An Efficient Deep Learning Approach To Detect Multiple Neurodegenerative Diseases Using Image Data

by

Shadman Rahman Sameen
20101341
Shahabuddin Ahmed Khan
20101340
Mir Md. Muktasif Sakib
20101343
Tushar Chowdhury
21301010

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
June 2025

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Shadman Rahman Sameen
20101341

Shahabuddin Ahmed Khan
20101340

Mir Md. Muktasif Sakib
20101343

Tushar Chowdhury
21301010

Approval

The thesis/project titled "An Efficient Deep Learning Approach To Detect Multiple Neurodegenerative Diseases Using Image Data" submitted by

1. Shadman Rahman Sameen(20101341)
2. Shahabuddin Ahmed Khan(20101340)
3. Mir Md. Muktasif Sakib(20101343)
4. Tushar Chowdhury(21301010)

Of Spring, 2025 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on June 20, 2025.

Examining Committee:

Supervisor:
(Member)

Md. Ashraful Alam, PhD
Associate Professor
Department of Computer Science and Engineering
Brac University

Thesis Coordinator:
(Coordinator)

Md. Golam Rabiul Alam, PhD
Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

In today's time early and accurate diagnosis of the neurodegenerative diseases (such as Alzheimer Disease (AD), Parkinson Disease (PD), Frontotemporal Dementia (FTD)) is essential to timely intervention and is quite laborious since the Magnetic Resonance Imaging (MRI) scans of all of them carry same set of characteristics. The deep learning pipeline proposed in this study will utilize 2D sagittal sectioning of 3D MRI columns and will be focusing on such crucial regions as the hippocampus, substantia nigra, and forehead lobe. These pictures were downsized to the identical standard of 224x224x3 and heavily augmented to promote generalization.

We benchmarked the transfer learning classification performance of a number of pre-learning CNN models. Among the standard models, MobileNetV2 exhibited superior performance on the test set (94.36%), as opposed to EfficientNetB0 (85.47%), ResNet50 (68.29%), DenseNet121 (62.65%), and VGG19 (33.33%). However, MobileNetV2 exhibited a little overfitting. To address the challenges of accuracy problem of multiclass disease detection and overfitting issue we proposed SadNetV1 that was constructed by integrating MobileNetV2, Squeeze-and-Excitation (SE) blocks and Spatial Attention mechanism to facilitate effective dimensionality reduction in capturing channel and spatial dependence.

The proposed SadNetV1 model showed improved performance of 96.15% test, 96.84% train, and 97.11% validation accuracy and can offer generalization in complex MRI images and could differentiate between AD and PD, and FTD. These results give credence to the potential of attention-augmented lightweight networks in effective categorization of neurodegenerative diseases in clinics.

Keywords: Deep Learning, Computer Vision,Neurodegenerative disease, Alzheimer, Parkinson, Fronto Temporal Dementia, MRI, Attention Mechanism, MobileNetV2, SadNetV1, Medical Imaging.

Acknowledgement

First of all, we would like to thank Almighty Allah with all our heart and gratitude for the blessings and directions throughout our whole thesis journey. We would like to heartily acknowledge the support of our supervisor, Md. Ashraful Alam who has made this milestone possible through his constant support, professional guidance and valuable piece of advice. We also would wish to thank our parents so much as we know their prayers and encouragement equipped us with strength to survive. Last but not least we recognize the combined efforts of every author of this paper who have struggled and finished this research paper with utmost effort.

Table of Contents

Declaration i

Approval ii

Ethics Statement iii

Abstract iii

Dedication iv

Acknowledgment iv

Table of Contents v

List of Figures viii

List of Tables x

Nomenclature xi

1 Introduction 1
- 1.1 Background Information . . . 1
- 1.2 Problem Statement . . . 3
- 1.3 Research Objective . . . 3

2 Related Work 5
- 2.1 Alzheimer . . . 5
- 2.2 Parkinson's . . . 9
- 2.3 Frontotemporal Dementia (FTD) . . . 12

3 Working Plan 16

4 Dataset 18
- 4.1 Dataset Sources . . . 18
- 4.2 Data Overview . . . 18
- 4.3 Pre-Processing pipeline . . . 20
  - 4.3.1 Slice Extraction . . . 20
  - 4.3.2 Filtering and inspection of Data quality . . . 20
  - 4.3.3 Pixel Intensity Normalization . . . 21
  - 4.3.4 Gamma Correction . . . 21

4.3.5 Contrast Enhancement with CLAHE . . . 22
4.3.6 Resizing . . . 22
4.3.7 Imbalance Handling . . . 22
4.3.8 Sample for Explainability . . . 23
4.3.9 Data Augmentation . . . 24

5 Pre-trained Models 26
5.1 VGG19 . . . 26
5.1.1 Architecture Overview . . . 26
5.1.2 Application in Neurodegenerative Disease Detection . . . 27
5.1.3 Insights from Pretraining Phase . . . 27
5.2 ResNet50 . . . 30
5.2.1 Architecture Overview . . . 30
5.2.2 Application in Neurodegenerative Disease Detection . . . 30
5.2.3 Insights from Pretraining Phase . . . 31
5.3 EfficientNetB0 . . . 32
5.3.1 Architecture Overview . . . 32
5.3.2 Application in Neurodegenerative Disease Detection . . . 34
5.3.3 Insights from Pretraining Phase . . . 34
5.4 DenseNet . . . 36
5.4.1 Architecture Overview . . . 36
5.4.2 Application in Neurodegenerative Disease Detection . . . 37
5.4.3 Insights from Pretraining Phase . . . 37
5.5 MobileNetV2 . . . 39
5.5.1 Architecture Overview . . . 39
5.5.2 Application in Neurodegenerative Disease Detection . . . 40
5.5.3 Insights from Pretraining Phase . . . 40

6 Proposed Methodology 43
6.1 Overview . . . 43
6.2 Model Architecture . . . 43
6.3 Data Preprocessing and Augmentation . . . 46
6.4 Squeeze-and-Excitation Block (SEBlock) . . . 46
6.5 Importance of Channel Attention . . . 47
6.6 Spatial Attention Mechanism . . . 47
6.6.1 Average and Max pooling . . . 47
6.6.2 Concatenation and Convolution . . . 48
6.6.3 Recalibration . . . 48
6.7 Dropout for Regularization . . . 48
6.8 Learning Rate and Optimization Strategy . . . 49
6.9 Early Stopping . . . 49
6.10 Loss Function . . . 49
6.11 Explainable AI (XAI) . . . 50
6.12 Grad Cam . . . 50

7 Result and Analysis **52**
7.1 Model Evaluation Overview 52
7.2 Performance Comparison of CNN architecture 52
7.3 Custom Model (SadNetV1 Evaluation) 54
7.4 Experiment with Healthy patient 55
7.5 Predictability of our mode 56
7.6 Discussion and Insights 57

8 Future Work and Improvements **59**

9 Conclusion **60**

Bibliography **65**

List of Figures

3.1 Work Flow of Pre-Train Model 17
3.2 Work Flow of Custom Model 17

4.1 Visual Representation of Distribution Per Class 19
4.2 Sample Picture of Extracted Slices 20
4.3 Slice Indormation Per Class 21
4.4 Pre Processing flow chart 23
4.5 Data Distribution Pie-Chart 24
4.6 Train Augmentation Visualization 24
4.7 Data Augmentation Sample 25
4.8 Parameters of Augmentation 25

5.1 VGG19 Block Diagram 27
5.2 Train and Validation Accuracy, Train and Validation Loss 28
5.3 VGG19 Confusion Matrix of Test and Validation 28
5.4 Block-wise specification of the VGG19 architecture 29
5.5 ResNet50 Block Diagram 30
5.6 ResNet50 Accuracy and Loss 31
5.7 ResNet50 Confusion Matrix of Test and Validation 31
5.8 Block-wise specification of the VGG19 architecture 32
5.9 EfficientNetB0 Block Diagram 33
5.10 EfficientNetB0 Accuracy and Loss 34
5.11 EfficientNetB0 Confusion Matrix of Test and Validation 35
5.12 Block-wise specification of the EfficientNetB0 architecture 35
5.13 DenseNet Block Diagram 36
5.14 DenseNet Accuracy and Loss 37
5.15 DenseNet Confusion Matrix of Test and Validation 38
5.16 Block-wise specification of the DenseNet architecture 38
5.17 MobileNetV2 Block DIagram 39
5.18 MobileNetV2 Accuracy and Loss 41
5.19 MobileNetV2 Confusion Matrix of Test and Validation 41
5.20 Block-wise specification of the MobileNetV2 architecture 42

6.1 SadNetv1 Architecture Diagram 44
6.2 Layer Specification for SadNetv1 45
6.3 Flow of SE Block 47
6.4 Flow of Spatial Attention Block 48
6.5 Heatmaps From Grad Cam 51

7.1 Comparison of Different model with our proposed model 52
7.2 Precision, Recall and F1 score of test, train and validation of all models 53
7.3 Comparison of Proposed Model with Other Pre-Trained Model 54
7.4 SadNetV1 Accuracy and Loss 55
7.5 SadNetV1 Confusion Matrix of Test and Validation 55
7.6 Accuracy and Loss of Normal Paient 55
7.7 Predicted as AD 57
7.8 Predicted as PD 57
7.9 Predicted as FTD 57

List of Tables

4.1 Dataset Distribution Summary Per Class 19
4.2 Pixel Intensity range per class 21
4.3 Applied Gamma Value Table 21
4.4 Contrast Per Class Detection 22

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

AD Alzheimer Disease

$ADNI$ Alzheimer Disease Neuroimaging Initiative

DL Deep Learning

FTD Fronto Temporal Dementia

GAP Global Average Pooling

$GRAD-CAM$ Gradient weighted Class Activation Mapping

IDA Image and Data Archive

$LONI$ Laboratory of Neuro Imaging

ML Machine Learning

$NIFTD$ Neuroimaging in Frontotemporal Dementia

PD Parkinson's Disease

$PPMI$ Parkinson's Progression Markers Initiative

$SMOTE$ Synthetic Minority Over-sampling

Chapter 1

Introduction

1.1 Background Information

Neurodegenerative diseases are among the most complex and debilitating conditions affecting the human nervous system. Characterized by progressive loss of structure or function of neurons, these disorders have a profound impact on cognitive, behavioral, and motor functions [36], [50]. Among the most prevalent and devastating of these disorders are Alzheimer's Disease (AD), Parkinson's Disease (PD), and Frontotemporal Dementia (FTD). Each presents distinct pathological features and clinical manifestations, yet they often share overlapping symptoms that complicate accurate diagnosis and effective treatment [9], [18], [55]. In recent years, the global burden of these disorders has surged dramatically, especially with the aging population [12], [20], [49], thereby necessitating deeper exploration into their pathophysiology, early detection, and intervention strategies [8], [21], [22].

Alzheimer's Disease (AD)

Alzheimer's disease is widespread all over the world, making up 60-80 percent of all cases of dementia in vogue [4], [64]. It is a slowly spreading neurodegenerative disease, and one of the most crucial signs is cognitive deterioration, inability to remember, and inability to reason [65]. The pathophysiological features of AD are the accumulation of β-amyloid plaque and neurofibrillary tangles, which consist of hyperphosphorylated tau protein in the brain [65], [67]. The protein aggregates interfere with synaptic transmission, causing death of the neurons and subsequent brain atrophy, especially in the hippocampus and cerebral cortex areas–key areas in memory and cognition [63], [65]. Clinically, the disease is characterized by a continuous but insidious deterioration of memory, subsequently of language, executive skills, and visuospatial skills [4], [64]. Patients at advanced stages of the disease usually exhibit behavioral and psychological symptoms, which include depression, agitation, and apathy [58]. It takes about 8-10 years on average between the diagnosis and the death, but it may take longer [16]. The exact cause of AD is still unknown, even though it has been widely studied. Early-onset familial AD is linked to genetic mutations in the *APP*, *PSEN1*, and *PSEN2* genes, and the *APOE 14* allele is a well-known risk factor of late-onset sporadic AD [5], [24]. At the moment, there is no cure against Alzheimer's Disease, and such interventions as cholinesterase inhibitors, NMDA receptor antagonists help mildly improve its

symptoms [65]. Thus, it is crucial to diagnose and, potentially, utilize advanced neuroimaging and biomarker analysis to counter the state of the disease [63].

Parkinson's Disease (PD)

The second most popular neurodegenerative disease is Parkinson s Disease which is mostly motor-related to the loss of dopaminergic neurons in the substantia nigra pars compacta of the midbrain [11], [37]. Resting tremor, rigidity, bradykinesia, and postural instability are the cardinal motor symptoms of PD [11]. Moreover, frequent non-motor symptoms of most patients include depression, sleep disturbances, anosmia, constipation, and cognitive impairment, which may precede motor ones by years, still [37], [48]. Lewy bodies, also known as intracellular inclusions formed mainly of the 2-synuclein protein, are the pathological feature that defines PD [2], [3]. Precise PD etiology is yet to be discovered, but a combination of genetic and environmental factors is supposed to contribute to it [2], [14]. It is known that in inherited cases, genes like *SNCA*, *LRRK2*, *PARK2*, and *PINK1* have been identified as genetic abnormalities related to the condition [3], [14]. The pharmacological approaches, in particular, levodopa-based treatment offer high levels of symptomatic relief [1], [54]. Nevertheless, the extended usage is frequently related to the movement issues, such as dyskinesia [1]. Deep Brain Stimulation (DBS) is a surgical treatment of advanced PD, which is used in only a number of patients [54]. Notably, the timely diagnosis and treatment are essential to preserve the quality of life and minimize disability or decelerate the progression of the disease [11].

Frontotemporal Dementia (FTD)

Frontotemporal Dementia is a syndrome comprising a set of disorders that occur as a result of loss of the frontal and/or temporal lobes of the brain [6], [7]. In contrast to AD, in early stages, FTD affects memory to a limited extent but is characterized by various alterations in personality, behavior, and dysfunction of language tasks [17]. It normally develops at the age of 45 to 65 years and therefore is a major cause of early dementia [9], [17].

There are FTD clinical subtypes:

- Behavioural variant FTD (bvFTD) - is characterized by the weakening of inhibitions, apathy, compulsive disorder, and loss of social decorum [17], [31].
- Semantic variant primary progressive aphasia (svPPA) - this type is comprised of the loss of word meaning [7], [17].
- Non-fluent/agrammatic variant PPA (nfvPPA) - the disorder is characterized by speech production arduousness [7], [17].

Pathologies The pathological bases of FTD usually consist of tauopathies (deposition of tau protein) or TDP-43 proteinopathies. Observed genetic mutations in the *MAPT*, *GRN*, and *C9orf72* genes are strongly associated with familial cases of FTD [44], [59], [62]. FTD is often mistaken as a psychiatric issue or other forms of dementia due to being heterogenously presented, [7], [68].

As of now, there exists no cure, no special treatment that has been approved by the FDA to treat FTD. The management aims at supportive care and symptomatic relief [7], [45]. Early and proper diagnosis is very crucial due to the social and psychological impact it leads to patients and caregivers [7], [44].

1.2 Problem Statement

The main issue in the diagnosis and treatment of these NDs is that they all have overlapping clinical manifestations particularly in the early phases. Problems with cognitive abilities, apathy, depression, and motor fragmentation may occur throughout Alzheimer Disease (AD), Parkinson Disease (PD), and Frontotemporal Dementia (FTD), which may lead to incorrect diagnoses or a late diagnosis in a similar mold [61]. This slows down the process of effective intervention, overburdens the caregivers, and reduces the effectiveness of possible disease-modifying treatment methods [61].
In addition to that, the existing diagnosis gold standards, e.g., brain magnetic resonance imaging (MRI), PET scans, cerebrospinal fluid (CSF) biomarker tests, and genetic testing might not be universally available in low-resource settings. It is stated that in such scenarios, PET or CSF testing accessibility is low, and it is around less than 1% of the entire population [61]. This is compounded by the lack of specific initial-stage biomarkers which leads to a lack of targeted therapies and clinical trials.
There is thus a need to be cautious when dealing with such disorders, even in the point of diagnosis. Early behavior or cognitive symptoms are sometimes misinterpreted, which creates the potential to determine poor management approaches. Clinicians and researchers need to focus on creation of non-invasive, economically supportive and precise diagnostic tools that can help towards the early identification of these conditions. Neuroimaging analysis, artificial intelligence, and machine learning are on their way to filling this gap [29], [51], [66].

1.3 Research Objective

The goal of the main research is the construction of an efficient, optimized, scalable deep learning-based system that can accurately detect major neurodegenerative disorders such as Alzheimer's Disease (AD), Parkinson's Disease (PD), Frontotemporal Dementia (FTD) with the help of medical images 3D (MRI) Sagittal data at the early stages of development. The disorders, though having different pathophysiological structures, tend to show similar symptoms at the initial stages. There is thus a tendency to misdiagnose or a time lag in the diagnosis process, especially in patients presenting in an atypical or mixed mode. The majority of the previous studies have aimed at identifying these diseases independently, usually contrasted with healthy subjects or, at best, contrasted to a solitary condition. Moreover, clinical practice can include cases where a patient might come with co-morbid manifestations of various neurodegenerative diseases, or where they exist in a continuum whose separation is but hard. This gap in the research is what our study will address by identifying AD, PD, and FTD concurrently as part of a multilateral classification system with the application of the deep learning technique. Not only does the ap-

proach increase clinical relevance, but it also allows identifying possible additional pathologies, which are understudied nowadays.
In particular, the following are its major goals:

- To build a lightweight model which can detect Alzheimer, Parkinson and Fronto temporal dementia in the most efficient manner and which can be deployable for clinical work.
- To Enhance the MRI slices, we used 3D Sagittal slices in an efficient preprocessing pipeline so that it normalizes well and modality of three diseases are different and for that it makes overlapping specially in our case Alzheimer, Parkinson and Fronto Temporal Dementia all are overlapping. Hippocampus for Alzheimer is overlapping with Frontal lobe for Fronto Temporal Dementia again hippocampus of Alzheimer is also overlapping with Substantia Nigra of Parkinson. For this reason making an efficient preprocessing pipeline is a challenge for promising results.
- To build a Custom model which will give us the best result to detect this neuro degenerative disease effectively.
- To measure the performance by the conventional metrics, such as accuracy, precision, recall.
- In order to interpret our models to gain clinical understanding (Grad-CAM).

Chapter 2

Related Work

Diseases like Alzheimer disease, Parkinson disease, and Frontotemporal Dementia, are degenerative diseases that result in neuronal structure and functional decline. The most prevalent basis of dementia is AD, which mainly targets memory and cognition [21]. PD is a disorder of the movement manifested by motor symptoms (e.g., tremor and rigidity) sometimes also featuring non-motor symptoms (e.g. cognitive decline) [10]. Instead, FTD impairs language and behavioral abilities because of the degeneration of the frontal and temporal lobes [9]. These conditions are clinically diverse yet often have overlapping symptoms, leaving diagnosis rather problematic.

2.1 Alzheimer

Alzheimer's Disease (AD) is an irreversible neurodegenerative disorder that leads to dementia in human beings, making it the worldwide common type of dementia [43]. No cures have been found for the disease therefore early detection is advisable to minimize the extent of impairment or prevention. Among several imaging techniques for displaying the brain condition of affected people, MRI is the most important one. While employing the application of deep learning is a sub-field of machine learning, AD can be distinguished between healthy and affected patients.

