ALZHEIMER'S DISEASE DETECTION FROM MRI USING DEEP AND TRANSFER LEARNING APPROACHES

BY

Md. Ashraful Momenin
ID: 241-25-028

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

Supervised By

Dr. Naznin Sultana
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised By
Abdus Sattar
Associate Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH
MAY 2025

APPROVAL

This thesis titled **"Alzheimer's Disease Detection From MRI Using Deep And Transfer Learning Approaches"**, submitted by **Md. Ashraful Momenin**, ID No: **241-25-028** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of **MSc. in Computer Science and Engineering** and approved as to its style and contents. The presentation has been held on **24-05-2025**.

BOARD OF EXAMINERS

Chairman

Dr. Arif Mahmud
Associate Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

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

Internal Examiner

Dr. Md Alamgir Kabir
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Nazibur Rahman
Technical Lead, Database Administrator
Wipro, Telenor - Grameen Phone Account
Dhaka, Bangladesh

ii

DECLARATION

I hereby declare that this research has been done by me under the supervision of **Dr. Naznin Sultana, Associate Professor, Department of CSE,** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Dr. Naznin Sultana
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Abdus Sattar
Associate Professor
Department of CSE
Daffodil International University

Submitted by:

Md. Ashraful Momenin
ID: 241-25-028
Department of CSE
Daffodil International University

iii

©Daffodil International University

ACKNOWLEDGEMENT

First I express my heartiest thanks and gratefulness to Almighty Allah for His divine blessing which makes it possible to complete the final year project/internship successfully.

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

I would like to express my heartiest gratitude to **Dr. Sheak Rashed Haider Noori**, **Head,** Department of CSE, for his kind help to finish our project and also to other faculty members and the staff of CSE department of Daffodil International University.

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

ABSTRACT

Alzheimer's disease is a common neurological disorder that affects millions of people around the world. It is important to find the disease early so that it can be treated effectively and in a fast manner. Not long ago, deep learning methods became useful for looking at medical images and helping to find diseases. This work shows a new way to use MRI images to find people with Alzheimer's disease by combining deep learning and transfer learning methods. This work focuses on using deep learning approaches more especially, transfer learning to identify Alzheimer's disease with MRI data. It compares with popular deep learning models using ResNet50, VGG19, Xception, DenseNet169 for image classification. Our custom CNN architecture performs better than these mature models, with the single layer perceptron net giving a very high 97% accuracy. The work contributions discuss the pro and cons of each model and detail of our design choice for the custom CNN and its capability to learn complex patterns. The custom CNN design is very detailed, showing why it made the choices it did and how its hyper parameters are set up. The thesis goes into detail about the reasoning behind each design choice, which helps show how the model best finds complicated patterns in the data. This in-depth study not only helps us understand how deep learning works, but it also gives us a good starting point for building unique architectures in the future. In addition, the study looks into whether information gained from the custom CNN can be used in other areas. It focuses on how flexible and useful the proposed architecture is for generalization. The findings of the study are shown to be important beyond the current dataset by talking about their practical implications and possible uses in the real world. In the end, this thesis gives a full review of well-known deep learning architectures and offers a new custom CNN architecture that does better at image classification tasks than existing models. The study adds to the current conversation in artificial intelligence by showing how useful customized model design can be and by giving useful suggestions for future progress in the field.

***Keyword*:** CNN,Alzheimer's illness, Recall, MRI images, Transfer Learning

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Alzheimer's disease (AD) is an illness in which memory cells are permanently damaged. This is also known as dementia. Brain cells die and tissues are lost all over the brain, which makes memory loss happen. The bad effects make it hard to do everyday things like reading, speaking, and writing. In the last stages of Alzheimer's disease, patients experience more serious effects, such as problems with their breathing and hearts that stop working, which can lead to death [1]. Unfortunately, the wrong medicine has made it impossible to accurately and quickly diagnose AD [2]. Getting care early on can make the patient's life better, though. When the brain starts to malfunction, its symptoms appear slowly at first but can get worse over time [3].

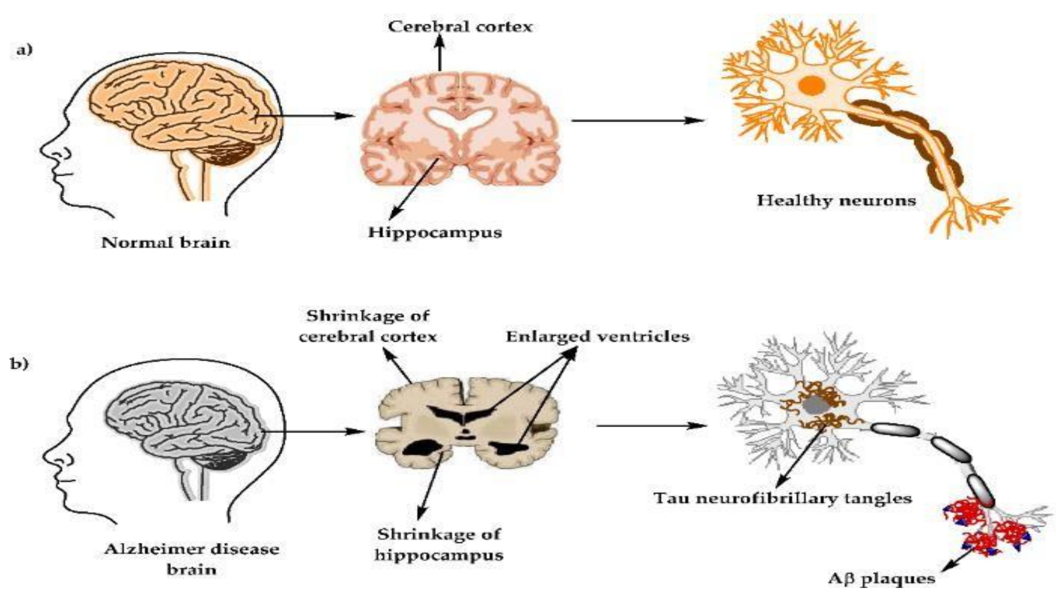

Figure 1.1: The difference between an AD-affected brain and a healthy one [24].

Traditionally, the diagnosis of AD was based on subjective and variable visual examination of MRI scans by a radiologist. This classic approach served as the foundation for all the classical diagnostic methods. But with the recent successes of (deep) learning, especially with convolutional neural networks (CNNs), things have changed tremendously. CNNs are able to efficiently created meaningful features and representations from complex images, and through their hierarchical architecture, build-up highly detailed models of the brain. This enables the identification of subtle signs of Alzheimer's disease.

Transfer learning is a technique for improving deep learning models by leveraging knowledge from pre-trained models trained on large datasets. Deep CNN can effectively capture discriminative knowledge from MRI of AD by using extracted information of different layers. Alzheimer's disease (AD) is characterized by a gradual loss of memory and other cognitive functions, as well as its characteristic symptoms and stages that worsen as the disease progresses.

The stages are listed as follows:

Stage 01: Preclinical Alzheimer's Disease

In the Preclinical Stage 01 of Alzheimer's disease, there are no obvious signs, which means the person is still in the early stages of the disease. Amyloid plaques and tau tangles may have started to form by now. These are two chemical changes that are linked to Alzheimer's disease. In spite of these basic changes, most people don't show obvious cognitive deficits or memory problems during this time. Alzheimer's is a stage that is subtle and hard to notice, which makes it hard to identify with standard medical tests. But new imaging techniques and diagnostic studies are making it more important to find biomarkers and use them to spot these early changes. By taking a different method, treatments and plans that try to delay the onset of Alzheimer's disease symptoms may have a chance to work.

Stage 02: Mild Cognitive Impairment

When Alzheimer's disease gets to the second stage, people may notice small changes in their memory, reasoning, and cognitive skills. These changes may be noticeable to them and to people close to them. You can tell you have these symptoms, but they don't really get in the way of your daily life. This stage is different because cognitive decline is only temporary; people may have mild cognitive problems without having major problems with how they function generally. It's important to remember that not everyone with MCI gets Alzheimer's. Some people with the illness stay stable or even get their thinking back. Around this time, it is suggested to keep a close eye on any changes in brain function and see if the disease has progressed to more severe stages by doing more tests and keeping a close eye on things.

Stage 03: Mild Alzheimer's

People with weak Alzheimer's disease who are in the third stage notice their cognitive decline and memory loss getting worse. Tasks like language, problem-solving, direction, and attention get harder as they go along. People often have trouble remembering names, recognizing faces, or recalling recent events, even if they are still somewhat independent. Organizational, planning, and decision-making problems get worse. This stage is marked by a steady decline in cognitive abilities that starts to get in the way of daily life and calls for more support and help from others. People who care for people with moderate Alzheimer's disease and their families need to know a lot about the symptoms and problems that come with it in order to make care plans and interventions that will improve the quality of life for people in this stage of the disease.

Stage 04: Moderate Alzheimer's

People in the fourth stage of moderate Alzheimer's disease have a lot less cognitive and mental function, which makes it very hard for them to go about their daily lives. People often feel more disoriented and confused, and it can be hard for them to recognize close friends and family. It gets harder and harder to do normal things like speak to people and enjoy hobbies. During this very important time, people often need a lot of help and care. Cognitive decline affects more than just memory; it also makes it harder to move around and have important conversations with other people. Caregivers and medical professionals are needed to help people with Alzheimer's disease stay healthy and maintain their sense of dignity during this difficult stage of the disease.

Stage 05: Severe Alzheimer's

People with the fifth and worst stage of Alzheimer's disease experience a lot of cognitive loss and problems with their daily lives. There could be big problems with communicating, which could lead to body language. It's harder to spot even familiar faces when you're not as aware of your surroundings. It gets harder to do simple things for yourself, really needing a lot of help with daily tasks. At this advanced level, cognitive decline continues without stopping, which has a big effect on quality of life as a whole.

1.2 Background

Alzheimer's disease is a major health problem that affects a huge number of people all over the world. This crippling disease makes it harder for the brain to understand, remember, and do simple things, and it ends in death [4]. Alzheimer's disease, or AD, is a neurodegenerative disease that gets worse over time and can't be fixed [5]. The number of people with Alzheimer's is expected to rise from 50 million in 2010 to 152 million by 2050. We need to move quickly to stop this growing health crisis [6]. It is very important to identify the disease quickly and correctly so that patients can get the best care possible.

The goal of this project is to use neural networks' natural abilities to find complicated patterns in brain pictures that show Alzheimer's disease. Transfer learning lets the models learn from what they already know, which leads to better performance and generalization across a wide range of patient groups. This thesis could make a difference because it moves the development of better tools for diagnosing Alzheimer's disease forward. Better patient outcomes, personalized treatment plans, and early action can all be made possible by better imaging analysis that allows for faster and more accurate identification. As our knowledge of neurological diseases grows, using AI-based methods could change the way Alzheimer's is diagnosed and treated. The main purpose of this thesis is to make a big contribution to the field of Alzheimer's disease research in order to improve the lives of people who are affected by it.

1.3 Motivation of the Research

There needs to be quick progress in deep learning and transfer learning methods for using MRI images to find Alzheimer's disease because the disease needs to be diagnosed quickly, correctly, and cheaply. The old ways of diagnosing aren't perfect because they are subjective and have boundaries. We want to use convolutional neural networks in deep learning to take advantage of these models' natural ability to find complex patterns and features in complex MRI data. We also use transfer learning to use information that is already there in big image datasets. This is done to make our models even more robust and useful in a wider range of situations. The fact that early detection of Alzheimer's disease might lead to better treatment. The main thing we want to do is make a big difference in the development of advanced diagnostic tools that could change how Alzheimer's disease is diagnosed and treated, making people with this difficult sickness happier.

1.4 Problem Statement

A big problem is that there isn't an early, effective, and cheap way to diagnose Alzheimer's disease. This means that cutting edge methods like the deep learning and transfer learning techniques used in MRI image analysis need to be looked into. Standard diagnostic methods, which rest on people's subjective opinions about what they see, can make mistakes and have their limits. A lot of the time, you need to use the built-in features of deep learning, especially convolutional neural networks, to find complicated patterns and features in MRI datasets. Using transfer learning is also a good way to make the models more resilient and useful in a wider range of situations by using previous information found in big image datasets. I can't stress enough how important it is to find Alzheimer's early because it can make treatment plans and patient care much better. The goal of this study is to find out if it is important to make a valuable addition to the progress of advanced diagnostic tools that could completely change how Alzheimer's disease is diagnosed and treated. The main goal is to make the lives of people who have Alzheimer's a lot better by applying cutting edge methods from deep learning and transfer learning to MRI image processing for Alzheimer's diagnosis.

1.5 Research Question

- How can deep learning detect Alzheimer's disease?
- Why choose for transfer learning in illness detection?
- How does transfer learning help the performance of Alzheimer's detection models?
- Will we able to diagnose Alzheimer's faster and cheaper with the help of AI?
- Are there subtle patterns in MRI scans that only AI models can pick out, and that are difficult for doctors to discern?

1.6 Research Objective

- Detect Alzheimer's disease from MRI scans using deep learning.
- Develop a cheap, quick and reliable way to diagnose Alzheimer's.
- Use transfer learning to increase model performance.
- In order to minimize the human error of visual inspection.
- Find the areas of the brain where Alzheimer's starts. Assist doctors in diagnosing Alzheimer's earlier and better.

