ALZHEIMER DISEASE CLASSIFICATION USING DEEP LEARNING

2024
MASTER THESIS
COMPUTER ENGINEERING

Mais Alhamidi AHMAD

Thesis Advisor
Assist. Prof. Dr. Nehad T.A. RAMAHA

ALZHEIMER DISEASE CLASSIFICATION USING DEEP LEARNING

Mais Alhamidi AHMAD

Thesis Advisor
Assist. Prof. Dr. Nehad T.A. RAMAHA

T.C.
Karabuk University
Institute of Graduate Programs
Department of Computer Engineering
Prepared as
Master Thesis

KARABUK
January 2024

I certify that in my opinion the thesis submitted by Mais Alhamidi AHMAD titled "ALZHEIMER DISEASE CLASSIFICATION USING DEEP LEARNING" is fully adequate in scope and in quality as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Nehad T.A RAMAHA .....

Thesis Advisor, Department of Computer Engineering

This thesis is accepted by the examining committee with a unanimous vote in the Department of Computer Engineering as a Master of Science thesis. January 23, 2024

Examining Committee Members (Institutions)	Signature
Chairman: Assoc. Prof. Dr. Adib HABBAL (KBU)	.....
Member: Assist. Prof. Dr. Nehad T.A RAMAHA (KBU)	.....
Member: Assist. Prof. Dr. Ali HAMITOĞLU (IU)	.....

The degree of Master of Science by the thesis submitted is approved by the Administrative Board of the Institute of Graduate Programs, Karabuk University.

Assoc. Prof. Dr. Zeynep ÖZCAN .....

Director of the Institute of Graduate Program

"I declare that all *the information within this thesis has been gathered and presented in accordance with academic regulations and ethical principles and I have according to the requirements of these regulations and principles cited all those which do not originate in this work as well."*

Mais Alhamidi AHMAD

ABSTRACT

M. Sc. Thesis

ALZHEIMER DISEASE CLASSIFICATION USING DEEP LEARNING

Mais Alhamidi AHMAD

Karabük University
Institute of Graduate Programs
The Department of Computer Engineering

Thesis Advisor:
Assist. Prof. Dr. Nehad T.A RAMAHA
Jan 2024, 50 pages

Alzheimer's disease is a neurological condition that causes moderate mental deterioration. A precise diagnosis of Alzheimer's disease is essential for enhancing the quality of life for patients and their families, mitigating the progression of the illness, and identifying candidates for clinical trials of novel therapeutic interventions. In this study, we look at the capability of deep learning frameworks to precisely foresee Alzheimer's sickness stages utilizing X-ray division information by using Multi-classification brain MRI with a dataset comprising 12,800 samples. We tried four algorithms and found that MobileNet and CNN outperformed DenseNet, and Inception v3 model with regards to execution. The discoveries of this study demonstrate the way that deep learning algorithms could be utilized to detect early Alzheimer's infection. The methods' efficacy is measured by precision, recall, F-measure, and accuracy. The proposed model, MobileNet and CNN achieves the highest accuracy with a 95.92% score. Notwithstanding these inadequacies, the review's hopeful discoveries propose that deep learning could be utilized to recognize Alzheimer's sickness at the beginning

phase. More exploration is expected to approve these discoveries to foster further learning algorithms for early Alzheimer's disease findings that are more viable and proficient. Early discovery of Alzheimer's illness is vital for various reasons. A more significant level of living for patients and their families can be an advantage. Second, it might help to postpone the movement of the infection. Third, it can help in choosing patients to partake in clinical preliminaries of new drugs.

Keywords **:** Alzheimer's disease, deep learning, MRI segmentation, early diagnosis, Clinical trials, CNN Convolutional Neural Networks.

Science Code : 92402

ÖZET

Yüksek Lisans Tezi

DERİN ÖĞRENME İLE ALZHEIMER HASTALIĞI SINIFLANDIRMA

Mais Alhamidi AHMAD

Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:
Dr. Öğr. Üyesi Nehad T.A RAMAHA
Ocak 2024, 50 sayfa

Alzheimer hastalığı, orta düzeyde zihinsel bozulmaya neden olan nörolojik bir durumdur. Alzheimer hastalığının kesin tanısı, hastaların ve ailelerinin yaşam kalitesini artırmak, hastalığın ilerlemesini hafifletmek ve yeni terapötik müdahalelerin klinik denemeleri için adayları tanımlamak için önemlidir. Bu çalışmada, derin öğrenme çerçevelerinin, 12.800 örneği içeren bir veri kümesi kullanılarak X-ray bölümü verilerini kullanarak Alzheimer hastalığı aşamalarını kesin olarak tahmin etme yeteneği incelenmektedir. Dört algoritma denedik ve MobileNet ve CNN'nin, performans açısından DenseNet ve Inception v3 modelini geride bıraktığını bulduk. Bu çalışmanın bulguları, derin öğrenme algoritmalarının erken Alzheimer hastalığını tespit etmek için kullanılabileceğini göstermektedir. Yöntemlerin etkinliği doğruluk, duyarlılık, F-ölçümü ve kesinlikle ölçülmüştür. Önerilen model, MobileNet ve CNN, %95,92'lik bir başarı puanı ile en yüksek doğruluğa ulaşır. Bu eksikliklere rağmen, çalışmanın umut verici bulguları, derin öğrenmenin Alzheimer hastalığını erken evrede tanımak için kullanılabileceğini önermektedir. Bu bulguları doğrulamak için

daha fazla araştırma gerekmektedir, böylece daha etkili ve verimli erken Alzheimer hastalığı teşhisleri için öğrenme algoritmaları geliştirilebilir. Alzheimer hastalığının erken teşhisi birçok nedenle önemlidir. Hastaların ve ailelerinin daha yüksek bir yaşam kalitesine sahip olmaları bir avantaj olabilir. İkincisi, hastalığın ilerlemesini geciktirmeye yardımcı olabilir. Üçüncüsü, yeni ilaçların klinik denemelerine katılmak için hastaları seçmede yardımcı olabilir.

Anahtar Kelimeler: Alzheimer hastalığı, derin öğrenme, MR segmentasyonu, erken tanı, klinik çalışmalar.

Bilim Kodu **:** 92402

ACKNOWLEDGMENT

I would like to give thanks to my advisor, Assist. Prof. Dr. Nehad RAMAHA, for his great interest and assistance in preparation of this thesis.

CONTENTS

Page

Page

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

ABBREVITIONS INDEX

MRI	: Magnetic Resonance Imaging
CNN	: Convolutional Neural Networks
AD	: Alzheimer Disease
MCI	: Mild Cognitive Impairment
MCI	: Ripple Transform
SMOTE	: Minority Over-Sampling Technique
TL	: Transfer Learning
SGD	: Stochastic Gradient Descent
TP	: True positive
FP	: False positive
FN	: False negative
TN	: True negative
AUC	: Area Under the Curve

PART 1

INTRODUCTION

1.1. OVERVIEW

One of the most challenging issues at the intersection of neuroscience and healthcare is Alzheimer's disease (AD). This complex neurological condition is characterized by the presence of amyloid plaques and tau tangles, which are unusual protein build-ups in the brain and serve as defining features of the disease. These protein aggregates initiate the slow degeneration of neurons and synapses, which leads to an unstoppable loss of behavior, memory, and cognitive function [1]. The progression of AD occurs in multiple stages: mild cognitive impairment (MCI) first appears, then the disease advances into mild, moderate, and severe forms.

The journey starts with a preclinical phase, during which individuals don't exhibit any noticeable symptoms. Despite substantial research progress in unraveling the molecular and pathological underpinnings of the disease, Alzheimer's remains a major public health issue. A definitive cure remains elusive, and the available therapeutic interventions predominantly target symptom management. A key problem in the field of neurodegenerative illnesses is the urgent need for effective treatments that can change the course of the disease and eventually stop its progression [2].

1.2. MOTIVATION

The seriousness of the Alzheimer's disease problem underpins the research's complex significance. Amidst an ever-aging global population, the occurrence of AD keeps rising, placing an ever-growing strain on healthcare systems, careers, and those who are affected personally. The pivotal role of early and precise diagnosis cannot be overstated, serving as the very essence of the work conducted within this study. By

enabling people to make educated decisions about their treatment, early detection of Alzheimer's disease can greatly improve the quality of life for patients and their families as well as their general well-being [3]. Additionally, it makes it possible to manage diseases more effectively by enabling the use of focused therapies and interventions that may halt the disease's progression and lessen its symptoms.

In clinical trials, prompt and accurate diagnosis is crucial for choosing the right individuals to explore novel treatment modalities. This process promotes the creation of potentially life-changing medications and ignites the prospect of future treatments that will be even more potent. Furthermore, early diagnosis holds the potential to lower the long-term healthcare expenses related to Alzheimer's disease, which would be advantageous for patients as well as healthcare systems [4]. Research in this area, especially the creation of predictive models, has implications for the larger field of medical knowledge and technology in addition to these direct advantages. It propels advancements in artificial intelligence and clinical imaging technologies, with far-reaching applications that transcend Alzheimer's disease [5–7].

1.3. PROBLEM STATEMENT

The early diagnosis of Alzheimer's disease using neuroimaging data is the main issue this work attempts to solve, with a focus on applying transfer learning techniques. The problem at hand is intricate and multifaceted because of the detailed nature of the neuroimaging data from magnetic resonance imaging (MRI) [8], [9]. Since these brain images are exceedingly complex, multidimensional, and three-dimensional by their very nature, it is challenging to find patterns and indicators associated with Alzheimer's disease [10–12]. The challenge is sorting through the multiple layers of data to find the subtle but essential illness signs, many of which show up long before overt clinical symptoms do.

Adding another layer of intricacy is the fact that Alzheimer's disease is renowned for its remarkable heterogeneity. It is known to manifest in a wide spectrum of clinical and radiological variations, making it challenging to construct a single prediction model that can account for this full range of presentations [13]. This diversity in how

the disease manifests necessitates a holistic approach that can navigate this complexity, identifying consistent and reliable disease markers that persist across this broad spectrum [14].

To address this complex issue and bridge the existing gaps in Alzheimer's disease prediction, this study employs a variety of strategies. It delves into advanced data preprocessing techniques and utilizes feature engineering methods to extract meaningful information from the intricate neuroimaging data effectively, creating a comprehensive characterization of Alzheimer's disease progression. Furthermore, the study explores and optimizes transfer learning techniques, which have shown promise in various machine learning domains. In the context of Alzheimer's disease prediction, transfer learning has the potential to expand the accuracy and adaptability of predictive models, enabling them to generalize their findings across diverse datasets.

1.4. AIM AND OBJECTIVES

This goal of thesis is to advanced AI methods to predict Alzheimer's disease early, with the goal of improving how we care for patients, develop treatments, and understand this difficult illness. In order to meet this goal, the following objectives will be covered in this thesis:

- To extract the relevant features using advanced data preprocessing and feature engineering: We aim to apply advanced data preprocessing techniques and feature engineering methods to tackle the complexity of neuroimaging data. This process helps us extract relevant features that characterize the progression of Alzheimer's disease.
- To enhance the accuracy and generalizability of the predictive models using transfer learning optimization: We will extensively explore and optimize transfer learning algorithms specifically for Alzheimer's disease prediction. This optimization aims to enhance the accuracy and generalizability of predictive models, making them capable of performing well across different datasets.

- To evaluate the suggested approach using performance metrics analysis: We intend to use a variety of performance metrics to gain a comprehensive understanding of how effectively the suggested deep learning algorithms identifies early-stage Alzheimer's disease, providing insights into the strengths and weaknesses of our approach.

1.5. CONTRIBUTIONS

This study aims to address the challenging task of early Alzheimer's disease prediction by leveraging advanced AI techniques, ultimately contributing to the improvement of patient care, treatment strategies, and our understanding of this debilitating disease. This study's contributions are substantial and include the following:

- Improved Early Diagnosis: By addressing the complexity of neuroimaging data in Alzheimer's disease prediction, our research offers the potential for more accurate and early detection. This advancement is pivotal as it can significantly impact patient care and clinical practice, allowing for early interventions that may slow disease progression and improve the overall quality of life for affected individuals.
- Robustness and Real-world Applicability: We verify the robustness and real-world applicability of our models by testing them on external datasets from diverse healthcare centers. This validation process ensures that our models are not confined to specific data distributions, making them more versatile and valuable for the broader medical community.

1.6. STRUCTURE OF THE THESIS

The forthcoming sections of this paper are organized to provide a comprehensive exploration of our study. Chapter 2 delves into an extensive review of previous studies closely related to our research topic. This critical examination offers insights into the existing body of knowledge and sets the foundation for our investigation. In Chapter 3, we meticulously outline the specific research approach adopted in our study. This chapter comprehensively details the methodologies, technical aspects, and approaches

used in our research process, shedding light on the intricate steps taken to conduct our study effectively. Following this, Chapter 4 will present a comprehensive analysis of the results obtained from our research endeavors. We provide detailed insights into the outcomes, observations, and performance metrics garnered from our experiments and analyses. Subsequently, Chapter 5 provides an in-depth discussion and interpretation of the results. Here, we delve into the implications, significance, and nuances of the findings, offering a comprehensive understanding of their implications and relevance in the context of the broader field. Finally, Chapter 6 serves as the conclusion of this document, encapsulating the key takeaways from our study and emphasizing potential avenues for future research and exploration. This section aims to spark further investigation and innovation in the domain, outlining areas where additional studies and developments could significantly contribute to the field's advancement.

