Detection of Prodromal Parkinson's Disease with fMRI data and deep neural network approaches

by

Farhan Shahriar
14201046
Amarttya Prasad Dey
16201081
Naimur Rahman
15321002
Zarin Tasnim
13101154
Mohammad Zubayer Tanvir
13321045

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

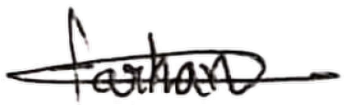
Department of Computer Science and Engineering
Brac University
June 2021

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:



Farhan Shahriar
14201046

Amarttya Prasad Dey
16201081

Naimur Rahman
15321002

Zarin Tasnim
13101154

Mohammad Zubayer Tanvir
13321045

Approval

The thesis/project titled "Detection of Prodromal Parkinson's Disease with fMRI data and deep neural network approaches" submitted by

1. Farhan Shahriar (14201046)
2. Amarttya Prasad Dey (16201081)
3. Naimur Rahman (15321002)
4. Mohammad Zubayer Tanvir (13321045)
5. Zarin Tasnim (13101154)

Of Spring, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on june 2, 2021.

Examining Committee:

Co-supervisor:
(Member)

Md. Tanzim Reza
Lecturer
Department of Computer Science and Engineering
BRAC University

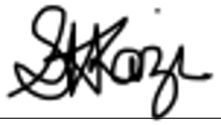
Supervisor:
(Member)

Dr. Muhammad Zavid Parvez
Assistant Professor
Department of Computer Science and Engineering
BRAC University

Program Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)



Dr. Sadia Hamid Kazi
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement (Optional)

This is optional, if you don't have an ethics statement then omit this page

Abstract

Parkinson's Disease is the second most common neurological disease after Alzheimer's Disease. The disease is incurable. However, if the disease can be detected earlier, then the consequences of it's effect can be relieved. The early phase of PD is called by Prodromal Parkinson's Disease. The symptoms of the Prodromal phase includes hyposmia, constipation, mood disorders, REM sleep behavior disorder, olfaction disorders etc. RBD or REM sleep behavior disorder is the most common symptom of Prodromal PD. In this study, we used various deep convolutional neural network architectures and trained them to detect Prodromal PD patients. We collected 20 Prodromal patients and 20 healthy control subject data from the PPMI website and applied CNN architecture mobilenet v1, incception v3, vgg19 and inception resnet v2 to achieve our goal. We ensembled inception resnet v2 and mobilenet v1 with the hope of getting a better result as well. However, we successfully carried out our training and with mobilenet v1 we gained the highest classification accuracy of 81.22%. Inception resnet V2, inception v3, vgg19 and ensemble model achieved respectively 75.30%, 62.55% and 63.32% accuracy.

Keywords: Convolutional Neural Network (CNN), Parkinson's Disease (PD), Neural Network (NN), fMRI, Deep Learning, Average Ensemble, VGG19, Inception-ResNet-v2, Inception-V3, MobileNet-V1, PPMI

Dedication (Optional)

We dedicate our thesis work to our supportive and loving family and friends

Acknowledgement

To begin with, we are thankful to the Almighty Allah for letting us continue our thesis even through these dreadful and scary Covid-19 days without any major interruption. We are ever so grateful that we were all in good health, along with our loved ones. Next, our thesis would have been unorganized and in disarray without the proper guidance and help from our supervisor Dr. Mohammad Zavid Parvez. His availability and kindness were a significant help during these past 4 semesters, of which 3 were online. Furthermore, we would like to thank our co-supervisor Md Tanzim Reza whose knowledge and expertise was invaluable in conducting our research. His feedback was very much appreciated. Finally, none of this would have been possible without the support of our beloved parents and family who stood beside us through even the longest of nights and were there to always pick us up.

Table of Contents

List of Figures

List of Tables

Nomenclature

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

CNN convolutional neural networks

CSF cerebrospinal fluid

DNN deep neural network

$fMRI$ Functional Magnetic Resonance Imaging

GM gray matter

$MFCC - GMM$ Mel-Frequency Cepstral Coefficients—Gaussian Mixture Model

MRI Magnetic resonance Imaging

MRI magnetic resonance imaging

$NMS - MRI$ Neuromelanin sensitive magnetic resonance imaging

PD Parkinson's disease

PLP Linear Signal Prediction

$PPMI$ Parkinson's Progression Markers Initiative

PSG polysomnographic

PWP People with parkinsonism

$RASTA - PLP$ Realitive Spectral PLP

RBD REM sleep disorder

RBF radial basis function

REM rapid eye movement

SNc substantia nigra compacta

SVM Support Vector Machine

TIV total intracranial volume

VBM voxel-based morphometry

WM white matter

Chapter 1

Introduction

PD which is the abbreviation of Parkinson's disease is the second most common neurodegenerative disease after Alzheimer's disease and it affects the regulation of movement and emotions due to the loss of dopaminergic neurons in the SN region of the brain. The loss of nerve cell neurons in the SN region, which is located in the basal ganglia area of the brain causes this disease. The neurons that are located in this part of the brain are responsible for producing a very important organic chemical called dopamine which helps other neurons to communicate between each other by acting as a neurotransmitter in the brain. If the amount of this dopamine produced neurons of the brain is somehow becomes insufficient, communication between the neurons that coordinate the body's movement get hampered and it results in the disease called PD. It has a variety of neurological and motor symptoms, such as tremor, balance problems, slow movements, gait disturbances, imbalance posture, speech problems etc [1][2][3]. Globally each year 6.2 million people get affected by it [4]. The symptoms of PD continue to get worsen gradually. To this day, even though studies has been going on for decades, there is no known cure for PD and no one is sure what causes Parkinsonism in the first place [5]. The time when Parkinson's disease is finally detected, all of the motor-symptoms are already visible and 50-70% dopaminergic neurons have already decayed [6]. So, researchers have been searching whether there were any non-motor symptoms of PD or not. Now it is well established facts that PD has many non-motor or pre-motor symptoms such as depression, constipation, pain, genitourinary problems, and sleep disorders [7]. Among all motor pre-motor sympotoms REM sleep disorder or RBD is one of most prominent sign of Parkinson's disease[8][9][10]. The pre-motor phase is formally known as Prodromal PD[11][12]. The main focus point of prodromal PD is REM sleep disorder and other sympotoms such as subtle motor dysfunction, olfactory loss and autonomic dysfunction. Even though motor symptoms starts to show at the ending phase of parkinson's disease, prodromal symptoms sometimes starts showing from decades earlier. So, it is possible to detect early PD by trying to detect the prodromal symptoms in a patients. In this study we applied some cnn deep learning architectures and tried to detect parkinson's disease at an early phase.

1.1 Motivation

Our main motivation for choosing this topic was to accurately detect the stage of this disease as early as possible, because in Bangladesh people are not very aware of

diseases such as PD, AD or schizophrenia and its consequences. As these diseases normally affects people at the later phase of life, whenever any elderly person show some kind of memory impairment or movement disorders, people normally assume it is due to their old age. This is a serious problem because Parkinson's disease cannot be cured and this ignorance can cause severe problem for these senior citizens as they continue to get older. In addition, Bangladesh has one of the poorest medical facilities in the world, and sometimes diagnosis are incorrect. So, we hope that we will be able to propose a model that can correctly detect this dangerous disease, so that it can extend the life of our elderly and give them a better lasting life.
However, researches regarding PD is being conducted throughout the world. Even in countries like the United States, billions of dollars are spent on PD related researches every year. So far, all the papers we have studied showed various methods and algorithms that can be used to detect PD and some at the early phase as well. However, our goal is to provide methodologies which are more cheaper and have more higher accuracy for detecting PD at an early phase than other related researches.

1.2 Contribution

The purpose of this study is to first detect Prodromal PD as accurately as possible. We have implemented the latest and most efficient CNN architecture to improve our results. The architectures we have used in this study are all taken from the CNN architectures used in the Imagenet challege, which is a global competition hold every year and participated by prominent tech companies like microsoft and google[13]. Each architecture we used are proven methodologies that has proved their consistency in various researches to be highly effective. We have used 4 such CNN models namely vgg19, inception v3, inception resnet V2 and mobilenet V1 architecture and then ensembled inception resnet V2 and mobilenet V1 models to achieve even higher accuracy.

1.3 Organization of paper

In this study, in the first chapter we gave an introduction to the paper which is followed by Chapter second that contains background information about our research and the used models and CNN. What follows is a description of the data set and the methodologies proposed in Chapters 3 and 4. Then we ended our work by providing the results, analysis and discussions in Chapter 5 and with the possible future works in chapter 6.

Chapter 2

Background

2.1 Literature Review

In this paper, they have adapted the latest speaker recognition system, which is called by x-vector, to use speech analysis to detect PD at an early stage. The X vector is an embedding extracted from DNN. When a large amount of training data is used, x-vector provide a reliable representation of the speaker and improve the speaker's recognition ability. They wanted to evaluate whether this x-vector technology outperforms the more standard MFCC-GMM classifier in the case of early detection of PD, and if so, under what conditions. They recorded 221 French-speaking people which were a mixture of newly diagnosed PD subjects and healthy control group with high-quality microphones and through the telephone network. They analyzed both men and women separately to build a more accurate model and evaluate possible gender effects. They tested Various experimental and methodological aspects to analyze x-vector's impact on classification performance. They evaluated the length of the audio clip, the data surge, the type of data set used for neural network training, the type of voice task and the impact of the back-end analysis. For non-text related tasks, x-vector technology provided better sorting performance than MFCC-GMM, and appeared to be particularly suitable for early detection of PD in women by 7-15%. These results were observed in both high quality microphone and telephone recording[14].
In this paper, they analyzed databases of PD patients and non-PD patients which contained speech records used to extract para-linguistic features. These data were used as input to machine learning models to predict the severity of PD. The research was managed by Sage Bionetworks and the data was available on its Synapse research portal and their final data set included 2,289 people including 2,023 controls, 246 PD and 15,227 voice tasks including 9,994 control tasks, 5,233 PD tasks. They used three popular models for binary classification which are L2 regular logistic regression, random forest and gradient weighted decision tree. The L2 regular logistic regression had a recall of 0.759, precision of 0.811 and F1 score of 0.784 with AUC of 0.85. random forest had a recall of 0.693, precision of 0.902 and F1 score of 0.783 with AUC of 0.83 and gradient weighted decision tree had a recall of 0.797, precision of 0.901 and F1 score of 0.836 with AUC of 0.88[15].
In this paper, they have examined a variety of vowel samples collected from PD patients and healthy subjects and used on them the latest signal processing algorithms, such as PLP and RASTA-PLP. They then used SVM model with four different ker-

nels to classify the set of extracted features. Their result performed 74% accurately. The data collected for this study belongs to twenty nine PWP including nine women and twenty men and thirty healthy people including ten woman and twenty men. SVM with RBF kernels had the highest accuracy of 74% with sensitivity of 70% and specificity of 77%. Polinomial kernel had the second highest accuracy of 72% with sensitivity of 80% and specificity of 12%[16].

In this study, healthy control subjects and PD subjects were tested by taking their sleep quality data through smart watches. Initially, the arm angle changes due to activities while sleeping was captured from the smart watch's triaxial accelerometry data and was used to estimate whether the subject was awake or asleep. Then, sleep metrics such as sleep efficiency index, total sleep time, sleep fragmentation index, sleep onset latency, and wake after sleep onset etc. were computed and was used to distinguish between controls and PD patients. Compared with the PSG-based ground truth in a clinical setting, the verification process of the proposed method yields comparable state estimates. In addition, data from 15 early PD patients and 11 healthy controls were used as the test set which also included 1376 valid sleep records in field environments. While comparing with the classification of healthy controls, univariate analysis of sleep indicators of early PD patients obtained an AUC as high as 0.77, and showed a statistically significant correlation, highest being 0.46 with the clinical counterparts of the PD sleep scale 2[17].