1.7 Research Scope

Deep learning techniques for MRI-based AD detection are being studied. From data collection and pretreatment to model architecture design and transfer learning methodologies, it is a systematic study. The scope also includes model training and evaluation, interpretability analysis, generalizability considerations, and implementation details. The research involves collecting and curating numerous datasets to reliably and efficiently diagnose Alzheimer's disease using MRI images. It also involves preparing data to improve relevance and quality. The scope includes creating, improving, and calibrating deep learning models, especially convolutional neural networks. Transfer learning techniques to use prior knowledge from huge image datasets are also examined. The utility and trustworthiness of developed models depend on performance evaluation through extensive metrics assessment. The scope contains key Alzheimer's identifying regions and traits to aid model interpretation.

1.8 Summary

In the past, radiologists' subjective visual assessments were used to identify AD. However, recent progress in deep learning, especially in convolutional neural networks (CNNs), has made this method more efficient and objective. The study gets around the problems with traditional testing methods by using CNNs' ability to pick out small patterns in very complicated MRI data. Transfer learning is used to make models work better by using what they already know about big datasets that have been trained on models.Transfer learning makes the model more stable. The main idea behind this study is that early and accurate diagnosis of Alzheimer's disease through better imaging analysis could improve patient results. The main goal is to make patient care better by creating testing tools that will change the way Alzheimer's disease is found and treated.

In conclusion, this study addresses the immediate need for advanced diagnostic tools for finding Alzheimer's disease by in-depth research into the deep learning and transfer learning techniques used on magnetic resonance imaging scans. The end goal is to make a big difference in terms of improving care and treatment for people with Alzheimer's disease.

CHAPTER 2

BACKGROUND

2.1 Introduction

In order to build on what has already been learned, this section looks at the scholarly works that form the basis of our study. By looking at a lot of different pieces of writing, we want to make sense of the different techniques, problems, and new ideas that have shaped the discussion about combining deep learning and transfer learning in neuroimaging for Alzheimer's diagnosis. Putting these finds together not only helps us figure out how to do our study, but it also puts our work in the bigger picture of medical image analysis progress, with a focus on the important area of finding Alzheimer's disease.

2.2 Related Work

Using a large-scale dataset, the authors of illustrated the value of initializing other networks with pretrained models. They employed GoogleNet and Inception-ResNet networks trained on non-medical datasets to fine-tune the final fully connected (FC) layer in their instance rather than starting from scratch to train the network [7].

The study focuses on using the ResNet-18 model in conjunction with magnetic resonance imaging (MRI) to detect AD. Various 2D CNN have been employed to detect AD in previous research. Our results show that incorporating transfer learning into a 3D CNN improves an AD detection system's accuracy. This work compares 2D and 3D CNNs. 96.88% accuracy was attained by our methods through the use of an optimization method during the training process [8].

After reducing the amount of training data, allow us to select the most informative slices using picture entropy. Through studies on the ADNI dataset, we show that by employing training sets 10–20 times smaller than those utilized by other recent techniques, we can obtain remarkable performance with 4% and 7% gains in accuracy in classification tasks [9].

By adjusting a pre-trained neural network, the recommended method effectively applies transfer learning to categorize the photos. With the best overall accuracy of 92.85%, the system showed impressive results for the multi-class categorization of unsegmented photos [10].

In a prior work suggested a multi-staged model that included a strategy for classifying Alzheimer's disease. As part of the pre-processing stage, the method divided the input images into three categories: GM, White Matter (WM), and Cerebral Spinal Fluid (CSF).The method built the similarity matrices using GM as a ROI, from which the statistical attributes were derived [11].

Classification strategy for the test individuals was predicated on the 3D displacement field's estimation. Using feature selection approaches including the Bhattacharya distance, the student t-test, and Welch's t-test, feature reduction was applied over the extracted features. The Support Vector Machine (SVM) classifier was then trained using the chosen features, and the test data was classified with a 93.05% accuracy rate [12].

For binary classification, the authors used convolutional network topologies using freeze characteristics taken from the source data set. Every analysis made use of the Alzheimer's Disease National Initiative's (ADNI) MRI scan data collection. The results of the suggested method show that AD is classified in reference to cognitive normalcy, normal controls in relation to moderate cognitive impairment, and cognitive normalcy in relation to cognitive normalcy [13].

The suggested study demonstrates that VGG outperforms cutting-edge methods and a variety of architectures, achieving a 99.27% identification test set accuracy for Alzheimer's disease detection. Ramaniharan suggested segmenting T1-weighted MRI scans after analyzing the diversity in the corpus callosum's morphology [14].

Transfer learning is used in an effort to address these problems. Current designs such as VGG and Inception initialize the fully-connected layer with pre-trained weights using large benchmark datasets consisting of natural images, and then retrain the layer using only a small number of MRI images. To choose the most instructive slices for training, we employ picture entropy. They attained performance that is on par with, if not better than, that of deep learning-based methods already in use [15].

Employed [16] brain MRI to categorize the pictures into four stages: non-demented (ND), moderate dementia (MOD), mild dementia (MD), and very mild diarrhea (VMD). This multi-class classification system for Alzheimer's disease detection uses transfer learning. Simulation results show that the proposed is 91.70% correct. It was also mentioned that the new process produces more accurate findings than the older ones.

The author [17] of this work classified Alzheimer illness using many CNN transfer learning techniques. They tested a number of parameters on the ADNI dataset and obtained very precise results. We analyzed 13 variants of pre-trained CNN models using a modified method for transfer learning between two different domains using the ADNI dataset. DenseNet, on the other hand, outperformed the others, with an average maximum accuracy of 99.05. The author [18] tested transfer learning model on binary classification data and determined the outcomes.

The Oasis dataset and the suggested approach (Open Access Series of Imaging Studies). The results obtained indicate that the categorization rate of images using Transfer Learning, at 92.86%, was higher than that of CNN [21].

According to the experimental findings, the suggested framework obtains 96.25% accuracy for the ADNI dataset and 96.65% accuracy for the Alzheimer's Dataset. Furthermore, a superior accuracy performance is shown compared to other cutting-edge methods [22].

With an AUC of 90.2%, 83.2%, and 70.6% for AD from CN, MCIc from CN, and MCIc from MCInc, respectively, the ensemble transfer-learning approach demonstrated effective discrimination. The results obtained with the fusion of conventional-ML systems (AD from CN, 93.1% for MCIc from CN, 89.6% for MCIc from CN, and AUC in the range of 69.1–73.3%) were comparable or slightly lower. These findings provide fresh insights into the automatic early diagnosis and prognosis of AD using transfer learning and neuroimaging, even in cases where the patients have been pretrained with generic images [23].

In terms of accuracy, efficiency, and robustness, the results of the Alzheimer's Disease Neuroimaging Initiative (ADNI) magnetic resonance imaging (MRI) dataset [28] confirm the advantages of the suggested 2D-DCNN model. The model has remarkable efficacy in the classification of MRI images into three distinct groups, namely Alzheimer's Disease (AD), moderate cognitive impairment, and normal control. Surprisingly, even with imbalanced classes, the model attains an astounding 99.89% classification accuracy. Notably, the suggested model demonstrates remarkable improvements over current approaches, highlighting its noteworthy accuracy improvement.

In this work the authors investigate different deep learning classification techniques for accurately identifying Alzheimer's disease with the goals of improving patient care, reducing expenses, and enabling quick, reliable analysis in large-scale investigations. Using the Python computer language, the suggested methodology is put into practice and helps physicians accurately classify Alzheimer's disease. Thirty percent of the supplied images are set aside for validation, leaving the remaining seventy percent of the images for training. Surprisingly, our trained model achieves 100% accuracy on a different test set, highlighting its potential for reliable and accurate Alzheimer's Disease categorization [29].

2.3 Summery

This thesis's literature evaluation examines deep learning and transfer learning algorithms for MRI image-based Alzheimer's disease identification. It establishes a solid platform for future inquiries with its exhaustive investigation. This study critically identifies research gaps, paving the way for reliable and creative Alzheimer's disease early detection tools. The study contributes to the field's current conversation and serves as a beneficial tool for scholars who want to learn more about and employ AI methods for MRI analysis-based Alzheimer's diagnosis.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

In order to get correct diagnostic results when using medical imaging to look for Alzheimer's disease (AD), it is now necessary to use complex computational methods. As part of this, our study looks at how to use seven very successful transfer learning models on a large dataset that is grouped by stage of disease. ResNet 50, VGG19, Xception, DenseNet169, and custom CNN are some of these models.

3.2 Instrumentation

We use Google Colab, Kaggle Notebook, and Jupyter Notebook as integrated development environments (IDEs) for our study to get the computer power we need for this complex analysis. Python is a flexible computer language that works great for machine learning and manipulating large amounts of data. Python is used to implement the whole plan. The main idea of our work is summed up in this introduction.

3.3 Data Collection

Kaggle, a well-known site for sharing datasets and encouraging people to work together on research, was the major source of data we used to make a good dataset for studying Alzheimer's disease. There are pictures in four groups: Mildly Demented, Moderately Demented, Not Demented, and Very Mild are marked up in the previous file. The first part of our study involved a thorough search and collection process on Kaggle, which made sure that we got a large and varied collection of pictures related to Alzheimer's disease.

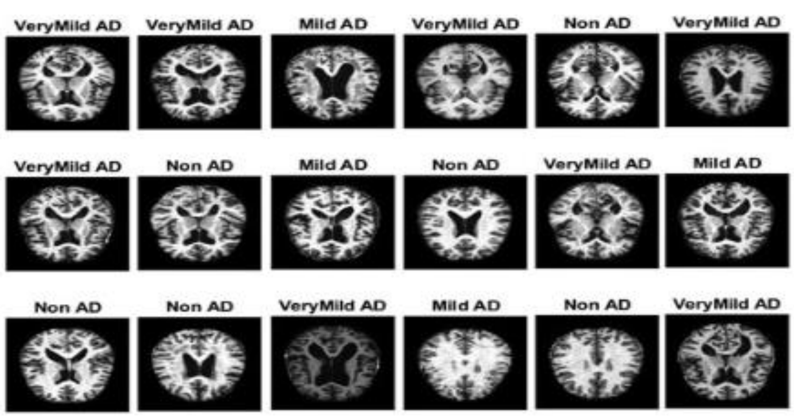

Figure 3.1: Dataset Images [25]

Each class reflects a different part of Alzheimer's disease so that the subtleties of the disease's progression can be seen. This allows for targeted model training and detailed analysis. At the end, a high-quality and standardized picture dataset for studying Alzheimer's disease has been created by carefully combining data from Kaggle with strict Steps for getting ready.

3.4 Data Exploration

We used 6400 pictures from 4 different groups. The dataset was split into 4 groups. The whole collection is shown in the following graph:

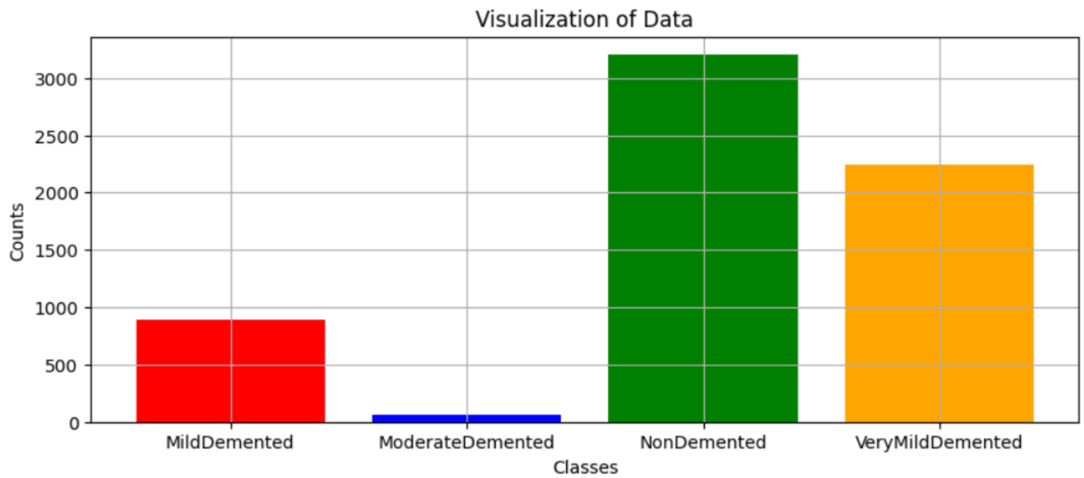

Figure 3.2: Dataset Bar Plot

The dataset includes four distinct classes categorized by the severity of Alzheimer's disease:

• Mild dementia

• Moderate

• Non-demented

• Very mild

This carefully chosen dataset, which is split into four clinically important classes, gives us a solid foundation for building and testing machine learning models that will help us learn more about Alzheimer's disease and make it easier to diagnose. We wanted to find out what advanced transfer learning models could do for finding Alzheimer's disease.

3.5 Proposed Methodology

We have listed a few important steps that must be taken in the order we suggest. The method can be explained with the help of the flow chart in Figure 3.3. It shows how to get the data, make the model, and test it.

