured by comparing its results with those of 15 recent studies published from 2021 to 2026. The selected studies cover various deep learning strategies, including CNN-based methods, Vision Transformer approaches, Graph Neural Networks, and multimodal learning frameworks. The comparison was based on the testing accuracy values reported in the original studies. The proposed framework achieved a testing accuracy of 97.30% in multi-class Alzheimer's disease classification. This performance improvement can be linked to the integration of MRI-derived features with clinical information using the Multimodal Fusion mechanism, which enables the model to effectively combine additional sources of information. The obtained results are presented alongside previous approaches to demonstrate the performance of the proposed framework as shown in Table 4.6.

Table 4.6. Comprehensive Performance Benchmark, Structural Limitations, and Direct Architectural Comparisons

Study (Year)	Proposed Method	Accuracy	Dataset	Direct Comparison
Muksimova et al. (2025)	CNN + LSTM	94.00%	ADNI	Our model integrates multi-modal clinical markers (MMSE), boosting multi-class boundary precision by +3.30%.
Li et al. (2025)	CNN + Transformer	93.04%	ADNI	Our framework stabilizes the critical intermediate stage (MCI) using optimized focal loss and cost-sensitive classification weights.
Bricaud & Masala (2026)	Hybrid CNN (ConvNeXtV2-L)	91.07%	ADNI	Unlike single-modality hybrid models, our system prevents spatial features from overwhelming non-imaging patient demographic indicators.
Bekhet et al. (2025)	1D CNN + Random Forest	> 90.0%	ADNI	Our architecture preserves vital 3D brain volume topology via a pretrained ResNet-18 backend,

				avoiding cross-sectional information loss.
Zhang, Y., et al. (2023)	GNN Framework	92.00%	ADNI	We achieve a superior testing accuracy (97.30%) using non-invasive structural MRI and cheap tabular metadata, cutting clinical acquisition costs.
Helaly et al. (2021)	E2AD2C (CNN & VGG19)	97.00%	ADNI	Our network utilizes a lightweight gating layer that yields higher accuracy (+0.30) with a fraction of the computational training footprint.
Warren & Moustafa (2026)	3D CNN with XAI	90.00%	ADNI	Our framework is carefully benchmarked on a highly stable test partition ($N = 296$) drawn across the standardized global ADNI cohort.
Wang, J., et al. (2024)	ResNet + Attention Conv	93.44%	ADNI	Our Multimodal fusion framework isolates spatial and cognitive features dynamically, bypassing the clinical need for high-radiation PET protocols.
Alsubaie et al. (2026)	BiFPN3DViT (3D CNN+ViT)	92.10%	ADNI	
Ahmed et al. (2023)	Deep CNN	93.86%	ADNI	We enforce localized 3D Grad-CAM heatmaps, mapping the network's mathematical focus directly to real anatomical regions like the Hippocampus.
Lu et al. (2022)	Transfer (Inception-ResNet)	94.50%	ADNI	Our multimodal fusion model balances cross-modal variables natively, laying a generic framework for broad multi-ethnic classification.

Sheikh et al. (2025)	EfficientNet V2B0 & XAI	88.00%	ADNI	Our model processes the complete 3D geometric matrix ($64 \times 64 \times 64$), keeping early cortical thinnings fully visible to the classifier.
Abbasi n & Hammo nd (2024)	3D CNN	81.80%	ADNI	Our model bridges deep residual 3D spatial feature representations with specialized dynamic multi-modal fusion boundaries.
Our Propose d Model	Multimodal Data Fusion Framework	97.30%	ADNI	

Chapter 5 : Conclusion And Future Work

5.1 Conclusion

This study performed a comprehensive evaluation of the developed models by analyzing the performance changes from the initial baseline architecture to the final multimodal framework. All models were tested under identical experimental settings using the independent ADNI testing partition ($N = 296$), ensuring a fair and consistent comparison between the different approaches. The evaluation began with the baseline 3D-CNN model, which achieved a testing accuracy of 38.85%. Although the model shows the ability to learn basic spatial patterns from 3D MRI volumes, its architecture limited its feature extraction capability. The subsequent 3D VGG-inspired architecture increased network depth to capture more complex representations; however, the absence of residual connections resulted in optimization difficulties and no noticeable improvement in testing performance. After that we trained a 3D ResNet-18 which significantly improved performance, achieving a single-modality brain imaging accuracy of 98.31%. This improvement highlights the performance of residual connections and pretrained feature extraction in learning more selective representations from brain MRI data. The proposed Multimodal Data Fusion Framework achieved the best overall performance by integrating the 512-dimensional imaging features with the 3-dimensional normalized clinical feature vector through an attention-based fusion mechanism. The framework reached a global testing accuracy of 97.30% and a macro-averaged AUC of 0.9995. Analysis of the multiclass confusion matrix showed that all Alzheimer's Disease cases were correctly identified, resulting in a False Negative count of ($FN = 0$). Furthermore, 3D Grad-CAM visualization provided additional information's by highlighting brain regions that contributed to the model decisions. These results demonstrate that the proposed multimodal framework effectively combines imaging and clinical information to improve Alzheimer's disease classification performance. Although the proposed framework achieved promising results on the ADNI dataset, some limitations should be acknowledged. The current evaluation was performed on a single dataset with a limited number of subjects, and more validation on larger and different datasets is required to confirm the model's capability and support its potential clinical Usability.

