ALSA-3: Customized CNN model through ablation study for Alzheimer's disease classification

BY

Arpa Saha
ID: 242-25-004

This Report Presented in Partial Fulfillment of the Requirements for
The Degree of Masters of Science in Computer Science and Engineering

Supervised By

Sheak Rashed Haider Noori
Professor & Head
Department of CSE
Daffodil International University

Co-Supervised By

Dr. Abdus Sattar
Associate Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

APPROVAL

This Project titled "**ALSA-3: Customized CNN model through ablation study for Alzheimer's disease classification**", submitted by **Arpa Saha**, ID No: **242-25-004** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfilment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on.

BOARD OF EXAMINERS

Chairman

Dr. Sheak Rashed Haider Noori
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Md. Zahid Hasan
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Naznin Sultana
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Mr. Nazibur Rahman
Head of IT Infrastructure
Networld Bangladesh PLC

©Daffodil International University ii

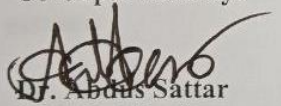
DECLARATION

I hereby declare that this research has been done by me under the supervision of **Sheak Rashed Haider Noori, Professor & Head, Department of CSE,** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

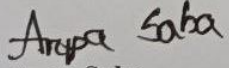
Sheak Rashed Haider Noori
Professor & Head
Department of CSE
Daffodil International University

Co-Supervised by:



Dr. Abdus Sattar
Associate Professor
Department of CSE
Daffodil International University

Submitted by:



Arpa Saha
ID: 242-25-004
Department of CSE
Daffodil International University

©Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartfelt thanks and gratitude to Almighty God for His divine blessing, which makes it possible to complete the final year project successfully.

I am grateful and wish to express my profound indebtedness to **Dr. Sheak Rashed Haider Noori, Professor & Head**, Department of CSE, Daffodil International University, Dhaka, deep knowledge & keen interest in the field of Machine Learning to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, and reading many inferior drafts and correcting them at all stages have made it possible to complete this project.

I would like to express my heartfelt gratitude to **Abdus Sattar**, **Associate Professor**, for his kind assistance in completing our project, as well as to the other faculty members and staff of the CSE department at Daffodil International University.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

Alzheimer's disease is a widespread neurological disorder which affects millions of people around the world and presents complex challenges for patients, families, and healthcare providers. A mix of non-pharmaceutical and pharmacological strategies are needed to manage AD in order to reduce symptoms and enhance quality of life during various phases of the disease. Since the course of AD varies greatly, personalized treatment strategies are vital, with early detection and stage-specific classification playing a key role. Medical imaging, particularly MRI, has become a cornerstone in diagnosing and tracking AD, though its interpretation can be slow and heavily dependent on clinical expertise. To address this issue, we propose a customized CNN model, called ALSA-3, which leverages deep learning for MRI-based AD classification. To refine image quality, methods such as PSNR, RMSE, MSE, and SSIM are applied, while an ablation study fine-tunes the architecture by adjusting hyperparameters and layer configurations. Experimental results demonstrate that ALSA-3 outperforms conventional approaches, reaching an impressive 99.50% accuracy along with precision, recall, and F1-scores of 100%, 99%, and 99% respectively. The model's robustness was further validated across different k-fold values, confirming its reliability. In addition, explainable AI techniques like Grad-CAM and Grad-CAM++ were integrated to make the decision-making process more transparent and interpretable for clinicians. All things considered, this study advances the field of AD categorization research and provides useful insights for patients, caregivers, and the medical community. ALSA-3 has the potential to provide quicker and more accurate diagnoses, which would eventually lead to improved treatment for people with AD because of its high accuracy and efficiency.

Keywords:

Deep learning, image processing, transfer learning, Alzheimer's disease, and ablation research

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	ii
Declaration	iii
Acknowledgements	iv
Abstract	v
CHAPTER	
CHAPTER 1: INTRODUCTION	**1-4**
1.1. Introduction	1-2
1.2. Motivation	3
1.3. Research Objectives	3
1.4. Research Questions	3
1.5. Expected Output	4
1.6. Project Management and Finance	4
1.7. Report Layout	4
CHAPTER 2: BACKGROUND	**5-10**
2.1. Introduction	5
2.2. Related Works	5-8
2.3. Research Gap	9
2.4. Challenges	9-10
CHAPTER 3: RESEARCH METHODOLOGY	**11-18**
3.1. Proposed Methodology	11
3.2. Data Collection Procedure	11-12
3.3. Image Pre-processing	12-13
3.3.1. Contrast Limited Adaptive Histogram Equalization	13
3.3.2. The Bilateral Filter	14
3.3.3. The Green Fire Blue	14-15
3.4. Quality of Images	15
3.4.1. MSE	15
3.4.2. RMSE	16
3.4.3. PSNR	16
3.4.4. SSIM	16
3.5. Evaluation of Quality	16-17

3.6. Methodology	17-18
CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION	**19-35**
4.1. Training Strategy	19
4.2. Model Comparison	19
4.2.1. MobileNetV2	19-20
4.2.2. ResNet50V2	20
4.2.3. InceptionV3	20-21
4.2.4. ALSA-3, the Proposed Model	21-23
4.3. Results and Discussion	23
4.4. Study of Ablation	23-25
4.4.1. Modifying the Flatten Layer	25
4.4.2. Modifying the Batch Size	25
4.4.3. Modifying Learning Rates and Optimizers	25
4.4.4. Final Setup Following Ablation Research	26
4.5. Evaluation Matrices	26
4.6. Performance Analysis	26-30
4.7. Examine Overfitting	30-31
4.8. Explain Ability Model using Grad-Cam	32-34
4.9. Comparison with Existing Works	34-35
CHAPTER 5: IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY	**36-37**
5.1. Impact on Society	36
5.2. Impact on Environment	36-37
5.3. Ethical Aspects	37
5.4. Sustainability Plan	37
CHAPTER 6: CONCLUSION AND FUTURE WORK	**38-39**
6.1. Summary of the Study	38
6.2. Conclusions	38-39
6.3. Implication for Further Study	39
REFERENCES	**40-49**

LIST OF FIGURES

FIGURES	PAGE NO
Fig. 3.1: Image pre-processing workflow	13
Fig. 3.2: Results and flowchart for image preparation	15
Fig. 3.3: A thorough evaluation and study of the Alzheimer's picture collection	18
Fig. 4.1: Custom CNN architecture is suggested	22
Fig. 4.2: A graph showing accuracy and loss	27-28
Fig. 4.3: Confusion matrix for model	29
Fig. 4.4: ROC curve	30

LIST OF TABLES

TABLES	PAGE NO
Table 3.1: Features of AD disease classifications	12
Table 3.2: The parameters of the image enhancing technique	14
Table 4.1: Performance metrics for every model.	27
Table 4.2: Cross-validation result of 3-Fold	31
Table 4.3: Cross-validation result of 5-Fold	31
Table 4.4: Cross-validation result of 7-Fold	31
Table 4.5: Visualization of four classes using grad-cam and grad-cam++	33
Table 4.6: Comparison of performance measures with previous research	35

CHAPTER 1
INTRODUCTION

1.1. Introduction

Alzheimer's is a neurological condition that worsens with time results in severe dementia-related symptoms and a steady reduction in cognitive ability [1–6]. Standard diagnostic methods, which usually include neurological tests, psychological tests, and physical tests, generally don't provide enough precision or efficiency. At the neuropathological level, AD is distinguished by the degradation of brain cells, particularly in the hippocampus and amygdala, which are essential for memory and decision-making processes [1–6]. The condition predominantly impacts adults aged 60 and older, hindering their ability to execute fundamental cognitive functions such as memory retention, problem-solving, and decision-making [1–6]. The global incidence of AD is frightening, with almost six million cases documented in the United States and over four million in India. This highlights the urgent need for better diagnostic tools and treatment options [7–8]. The condition usually gets worse slowly, starting with Mild Cognitive Impairment (MCI), which shows cognitive problems that are less severe but very similar to AD symptoms. Research shows that almost eight out of ten people with MCI will develop AD within seven years, making MCI a high-risk precursor state [9–10]. The MMSE, clinical examination, neurophysiological assessments, and health history evaluation are the key components of standard diagnostic techniques; nevertheless, these methods are time-consuming and often insensitive [11–15].

MRI is essential for the evaluation of neurological disorders, offering comprehensive insights into structural alterations linked to Alzheimer's Disease (AD) and other ailments, such as tumors and malignancies [16–17]. As computer techniques become more widely available, improved image processing has been useful for finding structural variations between cognitively healthy people, people with MCI, and people with AD. Imaging-based classification is becoming a more efficient and cost-effective alternative to traditional diagnostic methods. It may increase the diagnosis of AD's precision and

accessibility [18]. Traditional image processing techniques have considerable difficulties in managing the intricate and uneven pixel distributions linked to neurodegenerative brain alterations, often leading to misdiagnosis [19]. This restriction underscores the necessity for sophisticated computer methodologies capable of precisely interpreting such patterns. Modern methods not only speed up the diagnosis process, but they also show promise for finding biomarkers at earlier stages of disease. This would make it possible to manage and intervene in clinical situations before they get worse [18].

Because AD is so common, it's important to come up with new ways to diagnose it. Combining imaging technology with disease classification algorithms may help doctors find AD earlier, which would mean faster treatment and better clinical outcomes [18]. Ongoing improvements in MRI technology and image analysis have enhanced comprehension of AD's impact on the brain, especially in the discovery of biomarkers that facilitate preventive healthcare and tailored treatment approaches [18]. The changing field of Alzheimer's disease diagnostics shows how important modern neuroimaging and computational methods are in dealing with the growing global burden of this disease. In the field of medical research, unusual computational approaches have demonstrated significant potential. For instance, A. G. et al. [20] presented EfficientCovNet, a modified EfficientNet-based CNN for the diagnosis various lung conditions, such as COVID-19, while specifically, Tchuente et al. [21] created integrated learning models for pulmonary hypertension survival analysis. In a similar way, Ueslei et al. [22] came up with a new way to find tuberculosis using feed-forward neural networks with phase congruency features. Z. Sabir et al. [23] built smart solvers for HIV-1 systems that are associated to cancer progression. In a separate investigation, V. [24] utilized a hybrid Hopfield recurrent neural network (RNN) to forecast psychological diseases. Together, these contributions show how computational methods can change the way we diagnose and treat complicated disorders, such as AD.

1.2. Motivation

AD is a rapidly escalating worldwide health issue that profoundly affects healthcare systems, families, and patients. Early detect and precise staging are essential for effective disease management; however, conventional diagnostic techniques, such as MRI interpretation, frequently require substantial time and depend significantly on clinical expertise. These problems show how important it is to have advanced, automated, and reliable computer models that can improve the accuracy of diagnoses, shorten the time it takes to make decisions, and give clear information about how decisions are made. This will ultimately improve patient care and help medical professionals.

1.3. Research Objectives

- To create and validate a customized CNN model, for the classification of Alzheimer's disease using MRI data.

- To enhance image quality, optimize network architecture through ablation studies, and ensure interpretable clinical outcomes using explainable AI.

- To enhance diagnostic accuracy, reliability, and efficiency in Alzheimer's disease classification, which in turn provides meaningful benefits for patients, caregivers, and the broader medical community.

1.4 . Research Questions

- How can a customized CNN architecture improve MRI-based Alzheimer's disease classification compared to conventional methods?

- To what extent do image enhancement techniques contribute to improved diagnostic accuracy?

- What potential does customize lightweight CNN hold in reducing diagnostic time and supporting clinicians in real-world healthcare environments?

1.5. Expected Output

- A high-performing CNN model (ALSA-3) optimized for the categorization of AD based on MRI.
- Better picture quality thanks to sophisticated processing measures including SSIM, PSNR, RMSE, and MSE.
- A systematic evaluation of the effect of ablation studies on CNN architecture and performance.
- Superior classification accuracy (targeting ~99.50%) with high precision, recall, and F1-scores.
- Validation of model robustness across multiple k-fold values to ensure reliability.
- Integration of Grad-CAM and Grad-CAM++ for explainable AI, enhancing transparency in clinical decision-making.
- A framework that reduces diagnostic time while maintaining high accuracy.
- Practical contributions to patient care, caregiver support, and clinical workflow efficiency.

1.6. Project Management and Finance

The research work doesn’t get fund from any individuals or organization.

1.7. Report Layout

The study is organized into six clear sections so that it can fully cover the study. Section 2 carefully looks at related works, giving us a better understanding of what we already know. Again, section 3 describes the technique and system architecture that are suggested, including the new methods and techniques that were used. The training and model assessment techniques are thoroughly described in Section 4. This gives us a better idea of the methods used to create and test the models. Section 5 summarizes the study's results and the discussions that followed. Section 6 gives a brief overview of the study's conclusions and their implications.

CHAPTER 2
BACKGROUND

2.1. Preliminaries/Terminologies

Recent progress in medical imaging and artificial intelligence has made it possible to create advanced systems for finding and categorizing neurological disorders like Alzheimer's disease. In this context, magnetic resonance imaging (MRI) is an important way to get detailed structural information about the brain. Image processing techniques are then used to improve the quality and get useful features. MSE, RMSE, PSNR, and SSIM are some of the most common metrics used to check the quality of an image and help with preprocessing. Because of deep learning architectures, CNNs are now crucial for automated feature extraction. These networks focus on structural features like shape and texture to classify diseases correctly. Additionally, explainable AI techniques such as Grad-CAM and Grad-CAM++ enhance interpretability by emphasizing pertinent areas in MRI scans, thus reconciling computational predictions with clinical reasoning. The proposed ALSA-3 framework is based on these basic ideas. Its goal is to make Alzheimer's disease diagnosis better by using strong, accurate, and easy-to-understand classification methods.

2.2. Related Works

In the realm of AD investigations, approaches are continually evolving to automate diagnosis and utilize biomarkers for forecasting future disease issues. By employing transfer learning successfully, well-known convolutional neural network (CNN) models have been able to increase classification performance in numerous medical applications. This continual development shows a strong attempt to make it easier to diagnose Alzheimer's disease and make predictions about it, which is a very hard area of study.