This paper tried to detect PD based on the patient's handwriting capability. According to them, PD patients face writing disabilities and even before them, different researchers proposed different computer vision and machine learning methods to detect PD based on microscopic images and computer vision. But according to them these methods have two main problems. The first problem being model bias caused by unbalanced data, that is, machine learning models performing well in majority class, and performing poorly in a minority class. However, according to them, prior to their research, previous researchers did not discuss this problem and also did not take any steps to avoid it. So, in order to highlight the deviations in the built model and prove it in a practical way, they developed four different machine learning models. In order to alleviate the bias problem, they recommended using random undersampling methods to balance the training process. The second problem being the low accuracy of classification, which made the previous researches to have limited clinical significance. So, in order to improve the accuracy of PD detection, they proposed a cascaded learning system that cascades the Chi2 model and the adaptive boost model or Adaboost model. The Chi2 model can classify and select a subset of related features from the feature space, while the Adaboost model is used to predict PD based on the feature subset. The experimental results they got, confirmed that their proposed cascade system perform better than six other similar cascade systems using the latest machine learning models in six different states. Furthermore, they also observed that their proposed cascade system increased the strength of the conventional Adaboost model by 3.3% and reduced its complexity. In addition, their cascade system achieved a classification precision of 76.44%, a sensitivity of 70.94% and a specificity of 81.94%[18].

In this paper, they proposed a complex network-based method for accurately diagnosing early PD using MRI data. They claimed their method also would allow us to investigate that brain region that are most affected by the disease. At first they defined a network model of brain regions and associated appropriate connec-

tivity metrics with each region. Therefore, each brain was represented by a feature vector, which encoded the local relationships that inter weaved the brain regions. Then, random forest was used for feature selection and learning compact representations. In the end, they used support vector machines to combine complex network functions with typical clinical scores in the prodromal phase of PD, and provided diagnostic indicators. They collected their data from the PPMI database, which included mixed cohorts of 169 healthy controls and 374 PD patients. They obtained an AUC of 0.97, accuracy of 0.93, sensitivity of 0.93 and specificity of 0.92[19].
In this paper they have mentioned that, NMS-MRI is essential for the detection of abnormalities in the substantia nigra compacta region of the brain, also known as SNc region, as PD is characterized by the loss of dopaminergic neurons in the SNC. They mentioned that, current technologies use SNc contrast estimates which is visualized on NMS-MRI to distinguish PD patients from healthy controls. However, according to them extracting these brain features from these current technologies are very time-consuming and it has many manual labor included as well and also the fact that they provide low prediction accuracy. In addition, these do not consider the potentiality of subtle changes in PD in the central nervous system. To alleviate this situation, their work established a computer-based analysis technique that used convolutional neural networks or CNN to create prognostic and diagnostic biomarkers for PD from NMS-MRI. Their technique not only outperformed other technologies with a superior testing accuracy of 80% as compared to contrast ratio-based classification of 56.5% testing accuracy and radiomics classifier of 60.3% testing accuracy, but also supported discriminating PD from atypical parkinsonian syndromes with 85.7% test accuracy. In addition, their methodologies have the ability to locate the most distinctive area in the neuromelanin contrast image. These discriminatory activations indicated that, compared to the right SNc, the left SNc plays a key role in classification and was consistent with the concept of asymmetry in PD[20].
In this paper, they used voxel-based morphometry or VBM technology to compare morphological differences between PD patients and healthy controls in GM and WM region of the brain . The effects of the use of different covariates such as total intracranial volume or TIV , age, sex, and their combinations and two different hypotheses namely t-contrast and f-contrast was studied on the classification of PD from HC in this paper. They used the local difference between PD and HC and the two-sample t-test method to obtaine 3D masks for GM and WM tissues respectively. Then they used PCA to perform dimensionality reduction and SVM for classification. Their proposed method was evaluated between 40 PD and 40 HC subjects which they obtained from the ppmi data set. The results of their classification by f-contrast showed that GM, WM and the combination of GM and WM perform better than t-contrast. Furthermore, their experimental results showed that, compared to other covariate configurations, the use of TIV as a covariate provide more reliable results for PD classification. When they used TIV as a covariate and f-contrast for the construction of the model, they obtained the highest precision of PD and HC. For GM, WM and their combination, they obtained 73.75%, 72.50% and 93.7% accuracy respectively[21].
In this paper, they made a great emphasis on the early or pre-clinical diagnosis of PD because they claimed that when clinical symptoms start to appear, more than 60% of dopaminergic neurons is already lost and it's this loss of dopaminergic neurons that is the reason of PD. They claimed that before the appearance of

the classic motor symptoms, there is a pre-motor stage, which is characterized by clinical features, mainly non-motor, such as rapid eye movement sleep behavior disorder and loss of smell or olfactory loss. So, in this paper they used the non-motor features of RBD and loss of smell, as well as other important biomarkers, such as cerebrospinal fluid or CSF and dopaminergic imaging markers measurements from 183 normal healthy subjects and 401 early-stage PD subjects. They obtained their data from the PPMI database. They used methodologies such as naive bayes, SVM, boosting tree, and random forest classifier to deduce their task of detecting early PD patients. Their subjects were classified with normal people. They observed that the best performance came from the SVM classifier with 96.40% accuracy, 97.03% sensitivity, 95.01% specificity and 98.88% ROC area [22].
In this study, they used the widely used Movement Disorder Society-Unified Parkinson's Disease Rating Scale or MDS-UPDRS's Patient Questionnaire (PQ) part to develop models, that can detect early PD from healthy control. They used machine learning models such as logistic regression, random forest, boosted trees and SVM for their study. They performed subject wise and record wise verification to evaluate their machine learning techniques. They observed that the above techniques have higher accuracy and higher ROC and AUC for classifying between PD patients from normal healthy Control. In both cases they got output greater than 95%[23].
In this study, they seeked several deep convolutional neural network architectures and used it on magnetic resonance imaging T1 of the brain to identify PD. They used three different ensemble architectures that combined imagenet large-scale Visual Recognition Challenge or ILSVRC Convolutional Neuronal Network Models. They first model they ensembled, consisted of densenet, sufflenet and squeezenet, seconde model ensembled densenet, sufflenet, squeezenet and mobilenet and the third model ensembled vgg, sufflenet nad mobilenet. All these architectures that they proposed had excellent existing approaches to detect PD from MR images and they achieved up to 95% of test accuracy[24].

2.2 Convolutional Neural Network

Convolutional Neural Network or CNN is a type of Deep neural network which is usually used to process visual data. It is also known as ConvNet or CNN. CNN is very powerful neural network algorithm for image recognition and classification. It is a feed-forward network that takes in images, does some pre-processing, converts it into pixel arrays as input and applies various filters to it. The filters are known as convolutional layer and hence the name convolutional neural network.
A convolutional layer has several filters that perform different types of convolution operations. After the image passes through a filter, it is called a filtered image. After each convolution layer it outputs a higher dimentional tensor. Then an operation called max-pooling is applied to the tensor and the result gives us a smaller image. Due to the above process, the number of parameters gets reduced in each steps, which reduces the complexity and processing time as well. Then after pooling layer activation functions like Relu, sigmoid etc. are applied. After severel steps of applying convolution layer of different sizes the final higher dimentional tensor is flattened. Then these are fed into a fully connected feed-forward network each having its own sets of bias and weights. Then finally it is softmaxed into the output layer. To reduce overfitting of data, random dropouts are used as well.

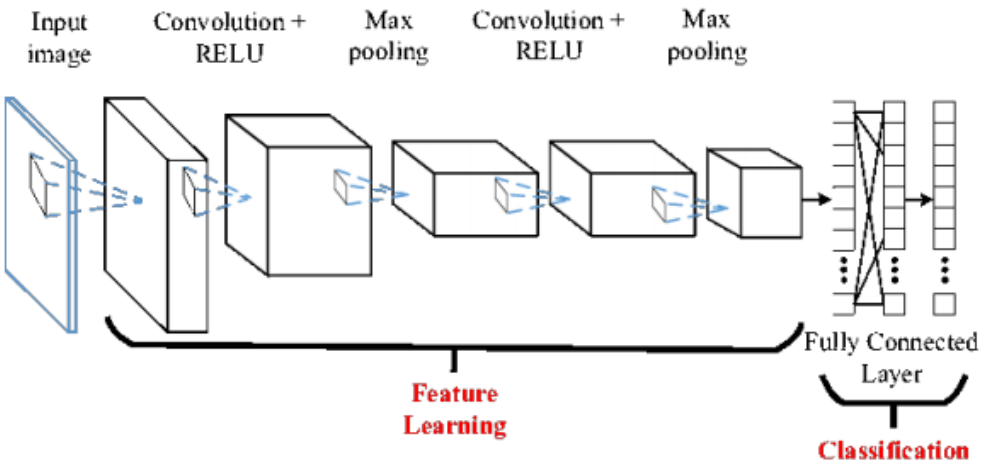


Figure 2.1: Block diagram of Convolutional Neural Network [25]

2.2.1 Convolution Layers

The convolutional layer is the first layer of CNN. It is mainly used for the feature extraction process. For the feature extraction process, a grid matrix which follows some patterns is dot product with the input image pixels. These grid matrices are called by filters. There can be various size of grid matrices as in 1x1, 3x3, 7x7 etc. All filters are dot product with each input image to create a feature map of the neuron. Here, a point is generated after the product operation between the filter and the input and the result of each convolutional layer becomes simpler and simpler. After one set of convolution operation is made we get an image with fewer shared parameters and with reduced complexity and improved model efficiency.

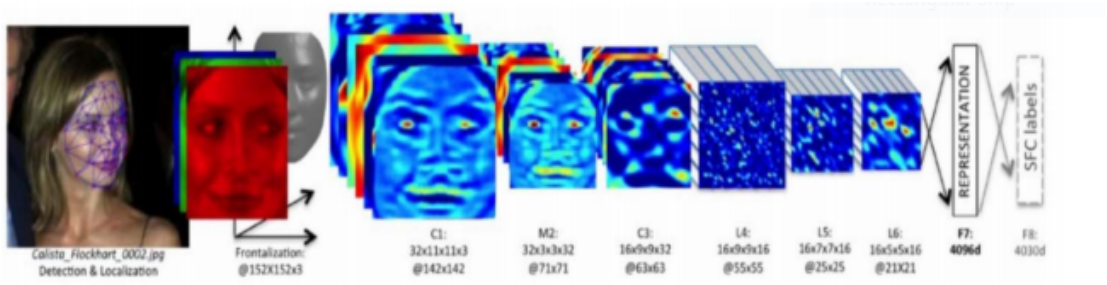


Figure 2.2: Visualizing Convolutional deep neural network layers [26]

2.2.2 Activation Functions

Activation functions are a very important component of Neural Network. This function decides whether the current cell will be activated for the calculation in the next cell or not. The output of the current layer is taken as input, and based on certain value of the current layer, it activates the current cell for the calculation in next layer. If it is activated, current layers value will be passed to next layer and if it is not activated, current values will not be passed. There are two types of activation function: linear activation function and non-linear activation function.

There are many activation functions like the sigmoid function which is a linear activation function or ReLU function which is a non linear activation function etc.