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that ultimately results in the death of millions of people globally, with symptoms such as cognitive impairment, memory loss, and other related complications. It is essential to detect AD as early as possible to intervene and control the disease's course. The article by Muhammad Raees and Vinu Thomas (2021) [43] discusses the use of deep learning methodologies in automatically identifying Alzheimer's disease through MRI images, based on the necessity of computationally efficient strategies with high precision in the early identification of the disease.
The paper [43] has emphasized the significant contribution of brain scans, especially MRI, in the diagnosis of Alzheimer's. The MRI scans give valuable structural details of brain atrophy that play a critical role in the diagnosis of AD in its initial phases. The authors [43] concentrate on how deep learning in the form of deep neural networks (DNNs) and transfer learning can be used to automatically clas-

sify MRI images into various categories: Alzheimer's disease (AD), Mild Cognitive Impairment (MCI), and Normal Controls (NC). Using a large MRI dataset of the Alzheimer's Disease Neuroimaging Initiative (ADNI), the authors demonstrate that deep learning models can achieve a high percentage (80-90%) of accuracy in classifying MRI images [43].

The paper's [43] first major contribution is that it compares deep learning models with traditional machine learning methods, including Support Vector Machines (SVM). The results indicate that the DNN models, including AlexNet, VGG-16, and GoogLeNet, show superiority over SVM models based on the classification accuracy because they present the best performance regarding all categories (AD, MCI, and NC) [43]. The significance of transfer learning usage, as already mentioned in the paper, lies in the possibility of using pre-trained models and fine-tuning them on medical imaging data, thus achieving higher prediction accuracy compared to large-scale labeled data requirements. This proved useful, especially in the setting of medical images where labeled data can be expensive and time-consuming to acquire. Also, the research illustrates the computing needs of training deep learning models, noting the convenience of GPU-based systems for efficiency. According to the authors, [43] deep learning models, although expensive in terms of computation needs, offer quicker training speeds as compared to the traditional machine learning approaches, hence can be applicable in clinical settings where time is crucial.

The paper [43] also recognizes the field difficulties that include data imbalance and the requirement for huge datasets that are annotated. Nonetheless, it concludes that, with additional development and clinical verification, the deep learning algorithms will become a powerful contribution to the early AD detection that will result in timely interventions and improved patient outcomes.

Overall, the research conducted by Raees and Thomas (2021) [43] supports the promising nature of deep learning methods in the automated identification of Alzheimer's disease based on MRI images. The study helps to reveal the recent attempt to provide a better diagnosis at an earlier stage and eventually manage the neurodegenerative diseases more clinically, by using the power of large-scale machine learning models and datasets.

Murugan et al. (2021) [41] proposed DEMNET, a novel deep learning framework that was explicitly trained to predict Alzheimer's and the various stages of dementia with high accuracy and early on structural Magnetic Resonance Imaging (MRI). Alzheimer's disease has been among the global health care issues, largely due to the modest nature of the onset of the disease, as well as the inability of the traditional diagnostic tools, which cannot provide the requested sensitivity and accuracy. The authors [41] detailed the need for computationally efficient, accurate, and interpretable automated diagnosis algorithms development, citing the importance of early intervention as a controlling mechanism of the disease process. DEMNET architecture concentrates on Convolutional Neural Networks (CNNs) and includes special layers, i.e., convolutional, pooling, dropout, and dense layers, which are arranged into regular "DEMNET blocks." The base, first significant contribution of the research is that it deliberately focuses on the imbalance of the datasets, frequent in the medical imaging domain, using the Synthetic Minority Oversampling Technique (SMOTE). The authors [41] could artificially inflate the classes that were

underrepresented in their dataset in this manner, achieving a balanced training set and providing a model that is more inclined to generalize when applied to the different kinds of dementia. The publicly available dataset used in the study [41] was obtained on Kaggle and described MRI scans in four categories of dementia: Non-Demented (ND), Very Mild Dementia (VMD), Mild Dementia (MD), and Moderate Dementia (MOD). After data augmentation, DEMNET outperformed current methods in diagnostic accuracy with an impressive overall accuracy of 95.23%, an Area Under the Curve (AUC) of 97%, and a large value of Cohen's Kappa of 0.93, which depicts an excellent inter-class agreement. Besides, the strength of DEMNET was also tested on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, in which the model attained an accuracy of 84.83% and an AUC of 95.62% [41]. Notably, the authors [41] also considered the issue of model interpretability and used the occlusion sensitivity maps to evaluate the vital areas of the brain that contribute to the classification of dementia based on the model. This method increases the transparency and helps clinicians comprehend and trust the decisions of the model, which is essential to clinical integration. In general, the DEMNET model offered by Murugan et al. [41] can be considered an important step towards AI-based early diagnosis instruments in the case of Alzheimer's disease and dementia. With the balance achieved between high diagnostic accuracy, interpretability, and computational efficiency, DEMNET has a great prospect of being adopted clinically, where it could lead to the introduction of interventions earlier in the process and to better patient outcomes. The authors' proposal of future directions involves the application of the model to multimodal imaging data and wider clinical validation in different populations.

In the article by Ejaz Ul Haq et al. (2023), [69] an efficient novel framework of deep learning using magnetic resonance imaging (MRI) data to distinguish among Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), and Normal Control (NC) is presented. The authors present the issue of the increasing need for effective diagnostic steps of the neurodegenerative disorders, and immediately Alzheimer's disease in particular, which has a slow but maybe barely visible progress [69]. A light multimodal fusion model, comprising Convolutional Neural Networks (CNN) combined with Long Short-Term Memory (LSTM) networks, is proposed in the article as an efficient solution to classification tasks on medical images. The provided architecture [69] is the first to employ CNN and LSTM networks, particularly to address both kinds of information: spatial and sequential. The CNN component performs the deep feature extraction on 2D slices of MRI, which are acquired by cutting the 3D MRI scans of the patients, thus reducing the volumetric data to a simplified 2D form. This transformation makes the data more aligned with the currently existing model of deep learning and makes it possible to compute the efficiency of the model without losing the precious spatial details. The CNN layers are followed by the LSTM layers, as the latter is specifically suggested to find the long-term dependencies in ordered data, and thus it will be more appropriate to learn the progressive nature of neurodegenerative diseases like Alzheimer's. Another aspect that the authors use [69] to improve the performance of the model is data augmentation through rotating, flipping, and translating the training dataset. This eases overfitting, which is a common problem with deep learning models, particularly

in the medical domain, where generally, large annotated datasets are not present. In order to determine the robustness and good performance, the model was trained and tested on the Alzheimer's Disease Neuroimaging Initiative-1 (ADNI-1) dataset of MRI images of AD, MCI, and NC, totaling 505 subjects. The model has several impressive aspects, one of them being the lightweight property. The model obtained 92.3 percent accuracy using fewer parameters and lower computation complexity than the currently available state-of-the-art models. Another important innovative idea is using Focal Tversky Loss (FTL) as a loss function, which is designed to reduce false-negative predictions and enables the model to concentrate on more difficult, atypical cases, which is especially significant in the scenario of Alzheimer's [69]. The study outcomes [69] are encouraging as the proposed model not only surpasses state-of-the-art deep learning models but also displays the inference time of only 1.7 milliseconds, which is appropriate to utilize in real-time. This is one of the main aspects to consider in the implementation of such models in clinical practice, where time and computer power can be limiting factors. Moreover, the model performance is evaluated against 26 diverse classifiers (both machine learning and deep learning models), and it has been seen that the proposed model is better in accuracy and efficiency. The contributions of the paper [69] to the overall body of knowledge about the diagnosis of Alzheimer's disease are quite solid, as the solution presented is more computationally efficient, without affecting the accuracy of the diagnosis. The capability of the suggested model to work with sophisticated 3D MRI data and convert it to 2D characteristics and consequently work on these characteristics using CNN and LSTM layers offers an optimistic way for further studies. According to the outcomes of the research, [69] the model holds great prospects of being used clinically, especially in environments where fast, efficient, and precise diagnosis is paramount. To sum up, the article by Ejaz Ul Haq et al. (2023) [69] provides an innovative and detailed solution to the diagnostic problem that Alzheimer's and MCI create. Lightweight multimodal fusion model based on CNN and LSTM network fosters not only the improvement of classification accuracy but also guarantees the shortening of processing times, which is an important aspect to consider in the model implementation into the real-world clinic. Nevertheless, additional verification on bigger, more varied datasets and the inclusion of other imaging modalities would be needed to harness the potential of the model in the realms of varied clinical settings completely.

In conclusion, deep learning is applicable as one of the methods that can aid in the diagnosis of AD with the help of MRI. Deep learning can get high accuracy with optimal computational power requirements in the system. Furthermore, One thing that was shared in this research is that they employed ADNI datasets. Among all of them, CNN's performance is optimal for identifying AD. Moreover, there is a need for more detailed research to enhance the systems' precision and the methods for acquiring knowledge.

2.2 Parkinson's

The advancement in artificial intelligence, and in particular, the utilization of machine learning (ML) and deep learning (DL) models, have been investigated in several modern research works focusing on the early diagnosis of Parkinson's disease (PD) to improve the diagnostic process's accuracy and dependability.

Wang et al. (2020) [33] introduced a powerful and contrasting framework of the early diagnosis of Parkinson Disease (PD) based on deep learning (DL) and machine learning (ML) methods with the usage of premotor biomarkers. Acknowledging the clinical relevance of PD diagnosis at its prodromal stage, the authors of [33] took advantage of the Parkinson Progression Markers Initiative (PPMI) data, including 401 participants with early-stage PD and 183 healthy control individuals. Thirteen biologically and clinically relevant characteristics were included in the research, such as REM Sleep Behavior Disorder scores (RBDSQ), loss of smell (UPSIT), cerebral spinal fluid (CSF) biomarkers (e.g., alpha-synuclein, A beta 1 42, tau proteins), and dopaminergic imaging markers as measured by SPECT scan.
They trained a feed-forward deep neural network (DNN) having two hidden layers and evaluated its performance against twelve standard ML algorithms, namely SVM, random forest, logistic regression, and various versions of boosting methods in [33]. One important impact of this paper [33] is the comprehensive comparison of the methods based on standardized metrics, which are accuracy, precision, sensitivity, specificity, F1-score, and AUC, on 100 stratified train-test splits. The deep learning ensemble model had the best average accuracy of 96.68% and came out better than all the models, and it had a good balance between sensitivity (97.17%) and specificity (94.84%) [33].
The study [33] is accompanied by an essential feature of being interpretable through feature importance analysis based on Boosted Generalized Additive Models (BOOST GAM). The left and right putamen imaging biomarker emerged as the strongest predictors of PD, followed by the UPSIT score. This is congruent with clinical knowledge of early deaths of dopaminergic neurons in PD. It is important to note that the deep learning models were not sensitive to the network structure and training epochs changes, which highlights their robustness when using small datasets.
Although deep learning yielded the best outcomes, the authors admitted the model limitations regarding interpretability and the modest sample size [33]. They also pointed out that, although DL methods held tremendous potential, enhancement methods like BOOST GAM offered almost similar performance along with transparent feature selection [33].
The research [33] is especially applicable because it emphasizes the possible role of data-driven models in the early detection of PD and provides a precedent for the successful integration of biologically plausible features and state-of-the-art learning techniques. It also shows that even when used on rather smaller medical datasets, deep learning can provide clinically meaningful and actionable insights [33].

Abdullah et al. (2023) [47] developed a novel architecture for the early diagnosis of Parkinson's Disease (PD) based on deep transfer learning and efficient feature

selection schemes. The paper discusses the weaknesses of the conventional handcrafted feature extraction-based methods, which usually have many limitations in terms of their accuracy, by proposing a powerful hybrid technique that uses three deep learning models, including ResNet50, VGG19, and InceptionV3. Such networks were used to obtain discriminative features on a publicly available dataset (NewHandPD), which consists of handwriting samples (spirals and meanders) of PD patients and healthy people.
The essential contribution to the research [47] is the use of a Genetic Algorithm (GA) with K-Nearest Neighbors (KNN) classifier when optimizing features. This procedure allows selecting and preserving the most informative attributes and eliminating the redundant information, thereby improving the classification rates and simplifying the computational structure. The authors [47] successfully used the idea of transfer learning by stripping the top classification layers of the pretrained models and replacing them with their dense layers to extract features efficiently, and then stacking the features of all three models into a 300-dimensional feature representation.
The performance of the framework was comprehensively tested with a vast amount of experimentation, which consisted of optimizing the parameters of the genetic algorithm (population size and number of iterations). On a population size of 20 with 200 iterations, the model went 95.29% accurate, 0.985 precision, 0.829 recall, and an AUC score of 0.901. The results were demonstrated to outperform a number of state-of-the-art models on a number of metrics, as presented in their comparison study. The effect of variation of hyperparameters, efficiency of feature selection, and model generalizability was also discussed in detail by the authors [47].
The innovative aspect of the work is the combination of several transfer learning models to avoid architectural bias and the use of genetic algorithms to select features adaptively and with high precision. It holds the promise of deep learning and heuristic optimization to build lightweight, accurate, and explainable diagnostic models of neurological disorders. The method helps to build the emerging field of AI-based medical diagnosis and demonstrates that the dynamics of handwriting can be a non-invasive and inexpensive Parkinson's disease biomarker [47].

The first breakthrough is a study by Quan et al. (2021) [42], who suggest a new deep learning (DL) framework to early diagnose Parkinson's disease (PD) on the basis of dynamic attributes of speech signals. In light of the constraints of the traditional machine learning (ML) methods based on the usage of mostly static acoustic features, the proposed research highlights the diagnostic opportunities that the time-series variations of human speech can offer. The research [42] is based on the idea that PD influences the motor speech control, causing observed disturbances in the articulation and phonation.
To extract and model dynamic articulation transition features of speech, the authors propose a Bidirectional Long Short-Term Memory (BiLSTM) network that is adapted to the task [42]. These characteristics are energy variations in transition between unvoiced and voiced segments and are said to correspond to neuromotor deficits typical of PD. Dynamic speech features were found to be significant when quantitatively comparing PD patients with healthy controls (HC), showing more articulation transitions and more irregular fundamental frequency contours in the

PD group ($p < 0.05$), thereby confirming the suitability of the investigated dynamic speech features [42].
To avoid sample overlap, two evaluation strategies were used: 10-fold cross-validation and splitting of the speaker-independent dataset [42]. The BiLSTM model trained on both protocols greatly surpassed the traditional classifiers, including Support Vector Machines (SVM), Decision Tree, and Convolutional Neural Networks (CNN), especially on continuous speech (e.g., short sentence utterances). Remarkably, the BiLSTM model attained an accuracy of 84.29% on short sentence inputs, making a significant increase over state-of-the-art results using static features, which were around 73% on average [42].
The methodology of the study [42] was also benchmarked against the end-to-end CNN models on spectrogram representations. CNNs were found to work relatively well on tasks of sustained phonation, but struggled with more natural, variable speech input, emphasising the better ability of the BiLSTM to capture temporal dependencies.
The piece of work has added value to the literature since it has revealed that time-variant speech features offer a more detailed and sensitive foundation to PD detection when compared to the use of static descriptors [42]. It paves the way to the implementation of DL models in telemedicine, especially in the case of non-invasive and speech-based early screening instruments. Such future directions as the extension to the multi-label classification of PD staging and the investigation of the hybrid architecture uniting the possibilities of BiLSTM and CNN were proposed [42].

Ali et al. (2024) [56] proposed an effective diagnosing system that would enhance the efficiency and reliability of Parkinson's Disease (PD) detection through voice recordings. There are two important deficiencies in the prior art that the study hopes to fill, namely the use of inappropriate validation schemes, especially k-fold cross-validation with subject overlap, and the inability to obtain high classification accuracy under unbiased testing conditions. To address these problems, the authors proposed a cascaded two-stage diagnostic system by using L1-regularized Support Vector Machine (SVM) to refine the features, followed by a Deep Neural Network (DNN) based classification [56].
The strength of the study [56] is that it is methodologically sound. Instead of splitting the subjects as in the previous studies, which tend to overfit the performance by duplicating the subjects in validation, this study proposes the application of Leave-One-Subject-Out Cross-Validation (LOSO-CV), which guarantees the generalization on unseen subjects. Moreover, the authors [56] reveal that the reduction of many samples on the subject to one composite sample is efficient in removing overlapping subjects and their biases.
Model training and testing were done using two openly available databases that contained both sustained vowel phonations and mixed speech tasks [56]. L1-SVM was applied to select features and produce a high-quality and sparse feature subset. Those refined features were in turn input to a DNN, the architecture of which (number of layers, neurons, dropout rate) was optimized through a Hybrid Grid Search Algorithm (HGSA). The hybrid method permitted the concurrent optimization of the hyperparameters of both the DNN and SVM blocks [56].
Experimental outcomes were convincing: the suggested L1SVM-DNN model at-

tained 100 percent classification accuracy under LOSO-CV on both datasets and up to 97.5 percent accuracy under 10-fold CV without subject overlap [56]. The model [56] showed good generalization when evaluated on a separate test set of unseen PD patients, achieving an accuracy of 96.42
Related to the previous methods, [56] both in terms of plain SVMs to the more intricate deep learning models, like CNNs and autoencoders, the L1SVM-DNN surpassed their performance, as well as the highest scores achieved in the previous literature. Notably, the experiment [56] verification shows that feature selection (through L1 regularization) plus feature extraction with DNNs indeed allows reducing noisy features, overfitting, and makes models more interpretable.
The work not only makes a contribution to PD detection but also leaves a validation framework that can be generalized to other biomedical diagnostic issues. The authors, [56], however, present limitations, which include failure to explore the issue of differentiating between idiopathic PD and atypical forms of the disease, such as MSA or PSP. They indicate [56] that the future research direction should be extended to multiclass classification with clinically stratified data.

2.3 Frontotemporal Dementia (FTD)

In this paper, Hu et al. (2021) [38] present a new deep learning (DL)-based method to categorize and visually demonstrate the differences between such diseases as Alzheimer's Disease (AD), Frontotemporal Dementia (FTD), and Normal Controls (NCs) via raw 3D T1-weighted MRI images. The paper focuses on the fact that distinguishing between AD and FTD can be a serious problem, as these two conditions can overlap each other in basic clinical symptoms, and it is hard to find an adequate diagnosis with the help of traditional tools. The authors emphasize how deep learning has a chance of making this process more straightforward, and it does not necessitate sophisticated preprocessing as was the case with neuroimaging studies in the past.
Among the fundamental contributions of the study [38] is the creation of a deep learning framework that applies raw MRI images directly. This has the benefit of lessening the reliance on specialist preprocessing, and makes the classification process more of an automated and repeatable process, better lending itself to clinical adoption. Also, the authors [38] take advantage of data augmentation to deal with MRI scan variations, including variations in slice thickness and pixel intensity. This guarantees that the model can be generalized or applied in a wide variety of imaging conditions, thereby making the model more robust.
The results of the carried out evaluation of the offered model indicate its effectiveness as it shows high accuracy, equal to 91.83%, in differentiating among AD, FTD, and NCs [38]. These encouraging findings support the possibility of the model enhancing the accuracy of diagnosis in a clinical environment. The study [38] also presents a guided backpropagating algorithm to provide visualization of the parts of the brain most responsible in the decisions made by the model. This interpretability plays an essential role in medical practices, where it enables clinicians to have insights into how the model comes up with its decision, facilitating trust and enabling clinical reasoning.