Gige et al. [25] developed a system during the early diagnosis of AD by examining MRI brain pictures to classify individuals into four categories: mild AD, moderate AD, no AD, and control. Their comparative analysis of VGG19 and DenseNet architectures demonstrated that VGG19 exhibited greater performance, effectively identifying 6000

images across the specified categories. In the same way, R. A. et al. [26] used DNN models to sort AD makes use of MRI data. They stressed that traditional image processing methods typically miss minor signals because brain architecture are so complicated. Their tests showed that the DenseNet-121 model was 86.55% accurate, but a mix of AlexNet and LeNet was much more accurate (93.58%) and cost less to run. Shojaei et al. [27] progressed this domain by utilizing a 3D-CNN model on ADNI MRI scans and integrating occlusion mapping with backpropagation explainability to pinpoint essential brain regions affecting categorization. When they used 5-fold cross-validation, their model got 87% accuracy, and when they focused on 29 brain regions, it got 93% accuracy.

In a separate study, Raees et al. [28] put forward a technique that sorted MRI pictures from 111 people into three groups: normal cognition, mild cognitive impairment (MCI), and Alzheimer's disease (AD). The model attained an accuracy range of 80–90% utilizing support vector machines (SVM) and deep learning techniques, indicating the feasibility of early Alzheimer's disease detection with considerable socio-economic advantages. Dar et al. [29] introduced a CNN-based architecture for the multi-class classification of AD stages utilizing structural MRI scans. Their approach successfully tackled dataset imbalance via resampling and augmentation, with an accuracy of 96.6% in distinguishing five AD cohorts. They also talked about how useful transfer learning and cloud-based monitoring tools are for real-time AD evaluation. In the same way, Bandyopadhyay et al. [30] used the OASIS dataset and suggested an ensemble-based method along with an ANN. Out of the eight machine learning methods, the ANN performed the best, with an accuracy of 91.9%, compared to 83.04% for the voting classifier and 85.7% for gradient boosting.

Balaji et al. [31] created a hybrid deep learning model by merging convolutional neural networks (CNN) with long short-term memory (LSTM) networks. Their approach was quite good in finding early MCI, with an accuracy of 98.5%, thanks to Adam optimization. The research investigated domain learning to enhance the forecasting of MCI advancement to AD. Jo et al. [32] used functional MRI (fMRI) data with stacked autoencoders to find features and classify them. They got 98.8% accuracy in finding AD and 83.7% accuracy in

predicting the conversion from MCI to AD. Deep learning models yielded similar outcomes with accuracies of 96% and 84.2%, respectively, underscoring the advantages of multimodal strategies that integrate neuroimaging and serum biomarkers. Hu et al. [33] introduced the VGG-TSwinformer, a hybrid model that integrates CNNs with transformers for longitudinal investigation of Mild Cognitive Impairment (MCI). The model utilized spatial and temporal attention mechanisms, achieving 77.2% accuracy in differentiating stable MCI (sMCI) from progressive MCI (pMCI), surpassing cross-sectional MRI-based studies with an AUC score of 0.8153.

Chang et al. [34] suggested a CNN-based approach utilizing T1-weighted MRI scans to distinguish between Alzheimer's Disease (AD), temporal lobe epilepsy (TLE), and normal controls. At 90.45% accuracy, their model demonstrated a strong ability to diagnose TLE. A long-term MRI framework that integrates CNNs and RNNs was suggested by Liu et al. [35] in order to record the temporal and spatial features of brain degeneration. Their approach attained 71.71% accuracy in distinguishing pMCI from sMCI and 91.33% accuracy in differentiating AD from normal controls. Shamrat et al. [36] performed a comparative investigation of five CNN models for MRI-based Alzheimer's disease classification, with InceptionV3 attaining 96.31% accuracy. They also suggested the AlzheimerNet model, which was improved with RMSprop and reached a record accuracy of 98.67%.

Ramya et al. [37] put forward a computer-assisted method that combines improved expectation maximization and adaptive histogram equalization approaches for the segmentation and classification of AD. Their model utilized MRI scans and integrated 2D-adaptive bilateral filtering, GLCM feature extraction, PCA-based dimensionality reduction, and logistic regression for classification, achieving an accuracy of 96.92% and surpassing previous methodologies. Rangaraju et al. [38] developed a octave with dual attention awareness convolutional network (DACN) for the identification of AD utilizing 3D MRI data. Their method, which included patch-based CNNs, octave convolution, and attention modules, got an amazing 99.87% accuracy on the dataset for ADNI.

Maskeliūnas et al. [39] used a vision transformer architecture to combine MRI and PET data at the pixel level in a way that expanded multimodal analysis. Their method reached 81.25% and 93.75% accuracy for MRI and PET modalities, respectively, by using discrete wavelet transformations, VGG16 transfer learning, and fused feature reconstruction. Odusami et al. [40] created a heuristic early feature fusion model that combined PET and MRI data using a modified ResNet18. This model was able to explain AI (XAI) procedures and got 73.9% accuracy. Additionally, Odusami et al. [41] utilized Adaptive learning with Pareto optimization using transposed convolution layers for the integration of MRI and PET, evaluating VGG11, VGG16, and VGG19 models on the ADNI dataset. Their findings illustrated the superiority of VGG19 in relation to SSIM, PSNR, and entropy-based metrics. In their subsequent research, Odusami et al. [42] presented optimal convolutional fusion utilizing multi-kernel architectures and instance normalization to synchronize MRI and PET data. When tested on the ADNI, OASIS, and AANLIB datasets, the method worked almost perfectly, with 99% accuracy, 99% specificity, and 98.44% sensitivity in telling the difference between AD and normal controls.

A considerable number of studies have introduced DL strategies in order to diagnose AD through neuroimaging data. Various architectures, including DenseNet, 3D-CNN, ensemble, VGG19 frameworks, LSTM as well as conventional CNN models, have been widely applied, achieving classification accuracies between 80% and 98.67% when distinguishing cognitively normal subjects, mild cognitive impairment (MCI), and AD [43–46]. However, a persistent limitation of many existing approaches lies in their inability to adequately manage noise and heterogeneity present in MRI scans, which often restricts their ability to generalize effectively across different datasets, thereby reducing diagnostic reliability. To address these challenges, the ALSA-3 framework incorporates advanced image preprocessing techniques alongside comprehensive evaluation metrics, ensuring that image quality is preserved and enhancing the robustness of stage differentiation in AD. Another critical shortcoming in prior studies is the lack of detailed optimization, as many models rely on generalized hyperparameters that fail to exploit the full potential of CNN architectures for Alzheimer's classification [47–50]. Model introduces a novel design that

has undergone extensive experimental validation, leading to a lightweight yet highly efficient network that maintains an optimal trade-off between performance and computational efficiency. Furthermore, the integration of Grad-CAM and Grad-CAM++ within ALSA-3 strengthens the interpretability of its predictions, producing transparent visual explanations. This feature is particularly vital in clinical practice, where interpretability not only supports accurate therapeutic decision-making but also fosters greater trust among healthcare professionals.

2.3. Research Gap

Despite significant progress in the application of deep learning and medical imaging for Alzheimer's disease classification, several critical gaps remain unaddressed. Existing CNN-based models often struggle with balancing accuracy, robustness, and interpretability, as many approaches achieve high performance but fail to provide transparent reasoning behind their decisions, limiting clinical adoption. Additionally, while MRI has proven to be a reliable modality, image quality variations caused by noise and acquisition conditions are frequently overlooked, leading to inconsistent diagnostic outcomes. Many studies also lack comprehensive ablation analyses to justify the necessity of specific architectural choices, which raises concerns about model generalizability and optimization. Furthermore, limited emphasis has been placed on integrating explainable AI techniques that can build clinician trust by highlighting relevant brain regions linked to disease progression. These shortcomings underscore the need for a customized, interpretable, and performance-driven model such as ALSA-3 that systematically addresses image quality, model architecture, and explainability to advance reliable Alzheimer's disease diagnosis.

2.4. Challenges

The development of an effective MRI-based Alzheimer's disease classification model faces several challenges that must be carefully addressed. One major challenge lies in the variability and complexity of MRI data, where differences in acquisition settings, noise, and artifacts can significantly impact image quality and subsequent model performance. Another difficulty is the high dimensionality of MRI scans, which requires powerful

feature extraction and optimization techniques to prevent overfitting while maintaining generalizability. Additionally, achieving a balance between high classification accuracy and interpretability remains a pressing concern, as most deep learning models function as "black boxes" that limit clinical trust. Ensuring robustness across diverse datasets and validation schemes adds further complexity, particularly when dealing with class imbalances across different stages of Alzheimer's disease. Moreover, integrating explainable AI methods without compromising model efficiency presents a technical hurdle. These challenges highlight the importance of designing a tailored and transparent solution, such as ALSA-3, that can effectively manage data variability, ensure accuracy, and provide reliable interpretability for clinical decision support.

CHAPTER 3
RESEARCH METHODOLOGY

3.1. Proposed Methodology/Applied Mechanism

In this part, we talk about the study's overall system design and the methods used. The approach incorporates previous methodologies, but its innovation resides in the distinctive refinement and integration of these techniques, particularly for AD categorization. The proposed framework improves the state of the art by using an optimized CNN architecture, strong image preprocessing methods, and explainable artificial intelligence (XAI) methodologies to get more accurate and reliable results. The ALSA-3 customized convolutional neural network use an ablation study to methodically modify the layer and hyperparameters configurations to make the model work better. Additionally, the report highlights the significance of various image quality assessment metrics—specifically SSIM, PSNR, RMSE, and MSE which guarantee the integrity of preprocessed pictures and enhance diagnostic accuracy. The combination of these preprocessing methods with a well-designed CNN architecture sets ALSA-3 apart from other models that have trouble with the complexity and variability of MRI information in Alzheimer's grouping [51–56]. Also strong point of that work is that it uses XAI methodologies, especially Grad-CAM and Grad-CAM++, which give clear visual explanations for the model's predictions. Such interpretability is crucial for clinical implementation, as it fosters trust, improves decision-making clarity, and provides more profound insights into the diagnostic process [57–58]. The design and methodological factors outlined in this study collectively create a thorough and pragmatic framework for the effective classification of Alzheimer's disease.

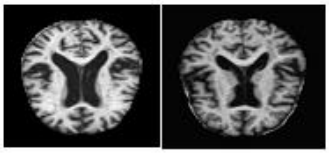
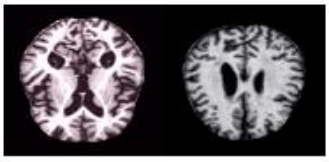
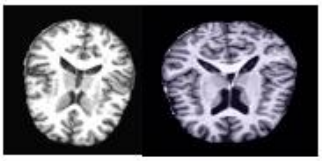
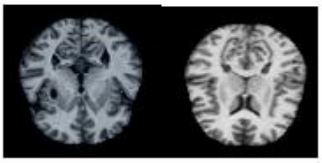
3.2. Data Collection Procedure/Dataset Utilized

The expanded Alzheimer MRI dataset [59], a publicly available online resource, is the source of the dataset used in this investigation. The open-source collection has over 6400 images, categorized into four categories: not demented, somewhat demented, very slightly demented, and mildly demented. Each set has huge number of images. This dataset was divided into 3 sections: test data, which made up the last 15% of the total photographs;

validation data, which made up 15%; and training data, which made up around 70% of all the photos. To assist you completely understand the qualities of the dataset, Table 3.1 includes a sample image of AD pathology along with some explanations.

Table 3.1:

Features of AD disease classifications.

No	Features	Example pictures
1	Mildly	
2	Moderately	
3	Very Mildly	
4	Non-Demented	

3.3. Image Pre-processing

In the field of DL, data preprocessing means cleaning and improving raw data so that it can be used to train DNN. The images in the dataset were resized to 224×224 pixels, which made training faster and made sure the computer could handle it. Rescaling makes training

models more efficient. Used methods are just a few of the preprocessing methods that make models more robust. The different ways that augmentation was done. Image preprocessing, which is very important for getting the best performance out of a neural network to remove noise. MRI images have a lot of noise and not much contrast. Motion correction fixes movement, which makes the image clearer, more consistent, and more accurate. Denoising cuts down on the bilateral filter and noise smooths out pixels while keeping the edges of regions sharp. Statistical evaluation measures like RMSE, MSE, SSIM, and PSNR are used to look at the effects of preprocessing [60]. Figure 3.1 depicts essential processes in the image preprocessing phase, safeguarding image quality.

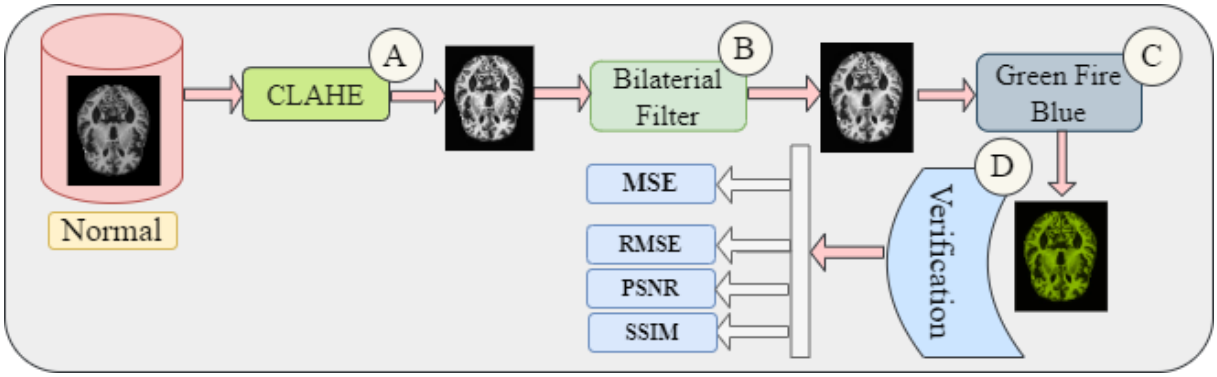


Fig. 3.1: Image pre-processing workflow.