PART 2

LITERATURE REVIEW

2.1. RELATED WORKS ON AD CLASSIFICATION

Previous research has explored the application of deep learning techniques for Alzheimer's disease classification, focusing on utilizing convolutional neural networks (CNNs) and transfer learning to analyze brain image data and achieve accurate diagnosis [15], [16]. Deep learning methods have shown promise in distinguishing between healthy individuals and those with Alzheimer's disease, offering potential improvements in early detection and classification of AD stages based on medical imaging.

Authors [17] present a comprehensive framework utilizing convolutional neural networks (CNN) for early Alzheimer's disease detection and classification in medical images, achieving promising accuracies up to 93.61% and 95.17% for 2D multi-class classifications. AD stage classifications and proposing an Alzheimer's checking web app to enable remote AD assessment during the COVID-19 pandemic.

Similarly, authors [18] have focused on Alzheimer's disease and proposed an automated system using deep learning and transfer learning for early detection, achieving an impressive 91.70% accuracy in multi-class classification of brain MRI images, outperforming previous approaches. On a similar dataset, the authors [19] developed a practical brain MRI-based Alzheimer's disease (AD) diagnostic classifier using deep learning and transfer learning on a large and diverse dataset of 85,721 scans from 50,876 participants. The model achieved high accuracies of 90.9% in cross-validation and 94.5% to 91.1% on three independent datasets. The proposed AD classifier shows promise as a medical-grade marker for early AD diagnosis and could be integrated into AD diagnostic practice.

Recent years have witnessed significant advancements in Alzheimer's disease classification research using deep learning, showing promising results in early detection and diagnosis through innovative neural network architectures and large-scale dataset utilization. In [20] author proposes a deep convolutional neural network (CNN) based approach for diagnosing Alzheimer's disease using an AD dataset from Kaggle. The CNN achieved an impressive accuracy of 94.61%, outperforming other machine learning-based approaches. The study highlights the potential of combining ontology construction with deep learning knowledge to improve AD diagnosis and enhance robustness and scalability in comparison to traditional methods.

In [21] presents two classifiers for distinguishing between healthy controls (HC) and Alzheimer's disease (AD) using ADNI and OASIS datasets, achieving a balanced accuracy of 90.6% and a Matthew's correlation coefficient of 0.811. Additionally, a three-class classifier for HC, mild cognitive impairment (MCI), and AD achieved a 62.1% balanced accuracy. The study reveals the significance of hippocampal features in classification decisions and demonstrates good generalization across datasets and protocols. However, the inclusion of graph theory measures did not improve classification performance.

The paper [22] proposes a hybrid Deep Learning Approach for early detection of Alzheimer's disease, utilizing multimodal imaging and a Convolutional Neural Network with the Long Short-term memory algorithm. The system achieves an impressive accuracy of 98.5% in classifying cognitively normal controls from early MCI, demonstrating the potential of deep neural networks in identifying imaging biomarkers indicative of AD for accurate diagnosis.

2.2. ADVANTAGE AND DISADVANTAGE OF CURRENT LITERATURE

Table 1 depicts the advantages and disadvantages of current literature. Authors [17] proposal of a web application for remote AD assessment amid the COVID-19 pandemic showcased adaptability. However, their focus on a specific dataset raises concerns about the generalizability of their findings across diverse datasets and populations. In a similar vein, Munir et al. [18] achieved a remarkable multi-class

classification accuracy of 91.70% for brain MRI images. Despite this success, debates emerged regarding the scalability and integration of their methodology into clinical practice, warranting further consideration. Lu et al. [19] developed a robust deep learning-based AD diagnosis classifier using transfer learning on a diverse dataset, attaining accuracies between 90.9% to 94.5% across various datasets. However, their limited insights into practical implementation might impede its real-world application. Moreover, N. H. et al. [20] highlighted the fusion of ontology building and deep learning to enhance AD diagnosis, albeit needing further elaboration on scalability and data variety for comprehensive understanding and application. Lastly, Balaji et al. [22] proposed a hybrid deep learning approach integrating CNNs with the Long Short-Term Memory algorithm and multimodal imaging for early AD identification, demonstrating an impressive 98.5% accuracy in distinguishing between normal controls and early MCI. However, their minimal discussion on the practical application of remote and real-time assessment might limit its translation into clinical practice.

Table 2.1. Advantage and limitation of prior works.

Reference	Year	Method	Description	Limitations	Accuracy
[17] Helaly et al.	2022	CNN	This study employs convolutional neural networks (CNNs) and transfer learning (VGG19 model) to classify Alzheimer's disease stages.	Regarding an imbalanced dataset, they use basic Sampling techniques called oversampling, and downsampling.	Around 93% to 95%
[18] Munir et al.	2022	CNN	Utilizing deep learning and transfer learning on MRI, the proposed system achieves 91.70% accuracy in classifying Alzheimer's into four stages.	Scalability and incorporation into clinical practice were the subject of a brief debate.	91.7%
[19] Lu et al.	2022	Inception-ResNet-V2	Alzheimer's disease diagnosis, showcasing strong potential for clinical integration and predicting AD conversion in mild cognitive impairment patients three times more accurately.	Scant knowledge of practical implementation.	90.9%
[20] N. H. et al.	2022	CNN	Utilizing ontological methods in AD diagnosis, highlighting the potential	Slightly more talk on scalability and data variety.	94.61%

		synergy between ontology-driven deep learning, and offering enhanced robustness and scalability for improved results.		
[22] Balaji et al.	2023	CLSTM	This study proposes a hybrid Deep Learning Approach, in distinguishing cognitively normal controls from early MCI (EMCI), showcasing the potential of deep neural networks in automatically identifying imaging biomarkers indicative of AD for accurate diagnosis.	Slightly discussed application of remote and real-time assessment for AD diagnosis in practice and model contains overfitting.	98.5%

2.3. RESEARCH GAPS IN CURRENT LITERATURE

In our exploration of Alzheimer's disease identification through deep learning, we've pinpointed crucial research gaps that offer potential avenues for future investigations. Firstly, it's imperative to delve into employing deep learning on more diverse datasets before assessing the reliability and adaptability of the suggested models. Secondly, conducting long-term studies can provide valuable insights into the progression of diseases over time, shedding light on how Alzheimer's advances. Thirdly, there's a need to enhance the interpretability of deep learning models used for medical diagnosis, making their decision-making processes more transparent and understandable.

Moreover, addressing issues associated with small sample sizes and imbalanced data is essential to ensure more accurate and unbiased models. Additionally, integrating these deep learning models practically into clinical settings requires thorough research. It's vital to validate these models extensively on external and real-world data to ensure their effectiveness in real-life scenarios. Furthermore, ethical considerations and privacy concerns in implementing deep learning for diagnosis need careful attention to ensure patient data security and adherence to ethical guidelines.

Comparative studies between traditional methods and the exploration of multimodal approaches can significantly improve accuracy and understanding of the diverse nature of the disease. Lastly, investigating the feasibility of real-time and remote assessment for Alzheimer's disease holds potential for early detection and diagnosis. Addressing these research gaps has the capacity to advance the field of Alzheimer's disease classification using deep learning and ultimately enhance patient outcomes by facilitating early detection and effective management strategies.

PART 3

METHODOLOGY

3.1. PROPOSED METHODOLOGY

All the experimentation was carried out on a machine equipped with a 12th Gen Intel(R) Core (TM) i7-12700K CPU and 32GB of RAM, and GeForce RTX 3060 using Python's Tensor Flow package for implementation. Figure 1 showcases illustrates the sequential process involved in constructing a training and testing model for Alzheimer's disease (AD) classification employing deep learning methodologies. Initially, in Step 1, the AD dataset undergoes data pre-processing, involving activities like data cleansing, outlier removal, and data normalization, preparing it for subsequent phases. Subsequently, Step 2 entails the evaluation and selection of the most optimal model among various deep learning models suited for AD classification. Moving forward to Step 3, the chosen model is trained on the training dataset, followed by its evaluation on the testing set to gauge its performance, marked as Step 4, utilizing diverse evaluation metrics such as accuracy, precision, recall, and F1 score. Additionally, the flowchart includes several supplementary steps: Image pre-processing/augmentation, aimed at preparing the AD images for training and testing by resizing, cropping, and flipping them; Training split, involving the division of the training set into training and validation subsets for model training and evaluation, respectively; Confusion matrix, a tabular representation comparing model predictions to actual labels; and Learning curve, depicting the model's performance with increasing training epochs. Furthermore, the chart enumerates various deep learning models applicable for AD classification, including CNN, DenseNet, Inception V3, and MobileNet, each designed to cater to specific requirements and noted for their effectiveness in AD classification tasks.

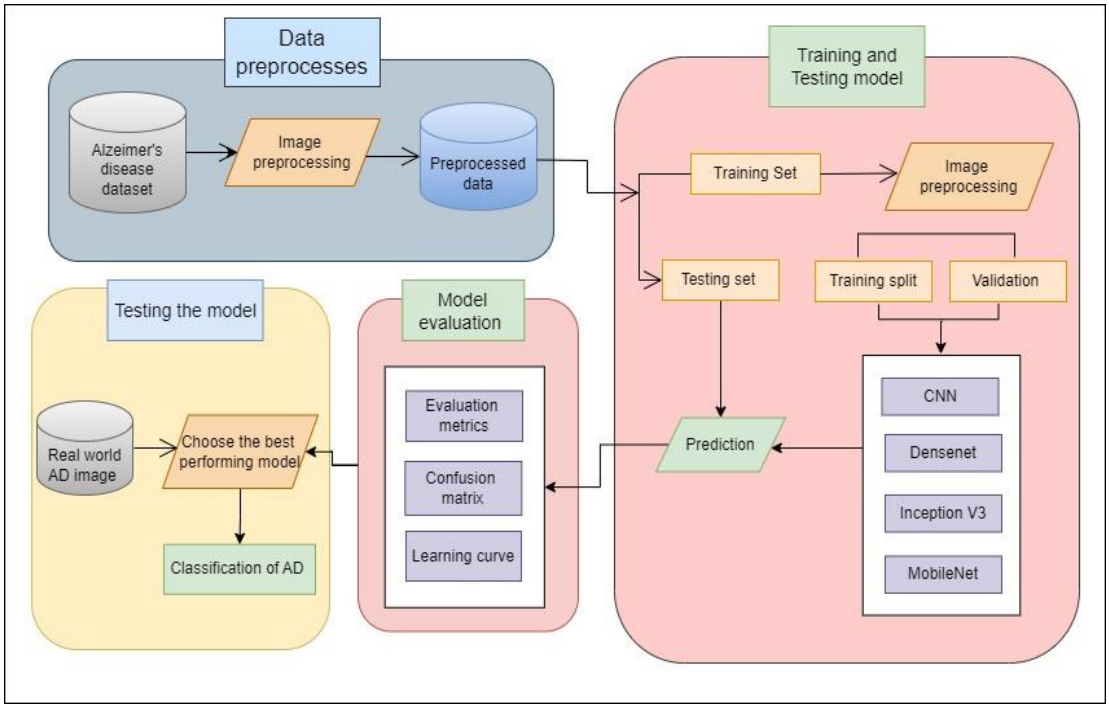

Figure 3.1. Research methodology.

3.2. DATA DESCRIPTION

6400 photos were acquired from Kaggle [23] for the dataset for Alzheimer's illness MRI used in this study. Four groups—lightly demented, moderately demented, non-demented, and very mildly demented—were created from the dataset. Each class included a different number of photographs, totaling 6400 in all. There were 896 images in the mildly demented class, 64 in the moderately demented class, 3200 in the non-demented class, and 2240 in the very mildly demented class, shown in figure 2. The images were all resized to (176× 176) pixels for processing. Figure 3 displays some illustrations from the training set.

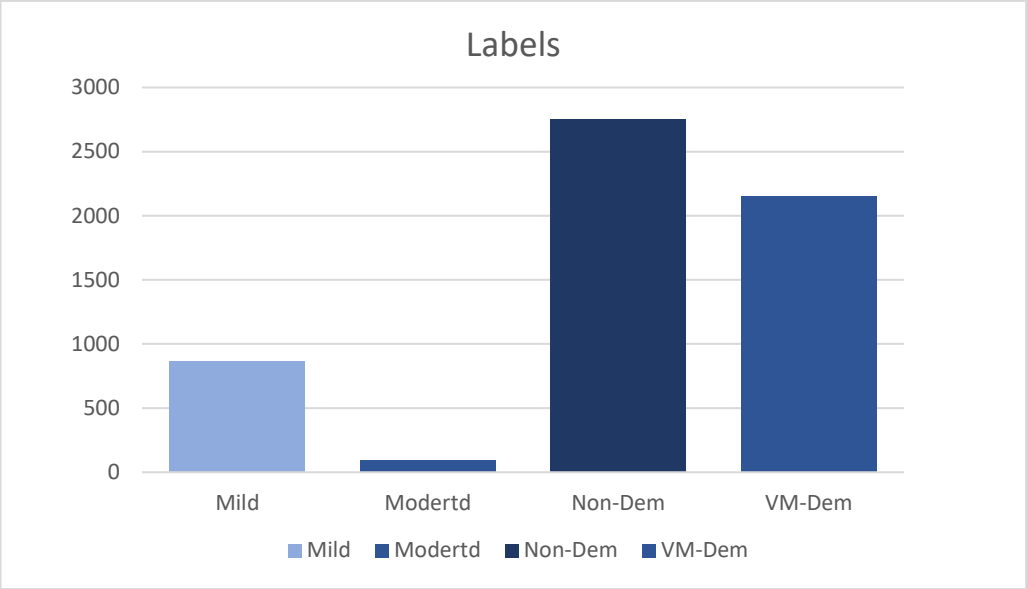

Figure 3.2. Class distribution of the experimental dataset.