Although the approach is strong, the paper has recognized the challenges that include data imbalance and variability of imaging protocols across multiple clinical centers. These may restrict the applicability of the model to be generalized and prevent its applicability to the real world. The authors, [38], however, deal with this by experimenting with the model on varied datasets, making sure that the suggested technique is capable of dealing with disparate variations in imaging. Another important point made by the study [38] is that although the model is effective in the controlled environment of an experiment, its implementation into daily clinical use continues to be hampered by limitations like small sample sizes, the absence of longitudinal tracking data, and so on.
To sum up, the work by Hu et al. [38] can be considered an important step towards using deep learning to diagnose AD and FTD. The study paves the way to introducing AI into the clinical pathway as it makes the diagnostic process less complex and offers interpretable results. Nevertheless, upcoming studies should address the challenges, including data imbalance, and should expand the size of the studies to make them more clinically validated.

Perez-Millan et al. (2023) [52] performed a study that aimed to determine how machine learning (ML) methods can be used to distinguish between Alzheimer's Disease (AD) and Frontotemporal Dementia (FTD) based on cross-sectional and longitudinal brain magnetic resonance imaging (MRI) data. The research problem is to use structural MRI scan images to learn the pattern of brain atrophy and utilize machine learning algorithms to achieve a higher diagnosis accuracy rate of these neurodegenerative disorders. The study particularly focuses on the problem of shared symptoms and brain signatures between AD and FTD, which means that the differential diagnosis represents a challenging clinical undertaking.
The architecture used by the authors [52] involved the ensemble of both unsupervised and supervised machine learning, where the Principal Component Analysis (PCA) dimensionality reduction technique was used alongside Support Vector Machine (SVM) classifiers to classify the diseases. The investigation employed a big sample size of 339 participants, 153 AD patients, 87 FTD patients, and 99 healthy controls, with some of the subjects chosen to undergo repeat MRI scanning in two years to enable longitudinal analysis [52]. The findings demonstrated that the ML algorithm presented good classification rates: 83.3 percent between AD and healthy controls, 82.1 percent between FTD and healthy controls, as well as 63.3 percent between AD and FTD at the cross-sectional level [52]. The accuracy was much higher when longitudinal data was incorporated and was 90% accuracy in AD vs healthy controls and 88% accuracy in FTD vs healthy controls. The findings support the significance of including the longitudinal data in order to enhance diagnostic accuracy [52].
An important part of this research [52] was interpreting the machine learning model. Then, by examining the weights connected with various brain parts, the researchers extracted major areas of atrophy that had the greatest impact on the classification decisions. These results were in line with published pathological correlates of AD and FTD, and the particular brain areas, including supramarginal and precuneus areas, were very essential in separating AD and FTD individuals and healthy controls. The longitudinal data use also assisted in disease trajectories capture, which is a more

refined way of understanding how these diseases evolve with time [52].
This combined approach of cross-sectional and longitudinal data sets in the study [52] provides a global method of enhancing the sensitivity of differential diagnosis of AD and FTD. Though the findings are encouraging, the research also admits to several limitations, such as the fact that the sample size of the longitudinal data used is rather small or that the research confines itself to structural MRI, which implies that in further research, it will be possible to include other neuroimaging modalities, such as diffusion tensor imaging (DTI) or amyloid PET scans, to make the model more powerful in terms of its diagnostic value.
Overall, the study by Pérez-Millan et al. (2023) [52] shows that machine learning on longitudinal MRI can lead to a substantial rise in the separation of AD and FTD, which can fuel beneficial practical implications. By doing so, not only does it push the accuracy of diagnoses, but it also provides interpretability, a necessary property to be used in the clinical setting. Nevertheless, prospective studies in larger sample sizes and wider varieties of imaging modalities are needed to corroborate the findings and adjust the algorithm to make it applicable to a wider spectrum of clinical practice.

In this paper, Ma et al. (2020) [30] propose a robust machine learning framework that aims to improve the differential diagnosis of Alzheimer's Disease (AD), Frontotemporal Dementia (FTD), and Normal Aging (NC) through the use of structural magnetic resonance imaging (MRI). Since it is stated that clinicians find it rather difficult to distinguish between AD and FTD as those disorders share most of the symptoms and the patterns of brain atrophy, the proposed research is expected to solve the problem of diagnosis in AD and FTD with the help of a new method of multi-scale and multi-type features extraction, deep neural networks, and data augmentation strategies.
They use Generative Adversarial Network (GAN) to augment MRI data, which is essential to address the problem of a small amount of training data in medical imaging [30]. The idea of this approach is to generate more feature vectors that can fool the network to classify them correctly and therefore generalize the model better. The researchers use a large amount of data, providing 1,954 MRI images obtained through the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Frontotemporal Lobar Degeneration Neuroimaging Initiative (FTLDNI) [30]. The network is trained using a combination of volume size and cortical thickness features obtained at various scales of MRI scans, which leads to better performance compared with the traditional methods.
The 10-fold cross-validation result indicates that the suggested approach attains a remarkable overall accuracy of 88.28%. Besides, multi-type and multi-scale features play an important role in increasing the discriminative power of the model between both AD and FTD and compared to other simpler classifiers such as Support Vector Machines (SVM) [30]. Remarkably, GAN-based data augmentation algorithm also increases the sensitivity of the classification, especially in the case of FTD that is more difficult to diagnose than AD. Another factor that leads to a robust and stable performance is the ensemble classifier approach in which many models are trained and their predictions averaged.
The paper [30] has three main contributions: it confirms that multi-scale multi-

type feature extraction of structural MRI scans can be an effective way of capturing the latent patterns to classify diseases; it confirms the usefulness of GANs in data augmentation to enhance the generalizability of the models; and it identifies the robustness provided by ensemble classifiers. The authors recommend future research using more imaging modalities and bigger, more varied datasets to realise the full potential of AI in clinical dementia diagnostics, even though the technology has already proven to be successful.
Nevertheless, it [30] has to be mentioned that the presented work is an important step towards machine learning-driven diagnostic tools in AD and FTD since it provides a high-accuracy framework based on deep learning with state-of-the-art feature extraction and augmentation methods. Nevertheless, the authors spell that although the method holds promise, it still requires additional validation and development before it may be generally used in medical practice [30].

Overall, current deep learning solutions to Alzheimer Disease (AD), Parkinson Disease (PD), and Frontotemporal Dementia (FTD) have demonstrated valuable performance in terms of accuracy and interpretability. Nevertheless, other issues related to overlapping clinical symptoms, data imbalance, and poor generalizability across variating datasets are still present, which underlines the necessity to develop more integrated, scalable, and clinically proven models.

Chapter 3

Working Plan

The initial approach was to gather MRI data in three publicly available sets that included the Alzheimer Disease Neuroimaging Initiative (ADNI) of Alzheimer Disease (AD), the Parkinson Progression Markers Initiative (PPMI) of Parkinson Disease (PD) and the Neuroimaging in Frontotemporal Dementia (NIFD) set of Frontotemporal Dementia (FTD). Once we got the datasets, we combined them and assigned classes as Alzheimer disease (AD), Parkinson disease (PD), and Frontotemporal Dementia (FTD). This data set was used to test and train a variety of Convolutional Neural Network (CNN) configurations.
The sagittal slices of 3D MRI images were used to train the model. We came up with the slice ranges of 60 to 89 per patient, which we determined manually and based on the guidance of our experts, and identified important locations in the brain of our study worth monitoring: the hippocampus (in AD cases), the substantia nigra (in PD cases), and the frontal lobe (in FTD cases). All the input images have been resized to 224x224 pixels, with the 3-color channels (RGB), therefore the input shape is 3 x 224 x 224.
The obtained dataset was separated into training (70%), validation (15%) and testing (15%) sets. Training data received immense data augmentation transformations such as rotation, translations, and cropping to augment model generalization so as to achieve consistent feature learning across diverse representations of brain slices.
We also tested a number of the most recently designed CNN architectures like VGG19, ResNet50, DenseNet121, EfficientNetB0, and MobileNetV2 to train the processed dataset. According to the results of comparative evaluations, the best results were obtained with MobileNetV2. We therefore came up with a unique study lightweight model called SadNetV1 that incorporates MobileNetV2 with a channel attention of Squeeze-and-Excitation (SE) block and Spatial Attention block to extract stronger spatial features.
The rationale of using this custom architecture was to resolve the tradeoff between the computational efficiency and accurate localization of disease-specific features. Since there was such a high inter-class similarity of the MRI slices of AD, PD, and FTD, we intend our model to pay more attention to discriminative spatial areas and potential related channels and enhance the performance of classification.

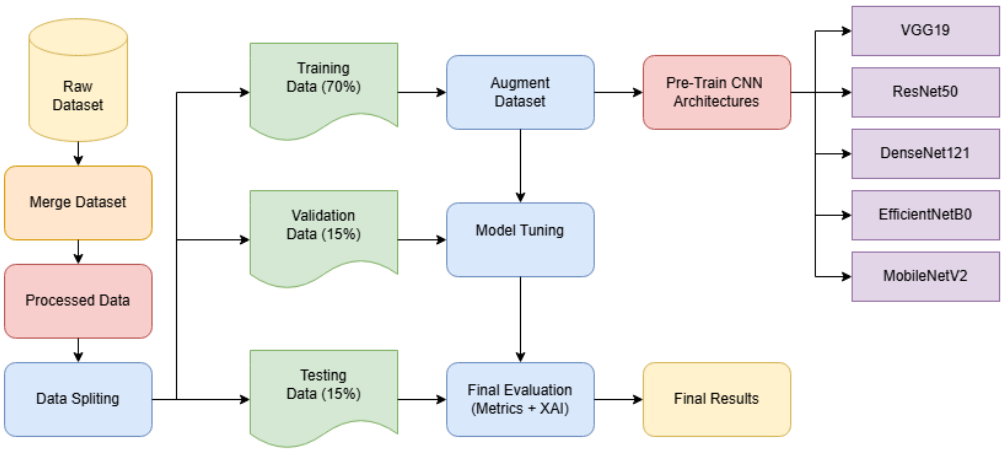

Figure 3.1: Work Flow of Pre-Train Model

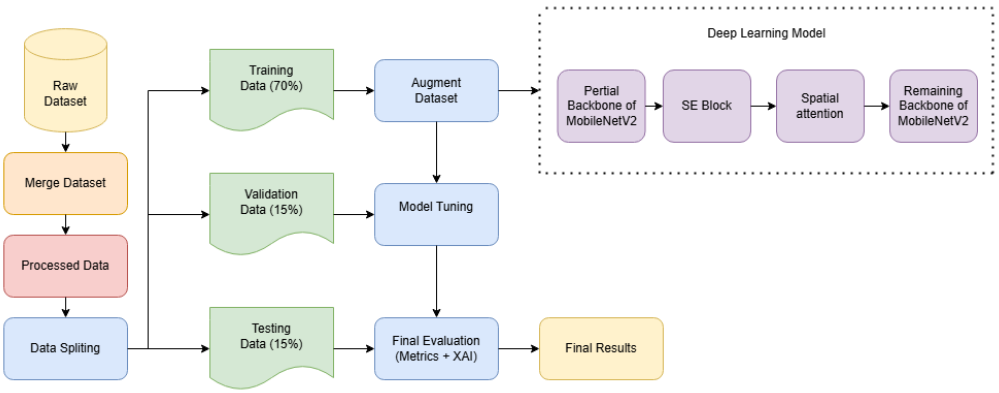

Figure 3.2: Work Flow of Custom Model

Chapter 4

Dataset

4.1 Dataset Sources

All the data generated via 3D MRI Sagittal scans in this study were acquired through the Image and Data Archive (IDA), which is maintained at the Laboratory of Neuro Imaging (LONI) at the University of Southern California. Datasets retrieved through this platform were as follows:

- ADNI Alzheimer's Disease Neuroimaging Initiative (ADNI).
- Parkinson Progression Markers Initiative (PPMI).
- Neuroimaging Frontotemporal Dementia Initiative (NIFTD).

To retrieve those datasets, it was necessary to apply to access the data by filling in a data use application, and was approved to use the data in scientific research after reviewing for several months. The data retrieved was T2-weighted 3D Sagittal MRI brain scans.

4.2 Data Overview

The total number of patients included in model training was 255, including slices with a rough average of 30 slices at the sagittal plane per patient. This led to a total of 60 to 89 good-quality slices per patient, or equal distributions of Alzheimer's Disease (AD), Parkinson's Disease (PD), and Frontotemporal Dementia (FTD). During preprocessing, some of the patients were dropped because their scans had been corrupted, incomplete, or of low resolution. Only scans taken with high-resolution sequences, e.g., MPRAGE and MGRAPPA, were retained to guarantee consistency and increase anatomical detail.
The 3 Tesla (3T) scanners used MRI scans with a slice thickness of 1.0 mm, with adequate spatial resolution to perform detailed structural analysis. The AD and FTD data sets included SIEMENS scanners, whereas the PD data set included PHILIPS Medical Systems scanners. Application of standardized imaging protocols, same field strength, and slice thickness enhanced cumulative quality and comparable data across these categories.
In the study, we aimed at the classification of three neurodegenerative diseases: Frontotemporal Dementia (FTD), Parkinson's Disease (PD), and Alzheimer's Disease (AD). In training to overcome the issue of imbalance in classes, various methods

Disease	Patient found	Re-visited Patient	Unique Patient	Patient Used
AD	240	148	92	85
PD	210	114	96	85
FTD	185	97	88	85

Table 4.1: Dataset Distribution Summary Per Class

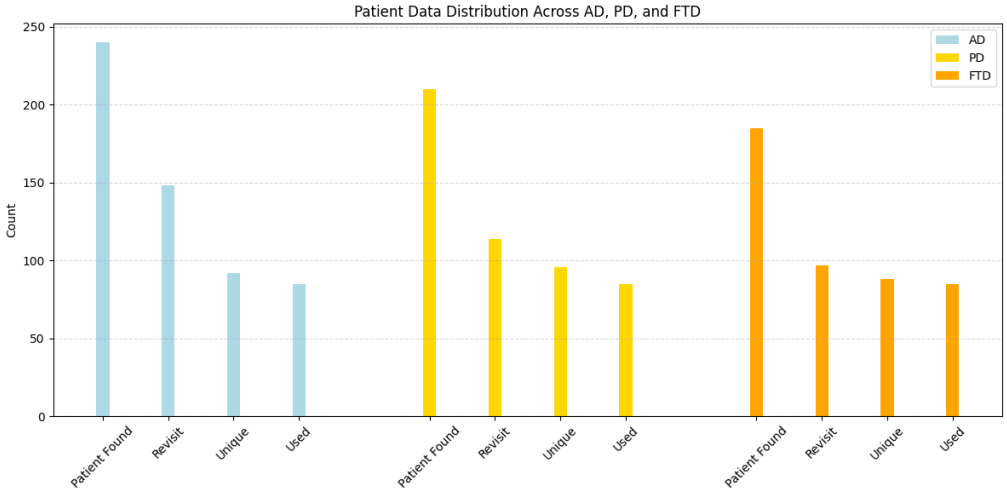

Figure 4.1: Visual Representation of Distribution Per Class

were tried, which included: Synthetic Minority Over-sampling Technique (SMOTE) and re-balancing of classes through the weight of classes. Nonetheless, following several testing sessions, the strategy of downsampling has proved to create the best and most consistent outcomes, in regard to model performance and generalization. To evaluate the model and comprehend it, nine test subjects (i.e. three in each category) were chosen and analyzed with the help of explainable AI (i.e., Gradient-weighted Class Activation Mapping (Grad-CAM)). The spatial visualizations enabled us to see which parts of the brain MRI slices made the most significant contribution to each of the predictions. The resulting heat maps showed that the model always paid special attention to the parts of the anatomy that are relevant to the disease, improving the clinical readability and reliability of the diagnosis results.

4.3 Pre-Processing pipeline

To maintain consistency, quality, and accuracy of input to train the convolutional neural networks, a collection of data preprocessing pipelines was put in place to process MRI data acquired by LONI Image Data Archive (IDA) on AD, PD, and FTD. The pipeline contained the following steps:

4.3.1 Slice Extraction

3D MRI volumes were selected, and 2D slices in a Sagittal plane were obtained. In each subject, about 30 contiguous slices were manually chosen in the middle portion of the brain in favor of anatomic relevance and reduction of background noise. The slice choice was cued to region, in that areas of interest related to a disease were visibly located in the clinical approach:

- Frontotemporal Dementia (FTD) - it was chosen in Slices 60-69 that cover the frontal lobe of the brain.
- The slices used to accommodate Alzheimer's Disease (AD) are 70-79, and the use of the hippocampus was very pronounced in these slices.
- The slices 80-89 of Parkinson Disease (PD) were chosen at the level of the substantia nigra.

The slices of each subject were read manually, and adequacy in terms of the anatomical coverage was measured according to the clinical modality and biomarkers that are pertinent to the disease.

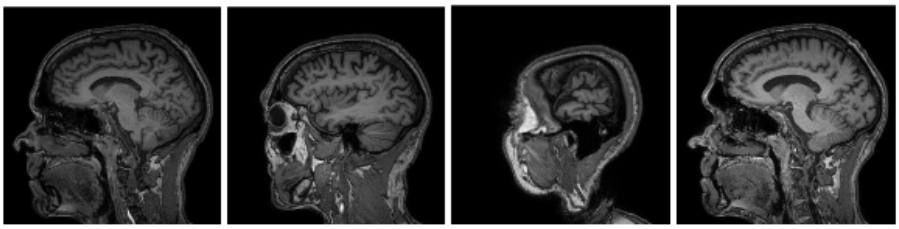

Figure 4.2: Sample Picture of Extracted Slices

4.3.2 Filtering and inspection of Data quality

- Black or blank slices, corrupt images, and slices with a low view of the anatomy were not considered.
- Only the high-resolution scans (e.g., MPRAGE, MGRAPPA) obtained using 3T MRI scanners (Siemens in the case of AD/FTD, Philips in the case of PD) were kept.
- Manual review was carried out to match the contrast, the pixel quality and intensity across the entire classes.

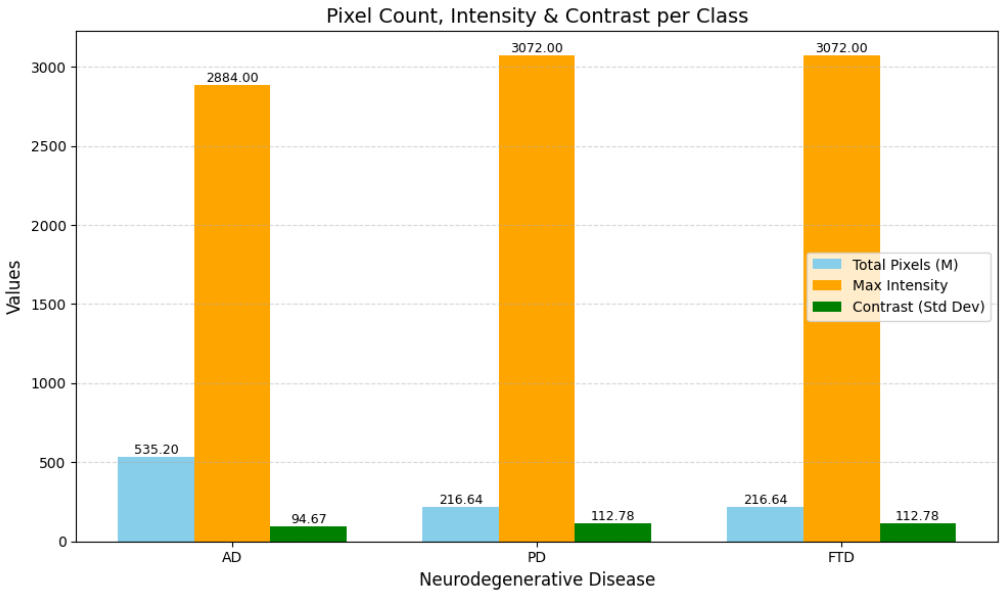

Figure 4.3: Slice Indormation Per Class

4.3.3 Pixel Intensity Normalization

We checked each class and found that each class had a different pixel intensity. The information is stated below:

Disease	**Pixel Intensity range**
AD	range: 0.0 to 2884.0
PD	0.0 to 3072.0
FTD	0.0 to 3072.0

Table 4.2: Pixel Intensity range per class

The need to balance all classes necessitated normalization (in this case with the use of min-max scaling) so that the values in the intensities were all within the same range [0, 1], which removed the grayscale mismatch among the scanner sources.