3.3.1. Contrast Limited Adaptive Histogram Equalization

CLAHE is used to lessen excessive contrast in order to equalize the total contrast [61]. Physicians may be able to diagnose general health issues by improving the local contrast in these pictures [62]. You need to set up the tileGridSize and ClipLimit parameters in order to use CLAHE. Each tile's size in terms of rows and columns is determined by the tilegridsize parameter, and the contrast value that will be used as a threshold is specified by the clipLimit argument. The optimal tileGridSize and clipLimit settings were determined to be (12,12) and 0.5 after testing with a number of combinations.

3.3.2. The Bilateral Filter

In image processing, a bilateral filter is a common non-linear filter type that smoothes out pictures while maintaining sharp edges [63]. It maintains borders and minute details by applying an average weighting to pixels based on the brightness along with spatial separation. The algorithm does its laborious job using the sigmaColor, dimension, and sigmaSpace parameters. The number of pixels in each neighborhood is indicated by the dimension, while the color space's sigma is shown by the sigmaColor. These parameters are very important for many image processing tasks, including blurring, smoothing, and improving photographs. The filter's operation may be changed to give you the expected results.

3.3.3. The Green Fire Blue

Last way to improve this study is to use a filter. Changing the colors of the pixels in a data image is what filtering it means. Using filters makes the contrast better and adds a lot of different visual effects to the pictures [64]. We did a number of tests to find the best filter for the dataset.

Table 3.2:

The parameters of the image enhancing technique.

Process	Algorithm	Parameter value
Image enhancement	CLAHE	clipLimit=2.0, tileGridSize=(12, 12)
	Bilateral Filter	sigmaColor=75, diameter =9, sigmaSpace=75
	Green Fire Blue	(green_filter, 0.8, fire_filter, 0.6, 0)

The procedures for enhancing a picture are listed in Table 3.2, including with the algorithms and parameter settings. The contrast-limited adaptive histogram equalization (CLAHE) employs a tileGridSize of (12, 12) and a clipLimit of 2.0 for optimal contrast enhancement. The bilateral filter's sigmaColor and sigmaSpace values are 75 and its

diameter is 9. This ensures that noise is minimized while maintaining edges. The green fire blue method uses a fire filter, a green filter, and three parameters (0.8, 0.6, and 0) to improve the appearance of colors.

The whole picture pre-processing employed in this work to enhance the image quality is shown in Figure 3.2. This image displays a comprehensive set of procedures, including bilateral filtering, green-fire-blue filtering, and CLAHE. The conditions are ideal for the finest outcomes, and every step is meticulously prepared.

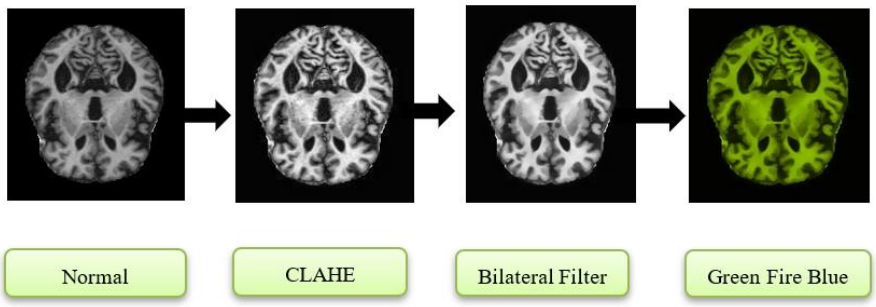


Fig. 3.2: Results and flowchart for image preparation.

3.4. Quality of Images

Different ways of preparing pictures were used, which could have changed the quality of the images. Calculating RMSE, MSE, SSIM, and PSNR, among other things, makes sure that the quality of the image stays the same.

3.4.1. MSE

The sum of the squared discrepancies between the pixel values in two photographs is shown by the mean squared error. An picture of higher quality has a lower MSE value. This is due to the fact that a lower MSE value indicates that the processed picture is very comparable to the original image. An image with no noise has a value of 0. Equation 1 shows what MSE means in math.

$$MSE = \frac{1}{pq}\sum_{t=0}^{m-1}\sum_{j=0}^{n-1}(O(m,n) - P(m,n))^2 \quad (1)$$

3.4.2. RMSE

The difference between the quality of the original picture and the changed image is measured by the root means square error. A lower RMSE, especially one that is getting closer to 0, means fewer mistakes and better image quality. You can find RMSE by using equation 2:

$$RMSE = \left[\sum_{j=1}^{N} \left(d_{f_i} - d_d\right)^2 / N\right]^{\frac{1}{2}} \quad (2)$$

3.4.3. PSNR

By comparing the maximum signal intensity to the power of noise, which reduces the clarity of an image after processing, one may get the peak signal-to-noise ratio. To find PSNR, you first need to find the value of MSE. Then, we use equation 3 to find PSNR:

$$\text{PSNR } = 20\log_{10}\left(\frac{(MAX)}{\sqrt{MSE}}\right) \quad (3)$$

3.4.4. SSIM

The degree to which pre-processing methods degrade picture quality is gauged by the structural similarity index. The estimate is between -1 and 1, with 1 meaning that the structures are exactly the same and 0 meaning that they are not at all similar. The formula 4 is used to figure out SSIM:

$$\text{SSIM } (x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_x^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (4)$$

3.5. Evaluation of Quality

A comprehensive analysis of ten pictures' quality utilizing RMSE, MSE, PSNR, and SSIM. You compare the original and processed images to determine how nice a picture is. Ten photographs selected at random, together with the outcomes of several image evaluation methods for each. By contrasting the modified picture with the initial random image, we were able to determine each image evaluation value. By examining the MSE and RMSE values, you may ascertain the average and root average discrepancies between the original and processed images. PSNR values show that the images are more accurate, while SSIM values show that the images are structurally similar. This careful evaluation gives a

quantitative look at how well image preprocessing techniques work, showing how they affect the quality of images everything throughout the data sources.

MSE results, which fall between 52.31 and 58.10, reveal how dissimilar with original and processed images are on a pixel-by-pixel level. The RMSE values, which range from 7.23 to 7.54, illustrate how big these discrepancies are on average. The PSNR values between 30.48 and 30.94 demonstrate that the image quality has gotten improved, while the degree of similarity between the original and modified photos is indicated by SSIM scores between 0.62 and 0.66. In general, these metrics give a good idea of how effectively the image processing or enhancement methods worked on the photos being looked at.

3.6. Methodology

Figure 3.3 shows a structured flowchart that shows the seven steps (A–G) that are used to evaluate and analyze an Alzheimer's disease picture dataset. Step A starts with the dataset, which has 6,400 MRI pictures that are divided into four clinical groups: non-demented, very mild dementia, mild dementia, and moderate dementia. In Step B, the preprocessing phase, advanced enhancement techniques including contrast-limited adaptive histogram equalization (CLAHE), bilateral filtering, and the blue, green, and fire filters are used to improve visual quality and bring out important features for further analysis. In Step C, established metrics including mean squared error (MSE), root mean squared error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM) are used to check the quality of the images. This makes sure that only high-quality images move on to the next steps. In Step D, model selection, three well-known convolutional neural networks—MobileNetV2, InceptionV3, and ResNet50V2—are trained on the processed dataset to define a baseline for performance. Step E introduces the ablation study, which systematically fine-tunes important hyperparameters such batch size, the configuration of the flatten layer, the choice of optimizer, the learning rate, the loss function, and the activation function to make the suggested ALSA-3 architecture work better. Step F is all about analyzing the results. It looks at a number of evaluation measures, such as classification accuracy, loss–accuracy curves, confusion matrices, and receiver operating characteristic (ROC) curves, to get a full picture of how well the model works in terms of accuracy, precision, reliability, and robustness. Finally, Step G talks about robustness

evaluation, which uses the K-fold cross-validation method to check how well the ALSA-3 model works on different datasets. This methodical pipeline guarantees a robust analytical framework and offers enhanced insights into the progression of Alzheimer's disease, facilitating greater diagnostic reliability and aiding the creation of more effective therapeutic decision-making tools.

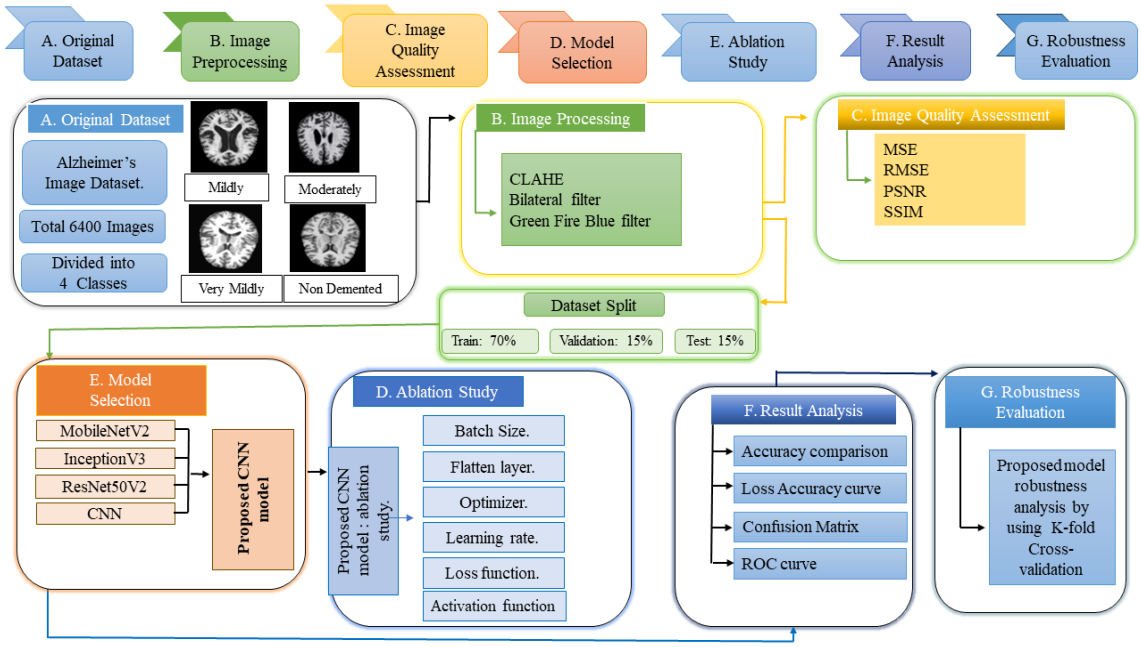


Fig. 3.3: A thorough evaluation and study of the Alzheimer's picture collection.

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4. Training and Model Evaluation

This part explains the full method used to train the models and test how well they work. For a thorough assessment, it contains details on model architectures, training techniques, and the measures that are used to compare them.

4.1. Training Strategy

To make the most of computing power and speed up model convergence, the model was trained in batches of 64 across 50 epochs. The Adam optimizer was used with a learning rate of 0.001 and the categorical cross-entropy loss function, which is great for jobs that involve classifying several classes. We used the rectified linear unit (ReLU) activation function to provide good estimates of class probabilities. The dataset was divided into three parts: training, testing, and validation. Each section got 70%, 15%, and 15% of the photos, respectively. The validation set was taken from the training set to keep an eye on how well the model generalizes. Three dedicated personal computers were used for the experiments. Each has a solid-state drive (SSD), 256 GB of DDR4 RAM, an NVIDIA GPU, and an Intel CPU, which was enough power to compare alternative models and hyperparameter settings.

4.2. Model Comparison

This study conducts comparative analysis among three deep learning models and the proposed model, emphasizing the assessment of their accuracy. The assessment utilized images measuring 224 × 224 pixels. Here is a description of these models.

4.2.1. MobileNetV2

A CNN is called MobileNetV2 (CNN) of 53 layers [65]. The model has been trained on more than a million images from the ImageNet dataset, which includes 100 different types of objects. This lets it learn a lot of different characteristics that may be used for image

classification tasks. There are 32 filters and 19 bottleneck layers in MobileNetV2's architecture, which are all arranged in fully convolutional layers. There are two primary blocks in the network, each with three layers. There are also 11 convolutional layers at the start and end of each block, all of which use 32 filters. This setup lets MobileNetV2 strike a compromise between fast computing and strong feature extraction, which makes it good for transfer learning in many areas of picture analysis.

4.2.2. ResNet50V2

There are 50 layers in the ResNet50 architecture. These include 48 convolutional layers and independent average and max pooling layers [66]. The reference says that the model has over 23 million parameters that can be learned. A lot of computer vision programs use ResNet50 because its design solves the disappearing gradient problem by adding residual shortcut connections that let gradients move more easily. The suggested model employs a stacked convolution strategy to amalgamate convolution and max pooling operations, a method that has shown highly effective in mitigating the issues related to vanishing gradients, therefore improving network stability and performance.

4.2.3. InceptionV3

One convolutional neural network (CNN) designed for picture sorting is the InceptionV3 architecture [67]. Several layers are stacked on top of one another in its structure, which progressively extracts and encodes valuable features from input pictures. The model initially normalizes the input photographs' pixel values to ensure consistency in the data. Convolution and spatial downsampling are carried out via the Conv2D and MaxPooling2D procedures in the first layers, respectively. This enables the network to detect local trends rapidly. The starting weights are established using the He normal initialization technique, and the Conv2D layers are made non-linear with ReLU activation functions. MaxPooling2D layers make feature maps smaller in space, which makes them faster to compute. Dropout layers randomly turn off some neurons during training to stop overfitting and let the learnt characteristics generalize. The flatten layer comes after the convolutional layers and turns multidimensional feature maps into a one-dimensional vector.

The recovered characteristics may then be used by the completely connected thick layers. Four neurons with a softmax activation function in the output layer forecast the likelihood that each input image falls into one of the classifications. ReLU activation is used in the thick layers as well.