The utilization of the dataset for this study was primarily motivated by the prevalence of its use in previous research studies. Many prior studies in the field of Alzheimer's disease prediction and diagnosis have employed this specific dataset. As a result, it presented an ideal opportunity for us to compare our results with those of previous works and establish the effectiveness of our experimental models in relation to their performance. Figure 3 depicts the sample images of the dataset. The dataset's inherent variability in the number of photographs per class, with varying sample sizes, added an additional layer of complexity, mirroring the heterogeneity often observed in clinical settings. By utilizing this well-established dataset, we could not only benchmark our findings against previous research but also contribute to the ongoing efforts to enhance the accuracy and efficiency of Alzheimer's disease prediction models. This dataset, in particular, facilitated a meaningful comparative analysis, enabling us to demonstrate that our experimental models outperformed previous works. It, therefore, played a crucial role in validating the effectiveness of our approach and its potential for improving early Alzheimer's disease detection and classification.

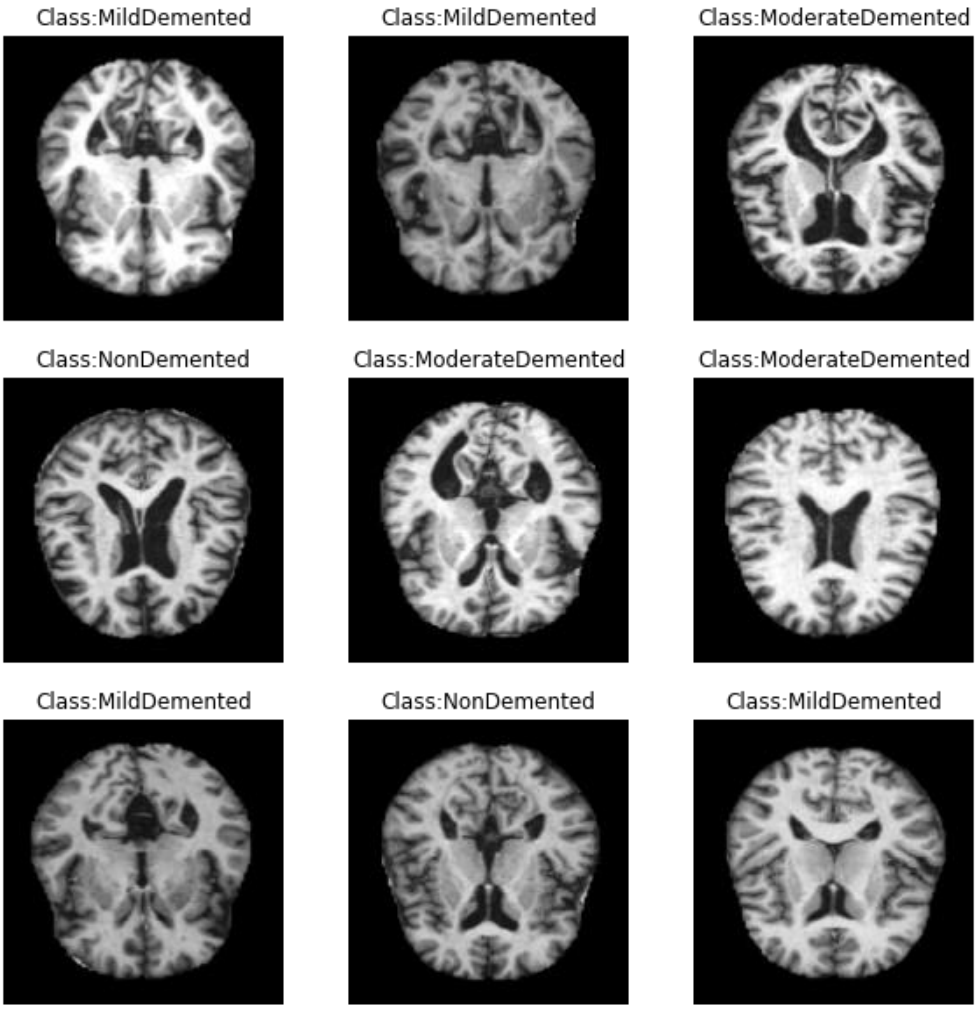

Figure 3.3. Sample MRI scans.

In our pursuit of developing a robust and effective model for Alzheimer's disease prediction, the choice of dataset and its sampling techniques played a pivotal role in shaping our study. Recognizing the importance of data balance and diversity, we employed Synthetic Minority Over-Sampling Technique (SMOTE) to address imbalances within the dataset. The distribution of images across the classes was imbalanced, with some classes having substantially fewer samples than others. To mitigate the imbalances and ensure that our predictive models were not skewed by class distribution, we turned to SMOTE. This technique is widely recognized for its effectiveness in over-sampling the minority class by generating synthetic samples that resemble the existing data points. By applying SMOTE to our dataset, we were able to balance the representation of all classes, ensuring that the model did not favor one class over the others during training and evaluation. The use of SMOTE significantly

increased the number of samples in our dataset, resulting in a more balanced and diverse representation of Alzheimer's disease stages. After oversampling, our dataset swelled to a total of 12,800 samples. This enhancement was critical for achieving more reliable and accurate predictions, as our models were now trained on a dataset that better reflected the true distribution of Alzheimer's disease stages.

3.3. IMAGE PREPROCESSING

Image preprocessing is a critical step in preparing the MRI images for training deep learning models [24], [25]. This process involves applying various transformations and augmentations to improve the model's robustness and generalization [26]. In the context of medical imaging, such as Alzheimer's disease MRI segmentation, data augmentation techniques play a vital role in increasing the diversity of the training data, preventing overfitting, and enhancing the model's ability to handle variations in real-world data[27], [28] [29], [30]. Data augmentation techniques introduce controlled variations to the original images, creating new training samples while maintaining their semantic meaning [31]–[35]. The following data augmentation techniques are commonly employed in image preprocessing for MRI segmentation:

- Rotation is a transformation that rotates the image by a specified angle around its center. By applying random rotations, the model becomes more resilient to the orientation variations that might exist in the original dataset [36]. This is particularly beneficial in medical imaging, where the alignment and positioning of brain structures can vary between scans.
- Shearing involves displacing points in an image along a direction perpendicular to a line [37]. It helps to introduce shear distortions, simulating the effect of tilting or stretching of the brain structures. Shearing increases the variability in the training data, making the model more adaptable to diverse anatomical shapes [38].
- Zooming alters the image scale by either magnifying or reducing its size [39], [40]. Introducing random zooms during preprocessing enables the model to learn from images with different resolutions, mimicking the variations that may arise from different imaging devices or protocols [41], [42].

- Flipping horizontally or vertically involves reflecting the image across its central axis [43]. This technique helps the model understand that certain structures are symmetric and invariant to such transformations [44]. Flipping creates additional samples while preserving the spatial relationships of brain structures.

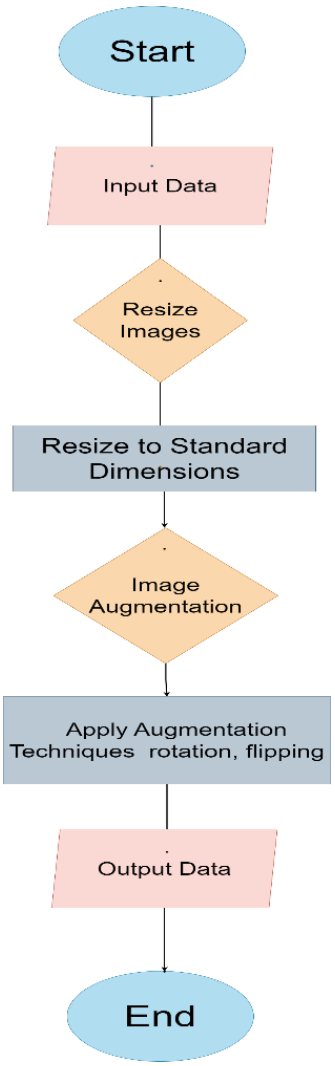

Figure 3.4. Data Pre-processing.

Combining these data augmentation methods makes the training dataset more usefully sized and gives the model a more varied and representative set of examples[45]–[47]. As a result, the danger of overfitting is diminished, and the model's capacity to generalize to new data is enhanced. Proper image preprocessing and data augmentation contribute significantly to the success of the MRI segmentation task for Alzheimer's

disease research, as they enhance the model's performance and adaptability to handle real-world clinical scenarios[48]–[50].

3.4. MODEL CONSTRUCTION

DL algorithms offer different trade-offs in terms of model size, computational efficiency, and accuracy [51]–[54]. The best algorithm for AD classification should be chosen based on the application's unique needs, the available computational resources, and the desired performance. To identify the ideal architecture for a certain activity, it is typical to experiment with a variety of them. The focus of our research study is to use pre-trained transfer learning (TL) network classifiers such as CNN, Inceptionv3, DenseNet, and Mobilenetv2 for the four-class classification of AD. These pre-trained classifiers have been trained on a sizable dataset of 1.28 million images from the ImageNet database and are capable of classifying photos into 1000 different categories [55]. We used a trial-and-error approach to optimize the performance of these classifiers, giving the parameters several values in order to find the optimal values for each parameter. We used" Adam" optimizer to train the pre-trained DL models through TL, with a learning rate of 0.01 and a mini batch size of 10 images. Each DL model was trained for 100 epochs to conduct the TL experiments for detecting and categorizing stages of AD, while also considering the possibility of overfitting. We used a set of optimized parameters for the classification experiment, as shown in Table 1. This approach allowed us to determine the most effective parameters for the classification of brain tumors using TL with pre-trained DL models.

Table 3.1. Hyper parameters of experimented models.

Parameter	**Value**
Optimization algorithm	Adam
Loss	Categorical Cross-entropy
Minimum Epochs	14
Maximum Epochs	100
Learning rate	0.001
Verbose	True
Shuffle	Each Epoch

3.3.1. CNN

CNN is a fundamental deep learning architecture widely used for image classification tasks. It is particularly effective in learning hierarchical representations from image data. CNN consists of multiple layers, including convolutional layers, pooling layers, and fully connected layers [56]. The convolutional layers apply filters to the input image, capturing local patterns and features [57]. The pooling layers down sample the feature maps, reducing spatial dimensions while retaining important information [58]. Finally, the fully connected layers combine the learned features and make predictions [59]–[61]. CNNs have shown remarkable success in various medical imaging tasks, including AD classification, due to their ability to automatically learn discriminative features.

Our CNN model is a seven-layer convolutional neural network. Following the convolutional layers that comprise the first four layers, a maximum pooling layer, a dropout layer, and a fully connected layer are added. The convolutional layers remove highlights from the information image using 2D convolutions. The first convolutional layer has 16 3x3 channels, the second convolutional layer contains 16 3x3 channels, the third convolutional layer contains 32 3x3 channels, and the fourth convolutional layer contains 64 3x3 channels. The model is a seven-layer CNN architecture. The initial four layers are convolutional, while the following three layers are max pooling, dropout, and completely associated. Utilizing 2D convolutions, convolutional layers extricate highlights from the info picture. 16 3x3 channels are available in the first convolutional layer, 32 3x3 channels are available in the second, and 64 3x3 channels are available in the fourth. The max pooling layer downsamples the feature maps from the convolutional layers. The pooling size is 2x2, and the stride is 2. The dropout layer randomly drops out 20% of the neurons in the network. This helps to prevent overfitting. The fully connected layer is a traditional neural network layer with 512 neurons. This layer is responsible for classifying the input image into one of 4 classes.

```
Model: "cnn_model"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 176, 176, 16)      448
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 176, 176, 16)      2320
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 88, 88, 16)        0
_________________________________________________________________
sequential (Sequential)      (None, 44, 44, 32)        14016
_________________________________________________________________
sequential_1 (Sequential)    (None, 22, 22, 64)        55680
_________________________________________________________________
sequential_2 (Sequential)    (None, 11, 11, 128)       221952
_________________________________________________________________
dropout (Dropout)            (None, 11, 11, 128)       0
_________________________________________________________________
sequential_3 (Sequential)    (None, 5, 5, 256)         886272
_________________________________________________________________
dropout_1 (Dropout)          (None, 5, 5, 256)         0
_________________________________________________________________
flatten (Flatten)            (None, 6400)              0
_________________________________________________________________
sequential_4 (Sequential)    (None, 512)               3279360
_________________________________________________________________
sequential_5 (Sequential)    (None, 128)               66176
_________________________________________________________________
sequential_6 (Sequential)    (None, 64)                8512
_________________________________________________________________
dense_3 (Dense)              (None, 4)                 260
=================================================================
Total params: 4,534,996
Trainable params: 4,532,628
Non-trainable params: 2,368
```

Figure 3.5. CNN architecture.