4.3.4 Gamma Correction

As we checked intensity we get to know that PD has the most intensity which suggests PD has brightest slices among three classes. Whereas FTD is least bright which makes AD fall in between them. For that we used conditional gamma correction.

Disease	Brightness	Gamma value	Reason
AD	Medium	1.0	Neutral
PD	High	1.2	Slight darkening
FTD	Darkest	0.6	Brightening

Table 4.3: Applied Gamma Value Table

$$I_{\text{out}} = 255 \times \left(\frac{I_{\text{in}}}{255} \right)^{\frac{1}{\gamma}} \tag{4.1}$$

Here we use gamma correction >1 for making slices dark and <1 for making slices bright and to keep neutral we keep the value 1.

4.3.5 Contrast Enhancement with CLAHE

Unlike intensity checking, we also checked each class contrast to understand how sharp each class was, and these were our findings,

Disease	Contrast (std dev, max, mean)
AD	min=0.16, max=197.32, mean=94.67
PD	min=3.14, max=285.62, mean=112.78
FTD	min=3.14, max=285.62, mean=112.78

Table 4.4: Contrast Per Class Detection

From here, we understood that:

- PD class has high contrast, which suggests that it is sharper than other classes
- FTD slices have a low contrast value
- AD class is in the middle of PD and FTD

To ensure balance among all classes, we used enhancement using CLAHE so that our dataset is balanced for better training.

4.3.6 Resizing

All slices was converted to 224x224 for feeding it to our model.

4.3.7 Imbalance Handling

Since the amount of data was relatively small and class imbalance, several methods to balance model training were tried. First, SMOTE (Synthetic Minority Oversampling Technique) was used to create synthetic samples on behalf of the underrepresented classes. Also, class weight changes have been tried by assigning more weight to minority classes (i.e., while training). Both methods, however, did not prove to deliver stable or better performance, perhaps owing to the high feature sensitivity of the MRI slices and noise amplification by oversampling.
Finally, the most reliable and accurate results were obtained with a downsampling method because the majority classes were downscaled to be equal to the minority class. The technique provided a balanced division of classes and maintained the quality of anatomical features, which resulted in an enhanced model generalization and operation.

4.3.8 Sample for Explainability

To visualize the discriminative areas of the model to be interpreted post-hoc, Grad-CAM was created on 3 test participants per class (AD, PD, FTD) to define the regions affecting the prediction of the model.

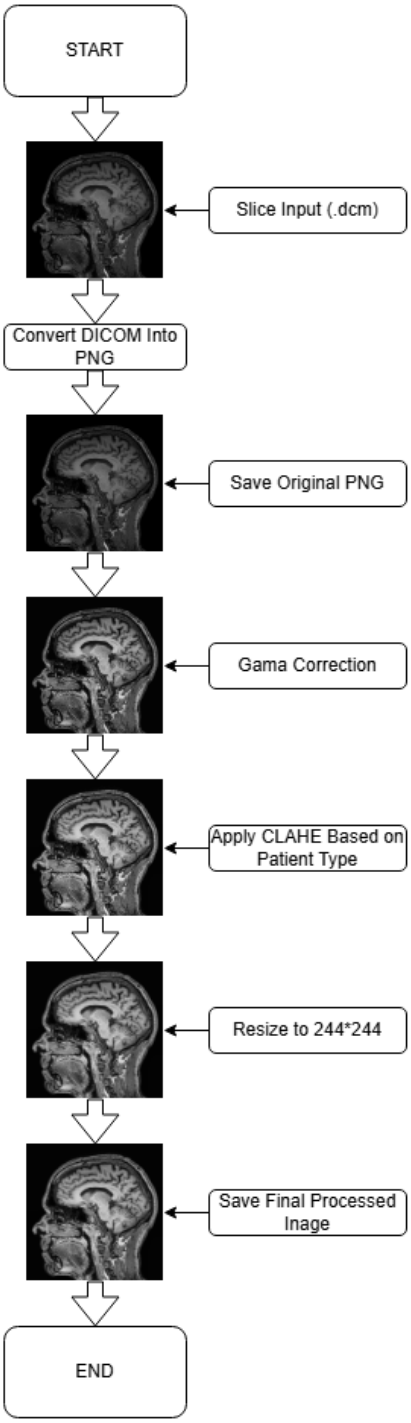

Figure 4.4: Pre Processing flow chart

4.3.9 Data Augmentation

The dataset was preprocessed, class balancing was done, and split into training, validation, and test sets under the stratified split, which maintains the ratio of classes. The final split ratio would be about 70 percent training, 15 percent validation, and 15 percent testing, visualized in fig.4.5

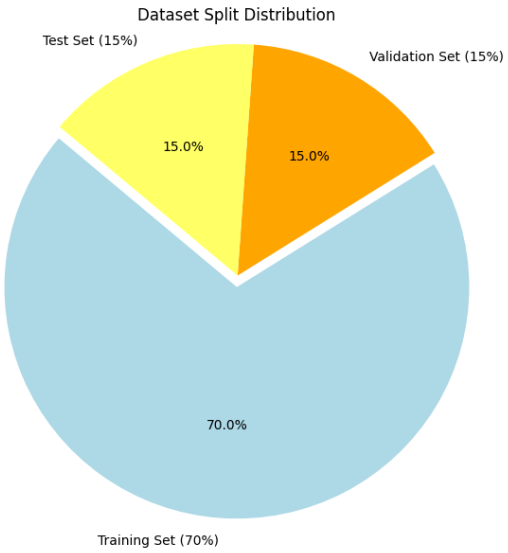

Figure 4.5: Data Distribution Pie-Chart

Data augmentation on the training set was only performed to enhance the generalization capacity of the model and eliminate overfitting given the small number of data in the dataset.

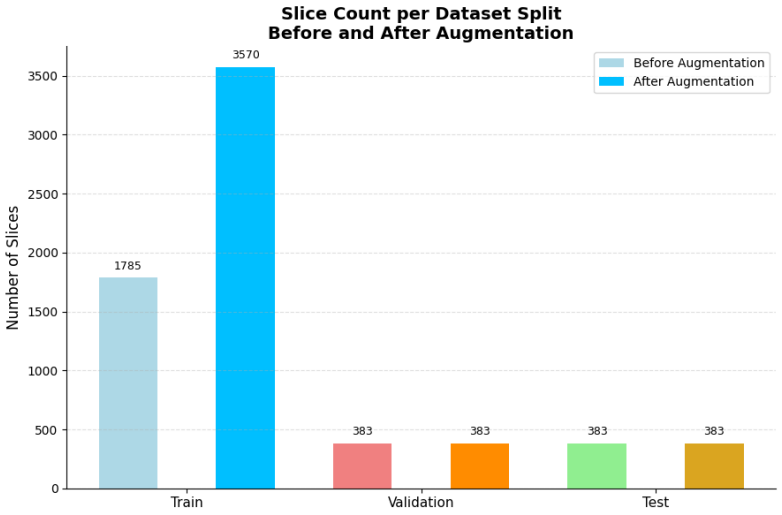

Figure 4.6: Train Augmentation Visualization

The augmentation methods involved: the small-angle rotations (10 o and 10 o), the zooming (0.9x-1.1x), the horizontal and vertical displacement (10x and 10y), and the intensity manipulation were used to induce variance of brightness and contrast. Such augmentations were well considered in such a manner as to preserve the clinical and anatomical integrity of the MRI slices. Fig.4.7 illustrates augmented slices that illustrate such transformations.

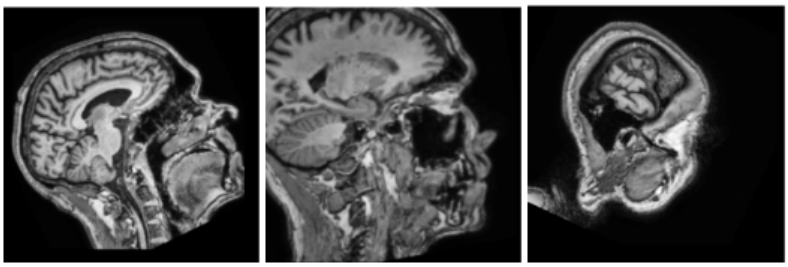

Figure 4.7: Data Augmentation Sample

The figure 4.8 illustrates Parameters of Augmentation.

Transform	Purpose
Resize((224, 224))	Resizes image to 224x224 for model input
ColorJitter(0.3, 0.3, 0.3, 0.3)	Adds variation in lighting and color
RandomRotation(20)	Adds rotation for orientation robustness
RandomAffine(degrees=20, translate=(0.1, 0.1), scale=(0.8, 1.1), shear=5)	Simulates translation, scaling, and distortion
RandomResizedCrop(224, scale=(0.7, 1.0))	Zoom-in effect and crop to reduce overfitting
RandomPerspective(distortion_scale=0.3, p=0.3)	Simulates perspective distortions
RandomGrayscale(p=0.1)	Converts 10% to grayscale for color invariance
ToTensor()	Converts image to PyTorch tensor format
Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])	Normalizes with ImageNet stats for consistency

Figure 4.8: Parameters of Augmentation

Chapter 5

Pre-trained Models

5.1 VGG19

VGG19 is a deep convolutional neural network (CNN) architecture introduced by the Visual Geometry Group at the University of Oxford. Simonyan and Zisserman proposed it in a 2014 paper. "Very Deep Convolutional Networks for Large-Scale Image Recognition" [13]. The model already has its fame in the area of uniform architecture and depth, and it becomes very useful in extracting the features in medical imaging, such as neurodegenerative diseases of the brain.

5.1.1 Architecture Overview

VGG19 has 19 layers in which the weights can be learned:

- 16 convolving layers
- 3 layers, which are fully connected
- ReLU activations between all convolutions
- Max pooling planes between each block
- Classification Softmax layer

All convolutional layers use 3×3 filters with a stride of 1 and padding to preserve spatial dimensions. The structure of the architecture is fixed, which is stated in fig:5.1

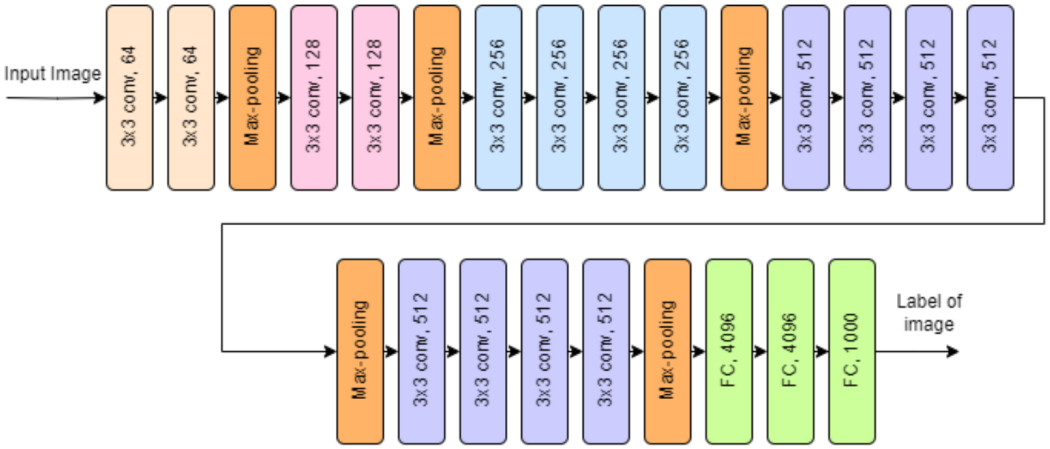

Figure 5.1: VGG19 Block Diagram

Due to its hierarchical capabilities in feature extraction, VGG19 can extract both low-level and high-level features (edges, textures, object shapes, and anatomical patterns), making it very useful in analyzing complex medical information, such as an MRI or a CT scan.

5.1.2 Application in Neurodegenerative Disease Detection

During the last years, VGG19 has been actively applied to the neuroimaging framework to identify such diseases as Alzheimer's Disease (AD), Parkinson's Disease (PD), and Frontotemporal Dementia (FTD). Its efficiency is proven by several studies:

- Swetha et al. (2023) trained VGG19 to classify Alzheimer's MRI data to reach an 89 percent accuracy rate, which made it competitive with more recent models, such as DenseNet169 [53].

- The accuracy of detection of VGG19 applied on PPMI MRIs to identify early situational conditions of Parkinson's Disease was at 91.5 percent, more accurate than other pre-trained networks [32] Florencia Esthero Zunino Sadiq, Imran, Noshin, Sulaiman, Khalid, and Malik (2020).

5.1.3 Insights from Pretraining Phase

During the pretraining phase using our dataset with this model, the train and validation accuracy, train and validation loss are illustrated in fig:5.2.

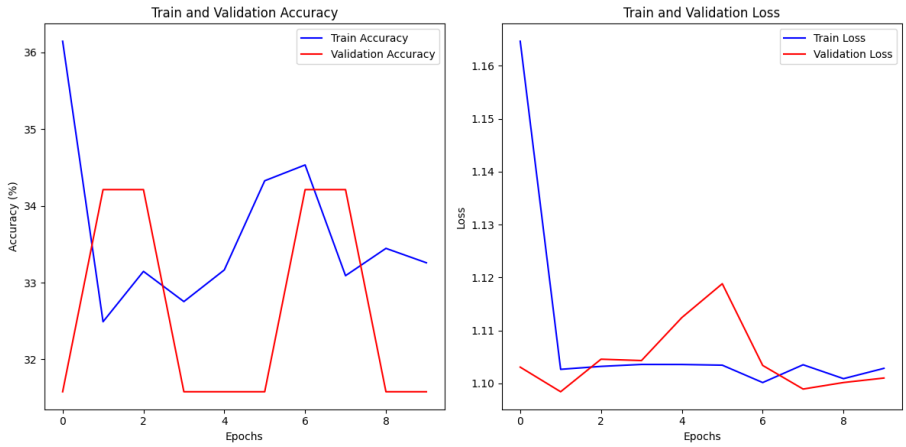

Figure 5.2: Train and Validation Accuracy, Train and Validation Loss

During the pretraining phase using our dataset with this model, the confusion matrix of test and validation is shown in fig:5.3.

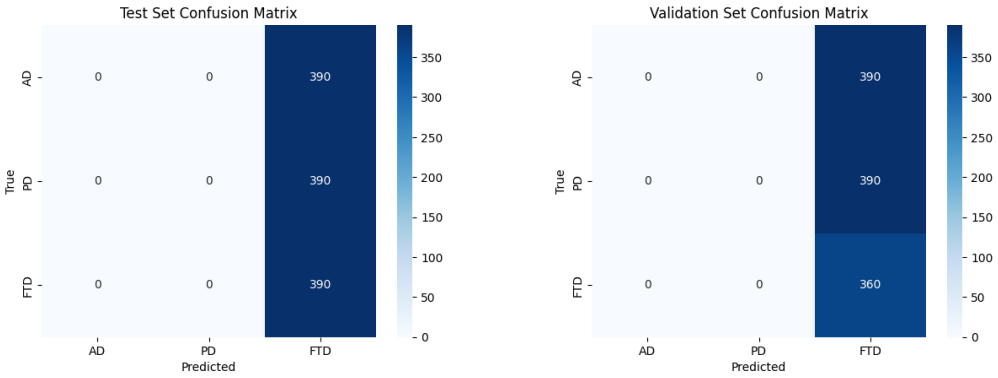

Figure 5.3: VGG19 Confusion Matrix of Test and Validation

Layer (type)	Output Shape	Parameters
input_1 (InputLayer)	[(None, 112, 112, 3)]	0
block1_conv1 (Conv2D)	(None, 112, 112, 64)	1792
block1_conv2 (Conv2D)	(None, 112, 112, 64)	36928
block1_pool (MaxPooling2D)	(None, 56, 56, 64)	0
block2_conv1 (Conv2D)	(None, 56, 56, 128)	73856
block2_conv2 (Conv2D)	(None, 56, 56, 128)	147584
block2_pool (MaxPooling2D)	(None, 28, 28, 128)	0
block3_conv1 (Conv2D)	(None, 28, 28, 256)	295168
block3_conv2 (Conv2D)	(None, 28, 28, 256)	590080
block3_conv3 (Conv2D)	(None, 28, 28, 256)	590080
block3_conv4 (Conv2D)	(None, 28, 28, 256)	590080
block3_pool (MaxPooling2D)	(None, 14, 14, 256)	0
block4_conv1 (Conv2D)	(None, 14, 14, 512)	1180160
block4_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block4_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block4_conv4 (Conv2D)	(None, 14, 14, 512)	2359808
block4_pool (MaxPooling2D)	(None, 7, 7, 512)	0
block5_conv1 (Conv2D)	(None, 7, 7, 512)	2359808
block5_conv2 (Conv2D)	(None, 7, 7, 512)	2359808
block5_conv3 (Conv2D)	(None, 7, 7, 512)	2359808
block5_conv4 (Conv2D)	(None, 7, 7, 512)	2359808
block5_pool (MaxPooling2D)	(None, 3, 3, 512)	0
flatten (Flatten)	(None, 4608)	0
dense (Dense)	(None, 2)	9218

Figure 5.4: Block-wise specification of the VGG19 architecture

5.2 ResNet50

5.2.1 Architecture Overview

ResNet-50 Architecture: He et al. (2015) [15] presented ResNet-50, a deep convolutional network of 50 layers, i.e.. 48 convolutional layers, 1 MaxPooling layer, and 1 AveragePooling layer. The architecture makes use of residual connections, which skip one or more of their layers, overcoming the problem of vanishing gradients and allowing deep networks to be trained effectively. The network commences on a 64 filter, 7x7 convolution layer with a stride of 2, followed by a 3x3 MaxPooling layer. Later layers include larger filters, such as 1x1 and 3x3 convolutions, increasingly fine-grained feature extraction, and allow a hierarchical learning of complicated patterns inside the input data. Deep architecture enables ResNet-50 to learn more and more abstract features of data as the network goes through various layers, shown in fig:5.5.

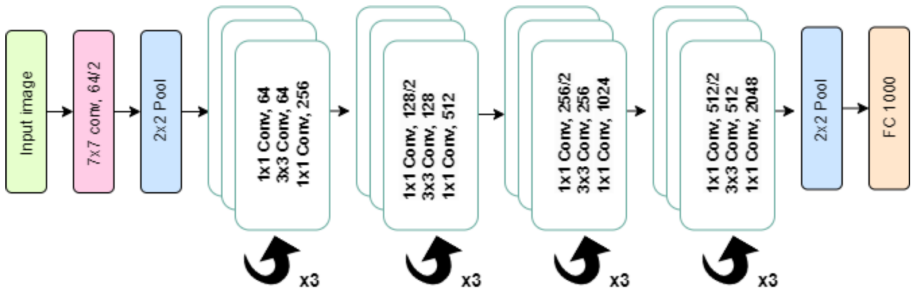

Figure 5.5: ResNet50 Block Diagram

Transfer Learning is one of the most useful benefits of ResNet-50 is the possibility to use transfer learning. This strategy enables pre-trained models on general-purpose large data sets, e.g., ImageNet, to be fine-tuned on specific medical tasks, e.g., the detection of Alzheimer's Disease (AD), Parkinson's Disease (PD), and Frontotemporal Dementia (FTD). Transfer learning assists in the transfer of information acquired in massive datasets to a smaller dataset that pertains to a field of interest. ResNet-50 is fine-tuned in terms of learning particular features in medical imaging by reusing weights that were previously trained to learn generic features of images.