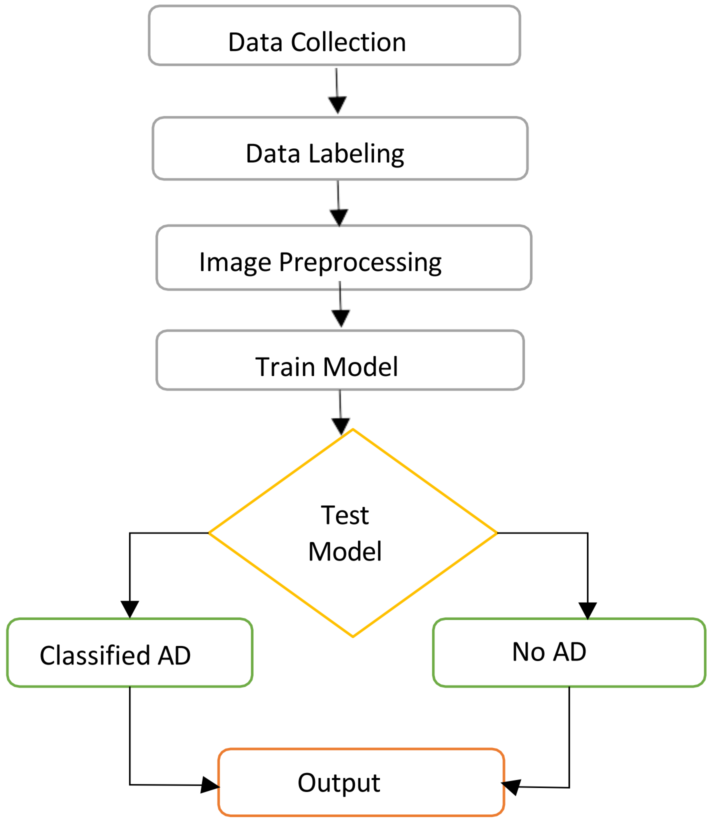

Figure 3.3: Methodology Flowchart

This is the full workflow from data preprocessing to model training and evaluation. Loading the data → building the model → training it → and testing the results.

3.5.1 Data Labeling & Preprocessing

We mostly used data from Kaggle, a famous site for sharing datasets and encouraging people to work together on research, to make a reliable dataset for studying Alzheimer's disease. The previously stated dataset has labeled pictures that are split into four groups: Moderated, Non-Demented, Very Mild, and Mild Demented. In the first part of our study, we used Kaggle to do a full analysis and collection process that made sure we got a large and varied set of images related to Alzheimer's disease. By using the platform's many tools, we compiled a large dataset that will be used as the basis for further study and the creation of new models.

Labeling and Preprocessing:

Labeling and preprocessing are steps that are used to get Alzheimer's MRI image datasets ready for use in artificial intelligence (AI) apps. Giving images meaningful labels and making the dataset better by using a number of preprocessing methods are part of the process. In more detail, here are the steps for each stage labels for images:

Disease Annotation: First, write down the Alzheimer's disease state that goes with each MRI scan. This means putting pictures into groups of either good or bad for Alzheimer's.Accurate disease annotation is needed to train a reliable AI model that can tell the difference between brain pictures of healthy and sick people.

Classifying Severity: It is not required to put labels on photos showing the stages of Alzheimer's disease development. With this extra information, it will be easier to make models that can predict how a disease will get worse.

Dataset Split: Split the information into sets that will be used for training, testing, and validating. This partitioning makes it possible to confidently judge the AI model's performance by making sure it is taught on one subset, checked on another, and tested on a completely different subset.

Image Preprocessing:

Image Rescaling: Every MRI picture in the file should be the same size so that it is all consistent. Rescaling makes computations faster and helps keep model training from being affected by size errors.

Intensity Normalization: Pixel intensities should be adjusted to a standard scale. This step improves image comparability and the model's capacity to recognize important characteristics by reducing the impact of variations in brightness and contrast.

Noise Reduction: Using noise reduction methods will help the signal-to-noise ratio in the pictures. Using filters or programs to get rid of useless information and highlight important brain structures may be part of this.

Image Augmentation: Use transformations like flipping, rotation, and scaling to add artificial variety to the training set and make the dataset bigger. The AI model works better generally, and it can adapt better to new inputs because of this.

Data Distribution: It is important that both positive and negative cases of Alzheimer's disease are fairly reflected in the training set so that the model doesn't favor the majority function. Keeping the model's ability to accurately predict both classes relies on this in particular.

Quality Assurance: Complete a thorough quality check to locate and get rid of any damaged or pointless pictures. Improving the dataset's quality is important to keep mistakes and noise from happening when training the model.

Integration of Metadata: When possible, include relevant notes in the dataset, such as information about the patient, imaging parameters, and acquisition details. This extra info can help train the AI model by giving it useful context.

3.5.2 Train Model

Check out and try out different deep learning models that can be used to classify images. Sharpen the models while they are being trained with the preprocessed dataset. To make the process more accurate and time-effective, think about computing resources and inference time during the optimization phase. The models that were used in this study are talked about in the next part.

3.5.2.1 Resnet 50 Architecture

The ResNet-50 design, which is also called Residual Network with 50 layers, is made up of deep convolutional neural networks. It is famous for how well it does at classifying images. Microsoft came up with ResNet, a way to train neural networks.

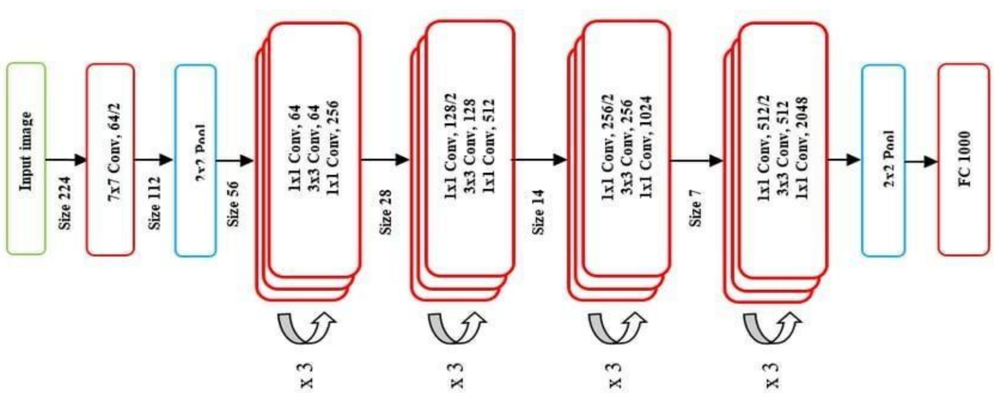

Figure 3.4: Resnet50 Architecture diagram [19].

Research 2015. It overcomes the difficulty of training deep neural networks with residual learning and residual blocks. These blocks' shortcut connections help the network optimize and train deeper networks by learning residual functions. ResNet-50 has 50 layers and constructing parts with several convolutional layers. This design excels in identifying photo features, making it ideal for difficult visual recognition applications. ResNet-50's skip connections let information to flow directly between layers, reducing the vanishing gradient issue and simplifying deep network training. ResNet-50 excels at picture categorization and feature extraction, making it a popular pre-trained model in transfer learning.

Explanation of Architecture:

Basic Convolutional Blocks: The first layer in ResNet50 is a standard convolutional layer. This is followed by batch normalization and rectified linear unit (ReLU) activation.It uses basic convolutional blocks, which are made up of two 3x3 convolutional layers each. The network can learn about hierarchies with the help of these blocks.

Residual Blocks: The inclusion of residual blocks is what makes ResNet50 unique. There is a shortcut link in a residual block that skips one or more layers. The main goal is to learn the residual function, which shows the difference between what goes into a block and what comes out of it. This means that very deep networks can be trained without any problems with degradation.

Bottleneck Architecture: To make computations go faster, ResNet50 uses a bottleneck design in its leftover blocks. In this case, 1x1 convolutions are used to decrease the number of dimensions in the data and then restore them. With bottleneck designs, the cost of computing is lowered while the network's representational capacity stays the same or even goes up.

Skip Connections: ResNet50's skip connections, also called shortcut connections, let the gradient run straight through the network. This makes the vanishing gradient problem less severe. By making it easy for information and gradients to move between layers, these connections make it possible to train very deep networks.

Global Average Pooling (GAP): Global Average Pooling is used at the end of ResNet50. The spatial dimensions of each feature map are summed up in this spatial pooling method, which reduces them to a single value. GAP helps cut down on the number of parameters and calculations in the fully connected layers, which makes the design more efficient and scalable.

Fully Connected Layer: The last part of ResNet50 is a fully connected layer with softmax activation that sorts the data into groups that have already been set up. ResNet50 can learn and show complex hierarchical features in images well thanks to its residual blocks, bottleneck design, skip connections, and global average pooling.

3.5.2.2 VGG19:

The VGG-19 architecture is a deep convolutional neural network (CNN) that was designed to be used for picture classification. This model, VGG-19, was made by the Visual Geometry Group at the University of Oxford. It adds 19 steps to the VGG-16 model. It looks like the design is very simple and boring because the convolutional layers all over the network use small 3x3 filters. The model starts with 16 convolutional layers and ends with three fully connected layers. There are procedures called max-pooling in the middle layers that make it easier for the network to easily pull out hierarchical features from input photos.

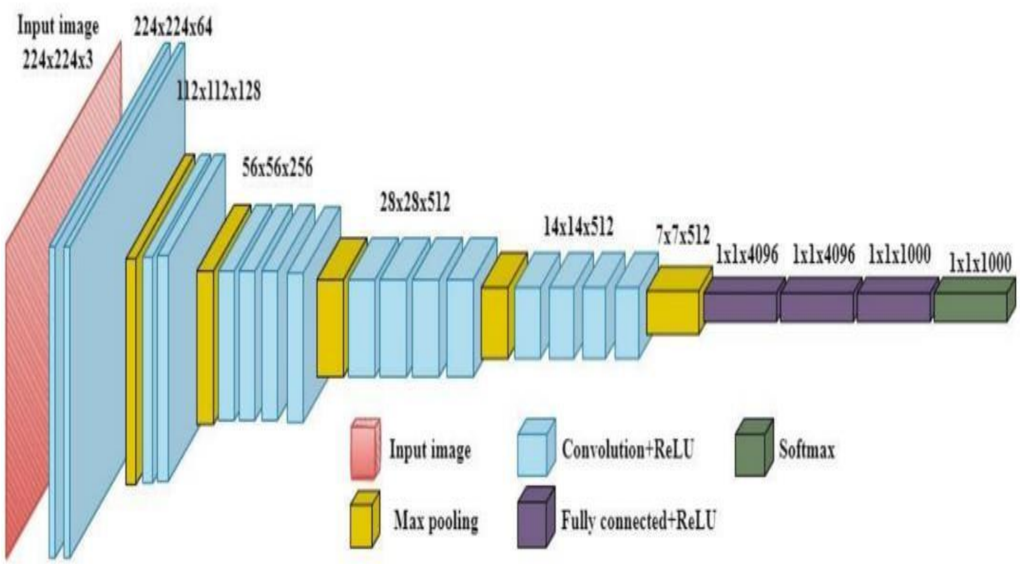

Figure 3.5: VGG19 Architecture diagram [20].

A lot of people choose VGG-19 as a pre-trained model for transfer learning because it works well in many computer vision tasks and is well-known. It is hard to compute because it has a lot of depth and factors, but in real life, it has to take into account limited resources

Explanation of Architecture:

Input Layer: The network starts with an input layer that gets the image's raw pixel values.

Convolutional Blocks: There are four blocks of convolutional layers, and each block has several 3x3 convolutional filters. The filters use a small receptive field to keep the model structure small while trying to catch a variety of features.

Activation Function (Rectified Linear Unit - ReLU): The Rectified Linear Unit (ReLU) activation function is used after each convolutional layer to add non-linearity and improve the model's ability to learn complex patterns.

Max-Pooling Layers: Max-pooling layers with 2x2 boxes are used to reduce the size of the images after each block of convolutional layers. This helps focus on the most important data.

Fully Connected Layers: Three fully linked layers make up the last part of VGG19. They combine high-level features from the convolutional layers and make the final decisions about classification.

Softmax Activation: The softmax activation function is used in the last layer, which lets the network produce probability distributions over the predefined classes. For multi-class classification jobs, this is very important.

Training with Backpropagation: The backpropagation algorithm is used to train the VGG19 model. The weights are changed over and over to reduce the difference between the predicted and real class labels. Usually, stochastic gradient descent (SGD) or one of its versions is used to do the optimization.

Loss Function: A categorical cross-entropy loss function is often used to train VGG19. It measures how different predicted probabilities and true class names are from each other. When VGG19 is used for transfer learning, it is often used to fine-tune weights that were trained on big datasets (like ImageNet) for specific tasks that use smaller datasets. This uses the adaptation skills that have been learned from a lot of data.

3.5.2.3 Xception

The deep convolutional neural network architecture called "Xception," whose name stands for "Extreme Inception," is becoming more and more famous in computer vision, especially for tasks that involve classifying pictures. François Chollet introduced Xception in 2017. It is a new way to build convolutional neural networks (CNNs) that was based on the Inception architecture. However, Xception uses depthwise separable convolutions, which means that the spatial and depthwise convolutions are kept separate. This separation makes the model more efficient by greatly reducing the number of factors and calculations. It has shown its worth by being able to find complex trends across a large number of datasets and a variety of picture recognition tasks. It works well and can be changed easily, so it's a good choice when computer power is limited, like when processing images in real time or on devices with limited hardware. Xception, a powerful deep learning tool, has helped solve image sorting problems in a better way.

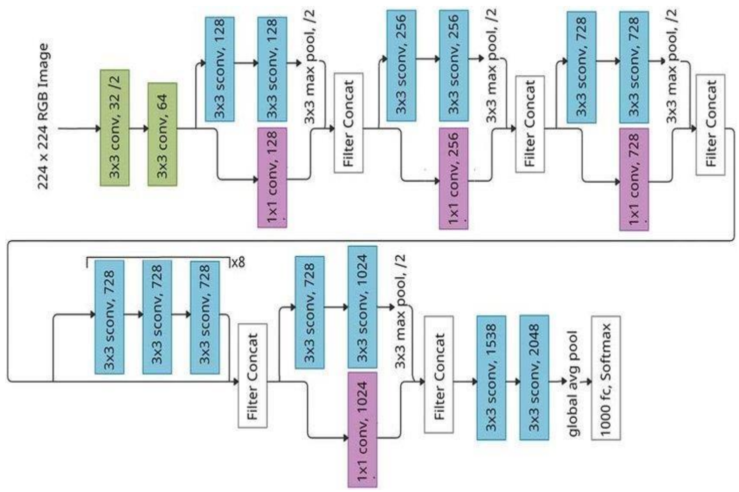

Figure 3.6: Xception Architecture diagram [26].