4.2.4. ALSA-3, the Proposed Model

The foundation of the ALSA-3 models' design is a technology that makes them lightweight and universal while maintaining consistently high performance. The suggested model is referred to as ALSA-3. CNNs do ablation research to determine what works. In this research, there are three convolution layer blocks. Figure 4.1 shows how to assemble the model. To choose the layers and parameters of the bespoke model, we carried out an ablation study. Making it easier to use and more precise was the aim. In an ablation study, certain network components are purposefully removed in order to gain understanding of how the network functions and identify the best layers for the suggested model. Adding a resealing layer is the first step in creating a bespoke CNN model, which ensures that the model always receives 224 x 224 images. This will assist maintain a consistent performance. The first convolutional layer then receives the rescaled picture. In order to identify various local elements in every picture, including edges, textures, and so on, it will use a 16-pixel filter size. The design has many convolutional layers. 16 filters and a 3x3 kernel make up the first convolutional layer with "same" padding. ReLU is used to turn it on. A pooling layer was then used. This preserved the most important portion while halving the spatial size. In order to identify the biggest value inside a constrained area—generally indicating the most prominent characteristic of that area a max-pooling layer with a 2x2 pool size was used in the study. 32 unique characteristics were extracted from the combined output, an additional convolution layer has been introduced in the following block. 32 is the filter size. Since the initial convolution only discovered simple characteristics, this stage is helpful for identifying more complex ones. A second max pooling layer is then used to exclude elements that aren't necessary or as significant. Because so many characteristics were eliminated, the model may have been overfitted or skewed toward certain patterns up until this point. To prevent overfitting, a dropout layer

has been included into the model. During training, this layer randomly changes 20% of the input unit to 0. This helps the model learn to generalize. The next block now has another convolution layer. This one can find things that are 64 different sizes. The proposed design used bigger filter sizes in the convolutional layers to make it easier to extract features. This was especially important because the units had a 20% random dropout, which might let the network learn certain patterns. Increasing the size of the filters lets the convolutional layers pick up on more features, which helps with generalization. A second max-pooling layer was added to choose the most important features from the complicated patterns that had already been found. To further reduce bias and overfitting, a dropout layer with a rate of 25% was added. This made sure that 75% of the learnt features were sent to the next dense layers for aggregation and learning. There are two fully connected dense layers in the network, one with 128 neurons and the other with 64. The first dense layer puts together and processes all the characteristics that the convolutional layers found. The second dense layer then fine-tunes and combines these representations for the final classification. The softmax activation function in the output layer lets the model put images into more than one class, which makes it good for multiclass classification problems.

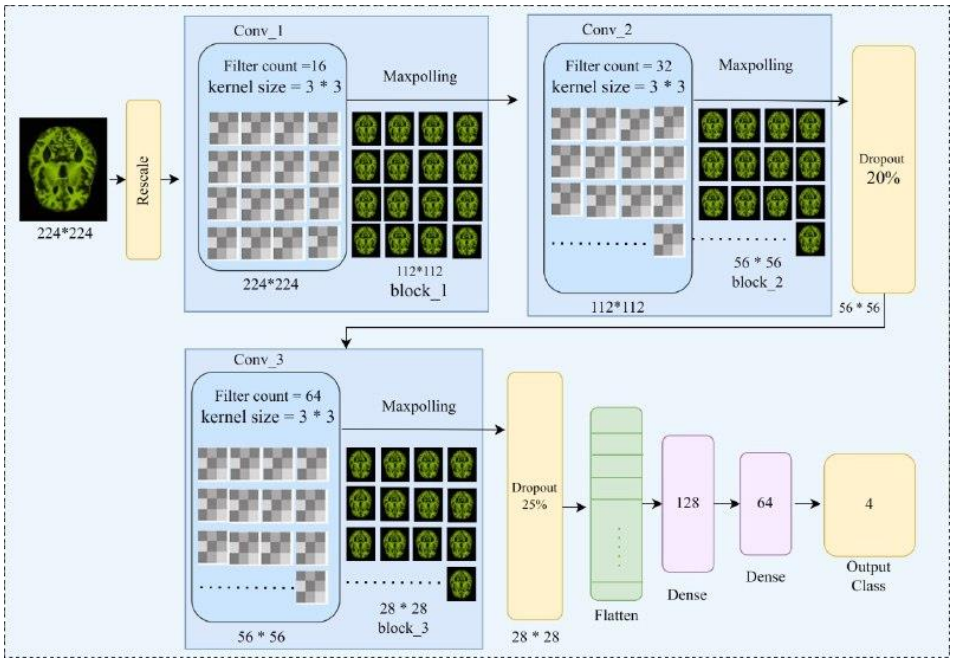


Fig. 4.1: Custom CNN architecture is suggested.

Modifications to certain features of a model's architecture may either enhance, worsen, or leave the network's performance unaltered. By modifying hyper-parameters like learning rates, optimizers, batch sizes, and loss functions, accuracy may often be raised. The overall effectiveness of the model is significantly impacted by changes made to its design. Analysis was done on the results of four case studies. The total accuracy increased, indicating that the strategy was effective.

4.3. Results and Discussion

This part goes into great detail on the research on ablation done on the model of ALSA-3, looking at how each part and hyperparameter affects the model's overall performance. The study looks at the accuracy and loss curves to give a clear picture of how the model is learning. A close look at the confusion matrix shows even more how well the model can tell the difference between different stages of Alzheimer's disease. We also used K-fold cross-validation to see how strong and generic ALSA-3 was. To fix the problem of class imbalance during training, class weights were used to make sure that each class had a fair share of the loss function. The model gave more weight to minority classes, which made these less common categories more important. This led to big gains in F1 score, recall, and precision, and it also reduced bias toward majority classes. This method made the model better at correctly identifying and classifying samples from less common classes, which led to more balanced performance measures. The ablation investigation was conducted via four consecutive experiments, methodically modifying essential elements of the ALSA-3 architecture, such as the optimizer, learning rate, batch size, loss function, and flatten layer. This methodical evaluation provided information about the unique contributions of each element, directing the model's improvement for reliable and accurate Alzheimer's disease categorization.

4.4. Study of Ablation

Because it methodically evaluates the ALSA-3 model to determine if each component is required and how it helps, the ablation study is essential to this investigation, which proves that the model is strong and works well for diagnosing Alzheimer's disease. Ablation

experiments are structured to isolate and evaluate the significance of particular components inside a model, guaranteeing that each piece substantively enhances overall performance [68–72]. In this study, we looked at important factors including the flattening learning rates, optimizer, layer, and batch size to see how they affected the performance of the model on their own and in combination. This thorough study not only backs up design choices, but it also helps us understand how different parts work together better, which helps us make a very accurate and reliable model with finding AD early also figuring out what stage it is in.

Because it converts the multi-dimensional output of convolutional layers into a one-dimensional vector that fully connected layers may use, the flattening layer is crucial. This layer's ablation experiments demonstrated how modifications to the model's configuration impact its ability to identify complex patterns, which in turn impacts the model's overall precision and forecasting ability. The study emphasizes the significance of maintaining spatial hierarchies and enhancing feature extraction for efficient Alzheimer's disease classification by examining many configurations [73–74].

We also looked at batch size to see how it affected the training of the ALSA-3 models stability, convergence rate, and overall performance. Batch size is a very important hyperparameter that has a big effect on gradient estimates, computational efficiency, and the trade-off between training speed and model correctness. The findings indicate that choosing a suitable batch size is essential for attaining optimal classification of AD stages [75–76].

The impact of optimizers and learning rates on parameter updates during training of the ALSA-3 model was also investigated. Tests with the Adam optimizer at different learning rates showed how these choices affect convergence, generalization, and avoiding local minima [77]. This investigation underscores the imperative of meticulously adjusting optimization parameters to guarantee swift, stable, and precise learning, that is essential for accurate classification of AD.

By identifying the key elements of the ALSA-3 model and refining its design, this ablation study significantly advances the research. The study guarantees optimal performance by systematically assessing the contribution of each aspect, resulting in a very dependable and precise approach for diagnosing the stages of Alzheimer's disease.

4.4.1. Modifying the Flatten Layer

The Flatten layer achieved the highest performance in the ALSA-3 model, with a validation accuracy (Val_Acc) of 98.75% and a test accuracy (Test_Acc) of 99.50%. At 98.569%, 75.292%, and 98.569%, respectively, other pooling techniques such as AveragePooling2D, MaxPooling2D, and GlobalMaxPooling2D fared worse. These findings demonstrate that the Flatten layer performs better at maintaining feature representation and aiding in the accurate classification of Alzheimer's disease stages.

4.4.2. Modifying the Batch Size

The choice of batch size affects how well the ALSA-3 model classifies things. To assess its effect, experiments utilized batch sizes of 32, 64, and 128. The results showed that both validation and test accuracy got better when the batch size went up from 32 to 64 and then 128. The lowest loss value was 0.036 and the highest test accuracy was 99.50% with a batch size of 64. For further testing, a batch size of 64 was used as it ensures that the model operates more regularly and dependably.

4.4.3. Modifying Learning Rates and Optimizers

In order to identify the best option for training the ALSA-3 model, this experimental research evaluated four optimizers: SGD, Adam, Nadam, and Adamax. Additionally, the research examined the effects of several learning rates (0.01, 0.001, and 0.0001) on the model's performance. Using a 0.001 learning rate and test and validation accuracy of 99.50% includes 98.75%, respectively, Adam was the best optimizer evaluated. Although the accuracy scores of other optimizers were comparable to Adam's, Adam was selected for further research since it consistently outperformed the others.

4.4.4. Final Setup Following Ablation Research

The correct calibration of the learning rate, loss function, batch size, optimizer, activation function, and dropout improved the model's accuracy, according to ablation studies.

4.5. Evaluation Matrices

The effectiveness of a statistical, machine learning, or deep learning model can be determined by analyzing its evaluation metrics. Assessing model performance requires the use of appropriate metrics, and it is essential to examine a study's model through multiple indicators. A true positive (TP) is when the model correctly identifies the positive class, while a true negative (TN) occurs when the model accurately recognizes the negative class. The model generates FP and FN when it incorrectly determines the positive or negative class. The f1-score, accuracy, precision, and recall are metrics used to assess this model's capacity for classification or prediction [78]. The equations (5), (6), (7), and (8) show the performance parameters.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

$$Recall\ or\ TPR = \frac{TP}{TP+FN} \quad (7)$$

$$F1 - Score = \frac{2 \times Recall \times Precision}{Recall+Precision} \quad (8)$$

4.6. Performance Analysis

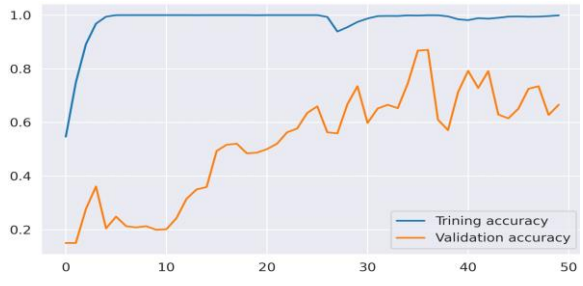
We utilize equations (5), (6), (7), and (8) to identify a number of objects for each model to see how well they function. The results of the performance metrics are shown in Table 4.1. With an accuracy score of 99.50%, our recommended ALSA-3 model provides the highest overall performance and accuracy. It also obtained a perfect score of 100% for precision, 99% for recall, and 100% for f1-score. InceptionV3, ResNet50, and MobileNetV2 aren't very good because their accuracy rates are only 97.50%, 98.78%, and 87.11%, respectively.

Table 4.1:

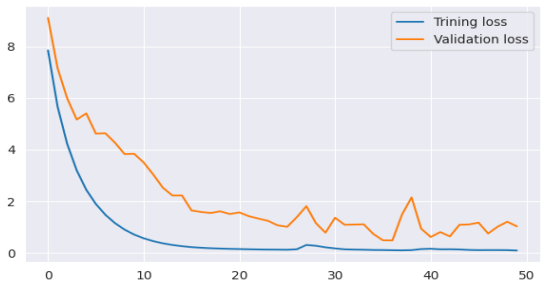
Performance metrics for every model.

Models	Recall (%)	Precision (%)	F1-score (%)	Accuracy (%)
MobileNetV2	86	94	89	87.11
InceptionV3	97	98	98	97.50
ResNet50	98	99	99	98.78
ALSA-3	99	1.00	1.00	99.50

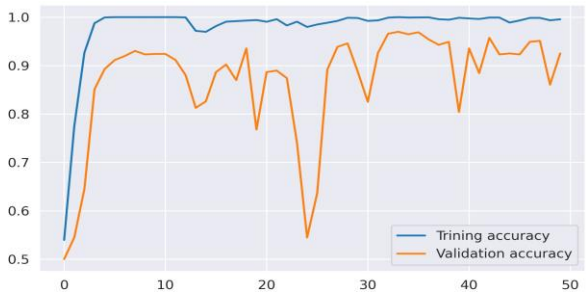
Graphs of accuracy and loss [79] might help you understand how the model works better. The graph illustrates the loss value for each epoch and how accurate the training and validation were. The accuracy graph shows the effectiveness of each epoch run for both validation and training. How much you lost throughout training and testing is shown on the loss graph. Figure 4.2 presents the accuracy and loss graphs for the ALSA-3 model, MobileNetV2, InceptionV3, ResNet50, and others. As illustrated in Figures 4.2(g) and 4.2(h), model accuracy improves with each epoch, while the loss value decreases. Moreover, compared to the other models, the proposed ALSA-3 model achieved the highest accuracy and the lowest loss.



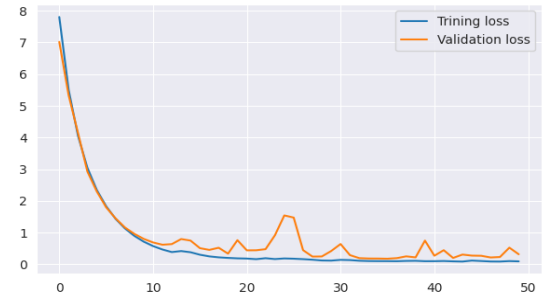
(a) MobileNetV2 accuracy graph.