5.2.2 Application in Neurodegenerative Disease Detection

- ResNet-50 has also been used in MRI scans in identifying brain atrophy, especially in parts like the hippocampus and the cortex, which are considered major indicators of the progression of Alzheimer's. ResNet-50 has been demonstrated to be very accurate in the classification of AD through the application of transfer learning [35] (Zhang et al., 2020).

- MRI scans of the PD patients have been analyzed with ResNet-50 in order to determine the structural changes in the brain. Such scans usually depict

degeneration of such structures as the Substantia Nigra, which contribute to controlling motions. These degenerative patterns are identified in the classification between PD and healthy controls with the assistance of ResNet-50 [27] (Liu et al., 2020).

- Frontotemporal Dementia (FTD): In the case of FTD, ResNet-50 has contributed to recognizing particular patterns of atrophy in the frontal and temporal lobes that are also not found in other neurodegenerative diseases [40] (Jiang et al., 2021).

5.2.3 Insights from Pretraining Phase

During the pretraining phase using our dataset with this model, the train and validation accuracy, train and validation loss are illustrated in fig:5.6.

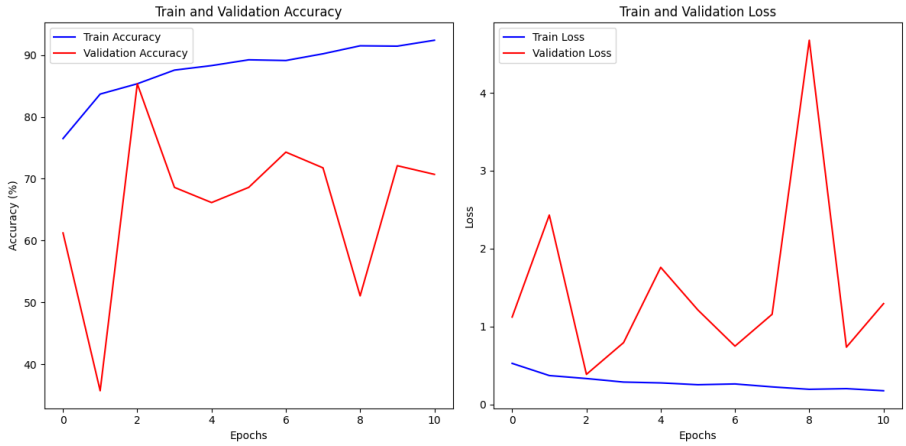

Figure 5.6: ResNet50 Accuracy and Loss

During the pretraining phase using our dataset with this model, the confusion matrix of test and validation is shown in fig:5.7.

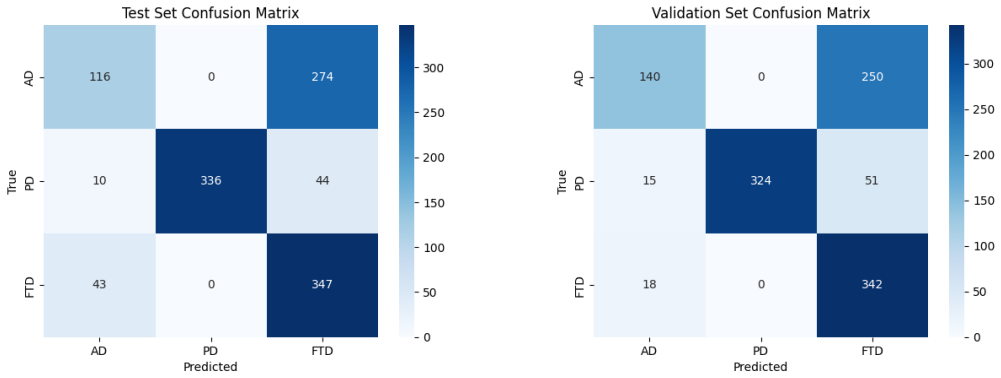

Figure 5.7: ResNet50 Confusion Matrix of Test and Validation

Layers	ResNet 50	Number of Layers
2D Convolutional Layer	7 x 7, 64, stride 2	1
2D Convolutional Layer	3 x 3 max pool, stride 2 [1 x 1, 64] x 3 [3 x 3, 64] x 3 [1 x 1, 256] x 3	9
2D Convolutional Layer	[1 x 1, 128] x 4 [3 x 3, 128] x 4 [1 x 1, 512] x 4	12
2D Convolutional Layer	[1 x 1, 256] x 6 [3 x 3, 256] x 6 [1 x 1, 1024] x 6	18
2D Convolutional Layer	[1 x 1, 512] x 3 [3 x 3, 512] x 3 [1 x 1, 2048] x 3	9
	Average Pool, 2	1

Figure 5.8: Block-wise specification of the VGG19 architecture

5.3 EfficientNetB0

5.3.1 Architecture Overview

Presented by Tan Le (2019), [26] efficientNet-B0 is a very efficient convolutional neural network that strives to ensure a high degree of accuracy and computational efficiency. It upscales all three parameters of the model (depth, width, and resolution) with a fixed set of scaling coefficients. This strategy is much better in performance with considerably fewer parameters compared to conventional CNNs.

- Compound Scaling: EfficientNet employs compound scaling, where the growth in depth, width, and resolution has been balanced. The model can be made to be more accurate with a lower number of parameters compared to conventional architectures such as ResNet and VGG.

- Mobile Inverted Bottleneck Convolutional (MBConv) The backbone comprises MBConv blocks and with depthwise separable convolutions and inverted residuals helps to reduce parameters significantly (Tan Le, 2019). This causes efficiency and preservation of model accuracy.

- Convolutional Layers: The architecture starts with a 3x3 convolution with 32 filters and a stride of 2, and then it continues with several MBConv blocks with successively more filter sizes. Every MBConv block contains a depthwise convolution, a pointwise convolution, and a skip connection to maintain the accuracy but minimize the parameters.

- Last Layers: Following the MBConv blocks, lastly, a Global Average Pooling (GAP) layer is then followed by a fully connected layer which predicts the output by passing the predictions calls upon classification tasks.

The EfficientNet-B0 model is computationally and memory efficient, which makes it appropriate to run in a mobile and cloud environment. The architecture of this model can be shown in fig:5.9

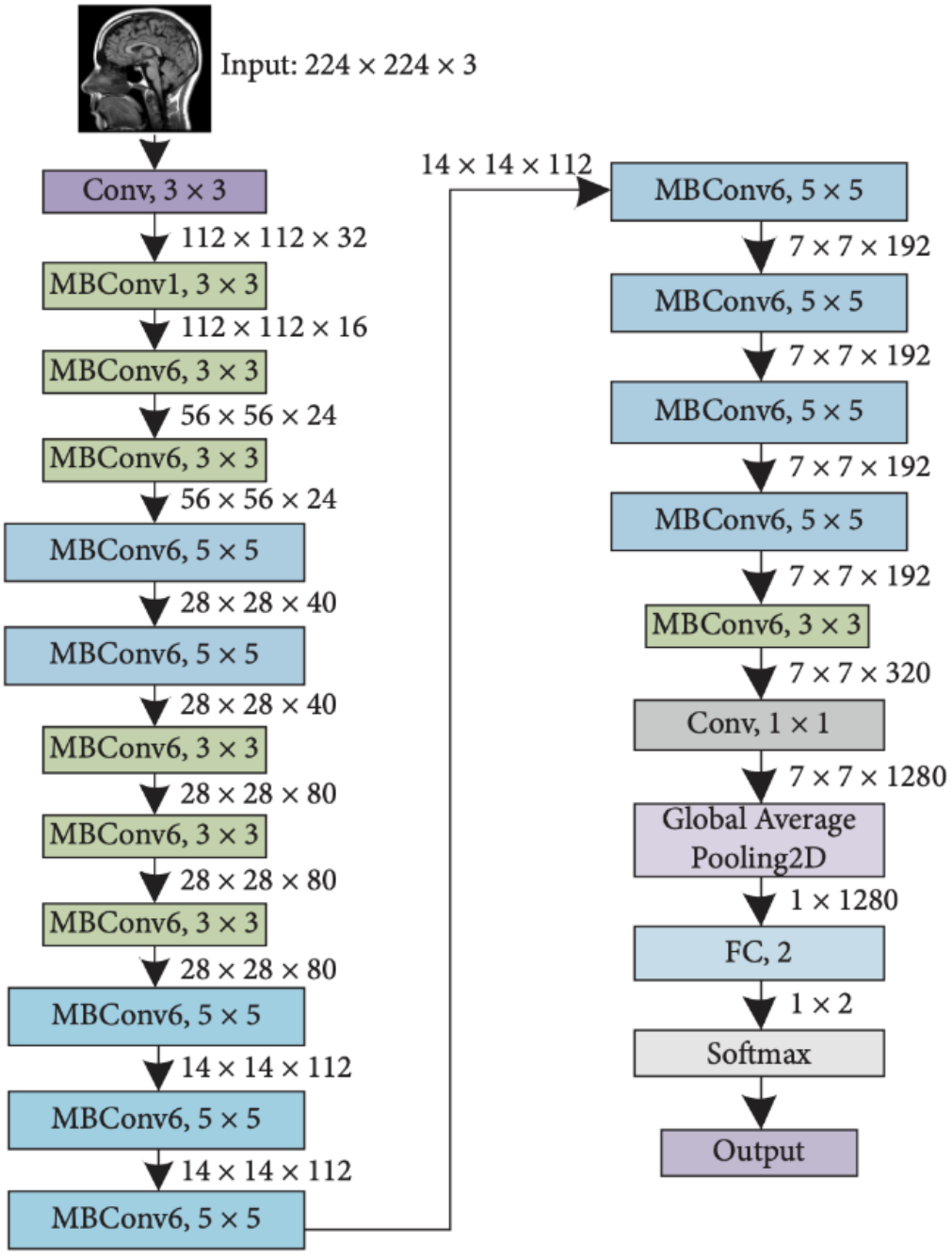

Figure 5.9: EfficientNetB0 Block Diagram

EfficientNet-B0 can utilize transfer learning, by first being trained at ImageNet and then being optimized on more domain-related tasks such as Alzheimer's Disease (AD) detection, Parkinson's Disease (PD) detection, and Frontotemporal Dementia (FTD) detection using MRI scans.

5.3.2 Application in Neurodegenerative Disease Detection

- Alzheimer's Disease (AD): EfficientNet-B0 is now utilized to identify the brain atrophy in cerebral parts such as the hippocampus and the cortex of the head in MRI scans. It has increased its powers to locate these important symptoms of Alzheimer's development through transfer learning [34] (Zhang et al., 2020).

- Parkinson's Disease (PD): MRI images have been classified with the model to diagnose Parkinson's disease in the brain, where the Sbstantial Nigra is the part of interest, which is vital in motor control and compromised in PD [28] (Liu et al., 2020).

- Frontotemporal Dementia (FTD): Using EfficientNet-B0, frontal and temporal brain atrophy can be detected, which allows differentiating Frontotemporal Dementia (FTD) and other neurodegenerative disorders [39] (Jiang et al., 2021).

5.3.3 Insights from Pretraining Phase

During the pretraining phase using our dataset with this model, the train and validation accuracy, train and validation loss are illustrated in fig:5.10.

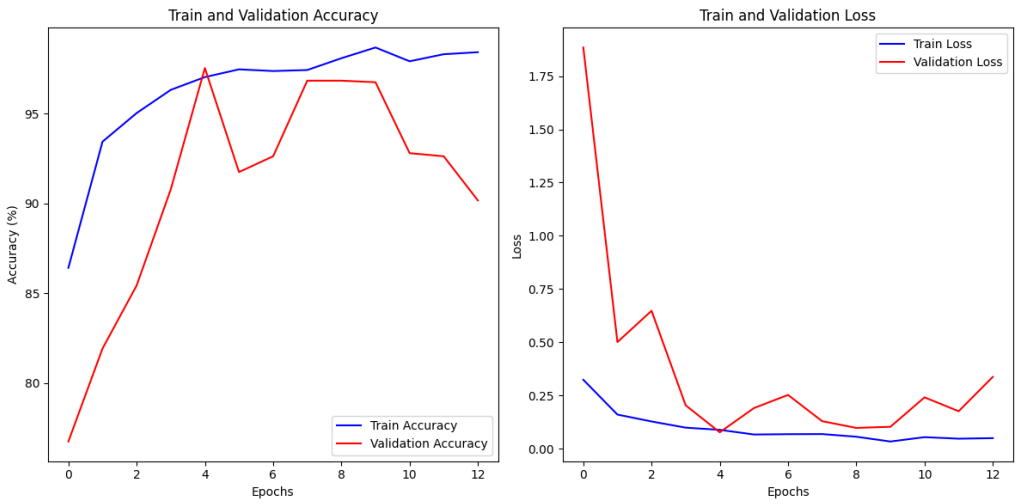

Figure 5.10: EfficientNetB0 Accuracy and Loss

During the pretraining phase using our dataset with this model, the confusion matrix of test and validation is shown in fig:5.11.

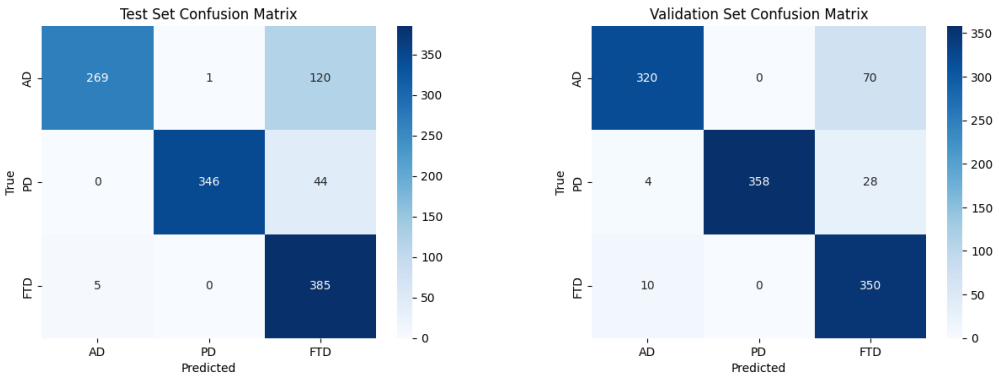

Figure 5.11: EfficientNetB0 Confusion Matrix of Test and Validation

Stage	Operator	Resolution	#Channels	#Layers
1	Conv3x3	224 × 224	32	1
2	MBConv1, k3x3	112 × 112	16	1
3	MBConv6, k3x3	112 × 112	24	2
4	MBConv6, k5x5	56 × 56	40	2
5	MBConv6, k3x3	28 × 28	80	3
6	MBConv6, k5x5	14 × 14	112	3
7	MBConv6, k5x5	14 × 14	192	4
8	MBConv6, k3x3	7 × 7	320	1
9	Conv1x1 & Pooling & FC	7 × 7	1280	1

Figure 5.12: Block-wise specification of the EfficientNetB0 architecture

5.4 DenseNet

A dense convolutional neural net (Densely-Connected Convolutional Network - DenseNet) is a neural net architecture presented by Huang et al. in 2017. In contrast to usual CNNs, where the connection between the layers follows a chain. DenseNet links all layers to each other in a feed-forward structure. The pattern of this connectivity leads to the reusing of its features, better gradient flow, and reduction of parameters as compared to comparable deep networks such as ResNet [19].

5.4.1 Architecture Overview

In a DenseNet, every layer has access to the feature maps of all previous layers as input and makes all its feature maps available to all subsequent layers. When there are l layers, then the direct connections are l(l + 1)/2.

- Dense Block: Block containing a set of layers and a concatenation of their outputs.
- Transition Layers: They are utilized to shrink feature map size and to impose pooling.
- Growth Rate (k): It regulates the number of filters being introduced by the layers. Standard ones are 12, 24, or 32.

The architecture of DenseNet is shown in fig:5.13

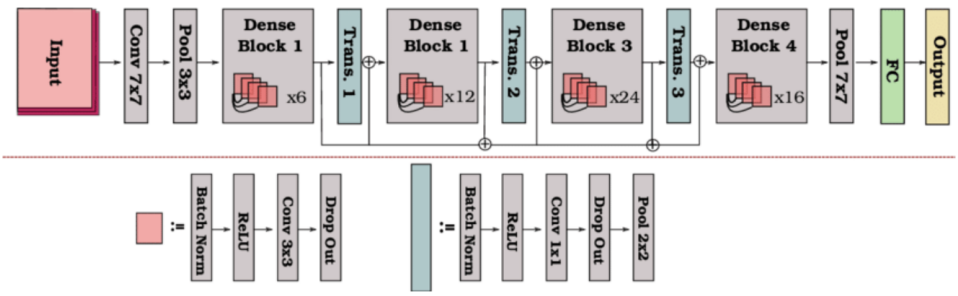

Figure 5.13: DenseNet Block Diagram

5.4.2 Application in Neurodegenerative Disease Detection

- Nguyen et al.(2022): proposed a deep grading + DenseNet ensemble to discriminate patients with AD, FTD, and CN based on structural MRI. The accuracy of their model was excellent on multi-center data and also resulted in explainable maps of atrophy [46].

- Kandhro et al. (2024): Hyperparameter Optimization of a Combined DenseNet and Grey Wolf Optimization (GWO) on MRI and SPECT data. Obtained a near-perfect modelling of the AUC of 1.0 between PD and CN groups [60].

5.4.3 Insights from Pretraining Phase

During the pretraining phase using our dataset with this model, the train and validation accuracy, train and validation loss are illustrated in fig:5.14.

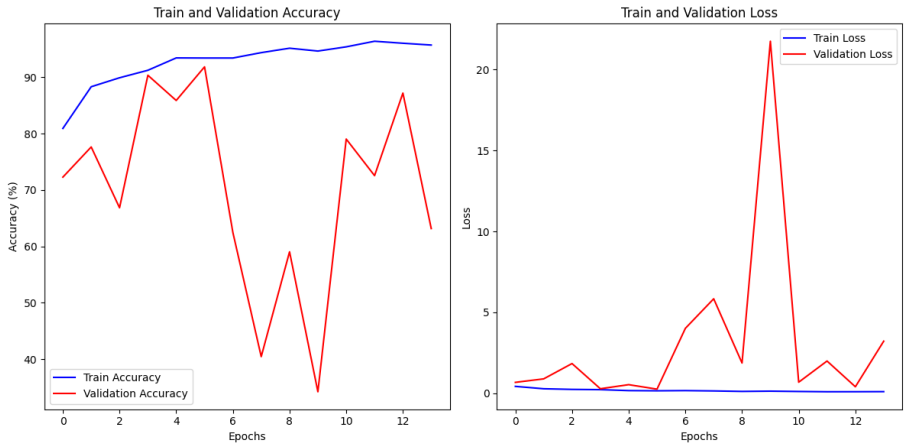

Figure 5.14: DenseNet Accuracy and Loss

During the pretraining phase using our dataset with this model, the confusion matrix of test and validation is shown in fig:5.15.