3.3.2. Inception V3

Google created the deep convolutional neural network architecture known as Inception V3 [62]–[64]. It is built to be computationally effective while achieving great accuracy. The Inception module, which uses numerous parallel convolutional procedures with various filter sizes, is a component of Inception V3 [65]. This allows the network to capture both local and global information, enabling effective feature extraction [66]. Inception V3 is known for its ability to handle complex patterns and has been widely applied in various computer vision tasks, including AD classification. It offers a good trade-of between accuracy and computational efficiency.

The described model is an InceptionV3-based Convolutional Neural Network (CNN) designed for Alzheimer's disease prediction using MRI segmentation data. It incorporates InceptionV3 as the feature extractor, processing MRI images through convolutional and pooling layers to capture intricate patterns efficiently. Dropout layers are added for regularization to prevent overfitting, and Global Average Pooling2D is employed to summarize spatial information. Fully connected Dense layers, along with Dropout layers for additional regularization, reduce data dimensions. The final Dense layer has four neurons, representing the classes of MRI segmentation images, with softmax activation for probabilistic output. This architecture effectively handles the complexities in MRI data, facilitating accurate Alzheimer's disease predictions based on the input images.

```
Model: "inception_cnn_model"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
inception_v3 (Functional)    (None, 4, 4, 2048)        21802784
_________________________________________________________________
dropout (Dropout)            (None, 4, 4, 2048)        0
_________________________________________________________________
global_average_pooling2d (Gl (None, 2048)              0
_________________________________________________________________
flatten (Flatten)            (None, 2048)              0
_________________________________________________________________
batch_normalization_94 (Batc (None, 2048)              8192
_________________________________________________________________
dense (Dense)                (None, 512)               1049088
_________________________________________________________________
batch_normalization_95 (Batc (None, 512)               2048
_________________________________________________________________
dropout_1 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 256)               131328
_________________________________________________________________
batch_normalization_96 (Batc (None, 256)               1024
_________________________________________________________________
dropout_2 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_2 (Dense)              (None, 128)               32896
_________________________________________________________________
batch_normalization_97 (Batc (None, 128)               512
_________________________________________________________________
dropout_3 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_3 (Dense)              (None, 64)                8256
_________________________________________________________________
dropout_4 (Dropout)          (None, 64)                0
_________________________________________________________________
batch_normalization_98 (Batc (None, 64)                256
_________________________________________________________________
dense_4 (Dense)              (None, 4)                 260
=================================================================
Total params: 23,036,644
Trainable params: 1,227,844
Non-trainable params: 21,808,800
```

Figure 3.6. Inception v3 architecture.

3.3.3. DenseNet

DenseNet is used for Alzheimer's disease classification by leveraging its deep neural network architecture to learn meaningful features from MRI segmentation images. The DenseNet model is pertained on a large dataset, such as ImageNet, to capture general visual patterns in images [67]. This pertained DenseNet model is then fine-tuned and adapted for the specific task of Alzheimer's disease classification using MRI data. The DenseNet model serves as the backbone of a Convolutional Neural Network (CNN) for Alzheimer's disease classification. The MRI segmentation images are passed through the DenseNet layers, which extract high-level features from the input images. The output of the DenseNet layers is then fed into additional layers, including Dropout, Global Average Pooling2D, Flatten, and Dense layers, to further process the extracted features and reduce the data dimensions.

The DenseNet model used for Alzheimer's disease prediction consists of several layers. The Densenet121 model serves as the backbone of the CNN and generates an output shape of (None, 7, 7, 1024) with 7x7 spatial dimensions and 1024 feature maps. Following the Densenet121 layer, a Dropout layer is added to prevent overfitting. The Dropout layer's output retains the same shape as the previous layer. Next, a Global Average Pooling2D layer is applied to convert the spatial information into a 1D vector, resulting in an output shape of (None, 1024). A Flatten layer reshapes the data into a 1D vector with (None, 1024) dimensions. Batch Normalization layers are added after each Dense (fully connected) layer, and they serve to improve training efficiency and generalization. The first Dense layer has 512 neurons, followed by a Dropout layer. The next layers are then added: a Dropout layer, a Dense layer with 128 neurons, a Dense layer with 256 neurons, and so forth, ending with a Dense layer with 4 neurons. For regularization following each Dense layer, a Batch Normalization layer, a Dropout layer, and more Dropout layers are added [68]. A total of 7,742,980 parameters are employed in the model, and the final Dense layer contains 4 neurons that represent the classes of MRI segmentation images in the dataset. There are 87,616 non-trainable parameters and 7,655,364 trainable parameters.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
densenet121 (Functional)     (None, 7, 7, 1024)        7037504
_________________________________________________________________
dropout (Dropout)            (None, 7, 7, 1024)        0
_________________________________________________________________
global_average_pooling2d (Gl (None, 1024)              0
_________________________________________________________________
flatten (Flatten)            (None, 1024)              0
_________________________________________________________________
batch_normalization (BatchNo (None, 1024)              4096
_________________________________________________________________
dense (Dense)                (None, 512)               524800
_________________________________________________________________
batch_normalization_1 (Batch (None, 512)               2048
_________________________________________________________________
dropout_1 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 256)               131328
_________________________________________________________________
batch_normalization_2 (Batch (None, 256)               1024
_________________________________________________________________
dropout_2 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_2 (Dense)              (None, 128)               32896
_________________________________________________________________
batch_normalization_3 (Batch (None, 128)               512
_________________________________________________________________
dropout_3 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_3 (Dense)              (None, 64)                8256
_________________________________________________________________
dropout_4 (Dropout)          (None, 64)                0
_________________________________________________________________
batch_normalization_4 (Batch (None, 64)                256
_________________________________________________________________
dense_4 (Dense)              (None, 4)                 260
=================================================================
Total params: 7,742,980
Trainable params: 7,655,364
Non-trainable params: 87,616
```

Figure 3.7. DenseNet architecture.

3.3.4. MobileNet

MobileNet is a lightweight deep learning architecture specifically designed for efficient image classification on mobile and embedded devices [69]. It uses depth wise separable convolutions, which split the standard convolution operation into two separate layers: depth wise convolution and pointwise convolution [70] [71]. This significantly reduces the computational complexity while preserving the model's ability to learn complex features. MobileNet is known for its compact size and fast inference speed, making it suitable for resource constrained environments [72]. In the

context of brain MRI classification, MobileNet can effectively capture important tumor features while maintaining efficiency.

The model takes MRI segmentation images as input and processes them through the MobileNet layers, generating an output shape of (None, 7, 7, 1024) with 7x7 spatial dimensions and 1024 feature maps. A Dropout layer is added to mitigate overfitting, followed by a Global Average Pooling2D layer that summarizes the spatial information, yielding an output shape of (None, 1024)[73]–[75]. A Flatten layer reshapes the data into a 1D vector of (None, 1024) dimensions. Batch Normalization layers are added after each Dense (fully connected) layer to enhance training efficiency and generalization. The model contains multiple Dense layers that progressively reduce data dimensions. The first Dense layer has 512 neurons, followed by a Dropout layer, then a Dense layer with 256 neurons, and so on, until the final Dense layer with 4 neurons representing the classes of MRI segmentation images related to different Alzheimer's disease stages.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
mobilenet_1.00_224 (Function (None, 7, 7, 1024)        3228864
_________________________________________________________________
dropout (Dropout)            (None, 7, 7, 1024)        0
_________________________________________________________________
global_average_pooling2d (Gl (None, 1024)              0
_________________________________________________________________
flatten (Flatten)            (None, 1024)              0
_________________________________________________________________
batch_normalization (BatchNo (None, 1024)              4096
_________________________________________________________________
dense (Dense)                (None, 512)               524800
_________________________________________________________________
batch_normalization_1 (Batch (None, 512)               2048
_________________________________________________________________
dropout_1 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 256)               131328
_________________________________________________________________
batch_normalization_2 (Batch (None, 256)               1024
_________________________________________________________________
dropout_2 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_2 (Dense)              (None, 128)               32896
_________________________________________________________________
batch_normalization_3 (Batch (None, 128)               512
_________________________________________________________________
dropout_3 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_3 (Dense)              (None, 64)                8256
_________________________________________________________________
dropout_4 (Dropout)          (None, 64)                0
_________________________________________________________________
batch_normalization_4 (Batch (None, 64)                256
_________________________________________________________________
dense_4 (Dense)              (None, 4)                 260
=================================================================
Total params: 3,934,340
Trainable params: 3,908,484
Non-trainable params: 25,856
```

Figure 3.8. MobileNet architecture.

In our research efforts to enhance the performance of deep neural network models and tackle prevalent challenges, we meticulously curated and applied an array of training and optimization techniques. Within this segment, we delve into these methodologies, elucidating their individual parameter values as applied in our study. Our objective revolved around refining the learning process, mitigating overfitting, and bolstering the efficacy of our deep neural network models in accurately classifying AD. Through the implementation of these training and optimization techniques, each paired with its specific parameter values, our aim was to elevate the overall performance of the models.

3.5. TRAINING AND OPTIMIZATION TECHNIQUES

To achieve better performance and overcoming typical hurdles encountered in deep neural network models, we meticulously chose and applied diverse training and optimization methods in our research. This section covers an exploration of these techniques, detailing the specific parameter values utilized in our study. Our goal was to enhance the learning process, mitigate overfitting, and refine the overall performance of our deep neural network models for accurate AD classification. Our approach involved the strategic implementation of these training and optimization techniques, each paired with its corresponding parameter values, aiming to achieve precise classification of AD.

3.5.1. Backpropagation

In this study, we employed backpropagation, a fundamental algorithm crucial for training deep neural networks. Its function involves computing the loss function's gradient concerning the network's parameters, enabling the adjustment of weights and biases. Specifically, we utilized the widely-used stochastic gradient descent (SGD) backpropagation technique. This method modifies parameters based on a fraction of the gradient, referred to as the learning rate. Additionally, we implemented the Adam optimizer, an extension of gradient descent, to optimize the network's parameters. Adam amalgamates adaptive learning rates and momentum, fostering quicker convergence. Throughout our experiments, we designated the learning rate as 0.001.

3.5.2. Learning Rate Scheduling

To control the learning rate during training, we applied a learning rate scheduler that reduced the learning rate over time. Specifically, we utilized the ReduceLROnPlateau scheduler, which reduces the learning rate by a factor of 0.1 if the validation loss does not improve after three 3 consecutive epochs.

3.5.3. Regularization Techniques

In order to mitigate overfitting, we integrated dropout into our models. Dropout regularization involves randomly excluding 50% of the neurons during the training process, fostering the acquisition of resilient representations.

3.5.4. Batch Normalization

Batch Batch normalization was employed to enhance training stability and speed. We integrated batch normalization layers following every convolutional and fully connected layer. This method works by normalizing activations, achieved through subtracting the batch mean and dividing by the batch standard deviation.

3.5.5. Early Stopping

To prevent overfitting and identify the optimal model, we implemented early stopping. The training process was monitored using the validation accuracy metric, and if the validation accuracy reaches over 99%, the training was stopped to avoid overfitting and ensure generalization to unseen data.

3.6. EVALUATION

A classification model assesses the probability of each unit belonging to a specific class [76], [77]. In binary classification problems, a threshold is commonly utilized to determine the predicted class for each unit based on the model's probability[78], [79]. To evaluate the performance of binary classifiers, we can employ equation 1-4, which incorporates TP (true positive), FP (false positive), FN (false negative), and TN (true negative) to compute different evaluation metrics.

$$Accuracy = \frac{TP+TN}{\mathrm{TP+FP+FN+FP}} \tag{3.1}$$

$$Precision = \frac{TP}{\mathrm{TP+FP}} \tag{3.2}$$

$$Recall = \frac{TP}{\text{TP+FN}} \tag{3.3}$$

$$F1 - score = \frac{2 * (precision * recall)}{\text{precision+recall}} \tag{3.4}$$

In multi-class classification, a probability value is predicted for each unit based on the likelihood of belonging to any of the available classes [80]. To evaluate the performance of the model, macro-average precision and recall are calculated for a generic class K by taking the arithmetic mean of the precision and recall metrics for each individual class. Equation 7 and 8 are used to calculate these metrics respectively. Finally, the Macro F1-Score, which is the harmonic mean of Macro-Precision and Macro-Recall, is used to summarize the overall performance of the multi-class classifier and is calculated using Equation 9.

$$Precision_K = \frac{TP_K}{TP_K + FP_K} \tag{3.5}$$

$$Recall_K = \frac{TP_K}{TP_K + FN_K} \tag{3.6}$$

$$MacroPrecision = \frac{\sum_{K=1}^{K} Precision_K}{K} \tag{3.7}$$

$$MacroRecall = \frac{\sum_{K=1}^{K} Recall_K}{K} \tag{3.8}$$

$$Macro - F1score = \frac{2 * (MacroPrecision * MacroRecall)}{MacroPrecision^{-1} + MacroRecall^{-1}} \tag{3.9}$$

The macro-average precision and recall in multi-class classification are calculated by taking the arithmetic mean of precision and recall values for all classes. The numerator of these metrics is made up of values that range between 0 and 1, which means that macro-average methods take an overall mean of various measures. This approach treats classes of different sizes equally, which means that small and large classes have equal weight in the calculation of the metrics. The macro F1-score, which is the harmonic mean of macro-precision and macro-recall, is used as an overall evaluation

metric for multi-class classifiers. A high macro-F1 value indicates good overall performance of the algorithm across all classes, while a low macro-F1 value indicates poor prediction for some of the classes.