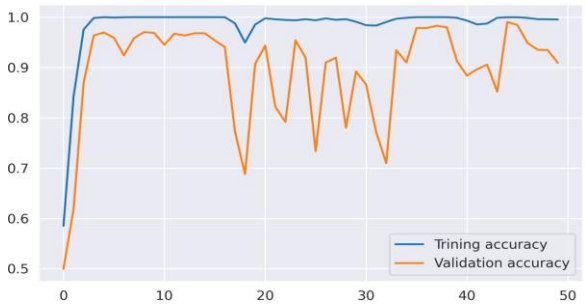
(b) MobileNetV2 loss graph.



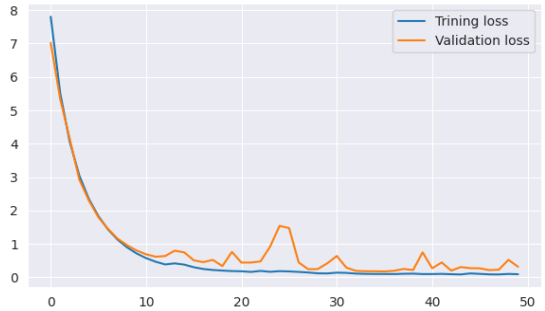
(c) InceptionV3 accuracy graph



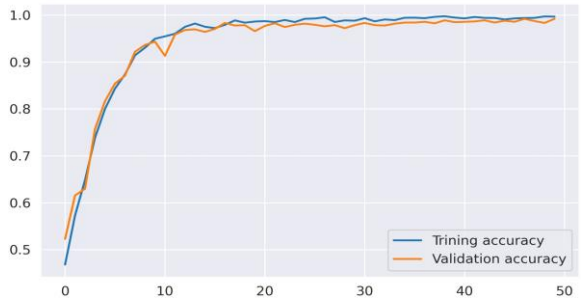
(d) InceptionV3 loss graph.



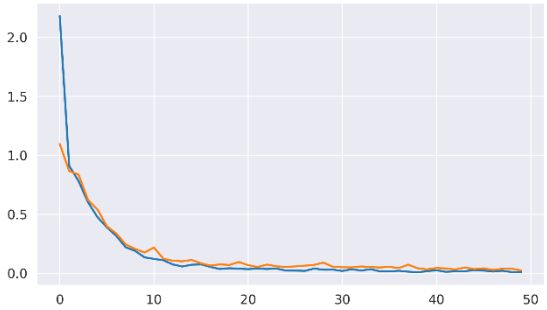
(e) ResNet50V2 accuracy graph



(f) ResNet50V2 loss graph.



(g) ALSA-3 CNN accuracy graph

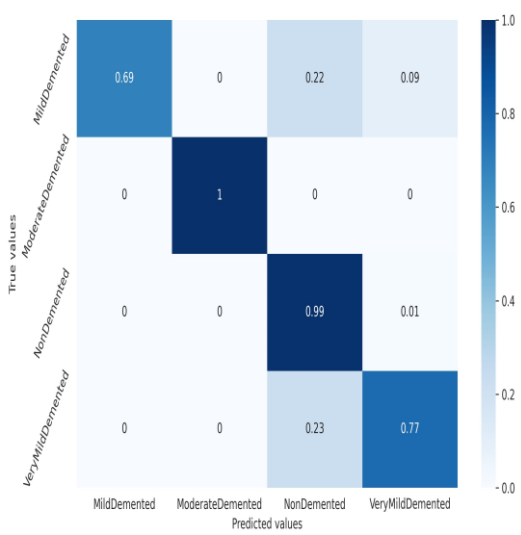


(h) ALSA-3 CNN loss graph.

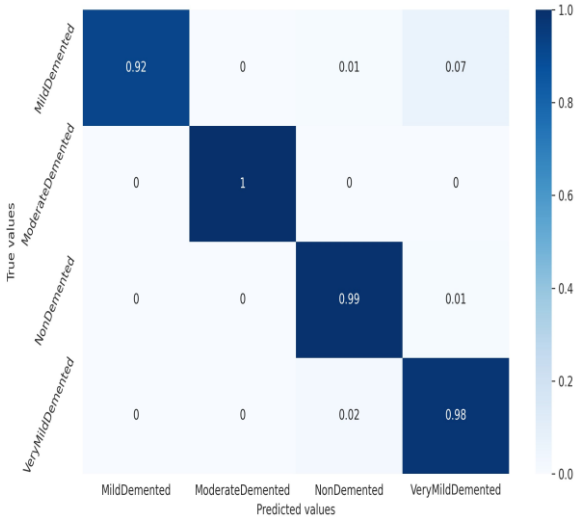
Fig. 4.2: A graph showing accuracy and loss.

Figure 4.3 displays the confusion matrix [80], which juxtaposes the performance of the ALSA-3 model with that of other models. The numbers in the rows show the genuine labels that were assigned to the test photos, while the numbers in the columns show the labels that were predicted for the test images. The percentage of test images that were correctly predicted is shown according to the confusion matrix's diagonal. Figure 4.3(d) shows that

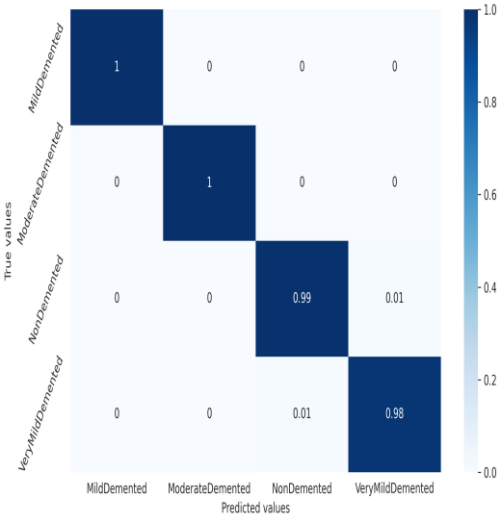
our proposed model predicts all four disease classes fairly evenly, without favoring any particular class. The results indicate that the model makes approximately the same number of correct predictions for each class.



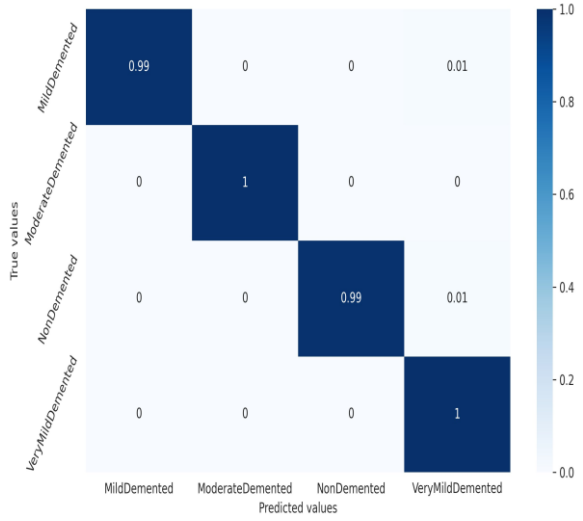
(a) MobileNetV2.



(b) InceptionV3



(c) Resnet50V2.



(d) Proposed ALSA-3.

Fig. 4.3: Confusion matrix for model.

The receiver operating characteristic (ROC) [81] and the area under the curve (AUC) [82] metrics were used to see how well the ALSA-3 model worked. The AUC measures how well the model can tell the difference between distinct classes. Values closer to 1 mean that the model does a better job at classifying. Figure 4.4 shows the ROC curve, which shows that the model has a high discriminative capability because the true positive rate is close to 1 and the false positive rate is close to 0. The ALSA-3 model has a perfect AUC score of 100%, which shows that it is very accurate and good at identifying Alzheimer's disease.

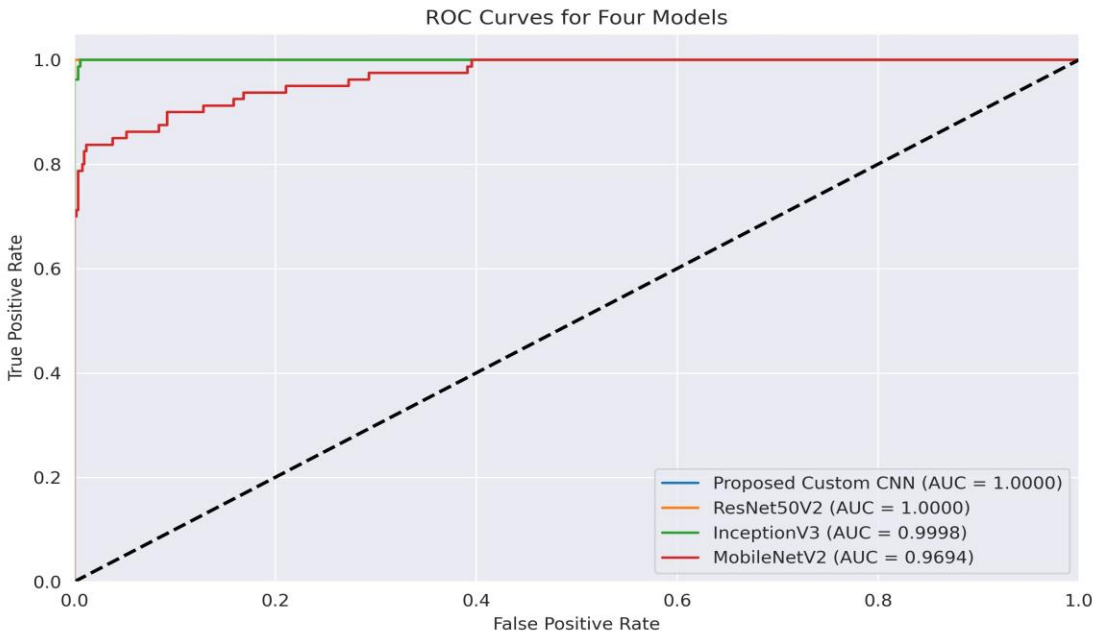


Fig. 4.4: ROC curve.

4.7. Examine Overfitting

Several techniques were used to reduce the likelihood of overfitting during the ALSA-3 model's training and testing. We used K-fold cross-validation and other cross-validation methods to test how well the model worked on different parts of the dataset shows in table 4.2, 4.3, 4.4. This made sure that the model was trained and tested on a wide range of cases, which improved its generalizability [83–86].

We also used regularization methods like dropout and weight decay to make the model more resilient. Dropout randomly turns off neurons during training to stop the model from relying too much on certain features. Weight decay punishes big weights to encourage models that work well with new data [87–88]. To further enhance generalization, participants from a variety of age ranges, genders, and nationalities, creating a sample representative of the broader population.

A confusion matrix was used to fully test the model's performance. It showed that ALSA-3 could accurately tell the difference between different levels of Alzheimer's disease severity [89–91], as shown by consistently excellent recall and accuracy in every class. Advanced visualization approaches, including heatmaps and grad-CAM, facilitated interpretability by emphasizing pertinent areas of MRI images that contribute to categorization, hence improving clinical application.

We utilized k-fold cross-validation [92] with k values of 3, 5, and 7 [93] to check for overfitting and generalization. This was shown by differences between the training and validation accuracies. Validation accuracies regularly surpassed 98% across all folds, and the training and validation metrics were well aligned, demonstrating that ALSA-3 delivers dependable performance without overfitting under diverse data shuffling scenarios.

Table: 4.2.
Cross-validation result of 3-Fold.

Accuracy (%)	F-1	F-2	F-3	Avg
Training	99.92	99.61	100	99.84
Validation	96.62	96.65	98.14	98.13

Table: 4.3.
Cross-validation result of 5-Fold.

Accuracy (%)	F-1	F-2	F-3	F-4	F-5	Avg
Training	100	100	99.80	99.80	99.80	99.88
Validation	99.22	99.25	99.27	99.27	99.27	99.25

Table: 4.4.
Cross-validation result of 7-Fold.

Accuracy (%)	F-1	F-2	F-3	F-4	F-5	F-6	F-7	Average
Training	100	100	99.89	99.89	100	99.91	100	99.95
Validation	99.39	99.39	99.39	99.39	99.59	99.51	99.63	99.47

4.8. Explain Ability Model using Grad-Cam

The purpose of explainable artificial intelligence (XAI) is to design AI systems that can deliver clear and understandable justifications for their predictions. In this study, Grad-CAM and Grad-CAM++ are highlighted as essential methods frequently applied to interpret model behavior [94–95]. Grad-CAM, a widely used approach in existing research, visualizes the decision-making process of deep neural networks. It effectively highlights the key regions of an input image that contribute to a specific model prediction. By quickly calculating gradients with respect to CNN feature maps, Grad-CAM accurately pinpoints the regions of interest. The generated heatmap serves as a valuable tool for researchers, as it reveals how the model attends to different image areas during classification. The essence of Grad-CAM can be summarized in the following mathematical equation.

$$\alpha_k^c = \frac{1}{Z}\sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^K} \tag{10}$$

$$L_{\text{Grad-CAM}}^c = \text{ReLU}\left(\sum_K \alpha_k^c A^k\right) \tag{11}$$

y^c = The network's class C pre-SoftMax activation score

A^k = Feature map activations

α_k^c = Neuronal weights

Z = The number of pixels in the feature map

The fundamental grad-cam technique is improved upon by Grad-Cam++. Its main goal is to make localization more accurate. By taking into account both positive and negative gradients, this evolution makes it possible for the CNN's foundational layers to absorb more information. Grad-CAM++ significantly improves the accuracy of finding unique features in an image by combining these gradients. This improvement, in turn, makes CNN models clearer and easier to understand. The following equation summarizes the formulation for

grad-cam++, which takes with more nuanced approach to improving localization than its predecessor.

$$W_k^c = \sum_i \sum_j \alpha_{i,j}^{kc} \text{ReL } U\left(\frac{\partial Y^c}{\partial A_{i,j}^k}\right) \quad (12)$$

W_k^c = Weights of neurons

A^k = Feature mapping activations

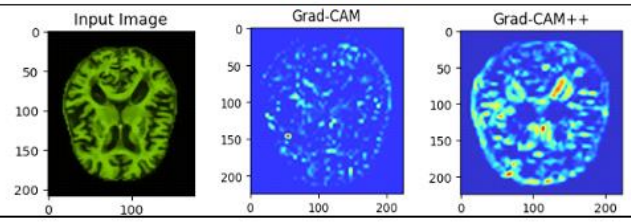
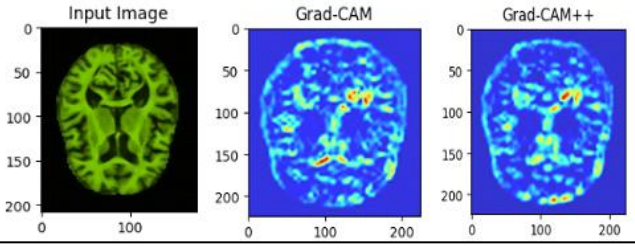
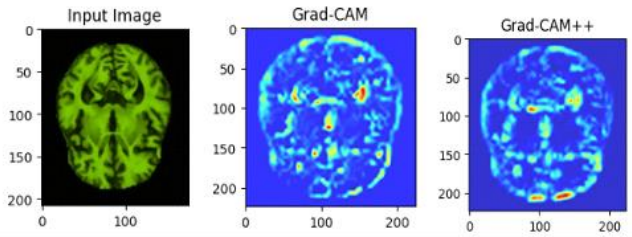
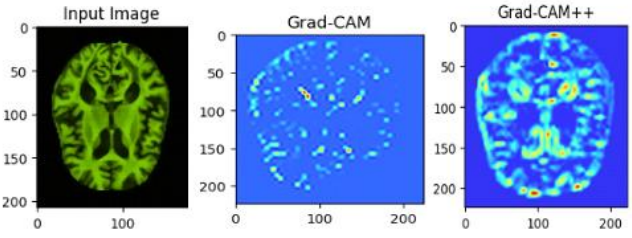
c = Class of targets

y^c = The network's pre-SoftMax activation score for Class C

$\alpha_{i,j}^{kc}$ = Emphasizes importance of location

Table 4.5:

Visualization of four classes using grad-cam and grad-cam++.