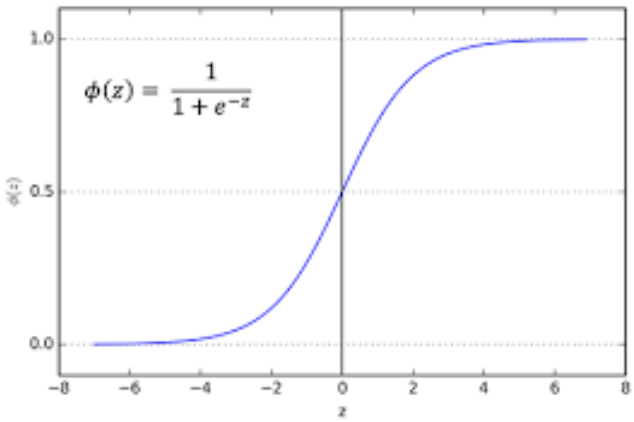


Figure 2.3: Graph of Sigmoid Function [27]

$$f(x) = \frac{1}{1+e^{-(x)}}$$

Figure 2.4: Sigmoid function Equation [27]

However, the rectified linear unit, or ReLU for short, is used as the most commonly used activation function in CNN because of it's non-linearity in the output. It does not accept any negative values, but will convert all negative values to zero, which will not activate the neuron. Due to this, the calculation becomes faster and efficient because fewer neurons get activated at a time.

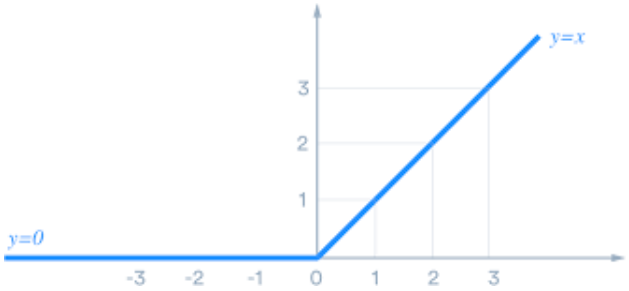


Figure 2.5: Graph of Relu Function [28]

2.2.3 Pooling

Once the convolution layer finishes applying it's filters, it generates a feature map, that contains all the summarized values of the input. This feature map has one

$$RELU(x) = \begin{cases} 0 \ if \ x < 0 \\ x \ if \ x >= 0 \end{cases}$$

Figure 2.6: Relu Function Equation [28]

disadvantage that is, the exact location of the input features are located. As a result, if the feature of the image gets changed by the slightest, the entire feature map gets dislocated. Even the slightest modification like rotation, cropping, flipping etc. can cause this problem. That is the reason after convolution layer, pooling layer was introduced. All the feature maps are processed separately by the pooling layer. As a result, it generates the same number of newly grouped feature maps. After adding this filter ,the size of the layer gets changed and its changes by 2 x 2 pixels with a step of 2 stride, so the feature map always get reduced by a factor of 2, as a result it reduces the number of pixels in each feature map. The most used pooling function is maxpooling function, where the number which has the maximum value in the specific feature map is selected and the output is subsampled to provide a smaller image.

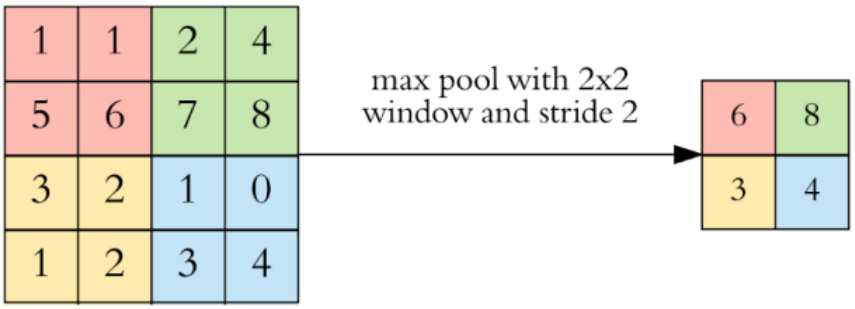


Figure 2.7: Output of Maxpooling Operation with stride 2 [29]

2.2.4 Dropout

Dropout method is used as an alternative to using regularization parameters. In dropout method, random values of neuron are dropped and it's value is made to 0. It includes both visible and hidden layers. In simple terms, it is randomly ignoring certain neurons during the training phase of a batch of neurons. Dropout majorly reduces overfitting and gives better results than other regularization techniques [30].

2.2.5 Batch Normalization

DNN training is very complicated, because in the training phase, the input distribution of each layer gets changed due to the change of the parameters of the previous layers like pooling and activation. So, it creates the need to standardize every batches. For that reason in each small batch, the input is standardized by

batch normalization technique, which is a DNN training method. Batch normalization not only stabilizes learning method but also reduces the number of training cycles required to train deep nets. Sometimes it reduces the number of epochs by 2 or even better in some cases. As a result our training speed becomes faster. It also provides regularization effect, thereby limiting the generalization error [31].

2.3 CNN Models

In this study, we have used 4 CNN architecture and one ensembled architecture to detect the Prodromal phase of PD. The 4 architectures are respectively vgg19, inception_resnet_v2, mobilenet_v1 and inception_v3. After applying each architecture individually, we ensembled inception_resnet_v2 and mobilenet_v1 to get optimized result as these two had the highest accuracy than the other two.

2.3.1 VGG19

VGG is the most simple and straightforward CNN architecture that we have used in our study. In vgg we continue to add increasing number of filters to it. For example, if layer 4 have 12 filters, layer 5 must have 12 or more filters. In vgg all filters have similar size of 3x3.
In vgg19, as the name suggest there are 19 total layers which includes the convolution, maxpooling, fully connected and softmax. It has 16 convolution layer, 3 fully connected layer, 5 maxpooling layer and 1 softmax layer. It's input image size is 224x224x3 [32].

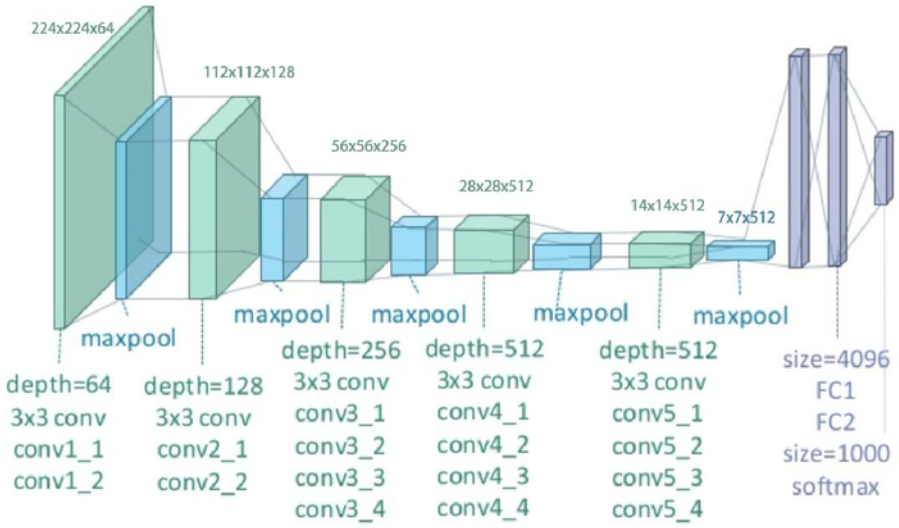


Figure 2.8: Block diagram of VGG19 architecture [33]

2.3.2 Inception V3

Inception architecture is one of the more complicated and complex architecture. V3 is the better version of it's predecessors. Here all the 7x7 filters from previous version have been replaced with multiple 3x3 filters. In inception v3 the layers are

respectively, 2 conv layer, 1 conv padded layer, pooling layer, 3 conv layer, 1 3x inception layer, 1 5x inception layer and 1 2x inecption layer, pooling layer, liner layer and softmax[34].

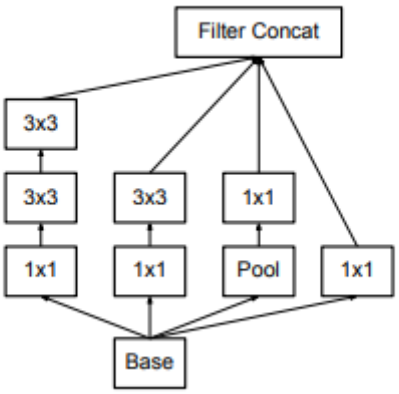


Figure 2.9: 3x Inception Layer architecture of inception v3 [34]

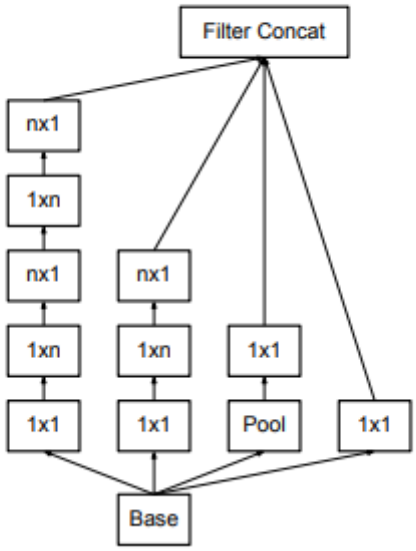


Figure 2.10: 5x Inception Layer architecture of inception v3 [34]

2.3.3 Inception Resnet V2

In machine learning problems after applying multiple layers, generated models used to get overfitted. So, even though DNN was so powerfull, it was not able perform to its full potential. This problem was solved by introducing residual blocks. It solved the problem called vanishing gradient which was responsible for this overfitting problem. Now, after applying the residual blocks DNN architectures can now have hundreds or even thousands of layer and still would not overfit.

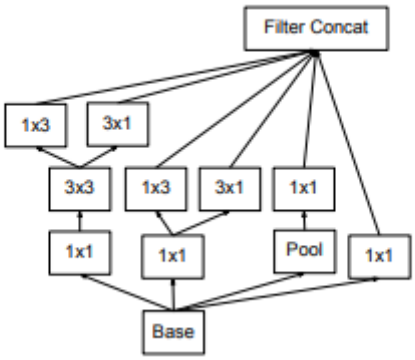


Figure 2.11: 2x Inception Layer architecture of inception v3 [34]

In inception resnet v2, this resnet residual block architecture has been introduced with the inception layer of the improved inception v4 version of the inception family. It is 164 layer deep neural network.

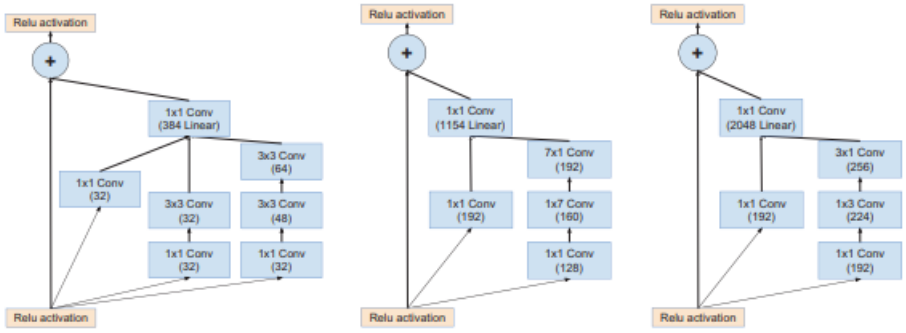


Figure 2.12: Block diagram of Residual operation in Inception Resnet V2 architecture [35]

2.3.4 MobileNet V1

MobileNet is one of most lightest CNN architecture out there. It uses a different approach for convolution operation. Except for the first layer, the mobile web relies on deeply separable convolutions operation which is slightly different than regular convolution operation. The first layer is a complete convolutional layer. It performs batch normalization and ReLU nonlinear processing after all layers. However, the last layer is a fully connected layer without any nonlinearity and is fed forwarded to the softmax for sorting. The total number of layers in the mobile network is 28.