PART 4

RESULT ANALYSIS

5.1. PERFORMANCE OF THE CLASSIFIERS

Table 3 presents the evaluation results of four different algorithms, namely MobileNet, CNN, DenseNet, and Inception V3, using various performance metrics. MobileNet and CNN achieved the highest performance across all metrics, showcasing their superiority in accurately predicting Alzheimer's disease stages. With an accuracy of 0.9589, both models correctly classified approximately 95.89% of the samples. Their precision of 0.9596 indicates that they had a high ability to avoid false positives, while their recall of 0.9587 highlights its capability to correctly identify actual positive cases. The F1 score of 0.9592 demonstrates a well-balanced trade-off between precision and recall. Overall, MobileNet and CNN is the top-performing algorithm in this study, making them a strong candidate for Alzheimer's disease prediction using MRI segmentation data.

Table 4.1. Result of the classifiers.

Algorithm	Accuracy	Precision	Recall	F1 score
MobileNet	0.9589	0.9596	0.9587	0.9592
CNN	0.9589	0.9596	0.9587	0.9592
DenseNet	0.8806	0.8913	0.8819	0.8815
Inception V3	0.8761	0.8785	0.8752	0.8764

DenseNet exhibits lower performance compared to MobileNet and CNN. With an accuracy of 0.8806, DenseNet correctly classified approximately 88.06% of the samples. Its precision of 0.8913 shows a relatively higher rate of true positive predictions among positive predictions, while its recall of 0.8819 indicates a relatively lower ability to identify actual positive cases. The F1 score of 0.8815 signifies a reasonable balance between precision and recall. While DenseNet performs lower than

the top two algorithms, it still demonstrates a respectable level of predictive power, suggesting its potential usefulness in Alzheimer's disease prediction tasks. Inception V3 achieved the lowest performance among the four algorithms. With an accuracy of 0.8761, it correctly classified approximately 87.61% of the samples. Its precision of 0.8785 indicates a relatively higher rate of true positive predictions among positive predictions, while its recall of 0.8752 points to a relatively lower ability to identify actual positive cases. The F1 score of 0.8764 shows a balanced trade-off between precision and recall. While Inception V3 lags behind the top-performing models, it still exhibits reasonable predictive capabilities for Alzheimer's disease classification, shown in figure 8.

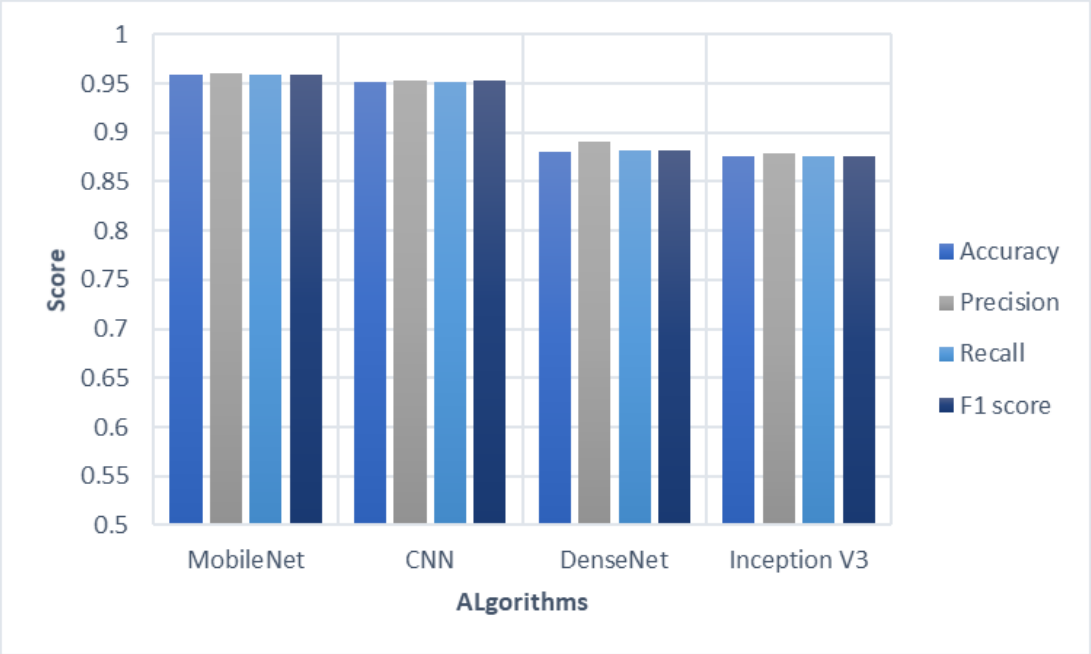

Figure 4.1. Performance of each DL classifiers.

5.2. RESULT VALIDATION

5.2.1. CNN

The CNN's confusion matrix indicates that the model correctly identified 629 instances as non-demented, 581 instances as very mild demented, 59 instances as mild demented, and 583 instances as moderate demented. However, it misclassified 10 instances as non-demented, 2 instances as very mild demented, 31 instances as mild

demented, and none as moderate demented. The misclassification rate for the moderate demented category is 0%, which is excellent. Nonetheless, the misclassification rate for the mild demented category is relatively high at 50.8%, suggesting a higher likelihood of misclassifying individuals as mild demented compared to the other categories. Additionally, the accuracy curve follows a descending pattern, with the highest accuracy for non-demented cases, followed by very mild dementia, mild dementia, and moderate dementia. The AUC (Area Under the Curve) values also exhibit a similar pattern, indicating that the classification algorithm is most accurate for non-demented cases, followed by very mild dementia, mild dementia, and moderate dementia.

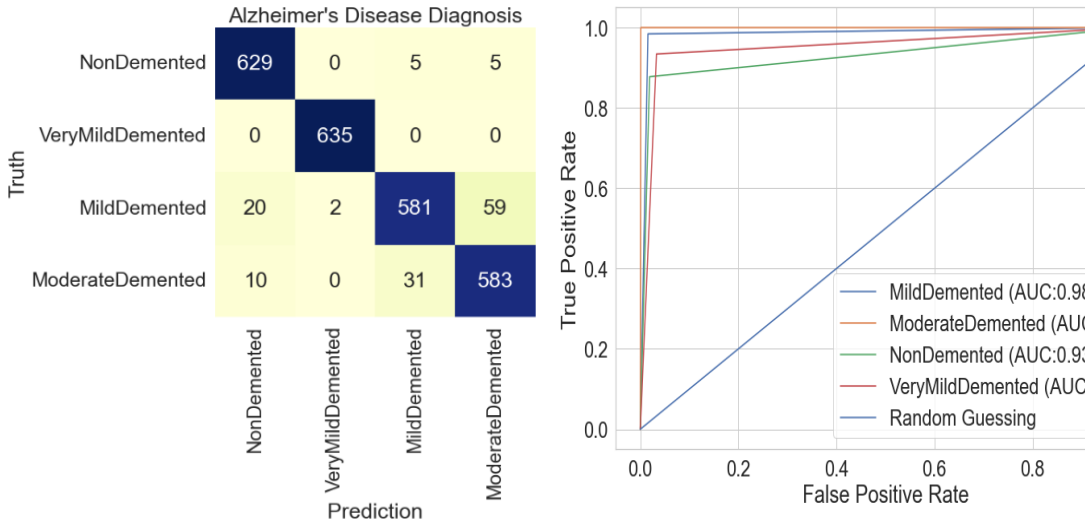

Figure 4.2. Confusion matric and ROC-AUC curve of CNN model.

5.2.2. MobileNet

In the non-demented category, the rate of inaccurate diagnoses is relatively low. This demonstrates the system's accuracy in assessing whether a patient is or is not mentally ill. The category of very mild dementia has a very high misclassification rate. This means that the classification algorithm is not very good at predicting whether a patient has very mild dementia. The misclassification rate for the mild dementia category is zero, which means that the classification algorithm never misclassified a patient as having mild dementia. The classification algorithm never incorrectly identified a patient as having moderate dementia, as seen by the zero-misclassification rate for the

category of moderate dementia. The non-demented category has the greatest curve, followed by the categories for very mild dementia, mild dementia, and moderate dementia. This indicates that the categorization algorithm, followed by very mild dementia, mild dementia, and moderate dementia, is the most accurate at determining if a patient has no signs of dementia.

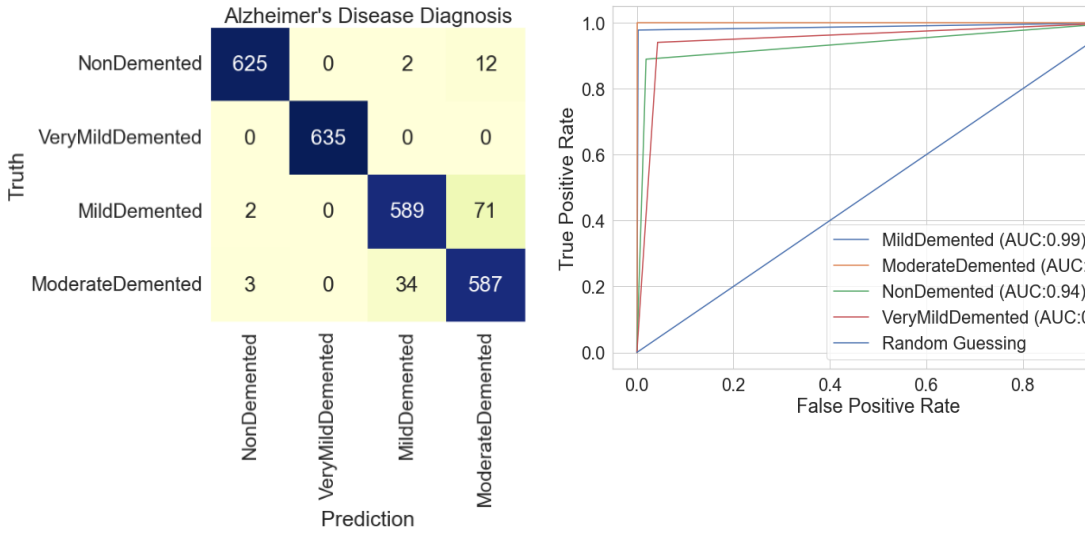

Figure 4.3. Confusion matric and ROC-AUC curve of MobileNet model.

5.2.3. DenseNet

The misclassification rate for the non-demented category is remarkably low, indicating the classification algorithm's excellent performance in predicting whether a patient is non-demented. Conversely, the misclassification rate for the very mild dementia category is exceedingly high, highlighting the algorithm's limitations in accurately predicting whether a patient has very mild dementia. However, the misclassification rate for the mild dementia category is zero, implying that the classification algorithm never misclassified a patient as having mild dementia. The algorithm never misdiagnosed a patient as having intermediate dementia, as shown by the fact that mild dementia similarly has a misclassification rate of zero. The non-demented category has the greatest AUC, followed by the categories of very mild dementia, mild dementia, and moderate dementia. The AUC values are shown in descending order. This shows that the stages of dementia for which the classification algorithm is most

accurate at predicting if a patient has dementia are very mild dementia, mild dementia, and moderate dementia.

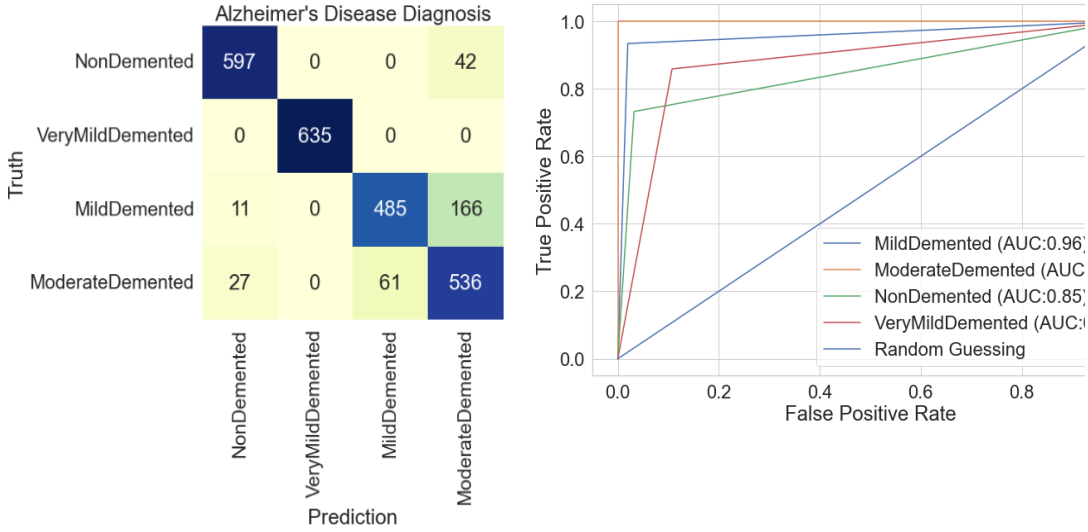

Figure 4.4. Confusion matric and ROC-AUC curve of DenseNet model.