Class	Visualization
Very Mild Demented	
Non-Demented	
Moderate Demented	
Mild Demented	

Grad-cam does a good job of highlighting disease areas and explaining important features, but grad-cam++ does an even better job by giving a more complete picture of important features in table 4.5. A careful look at the resulting heatmaps gives us important information about how the model makes decisions, which helps us understand why each prediction was made. Medical professionals can learn about how the model works by looking at these heatmaps. This may increase their confidence in the model's predictions and aid in their diagnostic decision-making. For individuals with very mild dementia and moderate dementia, for instance, the heatmaps may show activity in regions such as the entorhinal cortex and hippocampal regions that are associated with early indicators of cognitive deterioration. The model may display bigger regions in moderate to severe instances, indicating worsening sickness. The model helps doctors check its predictions and understand the underlying pathology by giving them these visual cues. The model's interpretability is of utmost importance, as it guarantees transparency and dependability, especially in situations where the model's decisions have a substantial impact on essential applications.

4.9. Comparing with Previous Works

The efficacy of the proposed technique for identifying AD from MRI scans is further assessed by comparing its performance with the findings of recent research. Table 4.6 shows the accuracies test of contemporary works, which are used in this comparison. The table reveals that the earlier study's results were less accurate, and the method hasn't been utilized in real life yet.

Our research utilized a compact, efficient model that surpassed prior studies in accurately classifying Alzheimer's disease. That study has also successfully employed a very effective methodology to make an Android app that would help doctors and nurses use their smartphones.

Table 4.6:

Comparison of performance measures with previous research.

Author and reference	Year	Models used	Best model with accuracy
Hazarika et al. [26]	2023	DenseNet-121, LeNet and AlexNet hybrid	Hybrid technique (accuracy: 93.58%)
Shojaei et al. [27]	2023	3D-CNN	3D-CNN (accuracy: 93% with 29 brain regions)
Dar et al. [29]	2023	CNN architectures	Proposed framework (accuracy: 96.6%)
Bandyopadhyay et al. [30]	2023	Ensemble methods, ANN	ANN (accuracy: 91.9%)
Jo et al. [32]	2019	DL, fMRI	DL (accuracy: 96%)
Hu et al. [33]	2023	VGG-TSwinformer	VGG-TSwinformer (accuracy: 77.2%)
Chang et al. [34]	2023	CNN	CNN (accuracy: 90.45%)
Cui et al. [35]	2019	CNN, RNN	Proposed framework (accuracy: 91.33%)
Shamrat et al. [36]	2023	Various CNN models, AlzheimerNet	AlzheimerNet (accuracy: 98.67%)
Our study	-	ALSA-3	ALSA-3 (accuracy: 99.50%)

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1. Impact on Society

The incorporation of the ALSA-3 model for the classification of Alzheimer's disease holds considerable promise for societal impact, especially in the healthcare domain. The model can cut down on the time between the start of symptoms and clinical intervention by making MRI-based diagnosis faster and more accurate. This will improve patient outcomes and quality of life. Its high accuracy and ease of understanding make it possible to provide expert-level diagnostic support to areas where there aren't many specialized neurologists, which would make healthcare morefair. Also, ALSA-3 can help lower the economic burden of Alzheimer's disease by cutting down on unnecessary tests and helping with more targeted treatment planning. As the model becomes more integrated into clinical workflows, it can help create personalized care plans that meet the needs of each patient. This will improve disease management and help people live longer, healthier lives. At the same time, it is still important to make sure that these new technologies are used in a fair and ethical way so that the benefits of ALSA-3 reach a wide range of people without making existing inequalities worse.

5.2. Impact on the Environment

Similar to many deep learning frameworks, the ALSA-3 model for classifying AD needs a lot of computing power for training and inference. This can lead to higher energy use and a big carbon footprint. Using high-performance GPUs for a long time to process a lot of MRI data makes this problem even harder. Still, improving the ALSA-3 architecture through ablation studies, better tuning of hyperparameters, and less data redundancy can make it less demanding on computers and save energy. Also, using the model in cloud infrastructures that run on renewable energy sources can help reduce environmental costs even more. The clinical use of ALSA-3 may significantly decrease dependence on extensive physical diagnostic tools, repetitive imaging sessions, and superfluous hospital visits, thereby indirectly reducing resource consumption and emissions linked to patient

travel. In the future, improvements in lightweight deep learning methods and green computing strategies could make ALSA-3 not only a useful tool for doctors, but also an environmentally friendly way to fight Alzheimer's disease.

5.3. Ethical Aspects

Using the ALSA-3 model to classify AD raises a number of ethical issues about patient privacy, data security, and the reliability of algorithms. MRI data is very private, so keeping patient trust means protecting their privacy and making sure that data is handled safely. Also, ALSA-3 needs to be clear about how it makes diagnostic decisions, especially since doctors need to be able to understand, question, and confirm the model's results. Another important problem is that datasets may be biased, which could lead to different levels of diagnostic accuracy for different demographic groups. If this isn't fixed, it could make healthcare disparities worse. To maintain ethical standards, ALSA-3 must be comprehensible, equitable, and resilient, guaranteeing that its implementation in clinical practice adheres to the principles of fairness, responsibility, and patient-centered care.

5.4. Sustainability Plan

To keep the ALSA-3 model for classifying AD up to date and clinically useful, a long-term plan must include constantly improving the architecture and adding new medical imaging insights. Healthcare professionals need ongoing training and support just as much as they need to be able to use the model as part of diagnostic workflows in a way that is both effective and ethical. To deal with concerns about sustainability, we should focus on making computers more efficient and using renewable energy sources for large-scale use. It will also be important to encourage partnerships with hospitals, research institutions, policymakers, insurance companies, and patient advocacy groups in order to create a long-lasting ecosystem. These kinds of partnerships can not only make sure that everyone has fair access to advanced diagnostic tools, but they can also build trust, accountability, and long-term value in clinical practice. This will allow ALSA-3 to play a big role in both improving healthcare delivery and responsible technological progress.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1. Summary of the Study

This research concentrated on the creation and assessment of ALSA-3, a tailored convolutional neural network intended for the classification of Alzheimer's disease stages utilizing MRI scans. The study focused on combining specialized convolutional blocks and attention mechanisms, which were tested through rigorous training and validation on a carefully chosen neuroimaging dataset. Performance metrics like accuracy, precision, recall, and F1-score showed how strong ALSA-3 was at telling the difference between different stages of Alzheimer's disease. The model's architecture, which was improved through a long ablation study, showed how tailored feature extraction and multi-level representation could improve classification performance. This makes ALSA-3 a promising tool for early and accurate diagnosis.

6.2. Conclusions

This study highlights the critical use of sophisticated DL methods, namely a customized CNN model and its variants, including EfficientNetB3, MobileNetV2, and ResNet50, in the MRI-based categorization of AD. The accuracy of the suggested ALSA-3 model was 99.50%, which was superior than that of conventional methods. It had 100% accuracy, 99% recall, and 100% F1-score, respectively. These results show that the ALSA-3 framework is more stable and reliable than other methods because it combines efficient image processing techniques with careful tweaking of hyperparameters. The study underscores the vital significance of early diagnosis and precise classification of Alzheimer's disease, facilitating prompt therapies that can markedly enhance outcomes for patients, caregivers, and healthcare providers.

Even though these results are promising, there are several limits that need to be recognized in order to help future research. The dataset's size and diversity pose a significant barrier, requiring initiatives to merge larger and more heterogeneous datasets to enhance the model's generalizability across diverse medical circumstances. Future research may

investigate the amalgamation of sources of multimodal data to attain a more thorough comprehension of the course of Alzheimer's disease. Additionally, making the model easier to understand is important for its use in the clinic. Using explainable AI methods would make it more transparent, build trust among medical experts, and help it be used more widely in real-world healthcare settings. These changes are necessary to make the ALSA-3 model more reliable, useful, and clinically beneficial for diagnosing Alzheimer's disease.

This work has important effects on society in addition to its scientific achievements. The ALSA-3 model's high accuracy and reliability can make it easier to diagnose Alzheimer's disease quickly and correctly, which can lead to earlier treatment, less stress for caregivers, and fewer healthcare expenses. Furthermore, progress in deep learning-based classification of Alzheimer's disease aids global health initiatives by enhancing comprehension of neurological illnesses and fostering the creation of novel diagnostic and therapeutic approaches. The socioeconomic significance of this research highlights its prospective influence on public health at both national and global scales.

6.3. Implication for Further Study

The results of the ALSA-3 study open up many new areas of research. There is potential to investigate the integration of multimodal neuroimaging data, such as PET or fMRI, to enhance diagnostic information and increase classification accuracy. Testing ALSA-3 in real-world clinical settings could give us useful information about how well it works in practice and show us where it could be improved. Moreover, additional scrutiny into the explicability of ALSA-3's predictions could guarantee that its decision-making is congruent with clinical reasoning, thereby enhancing trust and facilitating the progress of interpretable AI solutions in the diagnosis of Alzheimer's disease.

REFERENCES

[1] Alzheimer's Association, "2018 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 14, no. 3, pp. 367–429, Mar. 2018, doi: 10.1016/j.jalz.2018.02.001.

[2] Korolev, I.O. Alzheimer's disease: A clinical and basic science review. Med. Stud. Res. J. 2014, 4, 24–33.

[3] R. Donev, M. Kolev, B. Millet, and J. Thome, "Neuronal death in Alzheimer's disease and therapeutic opportunities," J Cellular Molecular Medi, vol. 13, no. 11–12, pp. 4329–4348, Nov. 2009, doi: 10.1111/j.1582-4934.2009.00889.x.

[4] S. W. Moon, B. Lee, and Y. C. Choi, "Changes in the Hippocampal Volume and Shape in Early-Onset Mild Cognitive Impairment," Psychiatry Investig, vol. 15, no. 5, pp. 531–537, May 2018, doi: 10.30773/pi.2018.02.12.

[5] J. Barnes, J. L. Whitwell, C. Frost, K. A. Josephs, M. Rossor, and N. C. Fox, "Measurements of the Amygdala and Hippocampus in Pathologically Confirmed Alzheimer Disease and Frontotemporal Lobar Degeneration," Arch Neurol, vol. 63, no. 10, p. 1434, Oct. 2006, doi: 10.1001/archneur.63.10.1434.

[6] R. A. Hazarika, A. K. Maji, S. N. Sur, B. S. Paul, and D. Kandar, "A Survey on Classification Algorithms of Brain Images in Alzheimer's Disease Based on Feature Extraction Techniques," IEEE Access, vol. 9, pp. 58503–58536, 2021, doi: 10.1109/ACCESS.2021.3072559.

[7] "Alzheimer's Disease Fact Sheet," National Institute on Aging. Accessed: Nov. 10, 2023. [Online]. Available: https://www.nia.nih.gov/health/alzheimers-disease-fact-sheet

[8] "Alzheimer's & Dementia Help | INDIA | Alzheimer's Association." Accessed: Nov. 10, 2023. [Online]. Available: https://www.alz.org/in/dementia-alzheimers-en.asp#diagnosis

[9] Y. Varatharajah et al., "Predicting Short-term MCI-to-AD Progression Using Imaging, CSF, Genetic Factors, Cognitive Resilience, and Demographics," Sci Rep, vol. 9, no. 1, p. 2235, Feb. 2019, doi: 10.1038/s41598-019-38793-3.

[10] "What Is Mild Cognitive Impairment?," National Institute on Aging. Accessed: Nov. 10, 2023. [Online]. Available: https://www.nia.nih.gov/health/what-mild-cognitive-impairment

[11] A. Burns and S. Iliffe, "Alzheimer's disease," BMJ, vol. 338, no. feb05 1, pp. b158–b158, Feb. 2009, doi: 10.1136/bmj.b158.

[12] "Diagnosing Alzheimer's: How Alzheimer's is diagnosed - Mayo Clinic." Accessed: Nov. 10, 2023. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/in-depth/alzheimers/art-20048075

[13] F. J. Huff, F. Boller, F. Lucchelli, R. Querriera, J. Beyer, and S. Belle, "The Neurologic Examination in Patients With Probable Alzheimer's Disease," Archives of Neurology, vol. 44, no. 9, pp. 929–932, Sep. 1987, doi: 10.1001/archneur.1987.00520210031015.