2.3.5 Ensemble

Ensembling is a method used in machine learning and statistics, where multiple algorithms are joined together to get a better outcome. It gives us more accurate data than a single architecture can give individually. In this study we have ensembled two architure namely inception_resnet_v2 and mobilenet_v1 and trained them together to get better accuracy.

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5\times$ Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
$5\times$ Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

Figure 2.13: Layer list of Mobilenet architecture [36]

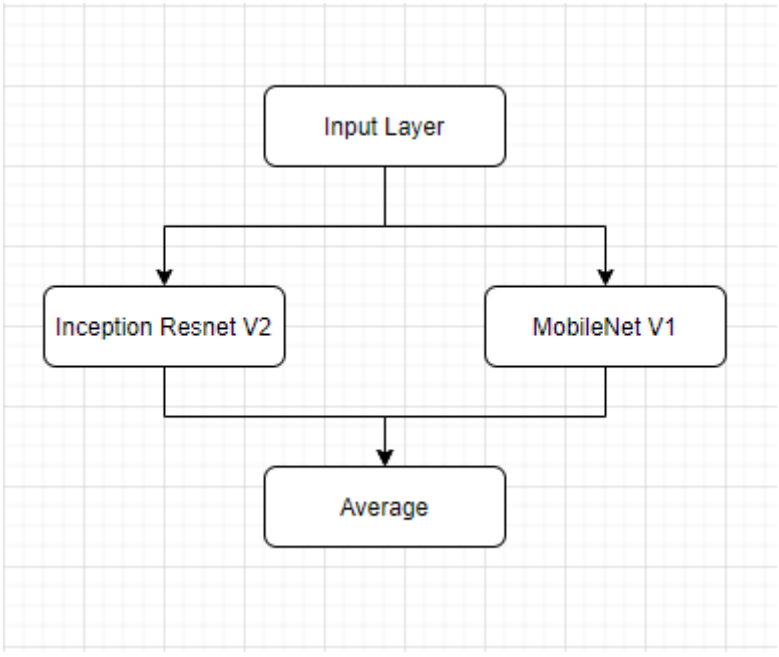


Figure 2.14: Ensemble architecture used in this study

Chapter 3

Dataset Description

3.1 Introduction

The PPMI which stands for Parkinson's Progression Markers Initiative was launched in 2010 for the sole purpose of helping researchers for Parkinson's Disease related issues. It is sponsored by The Michael J. Fox Foundation for Parkinson's Research and supported by various industry, non-profit and private partners. It has data from more than 1,400 individuals at 33 clinical sites in 11 countries.

3.2 Dataset Details

Now, PPMI have both MRI and fMRI data for research purpose. In this study, we are using fMRI data to conduct our research. It has 500+ fMRI subjects data for use. However, we are classifying prodromal PD patients from healthy control subjects. They had 20 control subjects and 20 prodromal subject available at the moment of writing this paper. The patients were between 40 to 80 years old and of both gender.

3.3 Dataset Acquisition Process

The dataset is open sourced however, we had to request for access of the dataset with our organization's identity and had to became a registered member of the PPMI website before accessing them. Once they approved our membership request and we became a registered member and we downloaded the subject information that was necessary for this study.

3.4 Dataset Processing

It took a lot of effort to process the data into our favourable environment. The data processing way is mentioned below.

3.4.1 Conversion from nifti to 2D image

Our downloaded data was in nifti image extension format. However, to work with our favourable architectures we needed to convert these nifti formatted 4D time

series image into simple 2D images. So, at first we had to use python library such as nibabel to convert this nifti extension data into 2D image data.

3.4.2 Grouping into classes

After converting all data into 2D format, we had group our data into 2 different sets of data. One in which all the control subject data was placed and one in which all the prodromal subject data was placed.

3.4.3 Train-Val-Test split

We divided our data in multiple ways. At first we went with the traditional approach. Using python split-folder library, We divided our entire dataset into 70% train, 15% validation and 15% test data. After we got the results of our research with these segmentation of the dataset, we segmented our data with a newer approach where train, test and validation had entirely separate subject data. That is among the 40 control and prodromal subject data, 32 subject data were entirely for training purpose, 4 subject data were entirely for validation purpose and 4 subject data were entirely for test purpose. However, these selections of subjects was random. Once we conducted our research on this random selection of subjects, we shuffled our subjects and experimented with the hope of getting a better test and validation accuracy that is we used newer sets of training, val and test data with 32 training patients, 4 validation patients and 4 test patients where the 8 subjects used in the current test and validation set were not same in the previous run.

Chapter 4

Proposed Methodology

4.1 fMRI Data Acquisition

We have collected our data from the PPMI website. It has fMRI data from over 500+ subjects. However, for this study we are mainly studying on detecting prodromal phase of PD and PPMI had only 20 subject data for Prodromal patients and 20 subject data of healthy controls. So, we collected in a total of 40 subjects data.

4.2 Data Preprocessing

In the dataset there were some data in dicom extension and some were in nifti extension. We had to convert all data into nifti extension to make further processing easier.

4.3 2D Labeling

We used nibabel python library to convert all nifti format data into simple 2D image as the CNN architectures that we used they require 2D images as input.

4.4 Grouping into classes

We then grouped our data into two class. One was healthy Control class and the other was Prodromal PD class. Then we divided our data into different sizes of test, validation and training dataset.

4.5 Classification using CNN models

First we used our first data segmentation where the dataset contained 70% training image, 15% validation and 15% test image. We applied vgg model vgg19 on it and then respectively inception v3 and inception resnet v2 model to train and see the result. Then we used our second dataset segmentation with distinct subjects in training, test and validation set. We used 32 subjects as training set, 4 for validation and 4 for test. The choice of subject was random here. Then respectively we applied inception v3, vgg19, inception resnet v2 and mobilenet v1 on this dataset

and observed it's result. Then we used our third dataset segmentation where we cross-matched every subject as in merry go round to check which set of train, val and test set gives us the maximum accuracy. We tested this merry go round with mobilenet architecture and finally we applied respectively mobilenet and inception resnet v2 to observe our outcome.

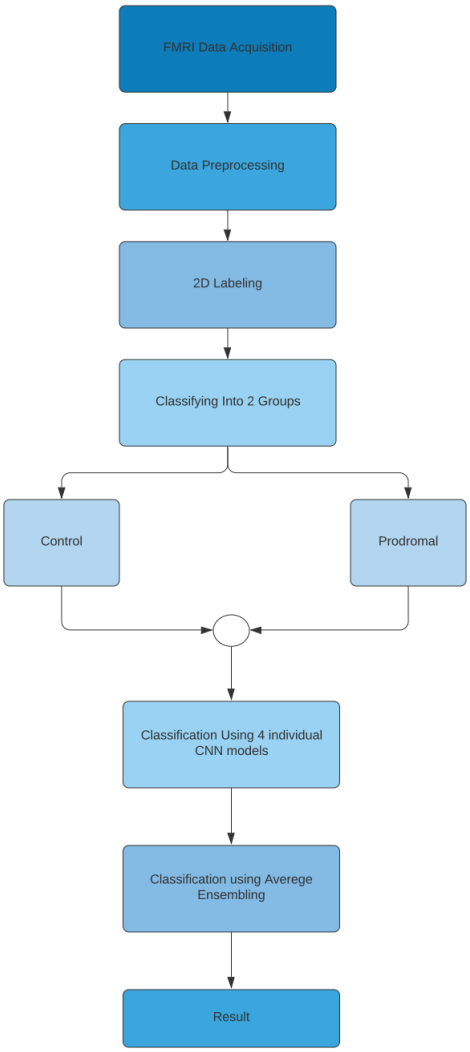


Figure 4.1: Block diagram of our working methodologies

4.6 Classification using Average Ensembling

We ensembled mobilenet and inception resnet v2 on the first and second type of our data segmentation as these two model gave us the highest accuracy on those dataset.

Chapter 5

Result Analysis and Discussion

5.1 Result Using Data segmentation 1

In data segmentation 1, all subject data were grouped together and was splitted into 70% training data, 15% validation data and 15% test data. The CNN models that we used here are vgg19, inception v3, inception resent v2, and ensemble between inception resent v2 and mobilenet v1.

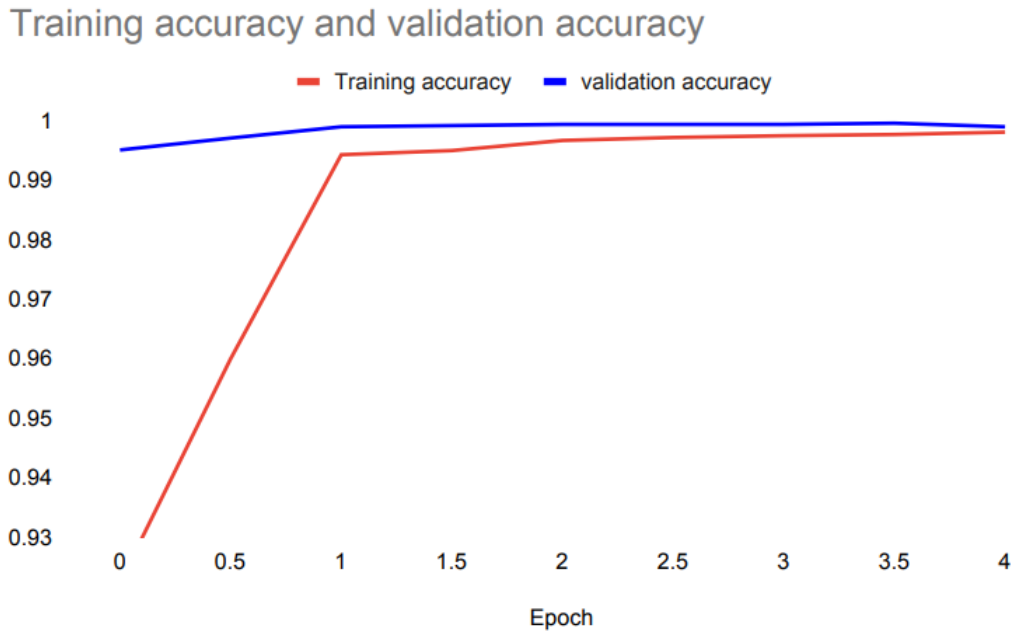


Figure 5.1: graph of vgg19 training and validation accuracy

Table 5.1: vgg19 precision, recall and f1-score for data segmentation 1

Subject	Precision	Recall	f1-score
Control	0.99	0.99	0.99
Prodromal	0.99	0.99	0.99

Figure 5.2: graph of vgg19 training and validation loss

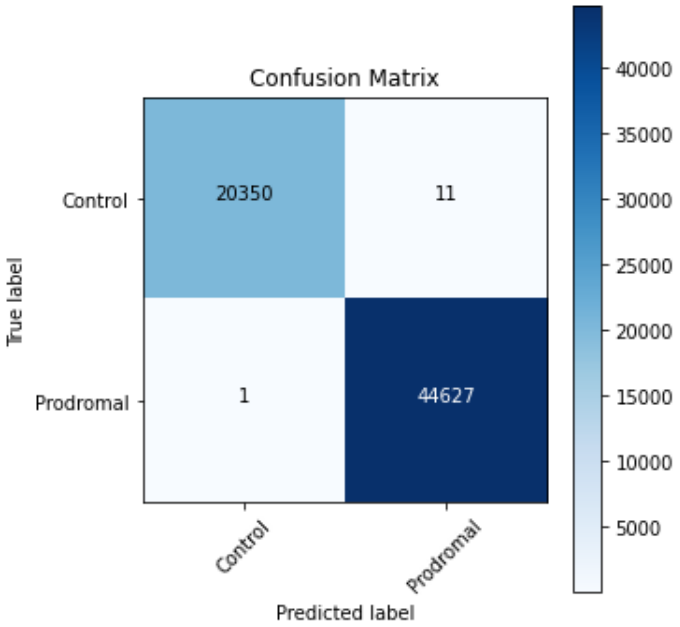


Figure 5.3: vgg19 confusion matrix

Table 5.2: inception v3 precision, recall and f1-score for data segmentation 1

Subject	Precision	Recall	f1-score
Control	0.99	0.99	0.99
Prodromal	0.99	0.99	0.99