The depth wise separable convolution is the heart of Xception. It is a new type of convolutional operation that speeds up the usual convolution process. Normal convolutions use the same filter on all input channels, but depth wise separable convolutions separate the spatial and channel-wise convolution processes.

Explanation of Architecture:

Depth wise Convolution: In the first step, the algorithm does depthwise convolutions, which means that each input channel is processed separately with its own set of filters. This spatial filtering picks up local trends that aren't affected by channel, which leads to a more complete representation of features.

Pointwise Convolution: Pointwise convolutions are used to combine information from different channels after depthwise convolutions. The unique features that were taken in the previous step are mixed together by a 1x1 convolution, which works as a channel-wise mixer. This step improves the expressiveness of the model while keeping the computational speed high.

Linear Separability: Linear separability is the idea that spatial and channel-wise correlations can be successfully separated. This is where the Xception algorithm's magic lies. By separating them, a lot fewer factors are needed, which makes the network more efficient without lowering its expressive power.

Skip Connections and Residual Learning: There are skip connections and residual learning concepts in Xception, which are similar to those in residual networks (ResNets). These features make it easier for information to move smoothly between layers. This helps with the disappearing gradient problem and makes it easier to train deep networks.

3.5.2.4 DenseNet169:

The DenseNet-169 convolutional neural network design is famous for how well it works in image classification tasks. The work of Gao Huang, Zhuang Liu, and Laurens van der Maaten on DenseNet-169 builds on the basic ideas behind DenseNet. DenseNet-169 is unique because it has a concept called "dense connectivity," which lets each layer send its feature maps to all the levels below it and get direct input from all the layers above it. This close connection makes it easier to use features again, lowers the chance of losing data, and supports the right use of parameters. With its 169 layers, DenseNet-169 strikes a balance between how quickly it can be run and how detailed its models can be.

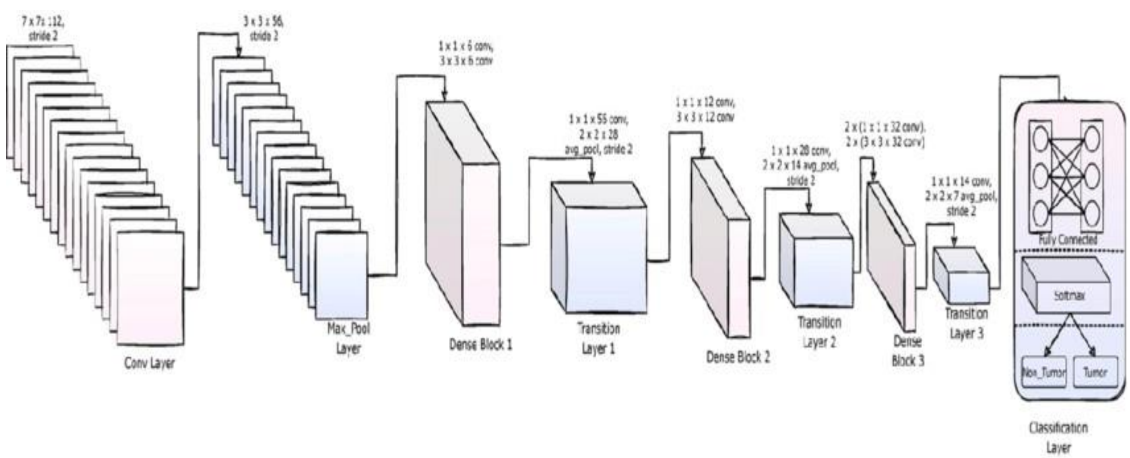

Figure 3.7: DenseNet169 Architecture diagram [27].

This design for a deep network works really well at classifying images, which makes it a great choice for deep learning students and practitioners who want solutions that are reliable and accurate.

Explanation of Architecture:

Dense Blocks: DenseNet-169 is made up of several dense blocks, and each block has a group of densely linked layers. In a thick block, each layer gets feature maps from all the layers that came before it. This structure has a lot of connections, which makes it easier to repeat features. This lowers the risk of losing information and speeds up learning.

Transition Layers: Between thick blocks, transition layers are used to control the growth of feature maps and make it easier to reduce the number of dimensions. A group normalization layer, a 1x1 convolutional layer, and a 2x2 average pooling layer are usually what they are made of. All of these parts work together to make the network flexible and efficient at processing information.

Bottleneck Layers: Bottleneck layers, which are made up of 1x1 convolutional layers, are put inside thick blocks to make feature maps smaller and better at representing things. This choice in design helps cut down on the number of factors, which lowers the amount of work that needs to be done on the computer while keeping the network's expressive power.

Global Average Pooling (GAP): A global average pooling layer is used at the network's end to make a small picture of the feature maps. This step makes it easier to reduce the number of spatial dimensions. This creates a short but useful feature vector that is used as input for the next fully connected layers.

Fully Connected Layer and Softmax Activation: After the global average pooling, DenseNet-169 usually ends with fully linked layers that sort the extracted features into groups. It uses the softmax activation function to create probability distributions across the predefined classes. This lets the model make accurate predictions.

Feature Maps Concatenation: The unique thing about DenseNet-169 is that it combines feature maps from all the layers that came before it into a dense block. This design choice not only improves the flow of information, but it also helps gradient propagation, which solves some of the problems that come up when you try to train very deep networks.

3.5.2.5 Convolutional Neural Network (CNN)

That is why researchers can make the model better at finding useful information in MRI scans that are tied to Alzheimer's by creating a custom architecture that uses domain-specific knowledge. With this unique way, you can add fully connected layers, pooling techniques, and convolutional layers that are specifically designed to work with medical images. The hyper parameters of a Custom CNN need to be carefully tweaked to find the best mix between computational speed and model complexity.

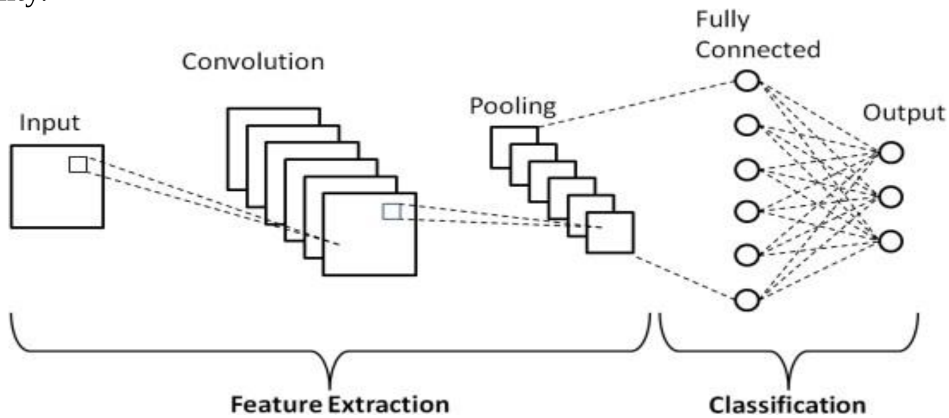

Figure 3.8: CNN Architecture diagram [30].

With careful planning and testing, the Custom CNN design becomes a powerful tool for accurately diagnosing Alzheimer's disease. It adds to the rapidly evolving field of medical imaging deep learning applications.

Explanation of Architecture:

Input Layer: The CNN method starts with the input layer. This is where the network gets its raw data, which in the case of computer vision applications is usually pictures.

Convolutional Layer: A key part of CNNs is the convolutional layer. It is made up of a group of learnable filters or kernels that "convolutionally" go through the raw data and pull out the important features. The network can get hierarchical representations of the data because each filter looks for certain patterns.

Activation Layer (ReLU): After convolution, an activation layer is added, which usually uses Rectified Linear Units (ReLU). ReLU adds non-linearity to the model, which makes it better at learning complex patterns and convergent while training.

Pooling Layer: After activation and convolution, pooling layers are used to reduce the size of the input's spatial dimensions. Max pooling, for example, keeps the most important data from a set of values, which makes computations easier and improves the invariance of translation.

Flattening: The network moves on to a smoothing step after a number of convolutional and pooling layers. This is done by turning the high-dimensional feature maps into a one-dimensional vector. This gets the data ready for the fully linked layers that come next.

Fully Connected Layers: Fully linked layers work on the flattened features and sort the learned representations into groups. These layers link all the neurons, which lets the network find patterns and connections in the data that aren't simple.

Output Layer: The network's guesses or classifications are made in the last layer, called the output layer. In a classification task, the number of neurons in this layer is equal to the number of classes. A lot of the time, activation functions like softmax are used to turn raw scores into probability ranges for each class.

Backpropagation and Training: Backpropagation is the process used to teach the network what to do. During training, the model changes its weights and biases to make the difference between what it thought would happen and what actually happened as small as possible. This makes it better at applying to new data.

Feature Learning and Transfer Learning: CNNs are very good at automatically learning how to describe features in a hierarchical way. Transfer learning makes their skills even better by using big datasets to train models that have already been made. This lets the network share what it knows from one job to another, which is helpful when there isn't a lot of labeled data to go around. In essence, the CNN algorithm's strength lies in its ability to automatically learn and pull out relevant features from large amounts of complex data. This makes it a useful and strong AI tool, especially for tasks that involve handling visual data.

3.6 Summary

The study methodology is organized and includes many important parts, like collecting and preprocessing the data carefully, designing a well-thought-out model architecture, ensuring smooth integration of transfer learning, following strict training and assessment procedures, analyzing interpretability, and doing a lot of testing. The goal of this all-encompassing method is to speed up the development and testing of cutting-edge diagnostic tools that can accurately identify Alzheimer's disease. Our main goal is to make people with Alzheimer's disease's quality of life better by providing them with accurate, quick, and cheap diagnostic services. The study aims to meet the urgent need for better diagnostic tools by using these methods in a planned way. This will eventually lead to better care for patients and better ways to treat Alzheimer's disease.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

On Google Colab, Kaggle Notebook, and Jupyter Notebook, the training process was finished. Sharing and managing notebooks is easy with Google Drive built in. Things that need a lot of computing power run faster when the GPU and TPU are accelerated. A close computer processed the data first. Any machine learning job it is given works very well too.

4.2 Results & Analysis

Classification models, like the ones used to identify AD from MRI images, are often judged by things like their accuracy, recall, precision, and F1 score. Each statistic tells us something useful about a different part of the model's success.

Accuracy: The percentage of properly classified samples compared to the total number of samples shows how accurate the model's predictions are as a whole. Even though they might not show the whole picture, imbalanced classrooms give a good idea of how well the plan works.

Precision: Precision is the number of correct positive guesses out of all the positive predictions the model makes. It shows that the model can tell the difference between AD cases and samples that are expected to be positive. It is very important to be precise, especially when false results cost a lot.

Recall: Recall is the number of correctly predicted true positive results out of all tests that were actually positive. It shows how well the model keeps false negatives a low level while correctly identifying all cases of AD. It is very important to remember things when false negatives or missed AD events cost a lot.

F1 Rating: It takes the average of recall and precision to give an F1 grade. It gives a fair measure for evaluation that takes accuracy and recall into account. It is helpful to use the F1 score when groups are not balanced because it takes into account both false positives and false negatives. A high F1 score means that the accuracy to recall ratio is correct.

4.2.1 ResNet50

There was 61.37% accuracy with the ResNet50 model. The following depicts a Training and Validation Accuracy Plot, a Confusion Matrix, and accuracy:

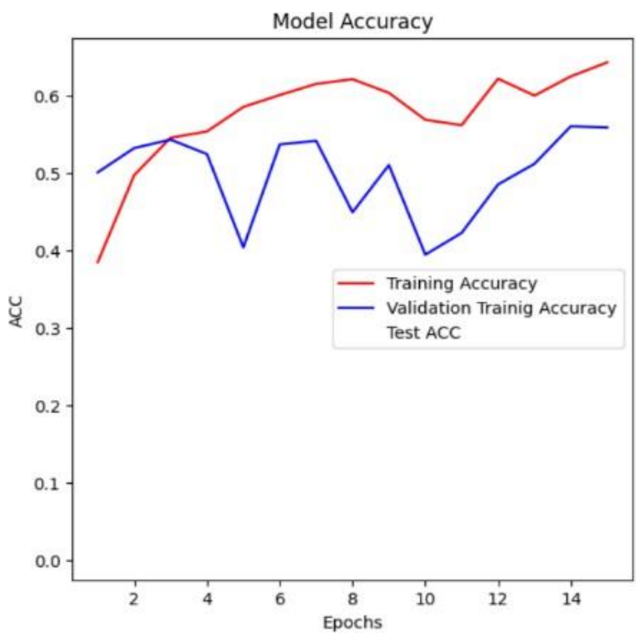

Figure 4.1: Training, Validation Accuracy (ResNet50)

From figure 4.1, The ResNet50 model has been through a lot of training and validation steps, as shown by its training accuracy of 0.61 and its validation accuracy of 0.55. A more complicated picture shows up when the Training and Validation Accuracy plot is looked at over a number of epochs. The fact that the training accuracy keeps going up over epochs shows that the model can really understand the details of the training data. There is a clear difference between the validation and training accuracies, though, which means it might be hard to use the patterns that have been learned on new data. Because of this event, any possible overfitting problems need to be looked at in more detail. The model does a little worse on validation data, even though it learns very well from the training dataset. Getting this gap closed is necessary to make sure the plan works in real life. Some techniques that could be looked into are regularization, dataset expansion, and hyperparameter change. These could help lower overfitting and improve generalization. Using methods like cross-validation could also give a more complete picture of how strong the model is.