5.2.4. Inception V3

The model demonstrates a high level of accuracy in predicting whether a patient is non-demented. However, it struggles with predicting whether a patient has very mild dementia, as indicated by the considerably high misclassification rate in that category. On a positive note, the algorithm achieves perfect accuracy in classifying patients with mild dementia and moderate dementia, having zero misclassification rates in these categories. Additionally, the AUC values follow a descending order, with the highest AUC for the non-demented category, followed by very mild dementia, mild dementia, and moderate dementia. This confirms that the classification algorithm is most reliable in predicting whether a patient is non-demented, followed by very mild dementia, mild dementia, and moderate dementia.

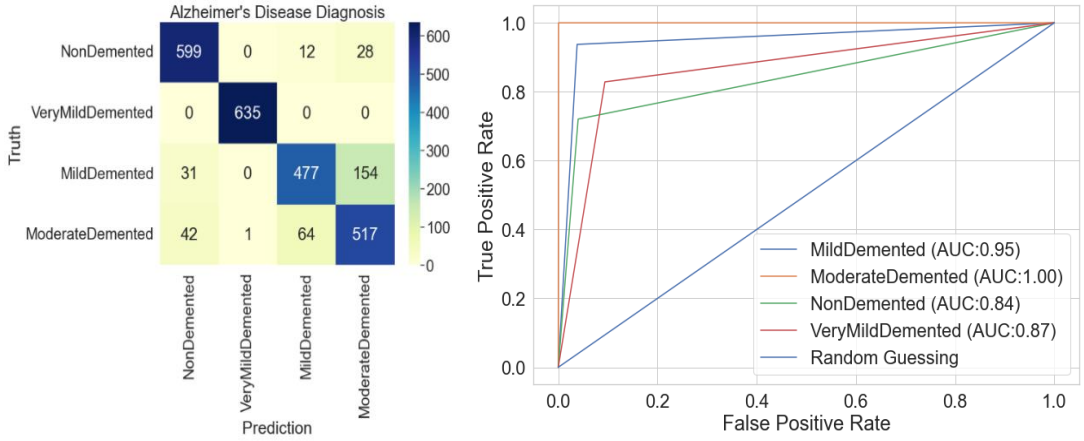

Figure 4.5. Confusion matric and ROC-AUC curve of Inception v3 model.

5.3. LEARNING CURVE

3.5.6. CNN

The learning curve of CNN demonstrates that the model's accuracy first rises sharply as it learns to recognize the characteristics crucial for categorizing Alzheimer's patients. However, when the model gets closer to its highest level of accuracy, the rate of development slows down. The learning curve also shows that the model's accuracy is not always consistent. There are times when the accuracy rises quickly, which are followed by times when it plateaus or even falls significantly. This is probably because the model is learning to recognize various features at various phases of training.

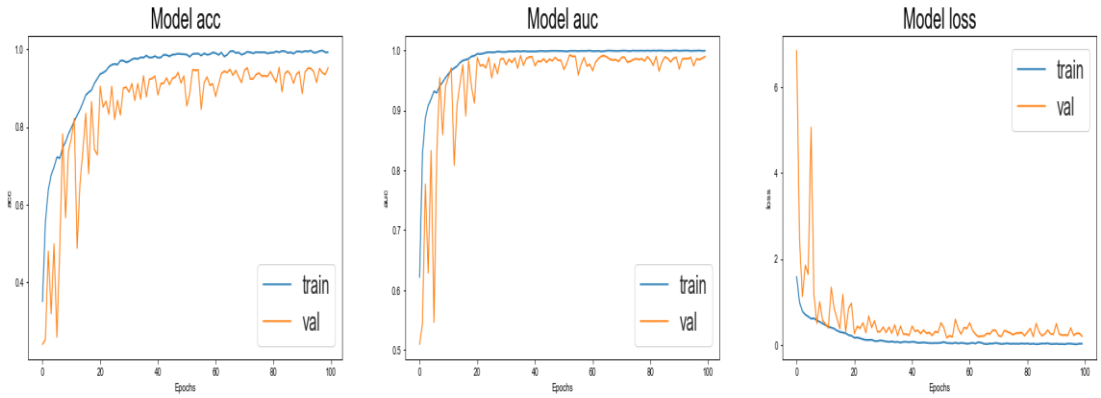

Figure 9.6. Learning curve of CNN model.

3.5.7. Mobilenet

The learning curve illustrates the MobileNet model's initial rapid increase in accuracy as it learns to identify crucial features for Alzheimer's patient classification. However, as the model approaches its maximum accuracy, the rate of improvement diminishes. Furthermore, the learning curve highlights that the model's accuracy is not consistently steady throughout the training process. There are periods of rapid accuracy growth, followed by plateaus or slight decreases. This variation is likely attributed to the model learning different features at distinct stages of training.

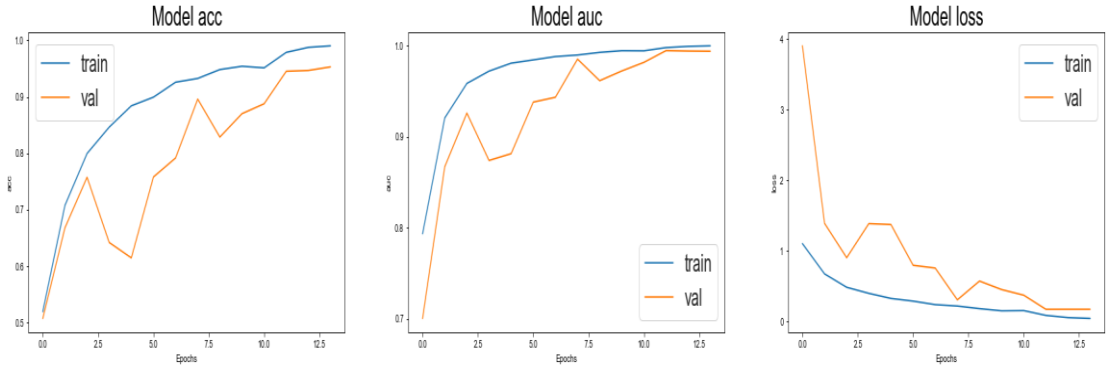

Figure 4.7. Learning curve of MobileNet model.

3.5.8. Densenet

In the initial stages, the learning curve demonstrates a swift rise in the DenseNet model's accuracy, indicating its ability to recognize crucial features for classifying Alzheimer's patients. However, as the model nears its maximum accuracy, the rate of improvement gradually decreases. The learning curve also reveals fluctuations in the model's accuracy, with periods of rapid increase, followed by plateaus or slight dips. This behavior is likely attributed to the model learning diverse features at various stages during training.

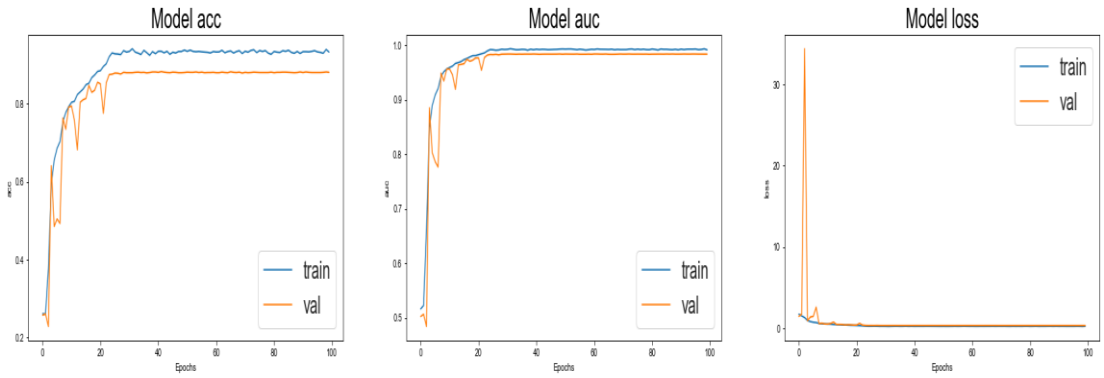

Figure 4.8. Learning curve of DenseNet model.

3.5.9. Inception V3

The learning curve demonstrates that the model's accuracy first rises sharply as it learns to recognize the characteristics crucial for categorizing Alzheimer's patients. However, when the model gets closer to its highest level of accuracy, the rate of development slows down. The learning curve also shows that the model's accuracy is not always consistent. There are times when the accuracy rises quickly, which are followed by times when it plateaus or even falls significantly. This is probably because the model is learning to recognize various features at various phases of training.

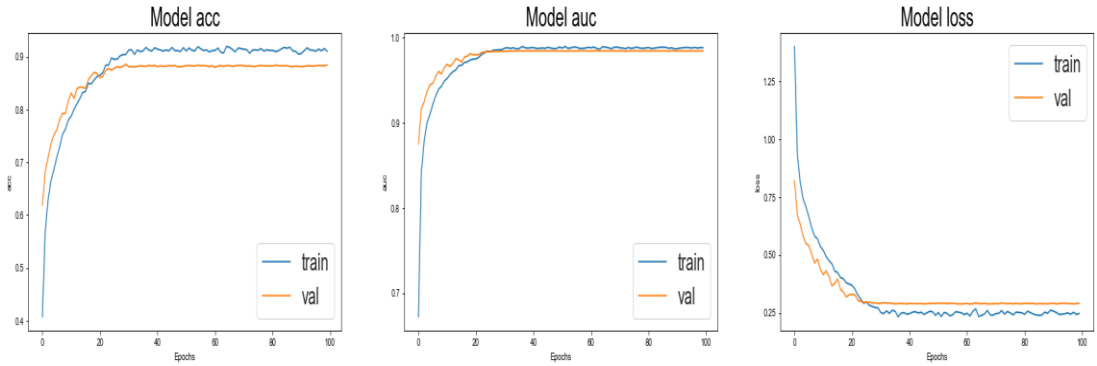

Figure 4.9. Learning curve of Inception v3 model.

PART 5

DISCUSSION

5.1. ALGORITHM PERFORMANCE

The results of this study demonstrate the potential of deep learning algorithms for accurately predicting Alzheimer's disease stages using MRI segmentation data. Authors [81] used deep learning algorithm for Alzheimer's disease early detection specifically focusing on CNN and MobileNet and achieved a significant result. For our research, four different algorithms were tested, and MobileNet was found to be the best-performing algorithm, followed by CNN, DenseNet, and Inception V3. This is likely due to the fact that MobileNet is a lightweight architecture that is specifically designed for efficient image classification. As a result, it is able to learn the important features from MRI segmentation images quickly and accurately. The confusion matrices and AUC values provided in the results section provide further confirmation of the relative performance of the four algorithms. The confusion matrices show that MobileNet has the lowest misclassification rates for all four categories, while Inception V3 has the highest misclassification rates for the very mild dementia category. The AUC values also show that MobileNet has the highest AUC for all four categories, while Inception V3 has the lowest AUC for the very mild dementia category. The learning curves also provide valuable insights into the performance of the four algorithms. They suggest that MobileNet and CNN are the best-performing algorithms, followed by DenseNet and Inception V3. This is probably because DenseNet and Inception V3 are simpler designs than MobileNet and CNN, which are more sophisticated.

5.2. IMPLICATIONS FOR EARLY AD DETECTION

This study's discoveries have various ramifications for the early identification of Alzheimer's infection. They initially recommend that using X-ray division

information, profound learning calculations can be used to anticipate the phases of Alzheimer's sickness exactly. This is a huge finding since it could prompt the improvement of novel, upgraded strategies for the early determination of Alzheimer's sickness. Second, the consequences of this study show that the best calculation for anticipating Alzheimer's infection from X-ray division information is MobileNet. This is vital in light of the fact that it raises the likelihood that MobileNet may be utilized to foster a convenient and proficient profound learning strategy for the early determination of Alzheimer's disease. In decision, the discoveries of this study show that utilizing a bigger dataset and preparing the calculations for a more extended timeframe can improve the viability of profound learning calculations for foreseeing Alzheimer's sickness. This is critical on the grounds that it recommends future upgrades in profound learning calculations' precision for identifying Alzheimer's sickness right off the bat. Generally, the review's discoveries are empowering and demonstrate that profound learning calculations may sometime be applied to the early recognizable proof of Alzheimer's infection. Further examination is expected to affirm these outcomes and work on the adequacy and utility of profound learning calculations for the early ID of Alzheimer's illness.

5.3. STATE-OF-THE-ART COMPARISON

This comparative analysis, shown in Table 4 indicates the effectiveness of our work and also provides valuable insights into how it fares against prominent studies in the field. Helaly et al. [17] conducted their research with a dataset comprising 6,000 samples, implementing a CNN-based approach that resulted in an accuracy of 93.61% and 95.17% for 2D. Munir et al. [18] expanded the dataset to 6,393 samples, using a CNN model that achieved an accuracy of 91.7%. B. Lu et al. [19] worked with 6,857 samples, deploying the Inception-ResNet-V2 architecture, and achieved an accuracy of 90.9%. N. H. et al. [20] tackled the problem with 5,121 samples, harnessing a CNN-based model that impressed with an accuracy of 94.61%. Balaji et al. [22] focused on a smaller dataset of 624 samples but employed a CLSTM model, resulting in an impressive accuracy of 98.5%.