[14] Arevalo-Rodriguez, I.; Smailagic, N.; i Figuls, M.R.; Ciapponi, A.; Sanchez-Perez, E.; Giannakou, A.; Pedraza, O.L.; Cosp, X.B.; Cullum, S. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst. Rev. 2015, 23, 107–120.

[15] J. L. Cummings, W. Ross, J. Absher, J. Gornbein, and L. Hadjiaghai, "Depressive Symptoms in Alzheimer Disease: Assessment and Determinants," Alzheimer Disease & Associated Disorders, vol. 9, no. 2, pp. 87–93, 1995, doi: 10.1097/00002093-199509020-00005.

[16] M. Symms, "A review of structural magnetic resonance neuroimaging," Journal of Neurology, Neurosurgery & Psychiatry, vol. 75, no. 9, pp. 1235–1244, Sep. 2004, doi: 10.1136/jnnp.2003.032714.

[17] M. F. Ijaz, M. Attique, and Y. Son, "Data-Driven Cervical Cancer Prediction Model with Outlier Detection and Over-Sampling Methods," Sensors, vol. 20, no. 10, p. 2809, May 2020, doi: 10.3390/s20102809.

[18] C. Ledig, A. Schuh, R. Guerrero, R. A. Heckemann, and D. Rueckert, "Structural brain imaging in Alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database," Sci Rep, vol. 8, no. 1, p. 11258, Jul. 2018, doi: 10.1038/s41598-018-29295-9.

[19] Y. R. Fung, Z. Guan, R. Kumar, J. Y. Wu, and M. Fiterau, "Alzheimer's Disease Brain MRI Classification: Challenges and Insights," 2019, doi: 10.48550/ARXIV.1906.04231.

[20] A. G. Argho, M. M. S. Maswood, Md. I. Mahmood, and N. Mondol, "EfficientCovNet: A CNN-based approach to detect various pulmonary diseases including COVID-19 using modified EfficientNet," Intelligent Systems with Applications, vol. 21, p. 200315, Mar. 2024, doi: 10.1016/j.iswa.2023.200315.

[21] G. Tchuente Foguem, L. Coulibaly, and A. Diamoutene, "Combined learning models for survival analysis of patients with pulmonary hypertension," Intelligent Systems with Applications, vol. 21, p. 200321, Mar. 2024, doi: 10.1016/j.iswa.2023.200321.

[22] A. Ueslei Da Fonseca et al., "A novel tuberculosis diagnosis approach using feed-forward neural networks and binary pattern of phase congruency," Intelligent Systems with Applications, vol. 21, p. 200317, Mar. 2024, doi: 10.1016/j.iswa.2023.200317.

[23] Z. Sabir, S. B. Said, and Q. Al-Mdallal, "Artificial intelligent solvers for the HIV-1 system including AIDS based on the cancer cells," Intelligent Systems with Applications, vol. 21, p. 200309, Mar. 2024, doi: 10.1016/j.iswa.2023.200309.

[24] V. Kamra, P. Kumar, and M. Mohammadian, "An intelligent disease prediction system for psychological diseases by implementing hybrid hopfield recurrent neural network approach," Intelligent Systems with Applications, vol. 18, p. 200208, May 2023, doi: 10.1016/j.iswa.2023.200208.

[25] Pradhan, A., Gige, J. and Eliazer, M., 2021. Detection of Alzheimer's disease (AD) in MRI images using deep learning. Int. J. Eng. Res. Technol. (IJERT), 10, pp.580-585.

[26] R. A. Hazarika et al., "An Approach for Classification of Alzheimer's Disease Using Deep Neural Network and Brain Magnetic Resonance Imaging (MRI)," Electronics, vol. 12, no. 3, p. 676, Jan. 2023, doi: 10.3390/electronics12030676.

[27] S. Shojaei, M. Saniee Abadeh, and Z. Momeni, "An evolutionary explainable deep learning approach for Alzheimer's MRI classification," Expert Systems with Applications, vol. 220, p. 119709, Jun. 2023, doi: 10.1016/j.eswa.2023.119709.

[28] P. C. Muhammed Raees and V. Thomas, "Automated detection of Alzheimer's Disease using Deep Learning in MRI," J. Phys.: Conf. Ser., vol. 1921, no. 1, p. 012024, May 2021, doi: 10.1088/1742-6596/1921/1/012024.

[29] G. Mohi Ud Din Dar et al., "A Novel Framework for Classification of Different Alzheimer's Disease Stages Using CNN Model," Electronics, vol. 12, no. 2, p. 469, Jan. 2023, doi: 10.3390/electronics12020469.

[30] A. Bandyopadhyay, S. Ghosh, M. Bose, A. Singh, A. Othmani, and K. Santosh, "Alzheimer's Disease Detection Using Ensemble Learning and Artificial Neural Networks," in Recent Trends in Image Processing and Pattern Recognition, vol. 1704, K. Santosh, A. Goyal, D. Aouada, A. Makkar, Y.-Y. Chiang, and S. K. Singh, Eds., in Communications in Computer and Information Science, vol. 1704. , Cham: Springer Nature Switzerland, 2023, pp. 12–21. doi: 10.1007/978-3-031-23599-3_2.

[31] P. Balaji, M. A. Chaurasia, S. M. Bilfaqih, A. Muniasamy, and L. E. G. Alsid, "Hybridized Deep Learning Approach for Detecting Alzheimer's Disease," Biomedicines, vol. 11, no. 1, p. 149, Jan. 2023, doi: 10.3390/biomedicines11010149.

[32] T. Jo, K. Nho, and A. J. Saykin, "Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data," Front. Aging Neurosci., vol. 11, p. 220, Aug. 2019, doi: 10.3389/fnagi.2019.00220.

[33] Hu, Z., Wang, Z., Jin, Y. and Hou, W., 2023. VGG-TSwinformer: Transformer-based deep learning model for early Alzheimer's disease prediction. Computer Methods and Programs in Biomedicine, 229, p.107291.

[34] A. J. Chang et al., "MRI-based deep learning can discriminate between temporal lobe epilepsy, Alzheimer's disease, and healthy controls," Commun Med, vol. 3, no. 1, p. 33, Feb. 2023, doi: 10.1038/s43856-023-00262-4.

[35] R. Cui and M. Liu, "RNN-based longitudinal analysis for diagnosis of Alzheimer's disease," Computerized Medical Imaging and Graphics, vol. 73, pp. 1–10, Apr. 2019, doi: 10.1016/j.compmedimag.2019.01.005.

[36] F. M. J. M. Shamrat et al., "AlzheimerNet: An Effective Deep Learning Based Proposition for Alzheimer's Disease Stages Classification From Functional Brain Changes in Magnetic Resonance Images," IEEE Access, vol. 11, pp. 16376–16395, 2023, doi: 10.1109/ACCESS.2023.3244952.

[37] J. Ramya, B. Uma Maheswari, M. P. Rajakumar, and R. Sonia, "Alzheimer's Disease Segmentation and Classification on MRI Brain Images Using Enhanced Expectation Maximization Adaptive Histogram (EEM-AH) and Machine Learning.," ITC, vol. 51, no. 4, pp. 786–800, Dec. 2022, doi: 10.5755/j01.itc.51.4.28052.

[38] B. Rangaraju, T. Chinnadurai, S. Natarajan, and V. Raja, "Dual Attention Aware Octave Convolution Network for Early-Stage Alzheimer's Disease Detection," ITC, vol. 53, no. 1, pp. 302–316, Mar. 2024, doi: 10.5755/j01.itc.53.1.34536.

[39] M. Odusami, R. Maskeliūnas, and R. Damaševičius, "Pixel-Level Fusion Approach with Vision Transformer for Early Detection of Alzheimer's Disease," Electronics, vol. 12, no. 5, p. 1218, Mar. 2023, doi: 10.3390/electronics12051218.

[40] M. Odusami, R. Maskeliūnas, R. Damaševičius, and S. Misra, "Explainable Deep-Learning-Based Diagnosis of Alzheimer's Disease Using Multimodal Input Fusion of PET and MRI Images," J. Med. Biol. Eng., vol. 43, no. 3, pp. 291–302, Jun. 2023, doi: 10.1007/s40846-023-00801-3.

[41] M. Odusami, R. Maskeliūnas, and R. Damaševičius, "Pareto Optimized Adaptive Learning with Transposed Convolution for Image Fusion Alzheimer's Disease Classification," Brain Sciences, vol. 13, no. 7, p. 1045, Jul. 2023, doi: 10.3390/brainsci13071045.

[42] M. Odusami, R. Maskeliūnas, and R. Damaševičius, "Optimized Convolutional Fusion for Multimodal Neuroimaging in Alzheimer's Disease Diagnosis: Enhancing Data Integration and Feature Extraction," JPM, vol. 13, no. 10, p. 1496, Oct. 2023, doi: 10.3390/jpm13101496.

[43] S. Venkatasubramanian, J. N. Dwivedi, S. Raja, N. Rajeswari, J. Logeshwaran, and A. Praveen Kumar, "Prediction of Alzheimer's Disease Using DHO-Based Pretrained CNN Model," Mathematical Problems in Engineering, vol. 2023, pp. 1–11, Jun. 2023, doi: 10.1155/2023/1110500.

[44] K. C. Sharmili, G. P. Suja, E. Pandian, Md. A. A. Walid, S. Arunachalam, and G. C. Babu, "An Effective Diagnosis of Alzheimer's Disease with the Use of Deep Learning based CNN Model," in 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India: IEEE, May 2023, pp. 443–448. doi: 10.1109/ICICCS56967.2023.10142306.

[45] A. M. El-Assy, H. M. Amer, H. M. Ibrahim, and M. A. Mohamed, "A novel CNN architecture for accurate early detection and classification of Alzheimer's disease using MRI data," Sci Rep, vol. 14, no. 1, p. 3463, Feb. 2024, doi: 10.1038/s41598-024-53733-6.

[46] J. Xin, A. Wang, R. Guo, W. Liu, and X. Tang, "CNN and swin-transformer based efficient model for Alzheimer's disease diagnosis with sMRI," Biomedical Signal Processing and Control, vol. 86, p. 105189, 2023, doi: https://doi.org/10.1016/j.bspc.2023.105189.

[47] J. Li et al., "What's Behind the Mask: Understanding Masked Graph Modeling for Graph Autoencoders," in Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Long Beach CA USA: ACM, Aug. 2023, pp. 1268–1279. doi: 10.1145/3580305.3599546.

[48] R. SinhaRoy and A. Sen, "A Hybrid Deep Learning Framework to Predict Alzheimer's Disease Progression Using Generative Adversarial Networks and Deep Convolutional Neural Networks," Arab J Sci Eng, vol. 49, no. 3, pp. 3267–3284, Mar. 2024, doi: 10.1007/s13369-023-07973-9.

[49] M. Hatami, F. Yaghmaee, and R. Ebrahimpour, "Investigating the potential of reinforcement learning and deep learning in improving Alzheimer's disease classification," Neurocomputing, vol. 597, p. 128119, Sep. 2024, doi: 10.1016/j.neucom.2024.128119.

[50] H.-D. Nguyen, M. Clément, V. Planche, B. Mansencal, and P. Coupé, "Deep grading for MRI-based differential diagnosis of Alzheimer's disease and Frontotemporal dementia," Artificial Intelligence in Medicine, vol. 144, p. 102636, Oct. 2023, doi: 10.1016/j.artmed.2023.102636.

[51] U. Sara, M. Akter, and M. S. Uddin, "Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study," JCC, vol. 07, no. 03, pp. 8–18, 2019, doi: 10.4236/jcc.2019.73002.

[52] U. Ejaz et al., "COMPARATIVE ANALYSIS OF LOSSY IMAGE COMPRESSION ALGORITHMS," PJOSR, vol. 3, no. 2, pp. 136–147, Dec. 2023, doi: 10.57041/pjosr.v3i2.1043.

[53] Y. Chen, R. Wang, P. Zong, and D. Chen, "Image Processing for Denoising Using Composite Adaptive Filtering Methods Based on RMSE," OJAppS, vol. 14, no. 03, pp. 660–675, 2024, doi: 10.4236/ojapps.2024.143047.

[54] K. Ohashi et al., "Applicability Evaluation of Full-Reference Image Quality Assessment Methods for Computed Tomography Images," J Digit Imaging, vol. 36, no. 6, pp. 2623–2634, Dec. 2023, doi: 10.1007/s10278-023-00875-0.

[55] Tasmim, S. C. Bakchy, H. I. Peyal, Md. R. Hoshen, and Md. A. I. Siddique, "Colon Cancer Detection Using A Lightweight-CNN With Grad-CAM++ Visualization," in 2024 3rd International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), Gazipur, Bangladesh: IEEE, Apr. 2024, pp. 1–6. doi: 10.1109/ICAEEE62219.2024.10561717.

[56] Y. Gao, J. Liu, W. Li, M. Hou, Y. Li, and H. Zhao, "Augmented Grad-CAM++: Super-Resolution Saliency Maps for Visual Interpretation of Deep Neural Network," Electronics, vol. 12, no. 23, p. 4846, Nov. 2023, doi: 10.3390/electronics12234846.

[57] K. Mridha, Md. M. Uddin, J. Shin, S. Khadka, and M. F. Mridha, "An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart Healthcare System," IEEE Access, vol. 11, pp. 41003–41018, 2023, doi: 10.1109/ACCESS.2023.3269694.

[58] K. Mridha, Md. M. Uddin, J. Shin, S. Khadka, and M. F. Mridha, "An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart

Healthcare System," IEEE Access, vol. 11, pp. 41003–41018, 2023, doi: 10.1109/ACCESS.2023.3269694.

[59]"Augmented Alzheimer MRI Dataset," www.kaggle.com. https://www.kaggle.com/datasets/uraninjo/augmented-alzheimer-mri-dataset. (accessed Nov. 28, 2023).

[60] V. Mudeng, M. Kim, and S. Choe, "Prospects of Structural Similarity Index for Medical Image Analysis," Applied Sciences, vol. 12, no. 8, p. 3754, Apr. 2022, doi: 10.3390/app12083754.