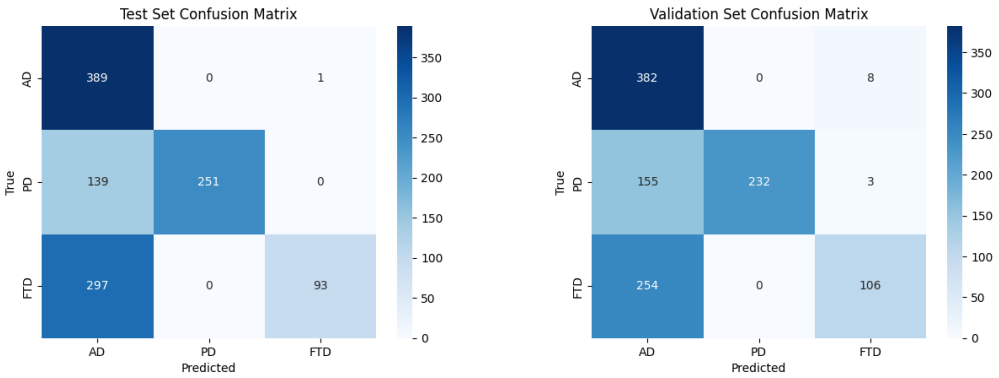

Figure 5.15: DenseNet Confusion Matrix of Test and Validation

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6
Transition Layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12
Transition Layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	[1 × 1 conv, 3 × 3 conv] × 24	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 48	[1 × 1 conv, 3 × 3 conv] × 64
Transition Layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	[1 × 1 conv, 3 × 3 conv] × 16	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 48
Classification Layer	1 × 1	7 × 7 global average pool			
		1000D fully-connected, softmax			

Figure 5.16: Block-wise specification of the DenseNet architecture

5.5 MobileNetV2

MobileNetV2 is a versatile and low-weight convolutional neural network architecture presented by Sandler et al. (2018). The MobileNetV2 was built specifically with mobile and resource-limited settings in mind, so it is very applicable in the scope of various medical imaging applications where the resources, fine-tuning flexibility of deployment, and overall performance of the model are vital considerations [23].

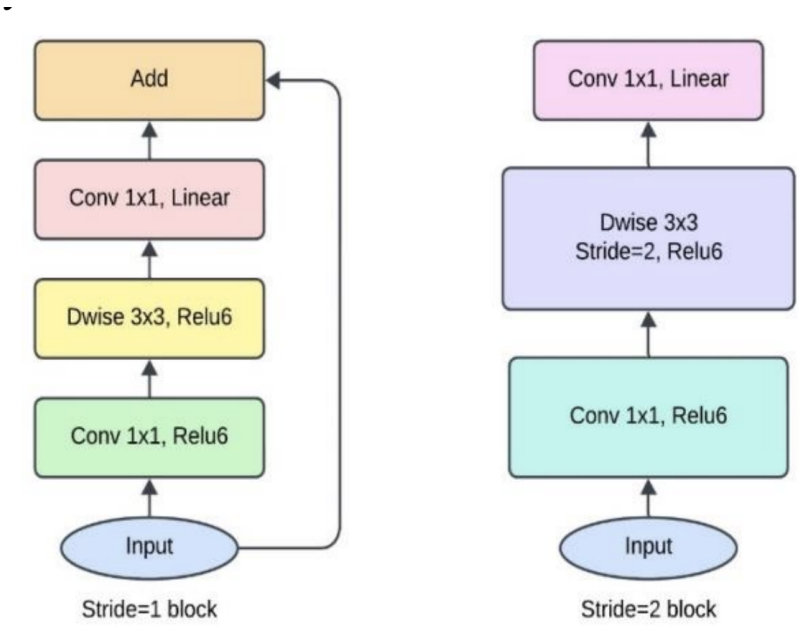

Figure 5.17: MobileNetV2 Block DIagram

5.5.1 Architecture Overview

MobileNetV2 shown in fig:5.17 is an extension of the original MobileNet that is made according to two main innovations:

Linear Bottleneck Inverted Residual Blocks

MobileNetV2 does not use conventional residual connections (e.g., ResNet) but inverted residuals:

- The input is expanded to a space of higher dimension.
- Depth-wise separable convolution is applied.
- Using a linear (non-ReLU) bottleneck, the result is regressed to the low-dimensional space.

By minimizing memory and computation, the structure loses little representational power.

Depthwise Separable Convolutions

This makes the convolution consist of two steps:

- Depthwise convolution (spatial filtering per channel)
- Pointwise convolution (1x1 convolution mix channels)

This reduces the computational expense by approximately 8 to 9 times compared to traditional convolutions, making the model incredibly efficient.

Summary of Architecture

- Much less than 3.4M parameters than ResNet50 or VGG19
- Input size: is usually 224x224x3
- Number of convolutional layers: ~53
- Compatible with transfer learning via ImageNet.

5.5.2 Application in Neurodegenerative Disease Detection

- A multi-person test conducted by Anagnostis et al. (2024) facilitated the training of a hybrid architecture that detects early Alzheimer's disease using MRI by utilizing MobileNetV2 and ResNet50. An accuracy of 96.19 per cent on ADNI and 99.82 per cent on the NACC dataset was achieved [57].
- Few works were found that discussed using MobileNetV2 in SPECT-based classification, and compared with ResNet and VGG16. It exhibited good performance because it infers faster and it is of similar accuracy [25].
- The direct applications of the MobileNetV2 in the context of FTD are not as widespread yet, but it appears effective in the process of AD vs CN identification, and it is therefore likely that it has great potential in terms of differentiating between the AD and the FTD, which are known to be hard to distinguish.

5.5.3 Insights from Pretraining Phase

During the pretraining phase using our dataset with this model, the train and validation accuracy, train and validation loss are illustrated in fig:5.14.

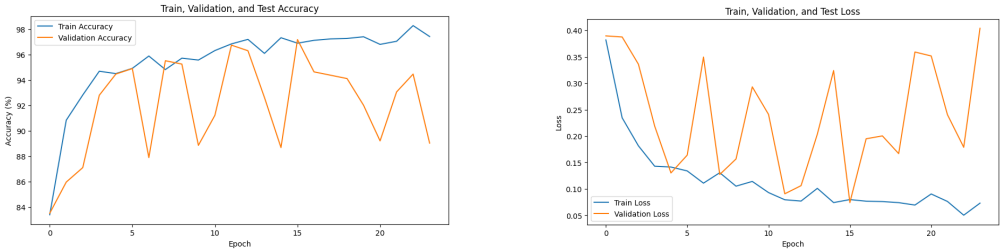

Figure 5.18: MobileNetV2 Accuracy and Loss

During the pretraining phase using our dataset with this model, the confusion matrix of test and validation is shown in fig:5.15.

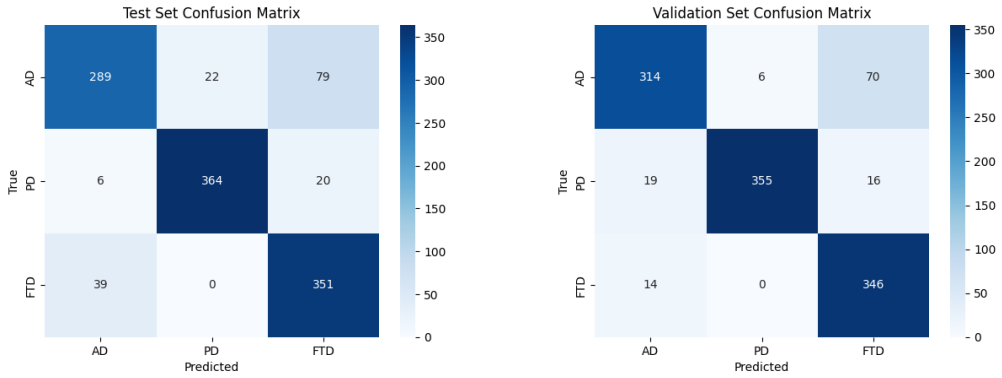

Figure 5.19: MobileNetV2 Confusion Matrix of Test and Validation

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	

Figure 5.20: Block-wise specification of the MobileNetV2 architecture

Chapter 6

Proposed Methodology

6.1 Overview

The model we are proposing is a customized deep learning architecture which we named SadNetV1. It is designed to distinguish Alzheimer's Disease (AD), Parkinson's Disease (PD), and Frontotemporal Dementia (FTD) from MRI data by using a single model which was never done before as of our knowledge. This new approach will be able to overcome the challenge of classifying multiple neurodegenerative disease which was not explored by most of the researchers. Our model was built by using MobileNetV2 as backbone which was later enhanced by Squeeze-and-Excitation Block (SEBlock) and the Spatial Attention Mechanism. MobileNetV2 was the best performing model when we were in our training pre-trained model phase and as a result it was chosen. Squeeze-and-Excitation Block (SEBlock) is used for channel wise attention and Spatial Attention Mechanism for focusing on spatial features of dataset. By combining these two components and incorporeting them, our model was able to differentiate diseases with almost similar characteristics. This customized model sets a bench mark when it comes to efficiency and effectiveness on classification of multi disease.

6.2 Model Architecture

MobileNetV2 architecture is known for usage of depthwise separable convulotions which helps to reduce number of parameter and its gives low computational cost which makes it ideal for medical image classification. This architecture have several inverted residual blocks to extract feature efficiently and keeps the model lightweight. Pretrained MobileNetV2 showed severe spike but in our proposed model, spike was minimized. In SadNetV1 a dropout layer applied after feature extraction to prevent overfitting and there is final classifier which maps the extracted features into our 3 output classes which are AD, PD and FTD. Our model takes MRI scan size of 3x224x224 and process it through series of layers. The spatial dimensions are reduced to 32x112x112 after initial layer of consists of Conv2d, BatchNorm and ReLU6 layer. Several InvertedResidual blocks helped to extract feature while progressively reducing spatial size and increasing the number of channel. Feature map is 1280x7x7 when the model reaches SEBlock and Spatial Attention Mechanism. These blocks focuses on spatial regions which are relevent and spatial size remains same. Then come the flatten layer and after that comes dropout for regularization. The final

layer which is Linear Classifier maps learned features to 3 output classes. By this structure Our model efficiently detects neurodegenerative disease.

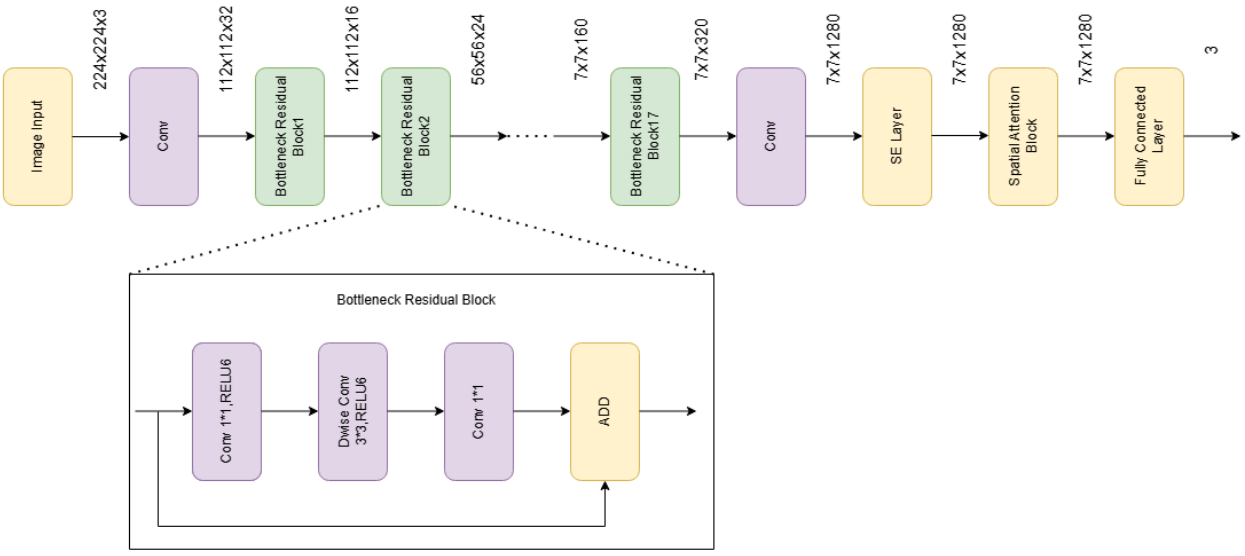

Figure 6.1: SadNetv1 Architecture Diagram

Layer Type	Input Size	Output Size	Kernel Size	Stride	Expansion Factor
Initial Conv	224x224x3	112x112x32	(3x3)	(2x2)	-
Inverted Residual Block	112x112x32	112x112x16	(3x3)	(1x1)	1
Inverted Residual Block x2	112x112x16	56x56x24	(3x3)	(2x2)	6
Inverted Residual Block x3	56x56x24	28x28x32	(3x3)	(2x2)	6
Inverted Residual Block x4	28x28x32	14x14x64	(3x3)	(2x2)	6
Inverted Residual Block x3	14x14x64	14x14x96	(3x3)	(1x1)	6
Inverted Residual Block x3	14x14x96	7x7x160	(3x3)	(2x2)	6
Inverted Residual Block x1	7x7x160	7x7x320	(3x3)	(1x1)	6
Final Conv	7x7x320	7x7x1280	(1x1)	(1x1)	-
Squeeze-and-Excitation Block	7x7x1280	7x7x1280	-	-	-
Spatial Attention Block	7x7x1280	7x7x1280	(7x7)	(1x1)	-
Dense Layer (Classifier)	7x7x1280	3	-	-	-

Figure 6.2: Layer Specification for SadNetv1

6.3 Data Preprocessing and Augmentation

In SadNetv1 model, data augrumentation serves as a critical technique allowing the increase in training data diversity and avoid overfitting. Data are augmented in a variety of ways; random rotations, affine transformations, color jittering, and perspective distortions. These transformations approximate various simulations of variations of the input data that in turn assists the model to better generalize across many possible input scenarios. The augmentation methods assist the model to be more tolerant with nominal difference image orientation, color, and perspective, which is significant as well to medical imaging tasks in which the variations are possible when medical scans differ in conditions or patient orientation. Also, the data set is divided into training (70 percent), validation (15 percent), and test (15 percent) categories. This division ensure that model trains on large subset of the dataset and still keeping enough data for validation and testing purpose. The training set is focused on optimization of the model parameters, the validation set assists in tracking performance of the model during the training, whereas the test set is applied to the assessment of the final model generalization capability. This practice aids in evaluating the strength of the model on unknown data, which will improve the generalization of the model to the practical situation.

6.4 Squeeze-and-Excitation Block (SEBlock)

The Squeeze-and-Excitation Block (SEBlock) improves the performance of a model as it re-calibrates the channel-wise feature responses so that the network can emphasize higher informative features. The SEBlock is works by globally average pooling (GAP) to the input feature map to produce a single value across the number of channels (i.e., to squeeze out the spatial dimensions of the feature map and aggregate the global signal across all the feature maps). These are then preceded by two fully connected layers which are channel attention mechanisms. Each of these layers generates an array of weights per channel, which are then sent back to the input feature map by performing element-wise multiplication so that channels containing the most insightful information become amplified and channels with low value deactivated. Mathematically the process can be explained as follows:

Global Average Pooling (GAP)

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{c,i,j} \tag{6.1}$$

Fully Connected Layers

$$s = \sigma\left(W_2 \cdot \mathrm{ReLU}\left(W_1 \cdot z\right)\right) \tag{6.2}$$

Recalibration

$$\hat{x}_c = x_c \cdot s_c$$
where $\hat{x}_c$ is the recalibrated feature map for each channel. (6.3)

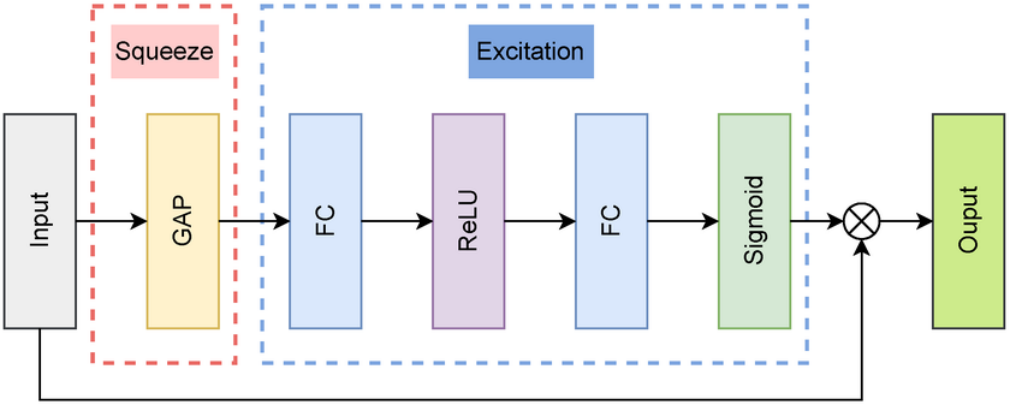

Figure 6.3: Flow of SE Block

6.5 Importance of Channel Attention

With the implementation of channel attention, the SEBlock enables the model to only highlight the critical information of the MRI scans. This feature is especially helpful in medical image classification, as in neurodegenerative diseases particular parts of the scans (i.e. parts of the brain) are more likely to signal the condition.It is possible to concentrate on these important features makes the model more resistance to noise or irrelevant data, because the attention mechanism will decrease the influence of less informative features. This will cause better performance of models, and it will teach the network more pertinent patterns and better generalize and come up with more accurate predictions.

6.6 Spatial Attention Mechanism

The Spatial Attention Mechanism works by emphasizing the spatial points of interest in the input MRI images to enable them (the model) to identify anatomical areas that are indicative of disease. The underlying of this mechanism is that, firstly, average pooling and max pooling are performed along the channel dimension of the feature map to retrieve the information in the spatial domain. The concatenation of these pooled features is then feed to 7x7 convolutional layer to form a spatial attention map. Attention map identifies the most relevant spatial areas and these identified areas are multiplied element-wise to the input feature map to adjust the spatial features. Mathematically this process is given by the following:

6.6.1 Average and Max pooling

Average pooling and max pooling are applied along the channel dimension to the input feature map $X \in R^{C \times H \times W}$, where C is the number of channels, H is the height, and W is the width of the feature map.

$$A_{avg} = \text{AvgPool}(X), \quad A_{\max} = \text{MaxPool}(X) \tag{6.4}$$

where A average and A max are the average and the max pooled feature maps respectively.

6.6.2 Concatenation and Convolution

he pooled feature maps along the channel axis are concatenated between the average and the max feature maps

$$A_{\text{concat}} = \text{Concat}\left(A_{\text{avg}}, A_{\text{max}}\right) \tag{6.5}$$

The concatenated map undergoes a 7x7 convolution operation to give the spatial attention map

$$A_{\text{spatial}} = \sigma\left(\text{Conv}\left(A_{\text{concat}}\right)\right) \tag{6.6}$$

where σ is a sigmoid activation function, and A_{spatial} is the resulting spatial attention map.

6.6.3 Recalibration

This output is re-calibrated by utilizing the element-per-element input feature map multiplied by the spatial attention map.

$$\widehat{X} = X \cdot A_{\text{spatial}} \tag{6.7}$$

where $\widehat{X}$ is the recalibrated feature map.