Table 5.1. State-of-the-art comparison.

Research	Year	Dataset Size	Algorithm	Result
[17] Helaly et al.	2022	6000 samples	CNN	Around 93% to 95%
[18] Munir et al.	2022	6,393 samples	CNN	91.7%
[19] Lu et al.	2022	6,857 samples	Inception-ResNet-V2	90.9%
[20] N. H. et al.	2022	5,121 samples	CNN	94.61%
[22] Balaji et al.	2023	624 samples	CLSTM	98.5%
Our Approach	2023	12,800 samples	MobileNet/CNN	95.92%

In contrast, our approach set out with a dataset comprising 12,800 samples, harnessing the MobileNet architecture to attain an accuracy of 95.92%. This exceptional performance surpasses the majority of state-of-the-art methods, highlighting our approach's potential in the early prediction of Alzheimer's disease stages. Notably, our approach achieves an accuracy that is on par with models designed for larger datasets, a testament to its efficiency and robustness. This outcome firmly positions our methodology as a competitive and promising contender in the field of Alzheimer's disease prediction, holding significant potential for the early detection and classification of this debilitating condition.

PART 6

CONCLUSION AND FUTURE WORK

In the course of our investigation, we have delved into the potential of deep learning systems to harness MRI segmentation data for the prediction of Alzheimer's disease stages. Among the four algorithms we scrutinized, MobileNet emerged as the standout performer, surpassing CNN, DenseNet, and Inception V3. This outcome marks a substantial stride in the utilization of deep learning for the early diagnosis of Alzheimer's disease. The results underscore the immense promise that deep learning algorithms hold for revolutionizing the early detection of Alzheimer's disease. They are not only heartening but genuinely thrilling. It's important to recognize that these promising results beckon for a more comprehensive examination to refine the efficiency and efficacy of deep learning algorithms in the context of early Alzheimer's disease diagnosis. Notwithstanding these positive outcomes, it is crucial to acknowledge the noteworthy constraints of our research, as they impact the appropriate interpretation of the data. First and foremost, our dataset is small even though it was carefully chosen. This limitation may have introduced some constraints on the accuracy of our findings. Additionally, the use of the same dataset for both algorithm evaluation and training introduce the risk of overfitting, potentially affecting the robustness of our algorithms. Furthermore, our exploration was restricted to just four alternative methods, leaving open the possibility that other algorithms may demonstrate even more remarkable performance. In light of these limitations, addressing these constraints represents a critical step as we endeavor to advance the field of early Alzheimer's disease diagnosis.

Our study lays the foundation for promising future research avenues. The journey ahead includes the imperative task of validating and refining our findings. For the suggested models to be reliable and scalable, access to bigger and more varied datasets is required. In addition, longitudinal research tracking the course of Alzheimer's

disease over an extended period of time could yield essential knowledge regarding the dynamic character of the illness. To make deep learning models for medical diagnosis more practically usable in clinical settings, we also need to prioritize improving their interpretability. In addition to these considerations, we are keen on addressing the challenges associated with small sample sizes and imbalanced data, and further exploring the practical integration of deep learning models into clinical practice, while being attentive to ethical considerations and the preservation of patient privacy. Comparative assessments of deep learning against traditional diagnostic methods, as well as the exploration of multimodal approaches, stand as promising paths to enhance both accuracy and our understanding of disease variability. Lastly, the prospect of real-time and remote assessment for Alzheimer's disease represents an exciting frontier for our future research. Achieving effective remote detection and diagnosis could transform patient care, revolutionizing early detection and management in the realm of this debilitating condition. Pursuing these research directions, we aim to bring us closer to more precise and timely diagnoses, ultimately improving patient outcomes and offering a brighter future for those grappling with Alzheimer's disease.

REFERENCES

[1] L. Piubelli, G. Murtas, V. Rabattoni, and L. Pollegioni, "The Role of D-Amino Acids in Alzheimer's Disease," *Journal of Alzheimer's Disease*, vol. 80, no. 2, pp. 475–492, Jan. 2021, doi: 10.3233/JAD-201217.

[2] 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. 457612, Aug. 2019, doi: 10.3389/FNAGI.2019.00220/BIBTEX.

[3] C. Laske *et al.*, "Innovative diagnostic tools for early detection of Alzheimer's disease," *Alzheimer's & Dementia*, vol. 11, no. 5, pp. 561–578, May 2015, doi: 10.1016/J.JALZ.2014.06.004.

[4] N. Yamanakkanavar, J. Y. Choi, and B. Lee, "MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey," *Sensors 2020, Vol. 20, Page 3243*, vol. 20, no. 11, p. 3243, Jun. 2020, doi: 10.3390/S20113243.

[5] H. A. Helaly, M. Badawy, and A. Y. Haikal, "Deep Learning Approach for Early Detection of Alzheimer's Disease," *Cognit Comput*, vol. 14, no. 5, pp. 1711–1727, Sep. 2022, doi: 10.1007/S12559-021-09946-2/FIGURES/15.

[6] S. Gao and D. Lima, "A review of the application of deep learning in the detection of Alzheimer's disease," *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 1–8, Jun. 2022, doi: 10.1016/J.IJCCE.2021.12.002.

[7] S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng, "Early diagnosis of Alzheimer's disease with deep learning," *2014 IEEE 11th International Symposium on Biomedical Imaging, ISBI 2014*, pp. 1015–1018, Jul. 2014, doi: 10.1109/ISBI.2014.6868045.

[8] A. P. Porsteinsson, R. S. Isaacson, S. Knox, M. N. Sabbagh, and I. Rubino, "Diagnosis of Early Alzheimer's Disease: Clinical Practice in 2021," *The Journal of Prevention of Alzheimer's Disease 2021 8:3*, vol. 8, no. 3, pp. 371–386, Jun. 2021, doi: 10.14283/JPAD.2021.23.

[9] G. Chetelat and J. C. Baron, "Early diagnosis of alzheimer's disease: contribution of structural neuroimaging," *Neuroimage*, vol. 18, no. 2, pp. 525–541, Feb. 2003, doi: 10.1016/S1053-8119(02)00026-5.

[10] J. Collinge *et al.*, “INHERITED PRION DISEASE WITH 144 BASE PAIR GENE INSERTION: 2. CLINICAL AND PATHOLOGICAL FEATURES,” *Brain*, vol. 115, no. 3, pp. 687–710, Jun. 1992, doi: 10.1093/BRAIN/115.3.687.

[11] N. Yamanakkanavar, J. Y. Choi, and B. Lee, “MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer’s Disease: A Survey,” *Sensors 2020, Vol. 20, Page 3243*, vol. 20, no. 11, p. 3243, Jun. 2020, doi: 10.3390/S20113243.

[12] D. S. Knopman, B. F. Boeve, and R. C. Petersen, “Essentials of the Proper Diagnoses of Mild Cognitive Impairment, Dementia, and Major Subtypes of Dementia,” *Mayo Clin Proc*, vol. 78, no. 10, pp. 1290–1308, Oct. 2003, doi: 10.4065/78.10.1290.

[13] S. S. Kundaram and K. C. Pathak, “Deep Learning-Based Alzheimer Disease Detection,” *Lecture Notes in Electrical Engineering*, vol. 673, pp. 587–597, 2021, doi: 10.1007/978-981-15-5546-6_50/COVER.

[14] S. Shojaei, M. Saniee Abadeh, and Z. Momeni, “An evolutionary explainable deep learning approach for Alzheimer’s MRI classification,” *Expert Syst Appl*, vol. 220, p. 119709, Jun. 2023, doi: 10.1016/J.ESWA.2023.119709.

[15] H. Hampel *et al.*, “Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic,” *Nature Reviews Neurology 2018 14:11*, vol. 14, no. 11, pp. 639–652, Oct. 2018, doi: 10.1038/s41582-018-0079-7.

[16] J. E. Wraith *et al.*, “Mucopolysaccharidosis type II (Hunter syndrome): A clinical review and recommendations for treatment in the era of enzyme replacement therapy,” *Eur J Pediatr*, vol. 167, no. 3, pp. 267–277, Mar. 2008, doi: 10.1007/S00431-007-0635-4/TABLES/2.

[17] H. A. Helaly, M. Badawy, and A. Y. Haikal, “Deep Learning Approach for Early Detection of Alzheimer’s Disease,” *Cognit Comput*, vol. 14, no. 5, pp. 1711–1727, Sep. 2022, doi: 10.1007/S12559-021-09946-2/FIGURES/15.

[18] S. Munir *et al.*, “Alzheimer Disease Detection Empowered with Transfer Learning”, doi: 10.32604/cmc.2022.020866.

[19] B. Lu *et al.*, “A practical Alzheimer’s disease classifier via brain imaging-based deep learning on 85,721 samples,” *J Big Data*, vol. 9, no. 1, pp. 1–22, Dec. 2022, doi: 10.1186/S40537-022-00650-Y/FIGURES/5.

[20] N. H. ; U. ; Rehman *et al.*, “Constructing Domain Ontology for Alzheimer Disease Using Deep Learning Based Approach,” *Electronics 2022, Vol. 11, Page 1890*, vol. 11, no. 12, p. 1890, Jun. 2022, doi: 10.3390/ELECTRONICS11121890.

[21] V. S. Diogo, H. A. Ferreira, and D. Prata, “Early diagnosis of Alzheimer’s disease using machine learning: a multi-diagnostic, generalizable approach,” *Alzheimers Res Ther*, vol. 14, no. 1, pp. 1–21, Dec. 2022, doi: 10.1186/S13195-022-01047-Y/FIGURES/4.

[22] 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 2023, Vol. 11, Page 149*, vol. 11, no. 1, p. 149, Jan. 2023, doi: 10.3390/BIOMEDICINES11010149.

[23] "Alzheimer MRI Preprocessed Dataset." Accessed: Aug. 08, 2023. [Online]. Available: https://www.kaggle.com/datasets/sachinkumar413/alzheimer-mri-dataset.

[24] S. P. Singh, L. Wang, S. Gupta, H. Goli, P. Padmanabhan, and B. Gulyás, "3D Deep Learning on Medical Images: A Review," *Sensors 2020, Vol. 20, Page 5097*, vol. 20, no. 18, p. 5097, Sep. 2020, doi: 10.3390/S20185097.

[25] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, "Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions," *J Digit Imaging*, vol. 30, no. 4, pp. 449–459, Aug. 2017, doi: 10.1007/S10278-017-9983-4/TABLES/4.

[26] L. Zhang *et al.*, "Generalizing Deep Learning for Medical Image Segmentation to Unseen Domains via Deep Stacked Transformation," *IEEE Trans Med Imaging*, vol. 39, no. 7, pp. 2531–2540, Jul. 2020, doi: 10.1109/TMI.2020.2973595.

[27] P. Celard, E. L. Iglesias, J. M. Sorribes-Fdez, R. Romero, A. S. Vieira, and L. Borrajo, "A survey on deep learning applied to medical images: from simple artificial neural networks to generative models," *Neural Computing and Applications 2022 35:3*, vol. 35, no. 3, pp. 2291–2323, Nov. 2022, doi: 10.1007/S00521-022-07953-4.

[28] A. Subasi, "Use of artificial intelligence in Alzheimer's disease detection," *Artificial Intelligence in Precision Health: From Concept to Applications*, pp. 257–278, Jan. 2020, doi: 10.1016/B978-0-12-817133-2.00011-2.

[29] M. Elgendi *et al.*, "The Effectiveness of Image Augmentation in Deep Learning Networks for Detecting COVID-19: A Geometric Transformation Perspective," *Front Med (Lausanne)*, vol. 8, p. 629134, Mar. 2021, doi: 10.3389/FMED.2021.629134/BIBTEX.

[30] Z. A. khalaf, and N. T. A. Ramaha, "Review Of Breast Diagnosis Detection and Classification Based on Machine Learning," *International Conference on Trends in Advanced Research*, vol. 1, pp. 222–230, Mar. 2023. [Online]. Available: https://as-proceeding.com/index.php/ictar/article/view/209

[31] V. Sandfort, K. Yan, P. J. Pickhardt, and R. M. Summers, "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks," *Scientific Reports 2019 9:1*, vol. 9, no. 1, pp. 1–9, Nov. 2019, doi: 10.1038/s41598-019-52737-x.

[32] O. S. Eyobu and D. S. Han, "Feature Representation and Data Augmentation for Human Activity Classification Based on Wearable IMU Sensor Data Using

a Deep LSTM Neural Network," *Sensors 2018, Vol. 18, Page 2892*, vol. 18, no. 9, p. 2892, Aug. 2018, doi: 10.3390/S18092892.

[33] P. Celard, E. L. Iglesias, J. M. Sorribes-Fdez, R. Romero, A. S. Vieira, and L. Borrajo, "A survey on deep learning applied to medical images: from simple artificial neural networks to generative models," *Neural Computing and Applications 2022 35:3*, vol. 35, no. 3, pp. 2291–2323, Nov. 2022, doi: 10.1007/S00521-022-07953-4.