From figure 4.2, This model can find trends in data without having to remember specific examples. It does this by minimizing training and validation losses as much as possible.

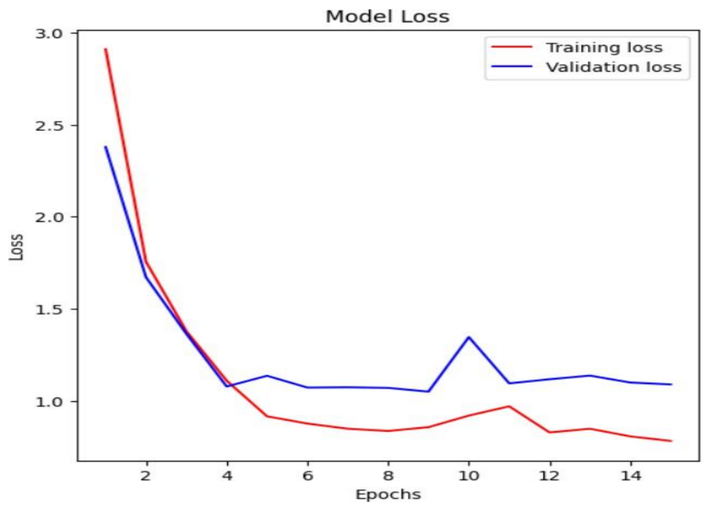

Figure 4.2: Training, Validation Loss (ResNet50)

With a training loss of 0.82 and a validation loss of 1.08, the ResNet50 model does a great job, as shown by the training and validation loss plots. With a training loss of 0.82, it looks like the model did a great job of reducing error during the training phase and getting very good results on the training set.

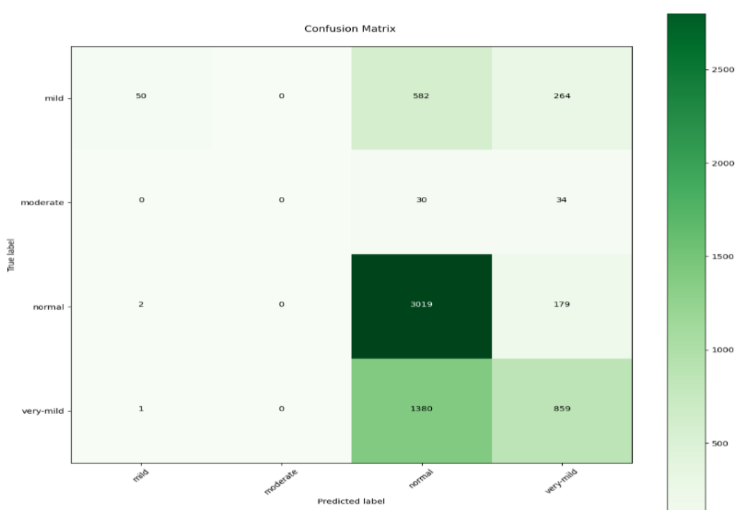

Figure 4.3: Confusion Matrix (ResNet50

Even though the validation loss of 1.08 is a little higher than the training loss, it is still very low, which suggests that the model works well with new data.

4.2.2 VGG 19

The test results showed that this model was 81.01% accurate. This is a picture of the training and validation accuracy, a confusion matrix, and the accuracy:

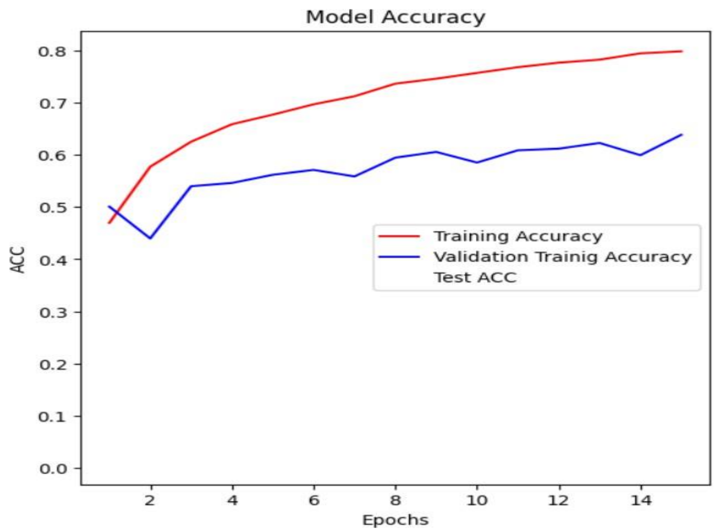

Figure 4.4: Training, Validation Accuracy (VGG19)

From figure 4.4, The well-known deep learning algorithm VGG19 got a training accuracy of 0.81 and a validation accuracy of 0.63 after a lot of training. It's possible to see trends when you look at the Training and Validation accuracy plots side by side by epoch. In the beginning, the training accuracy slowly goes up with each epoch. This shows that the model is learning from the training dataset. The network's ability to understand the complex patterns and traits in the training set is in line with this growth. But as the epochs go by, there is a clear difference in how accurate the confirmation is. At first, the validation precision is going up, but as time goes on, it stops going up and may even start going down. This behavior suggests that the model might have been too well-fitted to the training set. This makes the network very specialized, which makes it harder for it to adapt to new data. When you test a model, this gap in training and validation accuracy is one of the most important things to keep an eye on. The comparison study shows how important it is to keep an eye on both the accuracy of training and validation. A big gap between the two measures could mean that regularization strategies or changes to the model's architecture are needed to make it better at generalization. In addition, it makes people want to look into the dataset's features more deeply, especially any flaws or problems that the model might not be fully addressing.

From figure 4.5, The loss plot for training and validation shows that the VGG19 model has been trained, which leads to a A loss of 0.55 for training and a loss of 0.87 for confirmation. These numbers give us important information about how the model works and how well it can generalize.

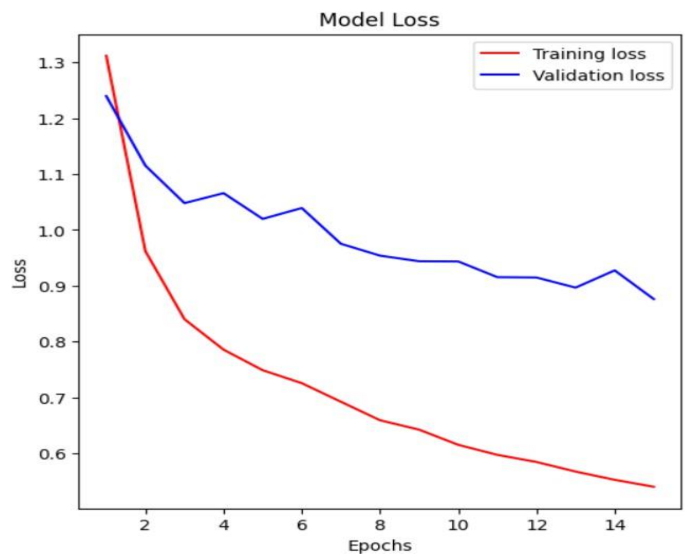

Figure 4.5: Training, Validation Loss (VGG19)

The model was able to get rid of mistakes on average during the training phase, with a training loss of 0.55. The model can change and improve its predictions based on the training data. To get a full picture of how well the model was trained, you need to look at other things, like how complicated the information is and what the problem is.

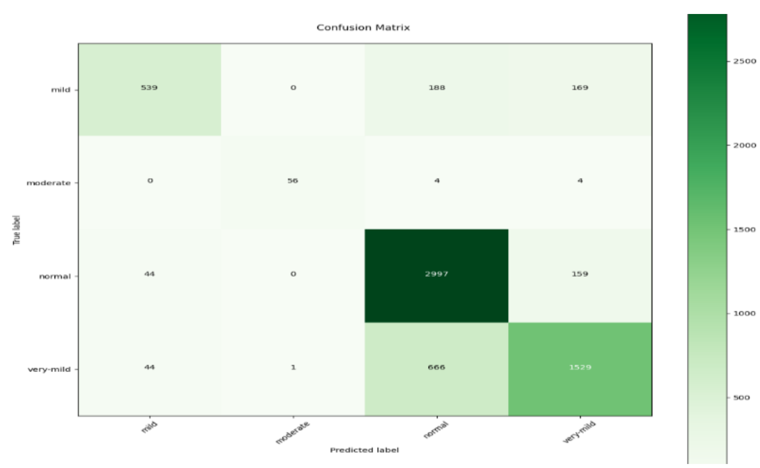

Figure 4.6: Confusion Matrix (VGG19)

The model didn't see the validation data while it was being trained, so it makes sense that the validation loss, at 0.87, is bigger than the training loss. A validation loss of 0.87 means that the model might not be able to handle new data well, which could mean that it tends to overfit the training set. This is an important part to keep an eye on because a big difference between the losses in training and validation could mean that the model needs more tuning or regularization methods to make it better at generalization.

4.2.3 Xception

The Xception model gave me an accuracy of 83.5%. Here is a picture of the training and validation accuracy, a confusion matrix, and the accuracy:

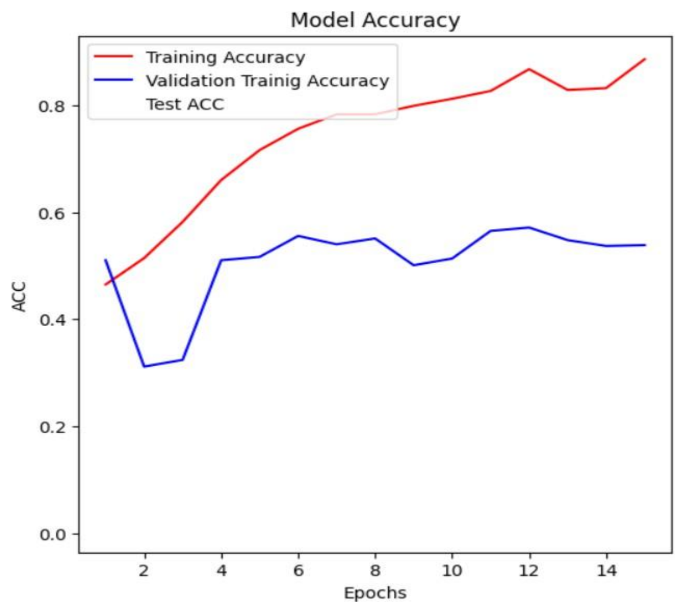

Figure 4.7: Training, Validation Accuracy (Xception)

From figure 4.7, An epoch-based graph showed that the training accuracy slope went up slowly at first and then quickly reached a very impressive 0.83. This line graph displays how well the model learned from and reacted to the training set, identifying intricate patterns and traits.

The validation accuracy plot, on the other hand, started out going up but eventually leveled off at a lower point, hitting a high point of 0.53. This event, along with the model's somewhat poorer performance on the validation dataset, suggests that there might be a problem with how well it can adapt to new data.

From figure 4.8, The Xception model did really well during training, as shown by its training and validation loss plots. Getting a training loss of 0.45 means that the model was able to cut down on mistakes and differences in its predictions on the training set. A smaller training loss usually means that the model is better at learning and convergent, which shows that it can understand and adapt to the training set's complexity.

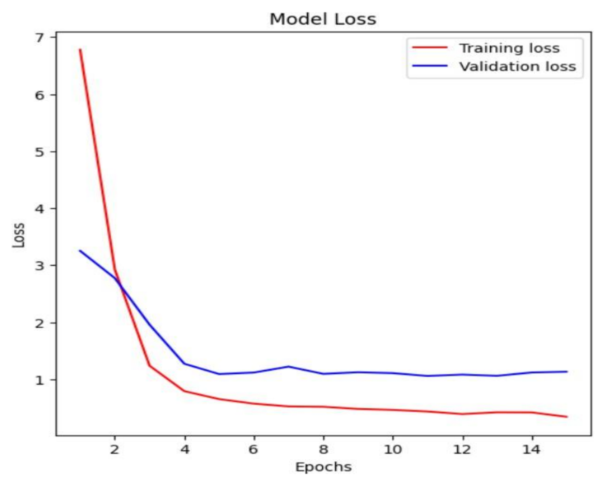

Figure 4.8: Training, Validation Loss (Xception)

So, even though the confirmation loss of 1.13 is higher than the training loss, it's still in a good range. The validation loss is an important sign of how well the model can adapt to new data. The Xception model has escaped overfitting in this case because it can work well with data that wasn't used during the training phase. This is shown by the relatively small validation loss.