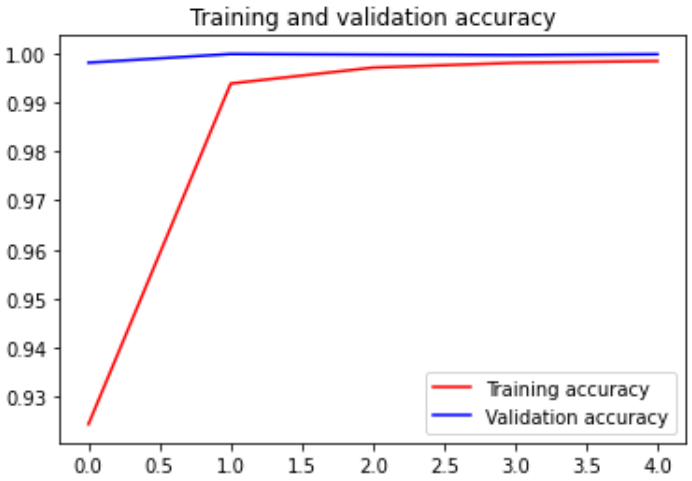


Figure 5.4: graph of inception v3 training and validation accuracy

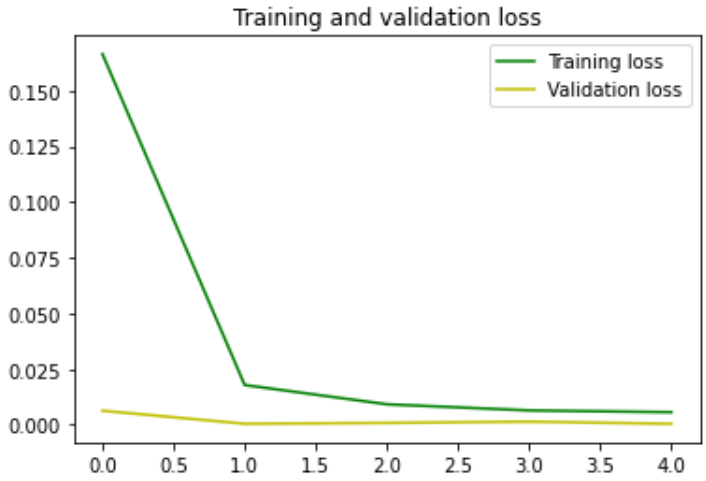


Figure 5.5: graph of inception v3 training and validation loss

Table 5.3: inception resnet V2 precision, recall and f1-score for data segmentation 1

Subject	Precision	Recall	f1-score
Control	0.99	0.99	0.99
Prodromal	0.99	0.99	0.99

Table 5.4: Ensemble Model precision, recall and f1-score for data segmentation 1

Subject	Precision	Recall	f1-score
Control	0.99	0.98	0.99
Prodromal	0.99	0.99	0.99

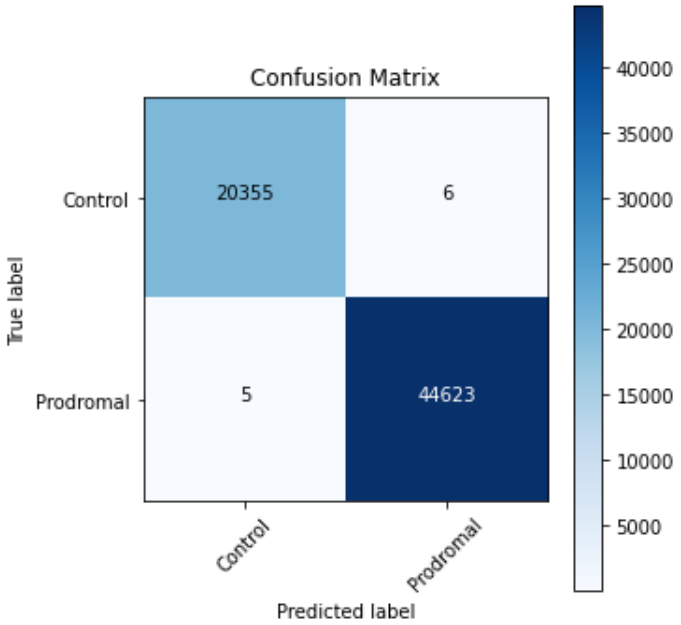


Figure 5.6: inception v3 confusion matrix

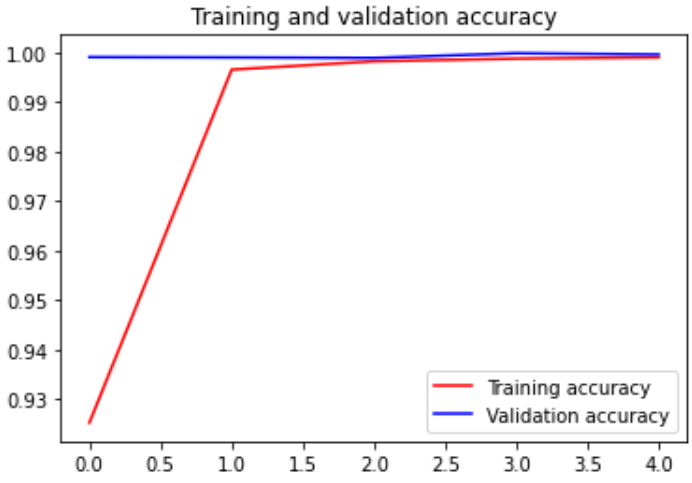


Figure 5.7: graph of inception resnet V2 training and validation accuracy

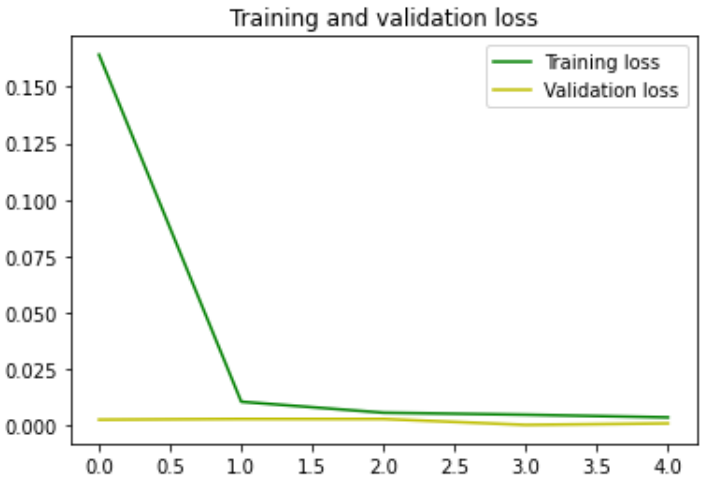


Figure 5.8: graph of inception resnet V2 training and validation loss

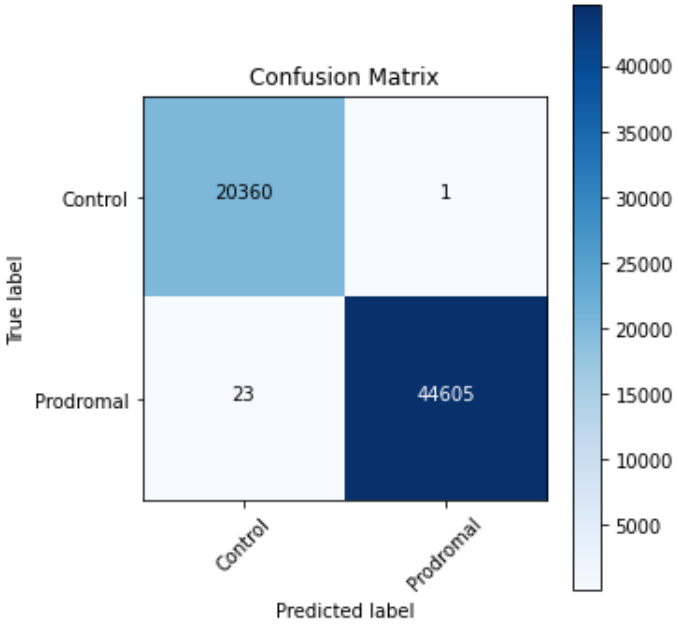


Figure 5.9: inception resnet V2 confusion matrix

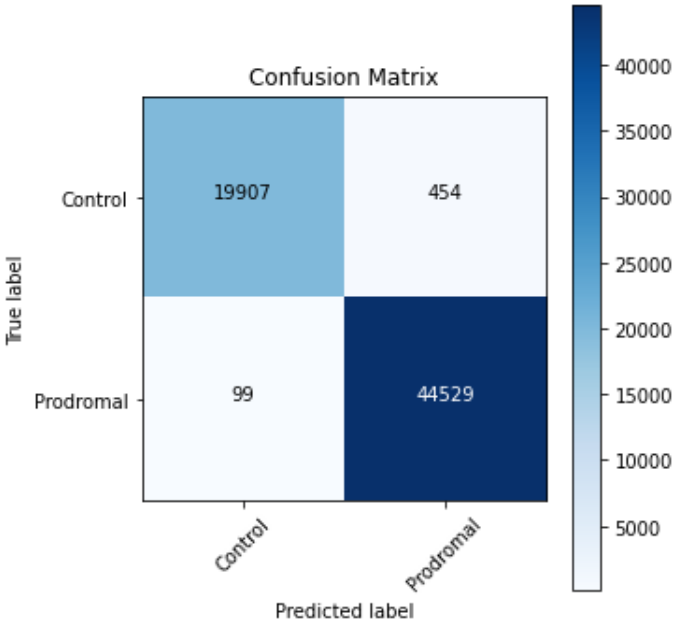


Figure 5.10: Data Segment 1 Ensemble Model confusion matrix

Table 5.5: Loss and Accuracy values for all used architecture after respective numbers of epochs for data segment1

Architecture	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy	Test Accuracy	Epoch
vgg19	0.0062	99.81%	0.0051	99.90	99.89%	5
inception v3	0.0057	99.84%	0.0004	99.98%	99.99%	5
inception resnet v2	0.0036	99.90%	0.0009	99.96%	99.96%	5
ensemble	0.2887	96.83%	0.2359	99.20%	99.07%	4

Discussion

So, we can see that all the models we used did significantly well and was able to accurately predict the data. All of them have test and validation accuracy of nearly 99.99%. Inception V3 almost have 100% test and validation accuracy. The precision and recall for all model was extremely good as well with 99.99% precision and 99.99% recall for all model.

5.2 Result Using Data segmentation 2

In data segmentation 2, we randomly selected 32 subjects and used all of their data as training set with 16 control and 16 prodromal subjects, 4 subjects as validation set with 2 control and 2 prodromal subjects and 4 subjects as test set with 2 control and 2 prodromal subjects. The CNN models that we used here are vgg19, inception v3, inception resent v2, mobilenet v1 and ensemble between inception resent v2 and mobilenet v1.