[34] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," *2018 International Interdisciplinary PhD Workshop, IIPhDW 2018*, pp. 117–122, Jun. 2018, doi: 10.1109/IIPHDW.2018.8388338.

[35] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning", doi: 10.1186/s40537-019-0197-0.

[36] G. Fanelli, J. Gall, and L. Van Gool, "Real time head pose estimation with random regression forests," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 617–624, 2011, doi: 10.1109/CVPR.2011.5995458.

[37] J. R. Rice, "Stresses Due to a Sharp Notch in a Work-Hardening Elastic-Plastic Material Loaded by Longitudinal Shear," *J Appl Mech*, vol. 34, no. 2, pp. 287–298, Jun. 1967, doi: 10.1115/1.3607681.

[38] O. Ecabert *et al.*, "Automatic model-based segmentation of the heart in CT images," *IEEE Trans Med Imaging*, vol. 27, no. 9, pp. 1189–1202, Sep. 2008, doi: 10.1109/TMI.2008.918330.

[39] T. Igarashi and K. Hinckley, "Speed-dependent Automatic Zooming for Browsing Large Documents".

[40] M. Bohan, D. S. McConnell, A. Chaparro, and S. G. Thompson, "The Effects of Visual Magnification and Physical Movement Scale on the Manipulation of a Tool With Indirect Vision," *J Exp Psychol Appl*, vol. 16, no. 1, pp. 33–44, Mar. 2010, doi: 10.1037/A0018501.

[41] C. Witharana *et al.*, "An Object-Based Approach for Mapping Tundra Ice-Wedge Polygon Troughs from Very High Spatial Resolution Optical Satellite Imagery," *Remote Sensing 2021, Vol. 13, Page 558*, vol. 13, no. 4, p. 558, Feb. 2021, doi: 10.3390/RS13040558.

[42] E. Luz *et al.*, "Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images," *Research on Biomedical Engineering*, vol. 38, no. 1, pp. 149–162, Mar. 2022, doi: 10.1007/S42600-021-00151-6/FIGURES/10.

[43] J. E. Murray, "Flipping and spinning: Spatial transformation procedures in the identification of rotated natural objects," *Mem Cognit*, vol. 25, no. 1, pp. 96–105, 1997, doi: 10.3758/BF03197287/METRICS.

[44] N. J. Mitra, M. Pauly, M. Wand, and D. Ceylan, “Symmetry in 3D Geometry: Extraction and Applications,” *Computer Graphics Forum*, vol. 32, no. 6, pp. 1–23, Sep. 2013, doi: 10.1111/CGF.12010.

[45] J. A. M. Sidey-Gibbons and C. J. Sidey-Gibbons, “Machine learning in medicine: a practical introduction,” *BMC Med Res Methodol*, vol. 19, no. 1, pp. 1–18, Mar. 2019, doi: 10.1186/S12874-019-0681-4/TABLES/5.

[46] S. Ammar, T. Bouwmans, and M. Neji, “Face Identification Using Data Augmentation Based on the Combination of DCGANs and Basic Manipulations,” *Information 2022, Vol. 13, Page 370*, vol. 13, no. 8, p. 370, Aug. 2022, doi: 10.3390/INFO13080370.

[47] R. Ranftl, K. Lasinger, D. Hafner, K. Schindler, and V. Koltun, “Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer,” *IEEE Trans Pattern Anal Mach Intell*, vol. 44, no. 3, pp. 1623–1637, Mar. 2022, doi: 10.1109/TPAMI.2020.3019967.

[48] A. Iqbal, M. Sharif, M. Yasmin, M. Raza, and S. Aftab, “Generative adversarial networks and its applications in the biomedical image segmentation: a comprehensive survey,” *International Journal of Multimedia Information Retrieval 2022 11:3*, vol. 11, no. 3, pp. 333–368, Jul. 2022, doi: 10.1007/S13735-022-00240-X.

[49] R. Wang *et al.*, “Applications of generative adversarial networks in neuroimaging and clinical neuroscience,” *Neuroimage*, vol. 269, p. 119898, Apr. 2023, doi: 10.1016/J.NEUROIMAGE.2023.119898.

[50] S. Wang *et al.*, “Advances in Data Preprocessing for Biomedical Data Fusion: An Overview of the Methods, Challenges, and Prospects,” *Information Fusion*, vol. 76, pp. 376–421, Dec. 2021, doi: 10.1016/J.INFFUS.2021.07.001.

[51] P. M. Tostado, B. U. Pedroni, and G. Cauwenberghs, “Performance Trade-offs in Weight Quantization for Memory-Efficient Inference,” *Proceedings 2019 IEEE International Conference on Artificial Intelligence Circuits and Systems, AICAS 2019*, pp. 246–250, Mar. 2019, doi: 10.1109/AICAS.2019.8771473.

[52] D. Preuveneers, I. Tsingenopoulos, and W. Joosen, “Resource Usage and Performance Trade-offs for Machine Learning Models in Smart Environments,” *Sensors 2020, Vol. 20, Page 1176*, vol. 20, no. 4, p. 1176, Feb. 2020, doi: 10.3390/S20041176.

[53] W. Guo, “Explainable Artificial Intelligence for 6G: Improving Trust between Human and Machine,” *IEEE Communications Magazine*, vol. 58, no. 6, pp. 39–45, Jun. 2020, doi: 10.1109/MCOM.001.2000050.

[54] M. Del Giudice and B. J. Crespi, “Basic functional trade-offs in cognition: An integrative framework,” *Cognition*, vol. 179, pp. 56–70, Oct. 2018, doi: 10.1016/J.COGNITION.2018.06.008.

[55] W. Liu, H. Han, and G. Han, “Transfer Learning with Deep Convolutional Neural Network for Automated Plant Identification,” *2022 7th International Conference on Image, Vision and Computing, ICIVC 2022*, pp. 555–560, 2022, doi: 10.1109/ICIVC55077.2022.9886149.

[56] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/S13244-018-0639-9/FIGURES/15.

[57] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, “Deep convolutional neural networks for hyperspectral image classification,” *J Sens*, vol. 2015, 2015, doi: 10.1155/2015/258619.

[58] M. A. Saleem, N. Senan, F. Wahid, M. Aamir, A. Samad, and M. Khan, “Comparative Analysis of Recent Architecture of Convolutional Neural Network,” *Math Probl Eng*, vol. 2022, 2022, doi: 10.1155/2022/7313612.

[59] S. Li, Y. Yao, J. Hu, G. Liu, X. Yao, and J. Hu, “An Ensemble Stacked Convolutional Neural Network Model for Environmental Event Sound Recognition,” *Applied Sciences 2018, Vol. 8, Page 1152*, vol. 8, no. 7, p. 1152, Jul. 2018, doi: 10.3390/APP8071152.

[60] Y. D. Zhang, Y. Zhang, X. X. Hou, H. Chen, and S. H. Wang, “Seven-layer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed,” *Multimed Tools Appl*, vol. 77, no. 9, pp. 10521–10538, May 2018, doi: 10.1007/S11042-017-4554-8/METRICS.

[61] N. T. A. Ramaha et al., "Brain Pathology Classification of MR Images Using Machine Learning Techniques," *MDPI*, vol. 12, p. 167, Aug. 2023, doi: 10.3390/computers12080167.

[62] A. Narin, C. Kaya, and Z. Pamuk, “Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks,” *Pattern Analysis and Applications 2021 24:3*, vol. 24, no. 3, pp. 1207–1220, Mar. 2020, doi: 10.1007/s10044-021-00984-y.

[63] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review 2020 53:8*, vol. 53, no. 8, pp. 5455–5516, Apr. 2020, doi: 10.1007/S10462-020-09825-6.

[64] V. Maeda-Gutiérrez *et al.*, “Comparison of Convolutional Neural Network Architectures for Classification of Tomato Plant Diseases,” *Applied Sciences 2020, Vol. 10, Page 1245*, vol. 10, no. 4, p. 1245, Feb. 2020, doi: 10.3390/APP10041245.

[65] H. Fırat, M. E. Asker, M. İ. Bayındır, and D. Hanbay, “Hybrid 3D/2D Complete Inception Module and Convolutional Neural Network for Hyperspectral Remote Sensing Image Classification,” *Neural Process Lett*, vol. 55, no. 2, pp. 1087–1130, Apr. 2023, doi: 10.1007/S11063-022-10929-Z/METRICS.

[66] M. A. Al-masni, D. H. Kim, and T. S. Kim, “Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification,” *Comput Methods Programs Biomed*, vol. 190, p. 105351, Jul. 2020, doi: 10.1016/J.CMPB.2020.105351.

[67] H. K. Gajera, D. R. Nayak, and M. A. Zaveri, “A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features,” *Biomed Signal Process Control*, vol. 79, p. 104186, Jan. 2023, doi: 10.1016/J.BSPC.2022.104186.

[68] A. Haghanifar, M. M. Majdabadi, Y. Choi, S. Deivalakshmi, and S. Ko, “COVID-CXNet: Detecting COVID-19 in frontal chest X-ray images using deep learning,” *Multimed Tools Appl*, vol. 81, no. 21, pp. 30615–30645, Sep. 2022, doi: 10.1007/S11042-022-12156-Z/FIGURES/25.

[69] X. Liu, Z. Jia, X. Hou, M. Fu, L. Ma, and Q. Sun, “Real-time Marine Animal Images Classification by Embedded System Based on Mobilenet and Transfer Learning,” *OCEANS 2019 - Marseille, OCEANS Marseille 2019*, vol. 2019-June, Jun. 2019, doi: 10.1109/OCEANSE.2019.8867190.

[70] L. Bai, Y. Zhao, and X. Huang, “A CNN Accelerator on FPGA Using Depthwise Separable Convolution,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 10, pp. 1415–1419, Oct. 2018, doi: 10.1109/TCSII.2018.2865896.

[71] H. Yuan, J. Cheng, Y. Wu, and Z. Zeng, “Low-res MobileNet: An efficient lightweight network for low-resolution image classification in resource-constrained scenarios,” *Multimed Tools Appl*, vol. 81, no. 27, pp. 38513–38530, Nov. 2022, doi: 10.1007/S11042-022-13157-8/METRICS.

[72] A. Kumar, A. Sharma, V. Bharti, A. K. Singh, S. K. Singh, and S. Saxena, “MobiHisNet: A Lightweight CNN in Mobile Edge Computing for Histopathological Image Classification,” *IEEE Internet Things J*, vol. 8, no. 24, pp. 17778–17789, Dec. 2021, doi: 10.1109/JIOT.2021.3119520.

[73] M. S. H. Shovon, M. J. Islam, M. N. A. K. Nabil, M. M. Molla, A. I. Jony, and M. F. Mridha, “Strategies for Enhancing the Multi-Stage Classification Performances of HER2 Breast Cancer from Hematoxylin and Eosin Images,” *Diagnostics 2022, Vol. 12, Page 2825*, vol. 12, no. 11, p. 2825, Nov. 2022, doi: 10.3390/DIAGNOSTICS12112825.

[74] 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.

[75] D. Brown and K. Bradshaw, “Deep Palmprint Recognition with Alignment and Augmentation of Limited Training Samples,” *SN Computer Science 2021 3:1*, vol. 3, no. 1, pp. 1–17, Oct. 2021, doi: 10.1007/S42979-021-00859-3.

[76] F. J. Ordóñez, D. Roggen, Y. Liu, W. Xiao, H.-C. Chao, and P. Chu, “Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition,” *Sensors 2016, Vol. 16, Page 115*, vol. 16, no. 1, p. 115, Jan. 2016, doi: 10.3390/S16010115.

[77] C. Ballabio and S. Sterlacchini, “Support Vector Machines for Landslide Susceptibility Mapping: The Staffora River Basin Case Study, Italy,” *Math Geosci*, vol. 44, no. 1, pp. 47–70, Jan. 2012, doi: 10.1007/S11004-011-9379-9/METRICS.

[78] G. Zhang, M. Wang, and K. Liu, “Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China,” *International Journal of Disaster Risk Science*, vol. 10, no. 3, pp. 386–403, Sep. 2019, doi: 10.1007/S13753-019-00233-1/FIGURES/10.

[79] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, “Machine learning: A review of classification and combining techniques,” *Artif Intell Rev*, vol. 26, no. 3, pp. 159–190, Nov. 2006, doi: 10.1007/S10462-007-9052-3/METRICS.

[80] R. Haque, N. Islam, M. Tasneem, and A. K. Das, “Multi-class sentiment classification on Bengali social media comments using machine learning,” *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 21–35, Jun. 2023, doi: 10.1016/J.IJCCE.2023.01.001.

[81] M. Alhamidi, and N. T. A. Ramaha, "Unveiling Alzheimer's Disease via MRI: Deep Learning Approaches for Accurate Detection," International Journal of Advanced Natural Sciences and Engineering Researches, vol. 7, pp. 418-422, Oct. 2023. [Online]. Available: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=Gw2-FFYAAAAJ&citation_for_view=Gw2-FFYAAAAJ:ufrVoPGSRksC

RESUME

Mais Alhamidi AHMAD she graduated from elementary in Qatar education and high school in Turkey. After that, she started the undergraduate program at The International Islamic University Malaysia in Malaysia Department of Information and Communication Technology specializing in information technology starting in 2017 and graduating in 2022. To complete M. Sc. education, she moved to Karabük University in 2022.