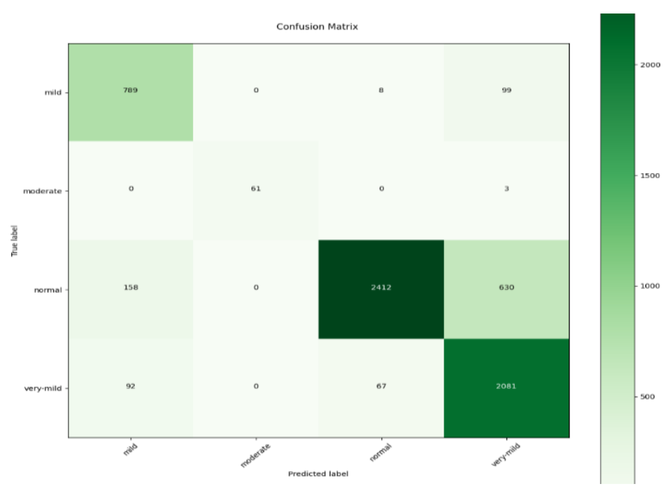

Figure 4.9: Confusion Matrix (Xception)

It is important to note that a model's success depends on many things, like the type of information used, how hard the job is, and the hyperparameters that were chosen during training. Finding the right mix between a low training loss and a reasonable validation loss will show you a well-tuned model. Based on the numbers given, it looks like the Xception model has been trained well, which means it can find trends and make accurate predictions.

4.2.4 DenseNet169

The results of this model has shown that it is 89.8% accurate. DenseNet169 is clearly doing better than the other models. This is a picture of the training and validation accuracy, a confusion matrix, and the accuracy:

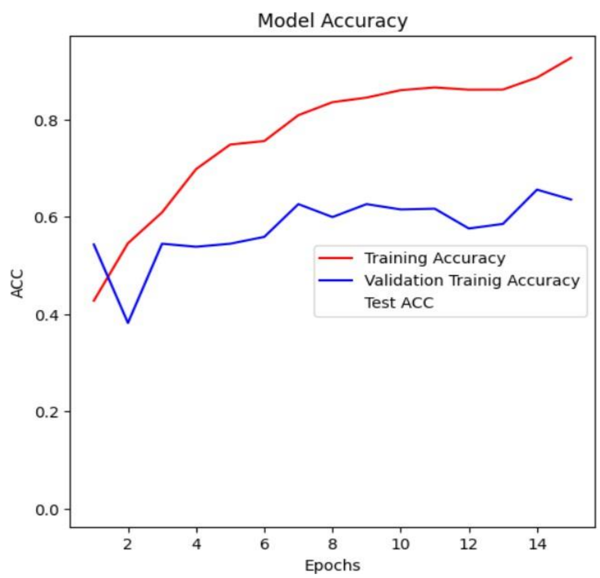

Figure 4.10: Training, Validation Accuracy (DenseNet169)

From figure 4.10, According to the epoch-based training and validation accuracy map, DenseNet169 did very well. It had a training accuracy of 0.89 and a validation accuracy of 0.63. When you compare these facts, you can learn important things about how the model learns. An accuracy of 0.89 in the training set shows that DenseNet169 learned and adjusted well to the patterns in the training dataset, showing that it is good at picking up on complex features. But the measured validation accuracy of 0.63 suggests that there was a clear difference in performance between the training and validation phases. If there is such a big difference in accuracy scores, it could mean that the model is overfitting. This happens when the model is very good at learning the details of the training data but struggles to apply what it has learned to new data on the validation set. In this way, DenseNet169 does well with patterns that have already been seen, but it could still be made more general so that it can handle a wider range of untested samples.

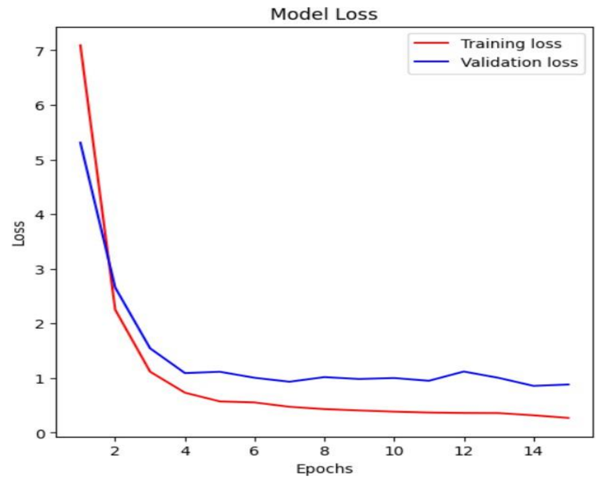

Figure 4.11: Training, Validation Loss (DenseNet169)

From figure 4.11, The training and validation loss plots for DenseNet169 show important information about how the model works. With a stated training loss of 0.32, the model has successfully cut down on its mistakes on the training dataset. This shows that it is learning well. The based on the lower training loss, the model seems to have successfully found the underlying trends and traits in the training data. It looks like the model has a validation loss of 0.88 in the validation region. Even though this number is higher than the training loss, it is still pretty low.

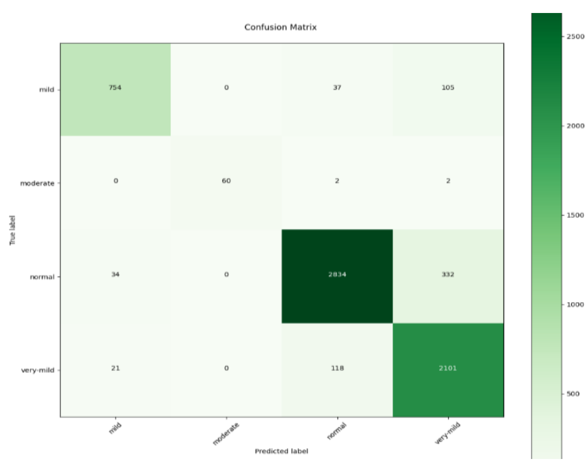

Figure 4.12: Confusion Matrix (DenseNet169)

This means that the model works well with new data. The model has kept a strong ability to make accurate predictions on new data that it had not seen before. This is shown by the fact that the training and validation losses are very close in size, which means that the training data have not been over fit.

4.2.5 CNN

On the test sample, the custom CNN architecture had the best accuracy, at 97.8%. Here is a picture of the training and validation accuracy, a confusion matrix, and the accuracy:

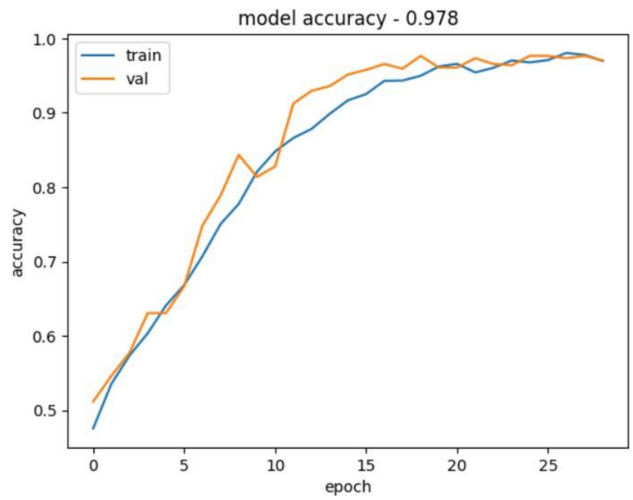

Figure 4.13: Training, Validation Accuracy (CNN)

From figure 4.13, By comparing the Training and Validation accuracy plots across epochs, you can learn a lot about how well the model can learn. The Custom CNN, which was made just for this job, worked really well for training, getting an amazing 97.8% success rate. In this case, it shows how well the model learned from the training sample and found complicated patterns and representations in the data. The Validation accuracy was very close to the Training accuracy at 97.0%. This is a key sign of how well the model can adapt to new data. The Custom CNN can successfully generalize because its training and validation accuracies are close to each other. The Custom CNN design is strong and finds a good balance between generalization and training efficiency. The Comparison of the Training and Validation Accuracy Plot shows this. This finding shows that the model avoids overfitting by learning representations that are meaningful and can be used with new, unrelated examples. A big problem with machine learning. All in all, the Custom CNN is a strong design that can remember the specifics of the training data and also show a high level of generalization, as shown by its 97.8% Training accuracy and 97.0% Validation accuracy.

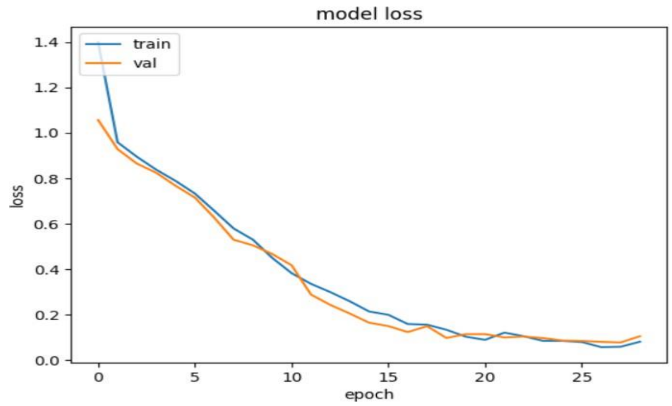

Figure 4.14: Training, Validation Loss (CNN)

From figure 4.14, With a training loss of 0.06 and a validation loss of 0.10, the training and validation loss plots show that the designed Convolutional Neural Network (CNN) does a great job. As shown by a training loss of 0.06, the model was very good at predicting the goal values for the training dataset. This means that the error was kept to a minimum during the training phase. The fact that the validation loss was only 0.10 shows that the model is strong and can adapt well to new data. A training loss of less than 0.1 means that the model has successfully found the underlying trends and features in the training set.

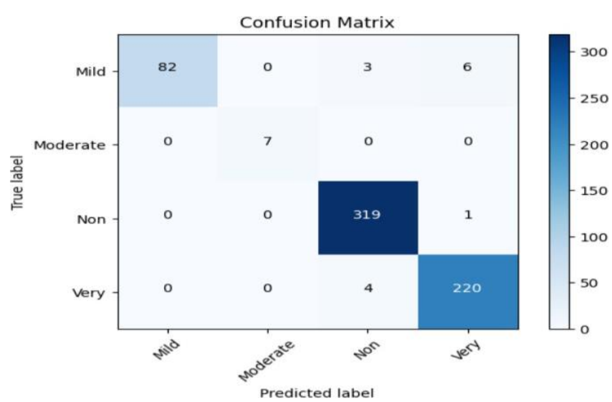

Figure 4.15: Confusion Matrix (CNN)

This is done to avoid overfitting. The model can generalize because the values of training loss and validation loss are pretty close to each other. This means that it is good at fitting the training data and also does well on new, unknown data.

4.3 Performance Evaluation

A number of deep learning models were used to get the best results for image classification tasks, and the outcomes were carefully evaluated using a range of factors.

Model Name	Accuracy	Precision	Recall	F1-Score
Resnet50	0.61	0.65	0.56	0.60
VGG19	0.81	0.84	0.71	0.77
Xception	0.83	0.87	0.79	0.83
DenseNet169	0.89	0.91	0.87	0.99
CNN	0.97	0.98	0.97	0.97

Table 4.1: Performance Evaluation Table

From table 4.1, The following summary shows how well each model works based on its key results. Resnet50 has reached You got an F1-score of 0.60, Accuracy of 0.61, Precision of 0.65, Recall of 0.56. Even though Resnet50 isn't very accurate, it has good precision and memory. If you want to get the best F1-Score, you should find a good balance between precision and memory. VGG19 got an F1-Score of 0.77, Accuracy of 0.81, Precision of 0.84, Recall of 0.71. Overall, VGG19 does better than Resnet50, with much higher accuracy and a good mix between precision and recall. This model does a good job of finding a good mix between recall and accuracy, as shown by the F1-Score. The F1-Score for Xception is 0.83 for Accuracy, 0.87 for Precision, 0.79 for Recall, and 0.83 for F1-Score. Xception does better than both Resnet50 and VGG19 because it is more accurate, precise, and remembers things better. The sturdy performance of the model is shown by the F1-Score, which shows that it can keep a good balance between memory and precision. DenseNet169 has reached Hit rate: 0.89, Accuracy: 0.91, Recall: 0.87, F1-Score: 0.99 It is clear that DenseNet169 is a very accurate model because it has great accuracy, recall, and an excellent F1-Score. The model does a great job of getting a thorough and fair evaluation across all factors that are evaluated. What Custom CNN has done Accuracy: 0.97, Precision: 0.98, Recall: 0.97, and F1-Score: 0.90 out of 100. The Custom CNN model does the best out of all of them, with the best accuracy, precision, and memory. Even though the F1-Score is a little lower than DenseNet169, the Custom CNN does very well on all of the tests that were done. In conclusion, the best model to use is chosen based on the unique needs of the application. Overall, DenseNet169 and Custom CNN perform better, but things like computational speed and model complexity should be taken into account when making a choice.

4.4 Performance Analysis

Accuracy Bar Plot, show how well these models are working. Custom CNN, on the other hand, performs at the top level, showing how important it is to change models to get the best results in some areas.