[61] F. J. M. Shamrat, S. Azam, A. Karim, K. Ahmed, F. M. Bui, and F. De Boer, "High-precision multiclass classification of lung disease through customized MobileNetV2 from chest X-ray images," Computers in Biology and Medicine, vol. 155, p. 106646, Mar. 2023, doi: 10.1016/j.compbiomed.2023.106646.

[62] P. Wang, J. Wang, Y. Li, P. Li, L. Li, and M. Jiang, "Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing," Biomedical Signal Processing and Control, vol. 65, p. 102341, Mar. 2021, doi: 10.1016/j.bspc.2020.102341.

[63] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), Bombay, India: Narosa Publishing House, 1998, pp. 839–846. doi: 10.1109/ICCV.1998.710815.

[64] S. Montaha et al., "BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated Using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images," Biology, vol. 10, no. 12, p. 1347, Dec. 2021, doi: 10.3390/biology10121347.

[65] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT: IEEE, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.

[66] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," 2015, doi: 10.48550/ARXIV.1512.03385.

[67] Xiaoling Xia, Cui Xu, and Bing Nan, "Inception-v3 for flower classification," in 2017 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, China: IEEE, Jun. 2017, pp. 783–787. doi: 10.1109/ICIVC.2017.7984661.

[68] T. Illakiya, K. Ramamurthy, M. V. Siddharth, R. Mishra, and A. Udainiya, "AHANet: Adaptive Hybrid Attention Network for Alzheimer's Disease Classification Using Brain Magnetic Resonance Imaging," Bioengineering, vol. 10, no. 6, p. 714, Jun. 2023, doi: 10.3390/bioengineering10060714.

[69] S. K. Tripathy et al., "Alzheimer's Disease Detection via Multiscale Feature Modelling Using Improved Spatial Attention Guided Depth Separable CNN," Int J Comput Intell Syst, vol. 17, no. 1, p. 113, May 2024, doi: 10.1007/s44196-024-00502-y.

[70] C. Qian and Y. Wang, "MMANet: A Multi-Task Residual Network for Alzheimer's Disease Classification and Brain Age Prediction," IRBM, vol. 45, no. 3, p. 100840, Jun. 2024, doi: 10.1016/j.irbm.2024.100840.

[71] V. Adarsh, G. R. Gangadharan, U. Fiore, and P. Zanetti, "Multimodal classification of Alzheimer's disease and mild cognitive impairment using custom MKSCDDL kernel over CNN with transparent decision-making for explainable diagnosis," Sci Rep, vol. 14, no. 1, p. 1774, Jan. 2024, doi: 10.1038/s41598-024-52185-2.

[72] X. Xing et al., "Efficient Training on Alzheimer's Disease Diagnosis with Learnable Weighted Pooling for 3D PET Brain Image Classification," Electronics, vol. 12, no. 2, p. 467, Jan. 2023, doi: 10.3390/electronics12020467.

[73] K. T. Chui, B. B. Gupta, W. Alhalabi, and F. S. Alzahrani, "An MRI Scans-Based Alzheimer's Disease Detection via Convolutional Neural Network and Transfer Learning," Diagnostics, vol. 12, no. 7, p. 1531, Jun. 2022, doi: 10.3390/diagnostics12071531.

[74] S. Montaha, S. Azam, A. K. M. R. H. Rafid, Md. Z. Hasan, A. Karim, and A. Islam, "TimeDistributed-CNN-LSTM: A Hybrid Approach Combining CNN and LSTM to Classify Brain Tumor on 3D MRI Scans Performing Ablation Study," IEEE Access, vol. 10, pp. 60039–60059, 2022, doi: 10.1109/ACCESS.2022.3179577.

[75] Y. AbdulAzeem, W. M. Bahgat, and M. Badawy, "A CNN based framework for classification of Alzheimer's disease," Neural Comput & Applic, vol. 33, no. 16, pp. 10415–10428, Aug. 2021, doi: 10.1007/s00521-021-05799-w.

[76] F. Zhang et al., "A Single Model Deep Learning Approach for Alzheimer's Disease Diagnosis," Neuroscience, vol. 491, pp. 200–214, May 2022, doi: 10.1016/j.neuroscience.2022.03.026.

[77] S. Tabarestani et al., "A Tensorized Multitask Deep Learning Network for Progression Prediction of Alzheimer's Disease," Front. Aging Neurosci., vol. 14, p. 810873, May 2022, doi: 10.3389/fnagi.2022.810873.

[78] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," 2020, doi: 10.48550/ARXIV.2010.16061.

[79] T. U. Ahmed, M. N. Jamil, M. S. Hossain, K. Andersson, and M. S. Hossain, "An Integrated Real-Time Deep Learning and Belief Rule Base Intelligent System to Assess Facial Expression Under Uncertainty," in 2020 Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan: IEEE, Aug. 2020, pp. 1–6. doi: 10.1109/ICIEVicIVPR48672.2020.9306622.

[80] N. D. Marom, L. Rokach, and A. Shmilovici, "Using the confusion matrix for improving ensemble classifiers," in 2010 IEEE 26-th Convention of Electrical and Electronics Engineers in Israel, Eilat, Israel: IEEE, Nov. 2010, pp. 000555–000559. doi: 10.1109/EEEI.2010.5662159.

[81] R. Kumar and A. Indrayan, "Receiver operating characteristic (ROC) curve for medical researchers," Indian Pediatr, vol. 48, no. 4, pp. 277–287, Apr. 2011, doi: 10.1007/s13312-011-0055-4.

[82] Jin Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," IEEE Trans. Knowl. Data Eng., vol. 17, no. 3, pp. 299–310, Mar. 2005, doi: 10.1109/TKDE.2005.50.

[83] C. M. Chabib, L. J. Hadjileontiadis, and A. A. Shehhi, "DeepCurvMRI: Deep Convolutional Curvelet Transform-Based MRI Approach for Early Detection of Alzheimer's Disease," IEEE Access, vol. 11, pp. 44650–44659, 2023, doi: 10.1109/ACCESS.2023.3272482.

[84] D. A. Arafa, H. E.-D. Moustafa, H. A. Ali, A. M. T. Ali-Eldin, and S. F. Saraya, "A deep learning framework for early diagnosis of Alzheimer's disease on MRI images," Multimed Tools Appl, vol. 83, no. 2, pp. 3767–3799, Jan. 2024, doi: 10.1007/s11042-023-15738-7.

[85] M. Sudharsan and G. Thailambal, "Alzheimer's disease prediction using machine learning techniques and principal component analysis (PCA)," Materials Today: Proceedings, vol. 81, pp. 182–190, 2023, doi: 10.1016/j.matpr.2021.03.061.

[86] K. M. M. Uddin, M. J. Alam, Jannat-E-Anawar, M. A. Uddin, and S. Aryal, "A Novel Approach Utilizing Machine Learning for the Early Diagnosis of Alzheimer's Disease," Biomedical Materials & Devices, vol. 1, no. 2, pp. 882–898, Sep. 2023, doi: 10.1007/s44174-023-00078-9.

[87] D. Guesmi, F. Salah, and Y. B. Ayed, "Recognition of Alzheimer's Disease Based on Transfer Learning Approach Using Brain MR Images with Regularization," in Advances in Computational Collective Intelligence, vol. 1864, N. T. Nguyen, J. Botzheim, L. Gulyás, M. Nunez, J. Treur, G. Vossen, and A. Kozierkiewicz, Eds., in Communications in Computer and Information Science, vol. 1864. , Cham: Springer Nature Switzerland, 2023, pp. 147–160. doi: 10.1007/978-3-031-41774-0_12.

[88] M. Masud, A. M. Almars, M. B. Rokaya, H. Meshref, I. Gad, and E.-S. Atlam, "A Novel Light-Weight Convolutional Neural Network Model to Predict Alzheimer's Disease Applying Weighted Loss Function," Journal of Disability Research, vol. 3, no. 4, Apr. 2024, doi: 10.57197/JDR-2024-0042.

[89]N. Raza, A. Naseer, M. Tamoor, and K. Zafar, "Alzheimer Disease Classification through Transfer Learning Approach," Diagnostics, vol. 13, no. 4, p. 801, Feb. 2023, doi: 10.3390/diagnostics13040801.

[90] A. Assmi, K. Elhabyb, A. Benba, and A. Jilbab, "Alzheimer's disease classification: a comprehensive study," Multimed Tools Appl, vol. 83, no. 27, pp. 70193–70216, Jan. 2024, doi: 10.1007/s11042-024-18306-9.

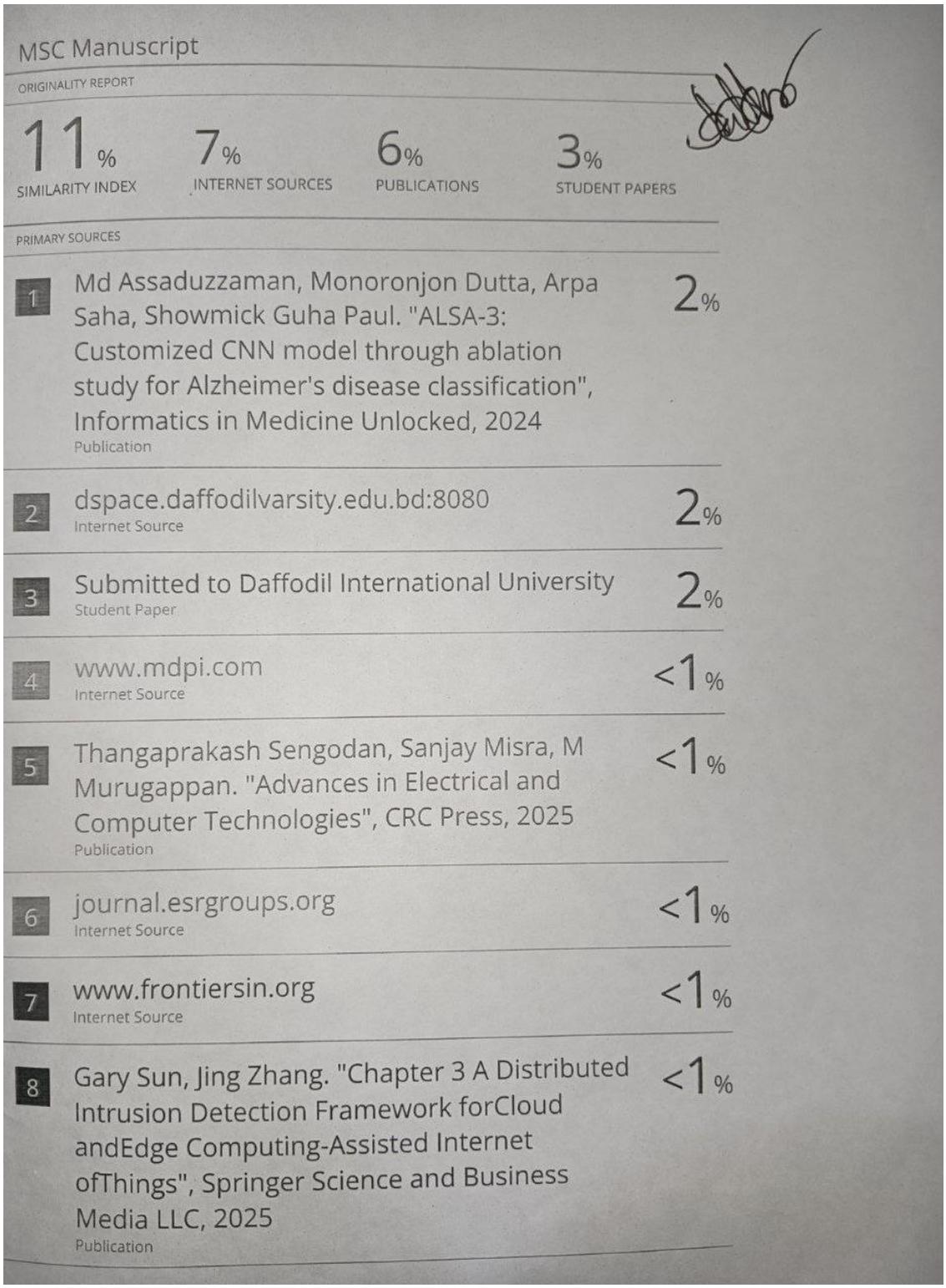
[91] A. Shukla, R. Tiwari, and S. Tiwari, "Alz-ConvNets for Classification of Alzheimer Disease Using Transfer Learning Approach," SN COMPUT. SCI., vol. 4, no. 4, p. 404, May 2023, doi: 10.1007/s42979-023-01853-7.

[92] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 3, pp. 569–575, Mar. 2010, doi: 10.1109/TPAMI.2009.187.

[93] T.-T. Wong and N.-Y. Yang, "Dependency Analysis of Accuracy Estimates in k-Fold Cross Validation," IEEE Trans. Knowl. Data Eng., vol. 29, no. 11, pp. 2417–2427, Nov. 2017, doi: 10.1109/TKDE.2017.2740926.

[94] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," in 2017 IEEE International Conference on Computer Vision (ICCV), Venice: IEEE, Oct. 2017, pp. 618–626. doi: 10.1109/ICCV.2017.74.

[95] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks," in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV: IEEE, Mar. 2018, pp. 839–847. doi: 10.1109/WACV.2018.00097.



MSC Manuscript

ORIGINALITY REPORT

11%	7%	6%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Md Assaduzzaman, Monoronjon Dutta, Arpa Saha, Showmick Guha Paul. "ALSA-3: Customized CNN model through ablation study for Alzheimer's disease classification", Informatics in Medicine Unlocked, 2024 Publication	2%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	2%
3	Submitted to Daffodil International University Student Paper	2%
4	www.mdpi.com Internet Source	<1%
5	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical and Computer Technologies", CRC Press, 2025 Publication	<1%
6	journal.esrgroups.org Internet Source	<1%
7	www.frontiersin.org Internet Source	<1%
8	Gary Sun, Jing Zhang. "Chapter 3 A Distributed Intrusion Detection Framework forCloud andEdge Computing-Assisted Internet ofThings", Springer Science and Business Media LLC, 2025 Publication	<1%