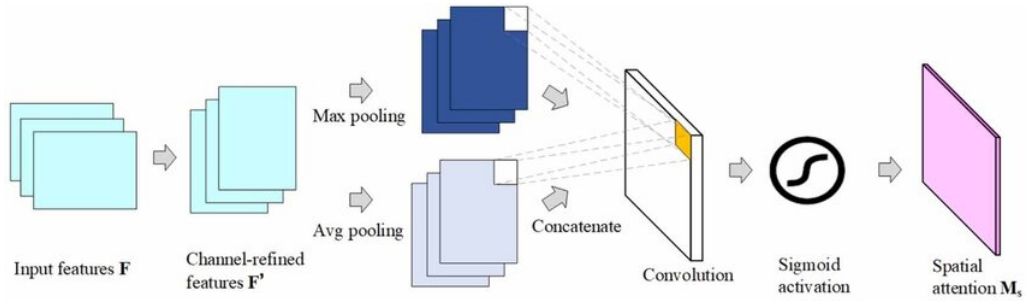

Figure 6.4: Flow of Spatial Attention Block

The Spatial Attention Mechanism directs the model to pay attention to useful spatial areas in the MRI reports and this is very crucial in locating an important anatomical area that signifies presence of diseases such as Alzheimer, Parkinson, and Frontotemporal Dementia. To take an example, there are regions of the brain that are unique in their way of atrophy, this atrophy is important in the diagnosis of neurodegenerative problems. Differentiating by areas thus makes the model ROBUST and accurate in its predictions.

6.7 Dropout for Regularization

In SadNetv1 classifier, Dropout layer is added to avoid overfitting by training at a rate of 50 percent. Dropout randomly discards (turns off) a certain percentage of the neurons of the network in every training iteration. This causes the model to depend on a wider range of features as opposed to OVERfitting on individual neurons or data patterns. Consequently, the Dropout procedures assist the network

to learn redundant representations, and consequently the network is better able to generalize well since training data are not seen. This is of particular importance in cases where working with small amounts of data, because it causes the model to avoid learning details of the training sample and ENCOURAGES learn more generalizable, transferable patterns. Dropout can make the model fit real world data set better by mitigating overfitting, so that the model does not become to specific to the training data distribution.

6.8 Learning Rate and Optimization Strategy

In SadNetv1, the learning rate of 1e-5 is adapted in the fine-tuning of pre-trained MobileNetV2. The reasoning behind it is since fine-tuning a pre-trained model involves minor changes in the weights and a low learning rate is useful in avoiding high updates that may corrupt the previously learned features. With a small learning rate, the model adapts the new task gradually without reverting to the lost knowledge gained in pre-training. The Adam variant of optimizer is selected because it is efficient in processing sparse gradients and learning rates are adjusted on individual parameters. Adam especially works well with deep learning applications as it incorporates both advantages associated with the momentum and adaptive learning rates so that it remains resistant to the noisy gradients and can have a faster convergence. Also, the ReduceLROnPlateau scheduling policy is applied to change the learning rate on training. Once validation loss stops improving, the scheduler will decrease the learning rate to make sure the model does not continue past the optimal point and make sure that the model settles at a good local minimum. This plan allows the model to more precisely fine-tune and prevents overspecification and enhances generalization.

6.9 Early Stopping

The SadNetv1 model employs pre-training and early stopping, where the patience value to halting training in the model is 10, with indication of poor performance on the validation loss after a certain number of epochs. This implies that in case the performance of the model on the validation set does not tend to improve after the 10th epoch, early stopping is employed. The first advantage of this method is to prevent overfitting which might happen when the model keeps finding particular patterns in the training set that do not generalize well to new data. Early stopping is useful to conserve the computations by eliminating the overtraining that the model does not perform significantly better. The method of stopping the process of training the model at the optimal point, improves its generalization, so that the model is not regarded as too long trained and overfit to the training set.

6.10 Loss Function

The loss function used in the our model is Cross-Entropy Loss because it is effective in multi-class classification task. Because Cross-Entropy Loss computes the difference between the predicted class labels and the actual one, it is specifically

appropriate in the problem domain in which the task is to classify an input to a single one of multiple categories. It penalizes wrong predictions by computing the logarithmic loss between the predicted and true distributions thus making the model to make more confident and accurate predictions. This makes the Cross-Entropy Loss a better fit to classify diseases, such as Alzheimer Disease (AD), Parkinson Disease (PD), and Frontotemporal Dementia (FTD), based on obtained MRI scans since the model directly focuses on predicting the most likely class of each given input.

$$L = -\sum_{i=1}^{C} y_i \log(p_i) \tag{6.8}$$

Where, L is considered as the single sample loss. C = classes (in your case 3 classes: AD, PD, and FTD). y i is the actual label of class i and it is one hot encoded. p i is the probability that is estimated by the model and assigned to the class i (that is the model output after putting the softmax on the logits).

6.11 Explainable AI (XAI)

Explainable AI (XAI) is a combination of tools and strategies that can help users understand more about the decision-making of the artificial intelligence (AI) system and become more transparent to humans. Although conventional machine learning systems, especially deep learning systems, can be termed as black-box as their decision process is highly obscure and complicated, XAI aims at giving clear explanations on how these systems make their predictions. This is especially relevant in the areas of health, finance, and law, where it is important to be able to trace how an AI made this decision in the context of trust, accountability and laws. Local explanation methods like LIME and SHAP are XAI methods that have practitioners visualize the features or areas of the input that were used to make the output of the model. Other forms of XAI include features that can visualize the input data to show which parts of the input were used (such as Grad-CAM, a visualization tool that is called a Gradient-weighted Class Activation Mapping.) XAI does not only aid users to be confident in the AI system, but can also offer important insights into how a model behaves, which should eventually result in higher-quality model improvements and more safe and reliable use in risk-sensitive tasks.

6.12 Grad Cam

Grad-CAM (Gradient-weighted Class Activation Mapping) has been introduced as a method of visualizing explanations of the convolutional neural networks (CNNs) predictions. It indicates areas of its input image that had the greatest influence in a model decision. Grad-CAM is composed of a gradient calculation of the target class score with respect to feature maps of the last convolutional layer. These gradients are average over the whole image to get a set of weights on each feature map, which areused to make a weighted sum of the feature maps. The outcome is a heatmap specifying what parts of the picture were most significant in classification. The heatmap is then passed through a ReLU operation keeping only the positive values,

as a result, it could be overlaid over the original image. Grad-CAM also contributes to the interpretability of a deep learning model because it visually shows what areas of an image contributed to its decisions, and thus it is quite beneficial in disciplines such as medical imaging.

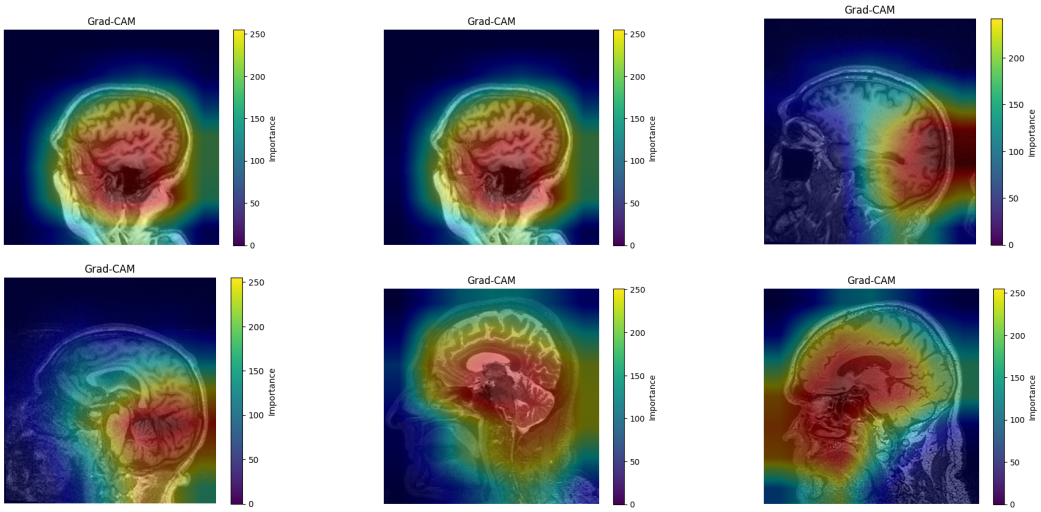

Figure 6.5: Heatmaps From Grad Cam

Chapter 7

Result and Analysis

7.1 Model Evaluation Overview

In order to test the effectiveness of different CNN structures, we used 3D sagittal MRI slices that had been labelled into three different classes namely Alzheimer Disease (AD), Parkinson Disease (PD) and Frontotemporal Dementia (FTD). The accuracy of each model according to the classification ability on the train, validation and test datasets was evaluated. The experiment was done to determine the best and effective model of neurodegenerative diseases identification.

7.2 Performance Comparison of CNN architecture

Various pre-trained networks were used, i.e., VGG19, ResNet50, DenseNet121, EfficientNetB0 and MobileNetV2. Both of the models were trained and tested with the same dataset that was divided consistently. Results are summarized in the following table-

Model	Test accuracy(%)	Train accuracy(%)	Validation accuracy(%)
VGG19.	33.33	33.26	31.58
ResNet50.	68.29	92.40	70.70
DenseNet12.	62.65	95.71	63.16
EfficientNetB0.	85.47	98.43	90.18
MobileNetV2.	94.36	98.30	94.47
SadNetV1	96.15	96.84	97.11

Figure 7.1: Comparison of Different model with our proposed model

	Test			Train			Validate		
Model	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
MobineNetV2	0.7947	0.7673	0.7788	0.9822	0.9840	0.9823	0.7593	0.7315	0.7440
ResNet50	0.5658	0.4106	0.4503	0.9219	0.9279	0.9200	0.5735	0.4298	0.4672
VGG19	0.3328	0.3378	0.3347	0.2156	0.3351	0.2164	0.3142	0.3194	0.3162
EfficientNet B0	0.6979	0.6249	0.6481	0.9824	0.9867	0.9836	0.6426	0.5953	0.6159
DenseNet121	0.6450	0.4757	0.5102	0.9542	0.9624	0.9550	0.6167	0.4519	0.4908
Proposed Model	0.8761	0.8574	0.8652	0.9625	0.9744	0.9635	0.8981	0.8835	0.8899

Figure 7.2: Precision, Recall and F1 score of test, train and validation of all models

In order to fully assess the performance of the classification of the CNN architectures, we used the precision, the recall, and the F1-score.
These are especially significant in our application which classifies neurodegenerative diseases using MRI slices where the spatial attributes are very important. Precision tells us that the positive predictions generated by the model (e.g. predicting Alzheimer or Parkinson) are valid, and we are unlikely to declare a false alarm. Recall is crucial as failure to detect a positive case (false negative) may be highly dangerous during early detections of the diseases. The F1-score combines the two of them giving a single quantification of model performance when classes are unbalanced or overlapping. Due to the complexity and the fine differences in the spatial domain of MRI data the accuracy predictor may be misleading, thus it was measured using these more discriminating metrics which assess the performance of the model in a holistic way. The formulas of precision, the recall and the F1-score are as follow-

Precision Formula

$$\text{Precision } = \frac{TP}{TP + FP} \tag{7.1}$$

Recall Formula

$$\text{Recall } = \frac{TP}{TP + FN} \tag{7.2}$$

F1-score Formula

$$\text{F1 } - \text{ score } = 2 \times \frac{\text{Precision } \times \text{ Recall}}{\text{Precision } + \text{ Recall}} \tag{7.3}$$

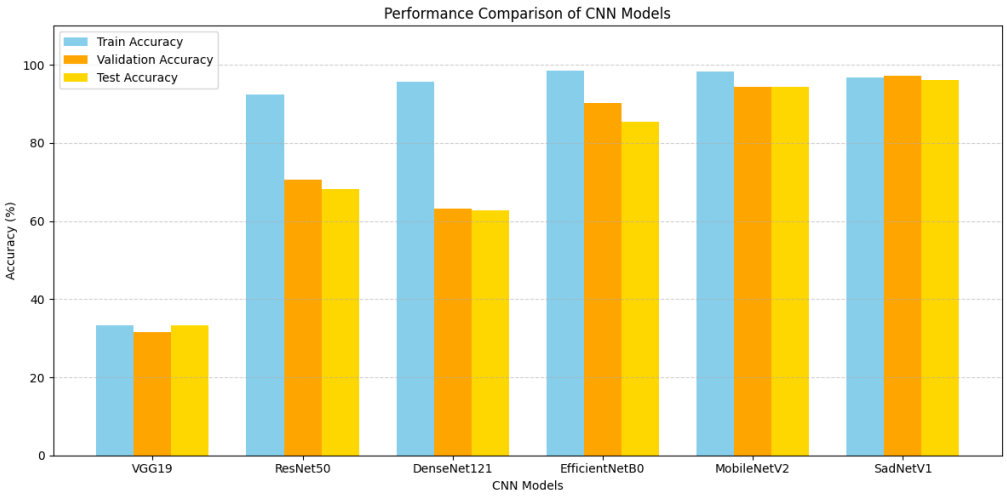

Figure 7.3: Comparison of Proposed Model with Other Pre-Trained Model

As we can see, MobileNetV2 performed better in all of the standard pre-trained models, but it led to overfitting. However, the SadNetV1 model proposed attained the optimal compromise between training, validation, and test accuracy.

7.3 Custom Model (SadNetV1 Evaluation)

SadNetV1 represents an extremely lightweight personalised CNN model, which is a modified version of MobileNetV2 with Squeeze-and-Excitation (SE) blocks and Spatial Attention modules. This enabled the model to actually pay considerations to disease-related spatial and channel-wise textures in the MRI slices. The analysis was able to attain:

- Train Accuracy: 96.84%
- Validation Accuracy: 97.11%
- Test Accuracy: 96.15%

Such outcomes imply that SadNetV1 does not overfit, and it can easily be used in clinical practice to classify multi-disease cases.
From observation we can say that the proposed SadNetV1 performed best out of all baseline models on training, validation and test data, registering test accuracy of 96.15 percent and effective generalization revealed by train accuracy of 96.84 percent and validation accuracy of 97.11 percent. It had a very small architecture and had attention building in it which enabled it to learn low-level and important spatial and channel-specific details, which is why it was the best and consistently good model during our experiments of multi-class classification of AD, PD, and FTD.
From observation we can say that the proposed SadNetV1 performed best out of all baseline models on training, validation and test data, registering test accuracy of 96.15 percent and effective generalization revealed by train accuracy of 96.84 percent and validation accuracy of 97.11 percent. It had a very small architecture and had attention building in it which enabled it to learn low-level and important spatial and channel-specific details, which is why it was the best and consistently good model during our experiments of multi-class classification of AD, PD, and FTD.

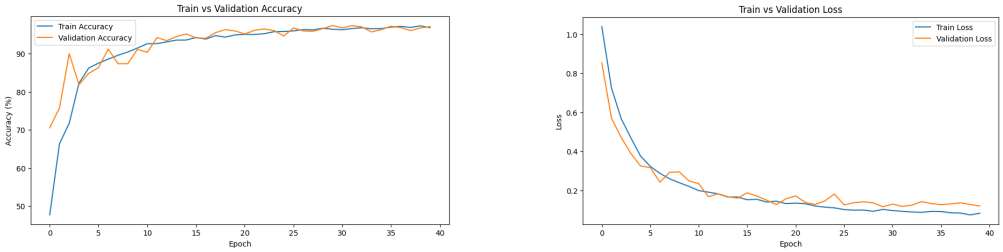

Figure 7.4: SadNetV1 Accuracy and Loss

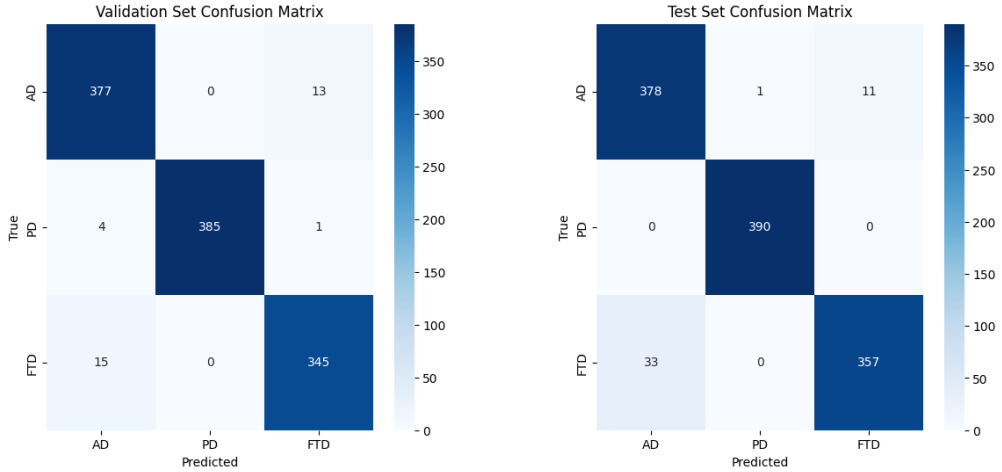

Figure 7.5: SadNetV1 Confusion Matrix of Test and Validation

7.4 Experiment with Healthy patient

To increase the flexibility of our model in terms of classification further, we added a fourth class to our dataset comprising a label representing Normal (CN) individuals, in addition to Alzheimer Disease (AD) and Parkinson Disease (PD). The addition of the normal samples however, created a large degree of imbalance in the classes and class overlap, especially because of the similarity of the early stage cases.
In reducing these difficulties and decreasing overfitting, we utilized some tactics:

- SMOTE (Synthetic Minority Over-sampling Technique): To over sample the normal samples and balance the minority classes synthetically.
- Class Weight Adjustment: To further discourage the model with drawing dis-

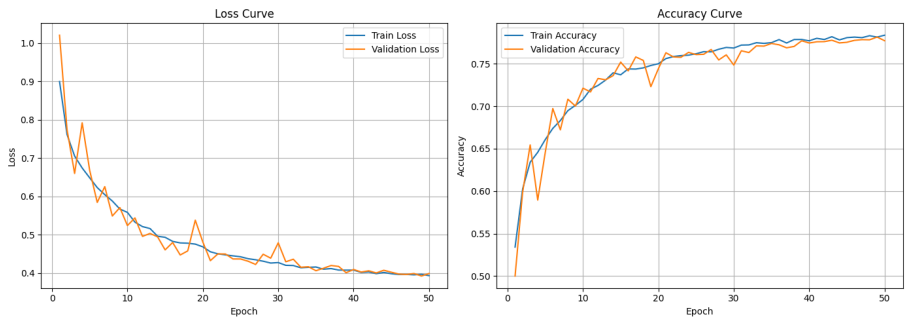

Figure 7.6: Accuracy and Loss of Normal Paient

torted conclusions on the results of training, with regard to misclassification with underrepresented categories

- Downsizing Major Classes: Majority classes have to be downsized to provide balanced training, and avoid bias due to the overrepresentation of disease classes.
- Slice-wise Patient Aggregation: We aggregated predictions across multiple slices per patient in order to increase robustness and consistency of classification; however, patient-wise evaluation was still done slice-wise.

Nevertheless, the best performing accuracy of the test was 78%, partly because of the insufficient amount of good healthy MRI data, and partly because features of early neurodegenerative cases were similar. This shows the need for more well-curated normal patient data and better methods of disambiguation of features that are very near each other. Our future research will focus on the increase of the dataset of healthy samples to highly reliable sample sizes and further combine it with the SadNetV1 model to reach a more extensive and clinically applicable scheme of early disease disease screening.