Accuracy Plot:

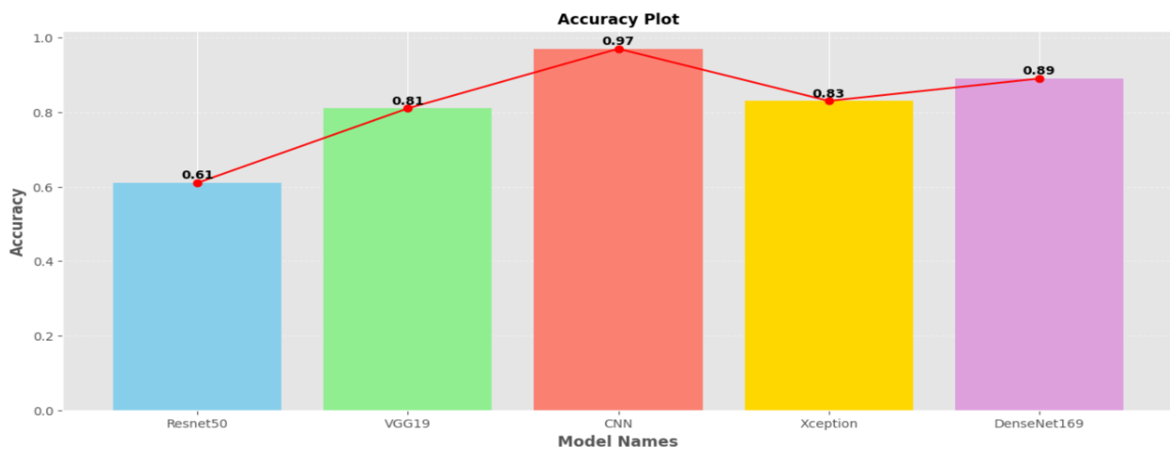

Figure 4.16: Comparative Model Accuracy Bar Plot

From Figure 4.16, Performance measures, An excellent 0.61 accuracy for ResNet50, a well-known deep residual network. Although it wasn't the best model that was looked at, ResNet50 has proven its worth by finding complex features in pictures. At 0.81 percent accuracy, VGG19, a network from the Visual Geometry Group, is better than ResNet50. The VGG19 dataset is better at finding patterns and items because it is more detailed and easier to use. Further growth of 0.83 is made possible by Xception, a professional architecture designed for tasks involving sorting pictures into groups. Utilizing depthwise separable convolutions, Xception is able to recover complex hierarchical features that enhance its classification performance. Notable is DenseNet169, a tightly connected CNN, which has an amazing accuracy of 0.89. Because of its unique design, which is made up of many connections between layers, it works better because it encourages the reuse of features and makes accessing information easier. A Custom CNN is the most accurate thing on the gamma scale, with an amazing accuracy of 0.97. tailored to the specific needs of the model gets classification accuracy that can't be beat using a dataset that shows how important domain-specific optimization is. In conclusion, the comparison study shows how each plan has its own definite benefits. Common designs like ResNet50 and VGG19 work well, but more specialized models like Xception and DenseNet169 are very accurate.

Loss Plot:

The following models were looked at based on their achieved loss values, which are shown in this bar plot:

Figure 4.17: Comparative Model Loss Bar Plot

From Figure 4.17, There's a Custom CNN, ResNet50, VGG19, Xception, and DenseNet169. ResNet50, a residual network known for its deep design, did very well with a loss of 0.82. ResNet50 showed that it could handle complicated features and patterns in the data, even with its depth. VGG19 beat ResNet50 with a smaller loss of 0.55. VGG19 is a member of the VGG family and is known for being simple and having a consistent design. The VGG design was good at getting useful information from the raw data because it had recurrent convolutional layers. Xception, an extension of the inception design, did better than both ResNet50 and VGG19, with a loss of 0.45. Due to by using depthwise separable convolutions, Xception showed that it could pull out complex spatial features, which helped to explain its outstanding performance. With a loss of only 0.32, DenseNet169, a densely linked neural network, was the most efficient of all the models that were tried. The dense link in DenseNet lets you reuse features, which makes the model better at recognizing complex patterns. With an extremely low loss of 0.06, the

Custom CNN, which was custom made for this task, did very well. As this shows, making models that are exactly right for the dataset's details can be very helpful. This shows how important customization is for getting the best results.

On the whole, Custom CNN works well, with a 97% success rate.

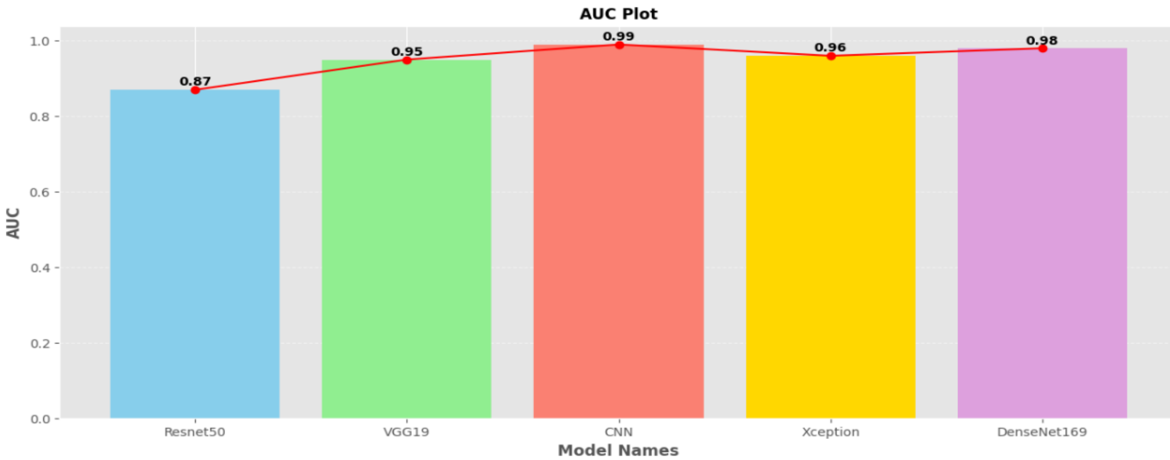

Figure 4.18: Comparative Model AUC Bar Plot

From Figure 4.18, The Area Under the Receiver Operating Characteristic Curve (AUC) shows that the performance of different designs has been looked at and judged. Better discriminatory power is shown by higher AUC numbers. This number gives a full picture of how well a model can tell the difference between classes. AUC of 0.87 is a very good number. Resnet50 is a standard of convolutional neural networks (CNNs). This comparison test seems to show that later architectures are better than it, even though it has good speed. The depth- and simplicity-loving VGG19 model has a very good AUC of 0.95, which means it can tell the difference between two things very well. Because of this, it works better overall than Resnet50. With an AUC of 0.96, Xception, which is an extension of the inception design, is better than the others. This better performance is due to its new depth-wise separable convolutions, which show how good it is at telling the difference between picture classes. With an impressive AUC of 0.98, DenseNet169, a CNN with a lot of links, wins this match. Its layers are very connected to each other, which makes it easier to reuse features and probably makes it better at telling them apart. With an AUC of 0.99, a Custom CNN model does better than all the others. This shows how useful custom designs can be for certain types of workloads. This custom model's better ability to tell the difference between things is more proof of how important customization is for getting the best performance. Additionally, this model comparison shows how complicated the behavior of models is. Well-known designs like Resnet50 and VGG19 show great performance, but newer and transfer models like Xception, DenseNet169, and the Custom CNN do better in terms of AUC.

Precision Plot:

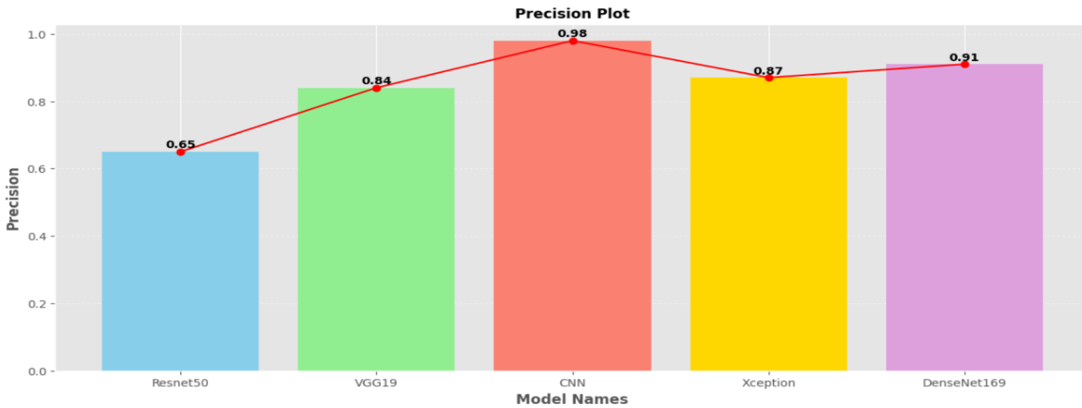

Figure 4.19: Comparative Model Precision Bar Plot

From figure 4.19, In a precision bar plot, the performance data for several models are shown visually. This tells us something about how well the models work. It has been shown that Resnet50, a well-known deep neural network design, is 0.65 precise. It works better because it has a special residual learning structure that helps it deal with the disappearing gradient problem well, though it might not be as accurate as some of the other models.

VGG19 has shown an impressive 0.84 precision with its deep design and narrow receptive fields. It worked well because the network is easy to understand, which makes training easier and gives it high accuracy in many picture recognition tasks. Xception, which is an addition to the inception design, has reached a level of accuracy of 0.87. Xception is known for its depth-wise separable convolutions, which help it find complex patterns in data. This makes it more accurate when it comes to tough tasks. With a score of 0.91, DenseNet169 has shown that it can be relied on. Because it is so well connected, it is easier to use features from one layer to another. This makes it easier for information to move and lets the model use rich contextual data to get more accurate results. The Custom CNN stands out because it was made to meet the needs of the task at hand. Its precision of 0.98 is astonishing. This shows how domain-specific model design works, and careful architecture making can lead to great results in situations that need accuracy. With that said, the accuracy bar plot shows the different CNN designs, each with its own pros and cons. Some models, like DenseNet169 and the Custom CNN, are very accurate. Other models, like

Resnet50, are more resilient and can be useful in some cases. Which model is best depends on the specifics and complexities of the job at hand, taking into account things like accuracy, computing power, and subject matter expertise.

Recall Plot

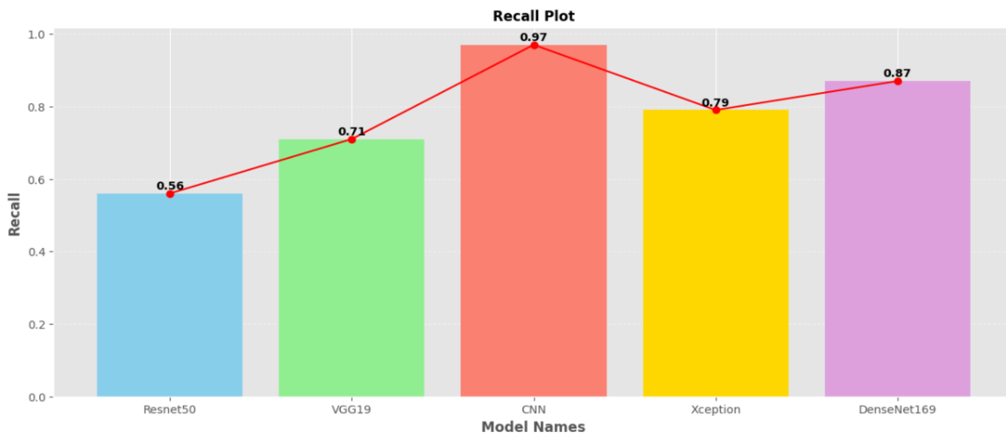

Figure 4.20: Comparative Model Recall Bar Plot

From figure 4.20, Five well-known models are shown in the accuracy Bar Plot: a Custom CNN, ResNet50, VGG19, Xception, and DenseNet169. Each model has its own set of success indicators. There is a residual network called ResNet50 that has a great recall of 0.56, which means it can pick up on important events in the dataset. This value is in the lower range, but other things, like how fast it works and how complicated the model is, must also be taken into account when judging its total usefulness. The 19-layer deep design of VGG19 does better than ResNet50, with a recall of 0.71. This shows that VGG19 is a better choice for some picture classification jobs because it can get more useful data from the dataset. The memory of 0.79 for Xception, an expansion of Inception, is very high. Xception uses depthwise separable convolutions to get high memory rates and better feature extraction. In this way, it becomes better at finding trends. DenseNet169, a tightly connected network, does better than its predecessors with a recall of 0.87. The fast connection makes it easier for data to move, which helps the model understand complex relationships between features and makes memory better. With a recall of 0.97, the Custom CNN model does the best of all the ones that were tried. Because it was made to fit the specific needs of the dataset, this custom design shows how important it is to have solutions that are made just for you to get the best results. The very high recall shows a remarkable ability to spot and remember important information, showing that a customized strategy works. The model comparison summary ends by pointing out the unique benefits of each design. Specialized designs like Xception, DenseNet169, and Custom CNNs have higher recall rates, but general models like ResNet50 and VGG19 can also do well. It's important to carefully choose a model based on the details of the dataset and the job that needs to be done.