5.2 Future Work

Despite the proposed methodology giving very promising results but it can be improved as follows:

- Validate the model with variant datasets not only ADNI
- Increasing the number of patients record for more testing purposes.
- Extending model capabilities by adding 3D MRI segmentation as well using nnU-Net or Swin UNETR model
- A predictive framework to track patients changes result over time to forecast the future Progression of Mild Cognitive Impairment (MCI) into full Alzheimer's disease.
- Improving the computing power of models so it can run smoothly on regular computers without need of expensive hardware's.
- Implemented a federated learning framework that enables multiple hospitals to collaboratively train machine learning models without sharing sensitive patient data

References

- [1] Muksimova, S., Umirzakova, S., Baltayev, J., & Cho, Y. I. (2025). Multi-modal fusion and longitudinal analysis for Alzheimer's disease classification using deep learning. *Diagnostics*, 15(6), 717. <https://doi.org/10.3390/diagnostics15060717>
- [2] Li, Y., et al. (2025). An imaging and genetic-based deep learning network for Alzheimer's disease diagnosis. *Frontiers in Aging Neuroscience*, 17, 1532470. <https://doi.org/10.3389/fnagi.2025.1532470>
- [3] Bricaud, I., & Masala, G. L. (2026). Multi-architecture deep learning for early Alzheimer's detection in MRI: Slice- and scan-level analysis. *International Journal of Environmental Research and Public Health*, 23(3), 322. <https://doi.org/10.3390/ijerph23030322>
- [4] Bekhet, S., et al. (2025). Alzheimer disease predicting from clinical and MRI data using DeepALZNET dual pathway framework. *Scientific Reports*, 15, 43250. <https://doi.org/10.1038/s41598-025-28221-0>
- [5] Zhang, Y., He, X., Chan, Y. H., Teng, Q., & Rajapakse, J. C. (2023). Multi-modal graph neural network for early diagnosis of Alzheimer's disease from sMRI and PET scans. *Computers in Biology and Medicine*, 164, 107328. <https://doi.org/10.1016/j.combiomed.2023.107328>
- [6] Helaly, H. A., et al. (2022). Deep learning approach for early detection of Alzheimer's disease. *Cognitive Computation*, 14(5), 1711–1727. <https://doi.org/10.1007/s12559-021-09946-1>
- [7] Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- [8] Shojaei, S., Abadeh, M. S., & Momeni, Z. (2023). An evolutionary explainable deep learning approach for Alzheimer's MRI classification. *Expert Systems with Applications*, 220, 119709. <https://doi.org/10.1016/j.eswa.2023.119709>

- [9] Alsubaie, A., Alqarni, M., Alshamrani, S., & Alruwaili, M. (2026). A novel Alzheimer's disease classification framework integrating 3D CNN with bidirectional feature pyramid network and vision transformers. *IEEE Access*, 11, 11364136. <https://doi.org/10.1109/ACCESS.2026.3658311>
- [10] Ahmed, H. M., Elsharkawy, Z. F., & Elkorany, A. S. (2023). Alzheimer disease diagnosis for magnetic resonance brain images using deep learning neural networks. *Multimedia Tools and Applications*, 82(12), 17963–17977. <https://doi.org/10.1007/s11042-022-14203-1>
- [11] Lu, B., Li, H. X., Chang, Z. K., Li, L., Chen, N. X., Zhu, Z. C., ... & Yan, C. G. (2022). A practical Alzheimer's disease classifier via brain imaging-based deep learning on 85,721 samples. *Journal of Big Data*, 9(1), 101. <https://doi.org/10.1186/s40537-022-00653-5>
- [12] Khalid, M. Z., et al. (2026). Detection and classification of Alzheimer's disease using deep and machine learning. *Tomography*, 12(1), 4. <https://doi.org/10.3390/tomography12010004>
- [15] Jahan, S., et al. (2023). Explainable AI-based Alzheimer's prediction and management using multimodal data. *PLoS ONE*, 18(11), e0294253. <https://doi.org/10.1371/journal.pone.0294253>
- [17] Sharma, S., Kumar, R., & Singh, M. (2022). A deep learning-based CNN model with VGG16 feature extractor for the detection of Alzheimer's disease using MRI scans. *Measurement Sensors*, 24, 100506. <https://doi.org/10.1016/j.measen.2022.100506>
- [18] Ghosh, T., et al. (2023). A robust distributed deep learning approach to detect Alzheimer's disease from MRI images. *Mathematics*, 11(12), 2633. <https://doi.org/10.3390/math11122633>
- [19] Abbasian, P., & Hammond, T. A. (2024). A comprehensive analysis of early Alzheimer disease detection from 3D sMRI images using deep learning frameworks. *Information*, 15(12), 746. <https://doi.org/10.3390/info15120746>
- [20] Warren, S. L., & Moustafa, A. A. (2026). Functional and clinical: An explainable deep learning model for multimodal Alzheimer's disease classification. *Brain and Behavior*, 16, e71240. <https://doi.org/10.1002/brb3.71240>

- [21] Payan, A., & Montana, G. (2015). Predicting Alzheimer's disease using 3D CNN. NeuroImage.
- [22] Odusami, M., Maskeliūnas, R., Damaševičius, R., & Krilavičius, T. (2021). Analysis of features of Alzheimer's disease: Detection of early stage from functional brain changes in magnetic resonance images using a finetuned ResNet18 network. *Diagnostics*, 11(6), 1071. <https://doi.org/10.3390/diagnostics11061071>
- [23] Wang, J., Wen, S., Liu, W., Meng, X., & Jiao, Z. (2024). Deep joint learning diagnosis of Alzheimer's disease based on multimodal feature fusion. *BioData Mining*, 17, 48. <https://doi.org/10.1186/s13040-024-00395-2>