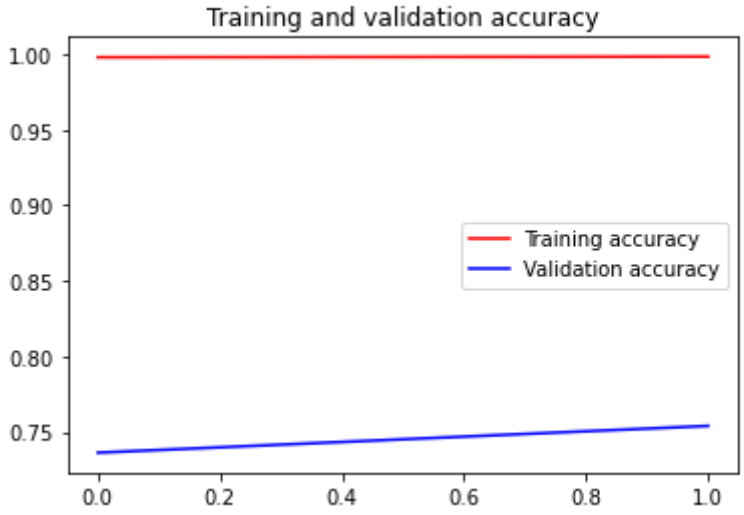


Figure 5.11: graph of vgg19 training and validation accuracy

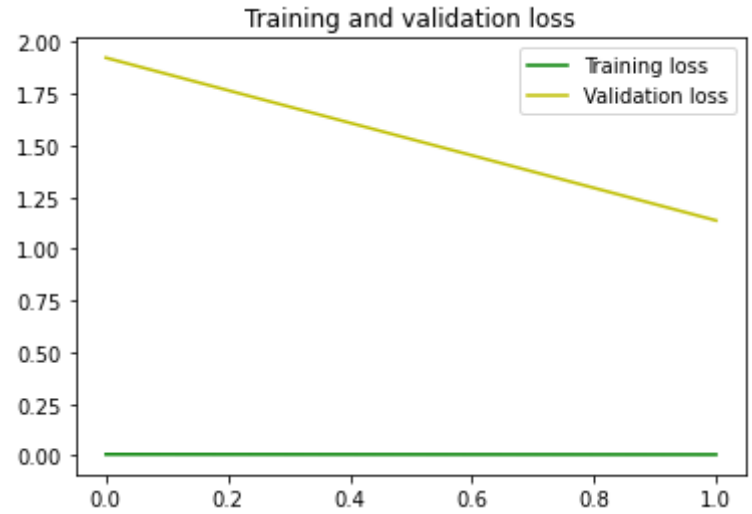


Figure 5.12: graph of vgg19 training and validation loss

Table 5.6: vgg19 precision, recall and f1-score for data segmentation 2

Subject	Precision	Recall	f1-score
Control	0.51	0.58	0.54
Prodromal	0.85	0.81	0.83

Table 5.7: inception v3 precision, recall and f1-score for data segmentation 2

Subject	Precision	Recall	f1-score
Control	0.57	0.62	0.59
Prodromal	0.87	0.84	0.86

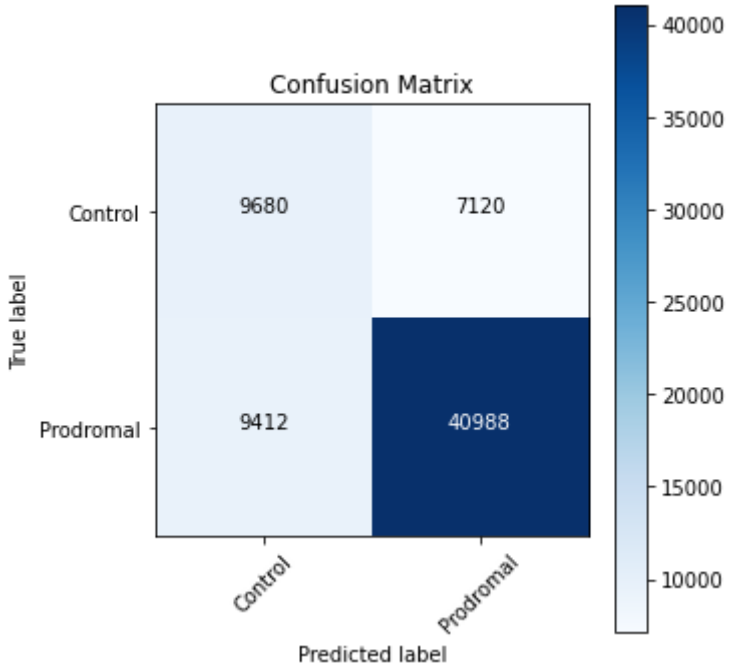


Figure 5.13: vgg19 confusion matrix

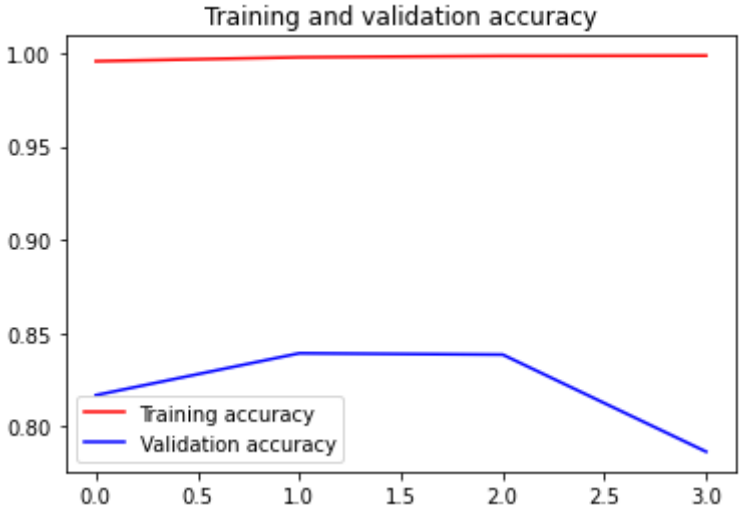


Figure 5.14: graph of inception v3 training and validation accuracy

Table 5.8: inception resnet V2 precision, recall and f1-score for data segmentation 2

Subject	Precision	Recall	f1-score
Control	0.66	0.57	0.61
Prodromal	0.86	0.90	0.88

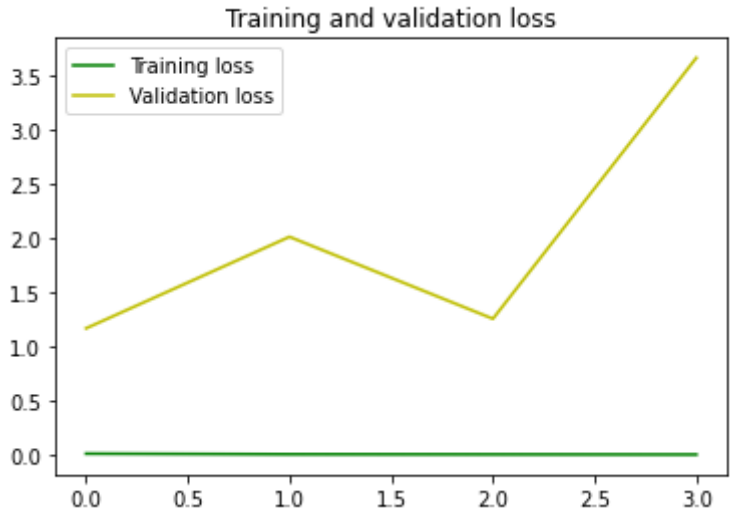


Figure 5.15: graph of inception v3 training and validation loss

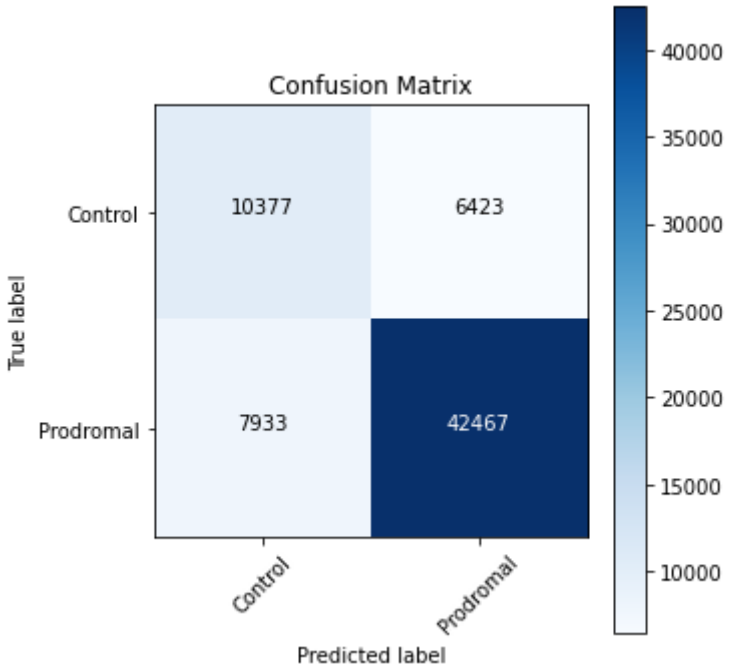


Figure 5.16: inception v3 confusion matrix

Table 5.9: mobilenet V1 precision, recall and f1-score for data segmentation 2

Subject	Precision	Recall	f1-score
Control	0.62	0.30	0.40
Prodromal	0.80	0.94	0.86

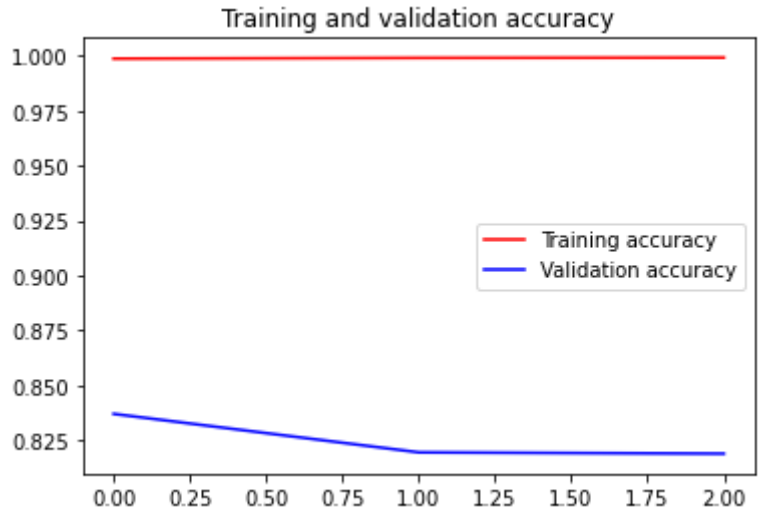


Figure 5.17: graph of inception resnet V2 training and validation accuracy

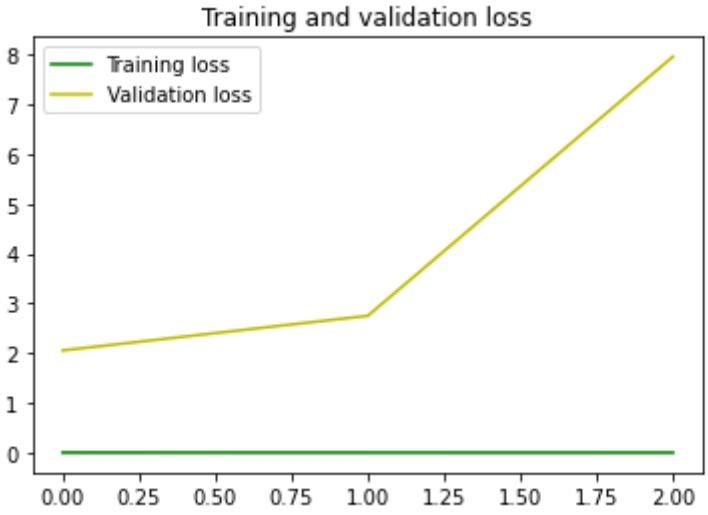


Figure 5.18: graph of inception resnet V2 training and validation loss

Table 5.10: Ensemble Model precision, recall and f1-score for data segmentation 2