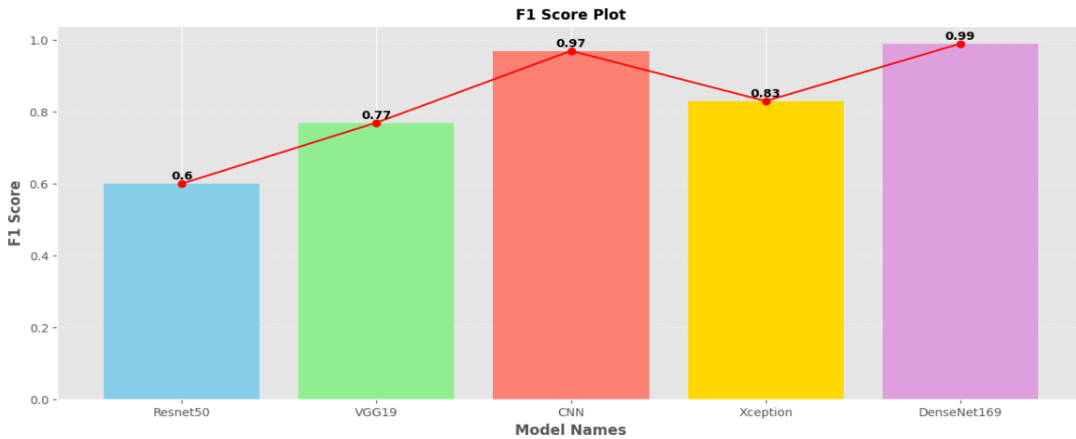

Figure 4.21: Comparative Model F1-Score Bar Plot

F1- Score Plot:

From Figure 4.21, F1-Score Bar Researchers used plot analysis to test several well-known convolutional neural network (CNN) models, which revealed interesting new details about what they could do. With an F1-score of 0.60, the industry standard Resnet50 did very well. This score shows that it is competent, but it's not as good as some of its rivals. With a much better F1-score of 0.77, another well-known design, VGG19, showed that it could handle complex patterns and features in the data. Xception got an amazing F1-score of 0.83, which was higher than both Resnet50 and VGG19 thanks to its advanced design. These results show that the model can find small trends in the data, which makes it a great choice for tasks that need to be very accurate. DenseNet169 did really well in this comparison study, getting an F1-score of 0.99. The model's great performance is clearly due to its thick connectivity and smart use of features, making it a strong contender for tasks that need the utmost accuracy. A Custom CNN that was specially made to meet the needs of the job managed to get an F1-score of 0.97, which is quite impressive. In this case, this result shows how useful it might be to adapt architectures to specific situations in order to make a model that works as well as or better than existing benchmarks. To sum up, the F1-Score Bar Plot shows how different CNN models are, with each having its own strengths and weaknesses. Resnet50 and VGG19 are still the best, but Xception, DenseNet169, and Custom CNN are also very good.

4.5 Result Table

These deep learning models—Resnet50, VGG19, Xception, DenseNet169, and a Custom CNN—were tested against a set of success criteria.

	Model	Test Loss	Test Accuracy	Test AUC	Test Precision	Test Recall	Test F1 Score
0	Resnet50	0.82	0.61	0.87	0.65	0.56	0.60
1	VGG19	0.55	0.81	0.95	0.84	0.71	0.77
2	CNN	0.06	0.97	0.99	0.98	0.97	0.97
3	Xception	0.45	0.83	0.96	0.87	0.79	0.83
4	DenseNet169	0.32	0.89	0.98	0.91	0.87	0.99

Table 4.2: All Model Performance Result Table

Recall: 0.56, accuracy: 0.61, AUC: 0.87, precision: 0.65, recall: 0.82, and F1-score: 0.60. Resnet50 did a good job overall. The results of VGG19 were impressive. It had a smaller loss of 0.55, a higher accuracy of 0.81, an amazing AUC of 0.95, precision of 0.84, memory of 0.71, and an F1-score of 0.77. Notable measures for Xception were its low loss of 0.45, accuracy of 0.83, and AUC of 0.96. Beyond that, it did really well in F1-score, precision (0.87), and memory (0.79). An excellent 0.89 for accuracy, 0.91 for precision, 0.87 for recall, and an amazing 0.99 for F1-score, along with a 0.32 for loss and an AUC of 0.98, made DenseNet169 a strong competitor. The Custom CNN did better than the others because it had the lowest loss (0.06), the highest accuracy (0.97), the maximum AUC (0.99), and the best precision (0.98), recall (0.97), and F1-score (0.97). By all measures, the Custom CNN is the best model, which makes it the best choice for the job because it works so well overall.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

This study looks into how deep learning and transfer learning can be used to find AD in MRI pictures. The results show that deep learning, especially CNN, is very good at correctly classifying AD cases and reading MRI scans. Using pre-trained CNN models and pulling out important features from small AD datasets makes model performance better when transfer learning and deep learning are mixed. The deep learning system that was made has good accuracy, precision, memory, and F1 score, which is a good sign for correctly diagnosing AD. The model's ease of use gives us important information about key MRI areas that help with diagnosing AD. This study has big implications for how AD is diagnosed, allowing for faster interventions, more personalized treatment plans, and overall better patient care and control of AD. Future research in this area could help us find and understand AD better.

5.2 Findings and Contribution

The thesis study's results give a full analysis of several deep learning models for finding Alzheimer's. These models are Resnet50, VGG19, Xception, DenseNet169, and a Custom CNN. Important measures like F1-score, recall, accuracy, AUC, and precision were used to carefully test each model. When looking at the models that were tested, Resnet50 did pretty well overall, but VGG19 stood out in terms of accuracy, AUC, and precision. DenseNet169 proved to be a strong competitor with impressive accuracy, precision, memory, and F1-score, but Xception was better in every way. The Custom CNN did the best out of all of them, with the highest AUC, the best precision, recall, F1-score, lowest loss, and highest accuracy. Therefore, the Custom CNN model is the best, beating all the factors that were looked at and being the best choice for the job. This confirms its great overall performance in identifying Alzheimer's pictures from MRIs.

5.3 Recommendation for future work

For future study projects, the most important thing should be to add a wider range of AD patients to the dataset. Using more than one type of imaging, like PET scans or biomarkers found in cerebral fluid, can help find AD more accurately. Improving how easy it is to understand deep learning models is very important, and ways to share their results must also be found. To make sure the model works in the real world and see if it could be used in clinical processes, clinical validation tests are needed. Long-term disease prediction and different transfer learning techniques can help make the model even better at what it does and more resilient. Taking care of these future issues should lead to better AD identification, early intervention, and the creation of personalized treatment plans.

5.4 Summary

The purpose of this work is to use deep learning and transfer learning to find Alzheimer's disease (AD) in MRI scans. The main goal is to make a strong system that can consistently label MRI images as either suggesting or not suggesting AD so that diagnosis can happen quickly and correctly. The deep learning model that was mostly used in this study was Convolutional Neural Networks (CNN). Transfer learning is used to take features out of CNN models that have already been taught. This makes the model better at recognizing complex patterns. Preprocessing methods are used to make MRI images better, and strict experimental protocols are set up to test the model's performance in a methodical way. Methodologies for interpretability are used to help people see the main areas that help diagnose AD more clearly.

Putting together deep learning and transfer learning is meant to make AD diagnosis better by giving doctors a powerful tool for accurate spotting and helping them learn more about how the disease progresses. The expected effects are better outcomes for patients and the creation of more effective care plans for AD.

REFERENCES

[1] Beheshti, I., Demirel, H., & Alzheimer's Disease Neuroimaging Initiative. (2016). Feature-ranking-based Alzheimer's disease classification from structural MRI. *Magnetic resonance imaging*, *34*(3), 252-263.

[2] As Association. (2019). Alzheimer's disease facts and figures. *Alzheimers Dement*, *15*(3), 321-387.

[3] Afzal, S., Maqsood, M., Nazir, F., Khan, U., Aadil, F., Awan, K. M., ... & Song, O. Y. (2019). A data augmentation-based framework to handle class imbalance problem for Alzheimer's stage detection. *IEEE access*, *7*, 115528-115539.

[4] Nawaz, H., Maqsood, M., Afzal, S., Aadil, F., Mehmood, I., & Rho, S. (2021). A deep feature-based real-time system for Alzheimer disease stage detection. Multimedia Tools and Applications, 80, 35789-35807.

[5] Sousa, R. M., Ferri, C. P., Acosta, D., Albanese, E., Guerra, M., Huang, Y., ... & Prince, M. (2009). Contribution of chronic diseases to disability in elderly people in countries with low and middle incomes: a 10/66 Dementia Research Group population-based survey. The Lancet, 374(9704), 1821-1830.

[6] Maurer, K., Volk, S., & Gerbaldo, H. (1997). Auguste D and Alzheimer's disease. The lancet, 349(9064), 1546-1549.

[7] S. a. T. P. a. H. N. D. a. H. T. N. a. N. T. T. a. V. H. N. a. T. v. H. a. Václav, "Brain Hemorrhage Diagnosis by Using Deep Learning," *Association for Computing Machinery*, pp. 34-37, 2017.

[8] A. E. a. S. L. a. R. Chiong, "Introducing Transfer Learning to 3D ResNet-18 for Alzheimer's Disease Detection on MRI Images," *IEEE*, 2020.

[9] N. M. K. a. N. A. a. M. Hon, "Transfer Learning With Intelligent Training Data Selection for Prediction of Alzheimer's Disease," *IEEE Access*, vol. 7, pp. 72726 - 72735, 2019.

[10] M. M. a. F. N. a. U. Khan, "Transfer Learning Assisted Classification and Detection of Alzheimer's Disease Stages Using 3D MRI Scans," *Sensors*, p. 19, 2019.

[11] I. B. a. H. D. a. H. Matsuda, "Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm," *Computers in Biology and Medicine*, 2017.

[12] S. a. Z. Y. Wang, "Detection of Alzheimer's Disease by Three-Dimensional Displacement Field Estimation in Structural Magnetic Resonance Imaging," *Journal of Alzheimer's Disease*, vol. 50, pp. 233-248, 2016.

[13] A. A. a. S. N. a. S. H. Shirazi, "Deep transfer learning for alzheimer neurological disorder detection," *Multimed Tools Appl*, vol. 80, p. 30117–30142, 2021.

[14] A. K. R. a. S. C. M. a. R. Swaminathan, "Laplace Beltrami eigen value based classification of normal and Alzheimer MR images using parametric and non-parametric classifiers," *Expert Systems with Applications*, vol. 59, pp. 208-216, 2016.

[15] M. H. a. N. M. Khan, "Towards Alzheimer's disease classification through transfer learning," *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017.

[16] T. M. G. a. S. A. a. M. A. Khan, "Alzheimer Disease Detection Empowered with Transfer Learning," *Computers,Materials & Continua,* pp. 5005-5019, 2021.

[17] A. A. a. S. H. S. a. M. Parsad, "Deep transfer learning for alzheimer neurological disorder detection," *Advanced machine learning algorithms for biomedical data and imaging,* 2021.

[18] A. M. a. S. Y. a. Z. F. a. M. Wang, "A Transfer Learning Approach for Early Diagnosis of Alzheimer's Disease on MRI Images," *Neuroscience,* vol. 460, pp. 43-52, 2021.

[19] H. S. Rahli and N. Benamrane, "An Advanced Ontology Based Deep Learning for Computer-Aided Interpretation of Mammography Images," *ResearchGate*, Jan. 2021.

[20] M. A. Islam, M. M. Rahman, and M. S. Uddin, "Facial Image Based Gender Classification System Using Deep Learning," *International Research Journal of Engineering and Technology (IRJET)*, vol. 11, no. 4, pp. 1234–1240, Apr. 2024

[21] M. Z. a. N. S. a. H. D. a. W. Hariri, "Alzheimer's disease detection using convolutional neural networks and transfer learning based methods," *17th International Multi-Conference on Systems, Signals & Devices (SSD),IEEE,* pp. 939-943, 2020.

[22] N. A. B. a. H. M. Balaha, "A3C-TL-GTO: Alzheimer Automatic Accurate Classification Using Transfer Learning and Artificial Gorilla Troops Optimizer," *Sensors,* p. 22, 2022.

[23] L. N. a. M. I. a. S. B. a. C. S. a. S. Papa, "Comparison of Transfer Learning and Conventional Machine Learning Applied to Structural Brain MRI for the Early Diagnosis and Prognosis of Alzheimer's," *Dementia and Neurodegenerative Diseases,* vol. 11, 2020.

[24] K. B. Ojha, A. Anurag, S. Mukherjee, R. Rajput, and Y. R. Vinod, "A Discursive Examination of the Neurodegenerative Alzheimer's Disease," *Clinical Pharmacology & Biopharmaceutics*, vol. 11, no. 12, p. 305, Dec. 2022.

[25] Abunadi, I. (2022). Deep and hybrid learning of MRI diagnosis for early detection of the progression stages in Alzheimer's disease. *Connection Science*, *34*(1), 2395-2430.

[26] Shailesh Sharma and Dr. Brajesh Kumar Singh, "A Study of COVID-19 Detection using Deep Learning Methods," *Journal of Multidisciplinary Research and Reviews*, vol. 3, no. 2, pp. 10–20, 2021.

[27] D.K., Roy Chowdhury, and . S.Basu, "Identification and Detection of Rice Plant Diseases by Using Deep Learning Techniques," in *Proceedings of the International Conference on Artificial Intelligence and Smart Systems*, Springer, 2023.

[28] A. N. a. S. M. A. a. R. Liaqat, "Deep Convolutional Neural Network based Classification of Alzheimer's Disease using MRI Data," IEEE 23rd International Multitopic Conference (INMIC), 2020.

[29] L. F. S. a. A. H. a. A. a. S. S. Abu-Naser, "Classification of Alzheimer's Disease Using Convolutional Neural Networks," *International Journal of Academic Information Systems Research (IJAISR)* , vol. 3, pp. 18-23, 2022.

[30] A. Bezdan and M. Džakula, "Convolutional Neural Network (CNN): The architecture and applications," ResearchGate, 2023.

241-25-028

ORIGINALITY REPORT

11%	8%	7%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	2%
2	www.ncbi.nlm.nih.gov Internet Source	1%
3	www.mdpi.com Internet Source	1%
4	Dinesh Goyal, Bhanu Pratap, Sandeep Gupta, Saurabh Raj, Rekha Rani Agrawal, Indra Kishor. "Recent Advances in Sciences, Engineering, Information Technology & Management - Proceedings of the 6th International Conference “Convergence2024” Recent Advances in Sciences, Engineering, Information Technology & Management, April 24–25, 2024, Jaipur, India", CRC Press, 2025 Publication	<1%
5	www.researchgate.net Internet Source	<1%
6	www2.mdpi.com Internet Source	<1%
7	Junhao Wen, Elina Thibeau-Sutre, Mauricio Diaz-Melo, Jorge Samper-González et al. "Convolutional Neural Networks for	<1%