7.5 Predictability of our mode

We kept some MRI data from all three classes to test if our model can classify accurately. Some images of PD were classified as AD but the number is very low and the image is very noisy. Below given some image of prediction.

Images predicted as AD

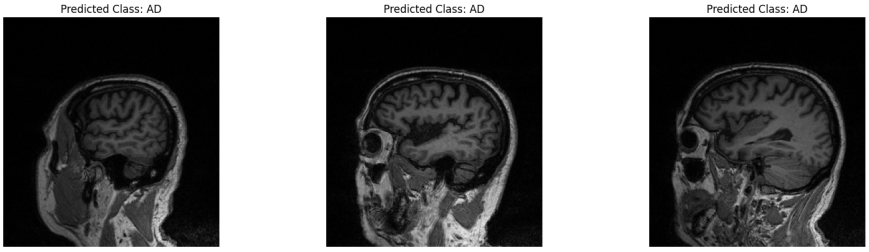

Figure 7.7: Predicted as AD

Images predicted as PD

Figure 7.8: Predicted as PD

Images predicted as FTD

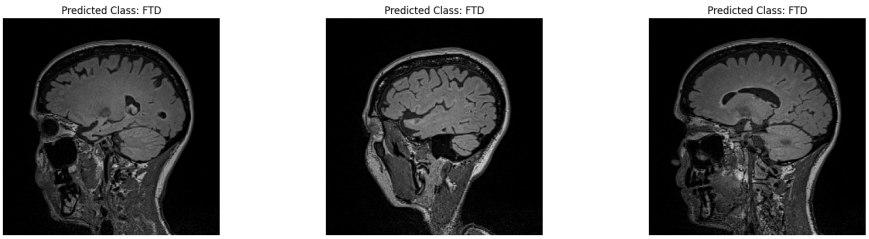

Figure 7.9: Predicted as FTD

7.6 Discussion and Insights

The experimental findings indicate that enormous variation in the performance exists among the studied CNNs architectures. Although VGG19 is a deep model, it did not manage to learn informative information of the MRI slices and did not perform well scoring underfitting and low training and test accuracy. ResNet50 and DenseNet121 yielded medium test accuracy and also significant overfitting in their models, that is, they possibly learned more than enough on the training set and did not generalize to unknown examples.

EfficientNetB0 performed more balanced in terms of good test accuracy (85.47%) with a rather stable validation accuracy, which means its compound scaling strategy contributed to making it fit better to the multi-class classification problem. The depthwise separable convolutions and lightweight architecture made MobileNetV2 achieve a test accuracy of 94.36 percent. Nevertheless, there is no significant difference between training (98.30%) and test accuracy, which means that there is a certain amount of overfitting.

As could be seen the SadNetV1 model efficiently managed this problem through the integration of SE blocks and Spatial Attention modules thus allowing the model to concentrate on the most relevant features on the channel and the spatial level. This was the best combination with the test accuracy 96.15%, validation accuracy 97.11% and training accuracy 96.84. These findings show better generalization and stability of the datasets.

Then there was an experiment run with normal patients with Alzheimer and Parkinson patients in which it resulted in overfitting because of the class imbalance. The checklist techniques SMOTE, class reweighting and downsampling were used to reduce this and a maximum test accuracy of 78% was reached with a slide-wise patient aggregation technique. This is lower but still indicates a positive direction of interest to future investigation in terms of improved sampling of normal patients and their incorporation into SadNetV1 pipeline.

On balance, the analysis indicates that attention mechanisms help enormously when it comes to enhancing the classification accuracy of lightweight architectures in complex medical image datasets whose features overlap.

Chapter 8

Future Work and Improvements

This study can be continued in the future to solve some of the most significant issues in order to classify neurodegenerative diseases like Alzheimer's Disease (AD), Parkinson's Disease (PD), and Frontotemporal Dementia (FTD) better by using MRI images. Firstly, we tried to include control normal as a class and went through three different approaches which are only taking control normal subject of AD, control normal subject of PD, control normal subject of FTD and merging all control normal subject of all three disease but we observed a lack of distinct control normal subjects, and as a result normal group was mixed with diseased individuals will be tackled by acquiring a diverse and more representative group of control normal subjects to create a better differentiation of the models. Also, in the dataset collection, it was learned that there were a lower number of unique subjects than expected, especially in the case of FTD. Future research will focus on an increased dataset and the statistical representation of the data through more unique subjects. Lastly, the combination of multi-modal data and genetic data will also be undertaken as future work to give a more complete picture of the diseases and enhance the classification performance using the complementary features across data modalities. These developments will result in a better and stronger deep learning model to diagnose neurodegenerative diseases.

Chapter 9

Conclusion

In this paper, we suggested a deep-learning-based model, which characterizes the neurodegenerative diseases, Alzheimer Disease (AD), Parkinson Disease (PD), and Frontotemporal Dementia (FTD) based on their sagittal slices of 3D MRI slices. MobileNetV2 was the best performing architecture in terms of test accuracy (94.36%) with tendencies of overfitting after evaluating several pre-trained architectures. To tackle this we proposed a lightweight deep model, SadNetV1, a variant of MobileNetV2 with Squeeze-and-Excitation (SE) and Spatial Attention blocks. The use of SadNetV1 greatly enhanced the performance of the classification process with the test set accuracy of 96.15%, training set accuracy of 96.84%, and validation set accuracy of 97.11%. We also tested the modalities of incorporating MRI results of normal patients together with AD and PD samples. Even though this has been the first factor to project overfitting, the model was able to identify 78 percent of the data correctly, which means that even better quality and more balanced normal patient data may further improve the results of a classification. We will further incorporate them into our own architecture and scale the dataset with highly labeled healthy controls and make SadNetV1 increase the understanding of the disease and become clinically useful as a web application that can perform predictions in real-time in practice and research settings with ease of access and interpretability in the future.

Bibliography

[1] A. E. Lang, "Parkinson's disease: Etiology, neuropathology and pathogenesis," *Neuron*, vol. 38, pp. 699–713, 2003.

[2] O.-B. Tysnes and A. Storstein, "Parkinson's disease," *Journal of Neurology*, vol. 257, pp. 157–165, 2010, PMID:20055267.

[3] D. G. Hernandez, X. Reed, and A. B. Singleton, "Genetic etiology of parkinson disease associated with mutations in the snca, park2, pink1, park7, and lrrk2 genes: A mutation update," *Human Mutation*, vol. 32, pp. C24–C37, 2011, PMID:3056147.

[4] V. Health, "Differences between dementia and alzheimer's," *Verywell Health*, 2012.

[5] Various, "The genetics and neuropathology of alzheimer's disease," *PMC*, 2013.

[6] C. Pottier, D. Hannequin, N. Brouwers, *et al.*, "An update on genetic frontotemporal dementia," *Journal of Neurology*, vol. 261, pp. 230–243, 2014.

[7] J. C. van Swieten and P. Heutink, "Clinical and pathological features of frontotemporal dementia subtypes," *The Lancet Neurology*, vol. 13, pp. 798–810, 2014.

[8] M. e. a. Tartaglia, "Frontotemporal dementia: A clinical and molecular review," *Neurology*, vol. 84, pp. 340–350, 2014.

[9] J. Bang, S. Spina, and B. Miller, "Frontotemporal dementia," *Lancet*, vol. 386, pp. 1672–1682, 2015.

[10] L. V. Kalia and A. E. Lang, "Parkinson's disease," *The Lancet*, vol. 386, no. 9996, pp. 896–912, 2015.

[11] ——, "Parkinson disease," *The Lancet*, vol. 386, pp. 896–912, 2015. DOI: 10.1093/brain/awv392.

[12] M. e. a. Prince, "World alzheimer report 2015: The global impact of dementia," *Alzheimer's Disease International*, 2015.

[13] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations (ICLR)*, 2015. [Online]. Available: https://arxiv.org/abs/1409.1556.

[14] C. N. Austin, "An overview of parkinson's disease," *Austin Journal of Clinical Neurology*, vol. 1, p. 1011, 2016.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[16] TIME, "Early-onset alzheimer's: What you need to know," *Time*, 2016.

[17] T. A. for Frontotemporal Degeneration, *What is bvftd? (behavioral variant ftd)*, online resource summarizing key bvFTD symptoms, 2017.

[18] O. e. a. Hardiman, "Amyotrophic lateral sclerosis," *Nature Reviews Disease Primers*, vol. 3, no. 1, p. 17 071, 2017.

[19] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *CVPR*, 2017, pp. 4700–4708. [Online]. Available: https://arxiv.org/abs/1608.06993.

[20] E. e. a. Dorsey, "Global, regional, and national burden of parkinson's disease, 1990–2016: A systematic analysis," *The Lancet Neurology*, vol. 17, no. 11, pp. 939–953, 2018.

[21] C. e. a. Jack, "Nia-aa research framework: Toward a biological definition of alzheimer's disease," *Alzheimer's & Dementia*, vol. 14, no. 4, pp. 535–562, 2018.

[22] M. Jucker and L. Walker, "Prevention and early detection of alzheimer's disease: A global challenge," *Nature Reviews Neurology*, vol. 14, pp. 497–498, 2018.

[23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 4510–4520. [Online]. Available: https://arxiv.org/abs/1801.04381.

[24] N. Bookshelf, "Alzheimer's disease: Etiology, neuropathology and pathogenesis," *NCBI*, 2019.

[25] M. Kim *et al.*, "Convolutional neural networks to classify alzheimer's disease severity based on spect images: A comparative study," *Journal of Clinical Medicine*, vol. 8, no. 11, p. 1828, 2019. DOI: 10.3390/jcm8111828.

[26] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 6105–6114. DOI: 10.1109/CVPR.2019.00772.

[27] X. Liu, Y. Zhang, K. Chen, Y. Wang, and X. Hu, "Parkinson's disease detection using deep learning and mri scans: A review," *Computers in Biology and Medicine*, vol. 121, p. 103 768, 2020. DOI: 10.1016/j.compbiomed.2020.103768.

[28] ——, "Parkinson's disease detection using deep learning and mri scans: A review," *Computers in Biology and Medicine*, vol. 121, p. 103 768, 2020. DOI: 10.1016/j.compbiomed.2020.103768.

[29] D. Ma, X. Zhang, and Y. Zhou, "Ftdi: A novel index for differentiating frontotemporal dementia from alzheimer's disease using structural mri," *Frontiers in Neuroscience*, vol. 14, p. 566, 2020. DOI: 10.3389/fnins.2020.00566.

[30] D. Ma, D. Lu, K. Popuri, L. Wang, and A. D. N. I. Beg Mirza Faisal and, "Differential diagnosis of frontotemporal dementia, alzheimer's disease, and normal aging using a multi-scale multi-type feature generative adversarial deep neural network on structural magnetic resonance images," *Frontiers in Neuroscience*, vol. Volume 14 - 2020, 2020, ISSN: 1662-453X. DOI: 10.3389/fnins.2020.00853. [Online]. Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2020.00853.

[31] G. Pinto *et al.*, "Repetitive behaviors in frontotemporal dementia: Compulsions or rituals?" *Frontiers in Neurology*, vol. 11, p. 467, 2020.

[32] [53], "Detection of parkinson's disease from neuro-imagery using transfer learning," Thesis, School of Computer Science, University of Nottingham, 2020. [Online]. Available: https://figshare.com/articles/thesis/Detection_of_Parkinson_s_Disease_from_Neuro-imagery_Using_Transfer_Learning/12773494.

[33] W. Wang, J. Lee, F. Harrou, and Y. Sun, "Early detection of parkinson's disease using deep learning and machine learning," *IEEE Access*, vol. 8, pp. 147 635–147 646, 2020.

[34] D. Zhang, Y. Wang, L. Zhou, T. Yuan, L. Huang, and D. Shen, "Alzheimer's disease detection using deep learning and mri scans: A comparison of efficientnet and other cnn models," *Journal of Alzheimer's Disease*, vol. 75, no. 4, pp. 1231–1243, 2020. DOI: 10.3233/JAD-200617.

[35] ——, "Alzheimer's disease detection using deep learning and mri scans: A comparison of resnet and other cnn models," *Journal of Alzheimer's Disease*, vol. 75, no. 4, pp. 1231–1243, 2020. DOI: 10.3233/JAD-200617.

[36] C. e. a. Abbate, "An overview of neurodegenerative diseases and therapeutic strategies," *Journal of Translational Medicine*, vol. 19, pp. 1–14, 2021.

[37] Y. J. Bae, J.-M. Kim, C.-H. Sohn, *et al.*, "Imaging the substantia nigra in parkinson disease and other parkinsonian syndromes," *Radiology*, vol. 299, no. 3, e210341, 2021. DOI: 10.1148/radiol.2021203341.

[38] J. Hu, Z. Qing, R. Liu, X. Zhang, P. Lv, M. Wang, Y. Wang, K. He, Y. Gao, and B. Zhang, "Deep learning-based classification and voxel-based visualization of frontotemporal dementia and alzheimer's disease," *Frontiers in Neuroscience*, vol. Volume 14 - 2020, 2021, ISSN: 1662-453X. DOI: 10.3389/fnins.2020.626154. [Online]. Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2020.626154.

[39] Z. Jiang, L. Wang, H. Wang, X. Zhang, and C. Yu, "Deep learning with efficientnet for frontotemporal dementia detection from mri images," *Frontiers in Aging Neuroscience*, vol. 13, p. 672 526, 2021. DOI: 10.3389/fnagi.2021.672526.

[40] ——, "Deep learning with resnet for frontotemporal dementia detection from mri images," *Frontiers in Aging Neuroscience*, vol. 13, p. 672 526, 2021. DOI: 10.3389/fnagi.2021.672526.

[41] S. Murugan, C. Venkatesan, M. G. Sumithra, X.-Z. Gao, B. Elakkiya, M. Akila, and S. Manoharan, "Demnet: A deep learning model for early diagnosis of alzheimer diseases and dementia from mr images," *IEEE Access*, vol. 9, pp. 90 319–90 329, 2021. DOI: 10.1109/ACCESS.2021.3090474.

[42] C. Quan, K. Ren, and Z. Luo, "A deep learning based method for parkinson's disease detection using dynamic features of speech," *IEEE Access*, vol. 9, pp. 10 239–10 252, 2021. DOI: 10.1109/ACCESS.2021.3051432.

[43] P. M. Raees and V. Thomas, "Automated detection of alzheimer's disease using deep learning in mri," in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1921, 2021, p. 012 024.

[44] K. Katisko *et al.*, "Serum total tdp-43 levels are decreased in frontotemporal dementia patients with c9orf72 repeat expansion or concomitant ftd-mnd phenotype," *Alzheimers Research & Therapy*, vol. 14, p. 151, 2022.

[45] H. Li *et al.*, *Mapt-related frontotemporal dementia – genereviews*, 2022.

[46] H. T. Nguyen, C.-H. Sun, J. Doshi, and et al., "Deep grading for mri-based differential diagnosis of alzheimer's disease and frontotemporal dementia," *NeuroImage: Clinical*, vol. 35, p. 103 106, 2022. DOI: 10.1016/j.nicl.2022.103106.

[47] S. M. Abdullah, T. Abbas, M. H. Bashir, I. A. Khaja, M. Ahmad, N. F. Soliman, and W. El-Shafai, "Deep transfer learning based parkinson's disease detection using optimized feature selection," *IEEE Access*, vol. 11, pp. 3511–3524, 2023.

[48] D. P. Agamanolis, "Parkinson's disease," *Neuropathology Web*, 2023, Chapter 9.

[49] Y. e. a. Hou, "Global burden of dementia: Estimates and projections from 1990 to 2050," *The Lancet Public Health*, vol. 8, no. 3, e210–e221, 2023.

[50] I. Moreno-Gonzalez and C. Soto, "Neurodegenerative diseases: A comprehensive overview," *Nature Reviews Neurology*, vol. 19, no. 2, pp. 75–93, 2023.

[51] B. R. Payne and C. L. Satizabal, "Machine learning applications in structural mri for alzheimer's and related disorders," *NeuroImage: Clinical*, vol. 38, p. 103 360, 2023. DOI: 10.1016/j.nicl.2023.103360.

[52] A. Pérez-Millan, J. Contador, J. Juncà-Parella, B. Bosch, L. Borrell, A. Tort-Merino, N. Falgàs, S. Borrego-Écija, N. Bargalló, L. Rami, M. Balasa, A. Lladó, R. Sánchez-Valle, and R. Sala-Llonch, "Classifying alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data," *Human Brain Mapping*, vol. 44, no. 6, pp. 2234–2244, 2023. DOI: https://doi.org/10.1002/hbm.26205. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.26205. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.26205.

[53] M. Swetha and V. Lakshmi, "Detection of alzheimer's disease using mri based on vgg19 & densenet169," *International Journal of Computer Trends and Technology (IJCTT)*, 2023. [Online]. Available: https://doi.org/10.14445/22312803/IJCTT-V71I1P104.

[54] M. H. Team, "How is parkinson's disease treated?" *Health.com*, 2023.

[55] J. e. a. Wong, "Neurodegenerative disease overlap: Diagnosis, biomarkers, and precision medicine," *Nature Reviews Neurology*, vol. 19, no. 5, pp. 255–270, 2023.

[56] L. Ali, A. Javeed, A. Noor, H. T. Rauf, S. Kadry, and A. H. Gandomi, "Parkinson's disease detection based on features refinement through l1 regularized svm and deep neural network," *Scientific Reports*, vol. 14, no. 1, p. 1333, 2024.

[57] N. Anagnostis *et al.*, "A robust distributed deep learning approach to detect alzheimer's disease from mri images," *Mathematics*, vol. 12, no. 12, p. 2633, 2024. DOI: 10.3390/math12122633.

[58] A. Assoc, "2024 alzheimer's disease facts and figures," *PMC*, 2024.

[59] M. Joy *et al.*, "Frontotemporal dementia: From genetics to therapeutic approaches," *Expert Review of Neurotherapeutics*, vol. 24, pp. 321–336, 2024.

[60] A. M. Kandhro *et al.*, "A grey wolf optimized densenet for parkinson's disease detection," *BMC Medical Imaging*, vol. 24, Article 45, 2024. DOI: 10.1186/s12880-024-01045-5.

[61] J. Samper-González, S. Cohen, and A. D. N. Initiative, "Ai for early detection of neurodegenerative diseases: Opportunities and limitations," *Nature Reviews Neurology*, vol. 20, no. 2, pp. 123–135, 2024. DOI: 10.1038/s41582-024-00988-1.

[62] C. Scialo *et al.*, "Plasma phosphorylated tdp-43 levels are elevated in patients with frontotemporal dementia carrying a c9orf72 repeat expansion or a grn mutation," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 95, pp. 684–691, 2024.

[63] N. I. on Aging, "Alzheimer's disease fact sheet," *NIA*, 2025.

[64] J. Huang, "Alzheimer disease," *MSD Manual Professional Edition*, 2025.

[65] ——, "Alzheimer disease," *MSD Manual*, 2025.

[66] M. Mirabian, H. Lee, and C. Yoon, "Artificial intelligence in eeg-based diagnostics of neurodegenerative disorders: A systematic review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 33, pp. 199–211, 2025. DOI: 10.1109/TNSRE.2025.3142990.

[67] Wikipedia, "Dementia," *Wikipedia*, 2025.

[68] ——, "Frontotemporal dementia," *Wikipedia*, 2025.

[69] E. U. Haq, Q. Yong, Z. Yuan, H. Jianjun, and R. U. Haq, "Multimodal fusion diagnosis of the alzheimer's disease via lightweight cnn-lstm model using magnetic resonance imaging (mri)," *Available at SSRN 4719918*,