Subject	Precision	Recall	f1-score
Control	0.62	0.30	0.40
Prodromal	0.80	0.94	0.86

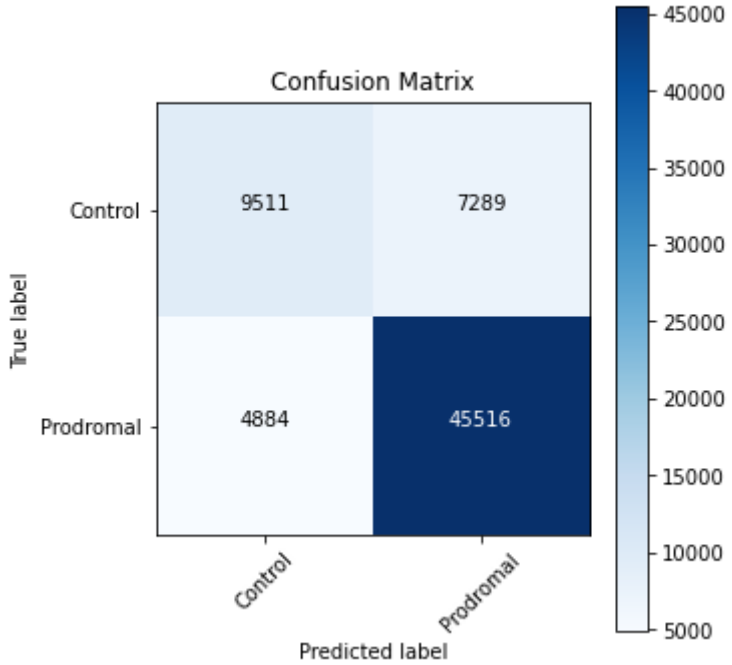


Figure 5.19: inception resnet V2 confusion matrix

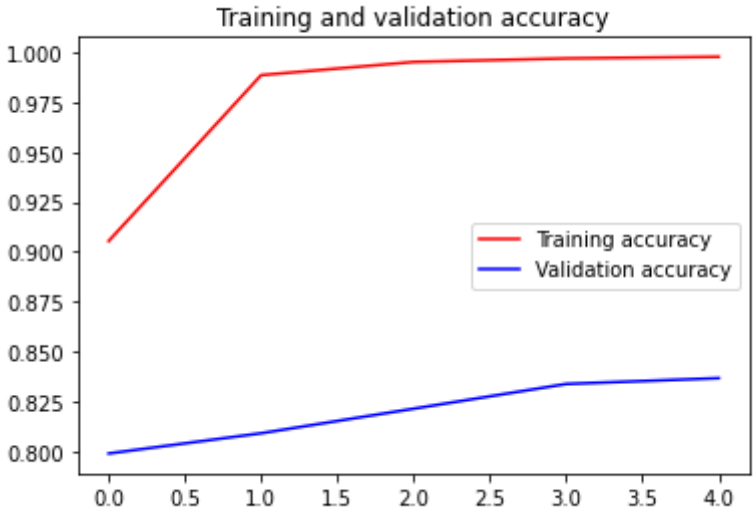


Figure 5.20: graph of mobilenet V1 training and validation accuracy

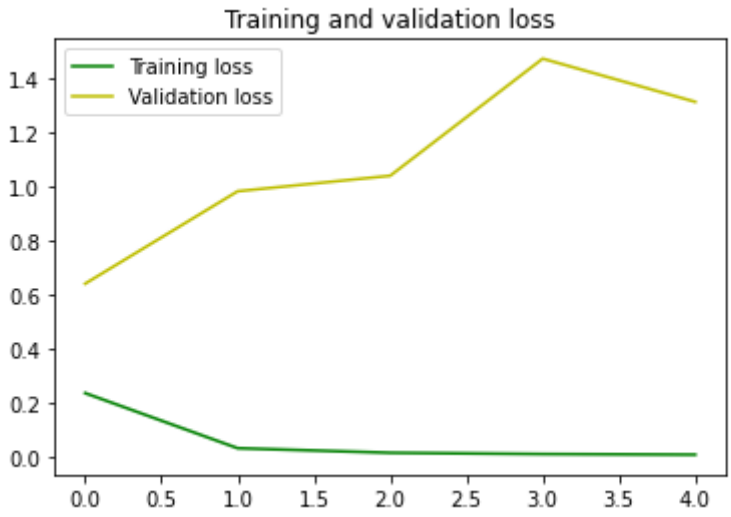


Figure 5.21: graph of mobilenet V1 training and validation loss

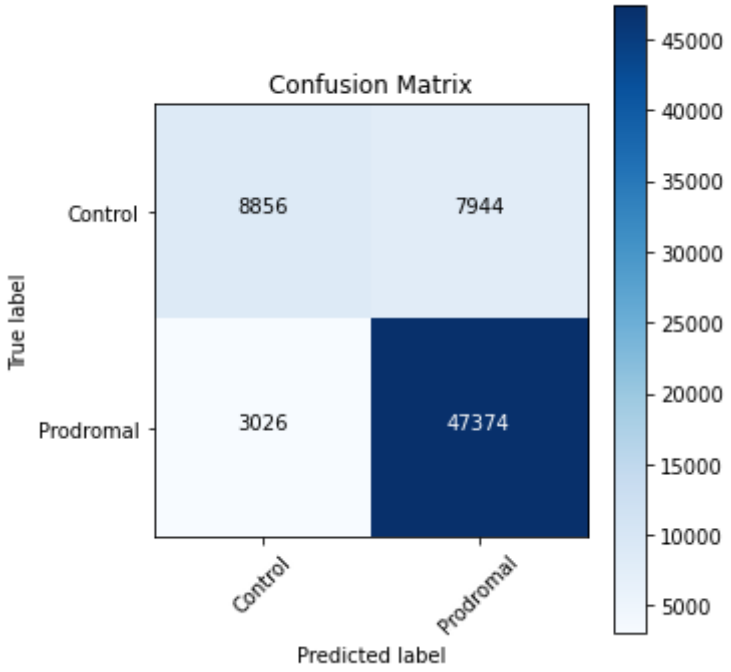


Figure 5.22: mobilenet V1 confusion matrix

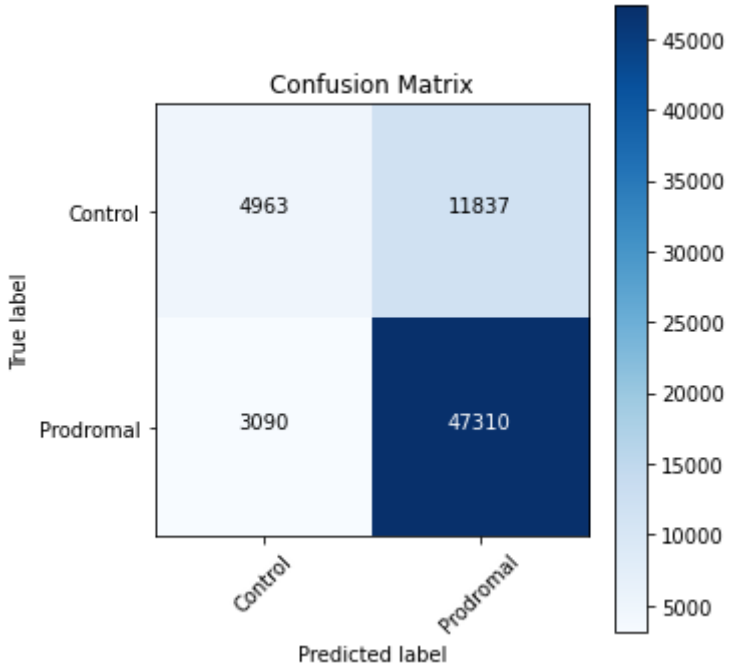


Figure 5.23: Data Segment 2 Ensemble Model confusion matrix

Table 5.11: Loss and Accuracy values for all used architecture after respective numbers of epochs for data segment2

Architecture	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy	Test Accuracy	Epoch
vgg19	0.0056	99.78%	1.1376	75.40%	58.36%	5
inception v3	0.0048	99.87%	3.6670	78.64%	62.55%	5
inception resnet v2	0.0029	99.92%	7.9570	81.89%	63.00%	5
mobilenet v1	0.0078	99.78%	1.3116	83.68%	67.20%	5
ensemble	0.2382	98.24%	0.9268	77.79%	63.32%	4

Discussion

So, for data segmentation 2 we can see that everyone has high training accuracy of almost 99%. However, even though validation accuracy is acceptable and test ac curacies are very poor. We can see that mobileNet v1 has the highest test and validation accuracy of 67.20% and 83.68%. However, if we observe closely we can see that precision and recall value of Control data was poor but precision and recall of prodromal data was significantly good. Ensemble and mobilenet v1 having highest prodromal recall with both having 94% accuracy and inception v3 having the highest prodromal precision of 87% accuracy and inception resnet v2 having 86% accuracy. Inception resnet v2 have the highest Control precision of 66.% and inception v3 having the highest Control recall of 62%. So, it is clear that the models are struggling to detect previously unseen Control Subjects.

5.3 Result Using Data segmentation 3

To achieve higher test accuracy, we continuously repeated what we did in data segmentation 2 and applied our models on to it. That is we continued to shuffled our subjects with newer sets of training, val and test data with 32 training patients, 4 validation patients and 4 test patients where the 8 subjects used in the current test and validation set were not same in the previous run. The CNN architectures that we used in this segment are inception resnet v2 and mobilenet v1.

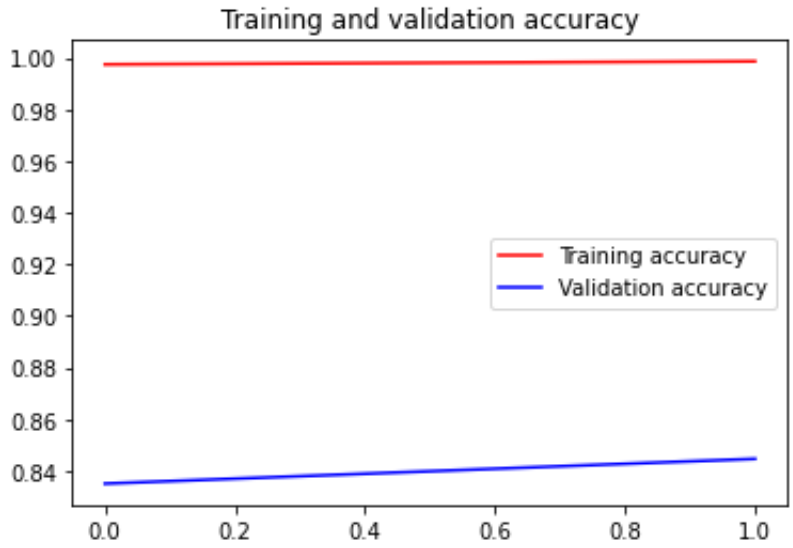


Figure 5.24: graph of inception resnet V2 training and validation accuracy

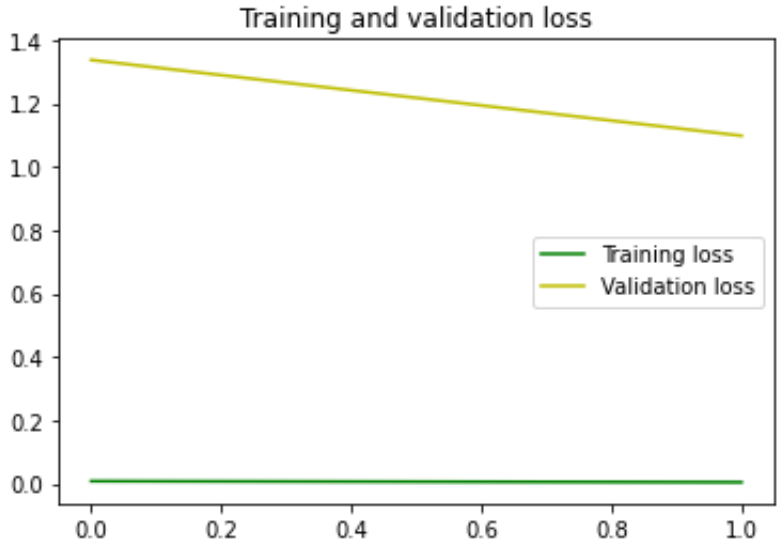


Figure 5.25: graph of inception resnet V2 training and validation loss

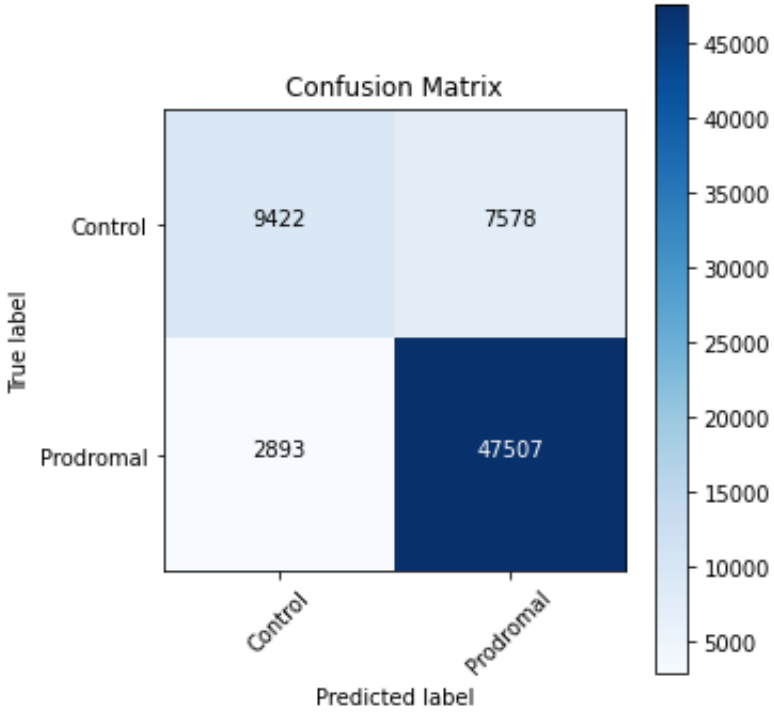


Figure 5.26: inception resnet V2 confusion matrix

Table 5.12: inception resnet V2 precision, recall and f1-score

Subject	Precision	Recall	f1-score
Control	0.77	0.55	0.64
Prodromal	0.86	0.94	0.90

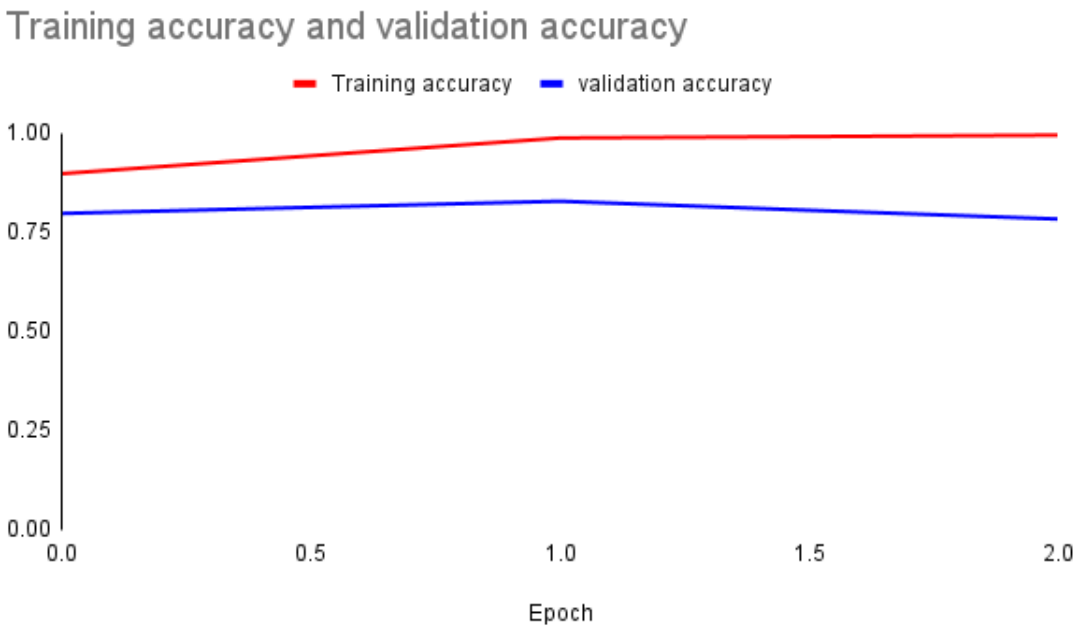


Figure 5.27: graph of mobilenet V1 training and validation accuracy

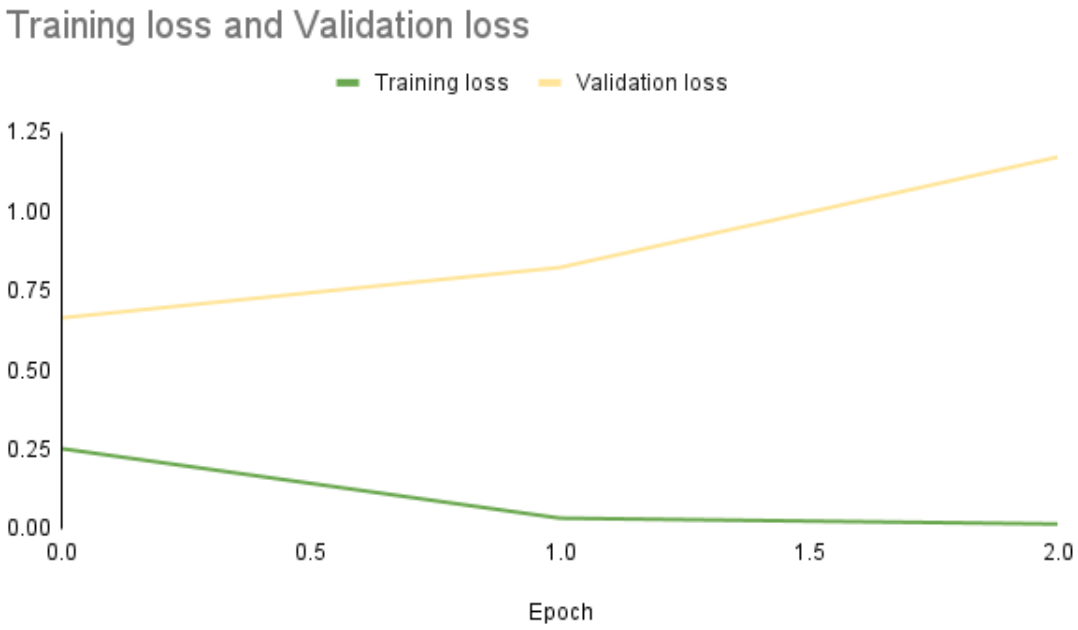


Figure 5.28: graph of mobilenet V1 training and validation loss

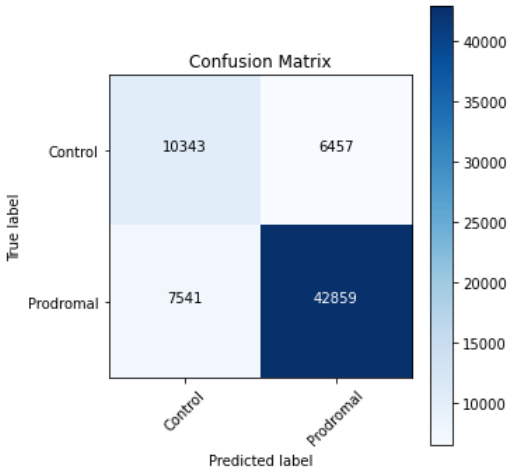


Figure 5.29: mobilenet V1 confusion matrix

Table 5.13: mobilenet V1 precision, recall and f1-score

Subject	Precision	Recall	f1-score
Control	0.58	0.62	0.60
Prodromal	0.87	0.85	0.86

Table 5.14: Loss and Accuracy values for all used architecture after respective numbers of epochs for data segment3

Architecture	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy	Test Accuracy	Epoch
inception resnet v2	0.0045	99.86%	1.0981	84.46%	75.30%	3
Mobilenet V1	0.0157	99.49%	1.1719	78.35%	81.22%	3

Discussion

So, after shuffling our dataset with all possible combination of different test and validation set, we indeed achieved proper test and validation accuracy. Where inception resnet v2 have val and test accuracy of 84.46% and 75.30% and mobilenet v1 have val and test accuracy of 78.35% and 81.22%. Also for Control patient, the precision and recall value is also higher than segment 2 dataset with inception resnet v2 having 77% precision and 55% recall and mobilenet v1 having 58% precision and 62% recall. However, in this case as well we got very high precision and recall for Predromal patient. So, regardless of which dataset we used our models are successfully classifying those those with Prodromal symptoms. As detecting the disease is more important than detecting healthy patients, our models are very successful regarding that task.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

Early detection of Parkinson's Disease is becoming very important day by day. With the help of deep learning techniques we can very accurately classify image datas. Even with previously unseen data. In this study, we have successfully classified Prodromal PD patients from healthy Controls with high accuracy of 81.22%.

6.2 Challenges/Limitations

The first challenge that we faced was the hardware problem. Due to the unwanted Covid situation we were unable to access our university facilities for training our data. We had two options to either use google colab or use a local machine. Fortunately, we found one local machine to process our work among our close acquaintances. However, after that we faced our second problem of making the cpu and gpu compatible with tensorflow gpu environment. It took us days to sort out this problem. We were being delayed to work in colab as well because, uploading dataset in google drive was an extremely slow process. However, finally, we did not have to work in the colab and we were able make our pc compatible with the environment and was able to finish our work. We had to struggle with reducing our overfitting problem as well because the features of Parkinson's disease are not as visible as some other diseases like tumor etc. and as a result, extracting feature maps from it was very difficult for our models.

6.3 Future Work

We are working on detecting the prodromal phase of Parkinson's Disease which is a pre-motor symptoms of PD. The most common problem that PD patient face during this phase is REM sleep behavior disorder or RBD. Now, RBD is an early syndrome of Alzheimer's disease [37] and schizophrenia [38] patients as well. So, there is a possibility that studying on RBD patients may give us more insight on how to detect all three of these brain disorder at the same time. So, we would like to apply the trained model that we have achieved in this study on prodromal Alzheimer and prodromal schizophrenia patients and see whether our models can detect them with significant accuracy or not.

Bibliography

[1] L. M. De Lau and M. M. Breteler, "Epidemiology of parkinson's disease," *The Lancet Neurology*, vol. 5, no. 6, pp. 525–535, 2006.

[2] W. Dauer and S. Przedborski, "Parkinson's disease: Mechanisms and models," *Neuron*, vol. 39, no. 6, pp. 889–909, 2003.

[3] A. E. Lang and A. M. Lozano, "Parkinson's disease," *New England Journal of Medicine*, vol. 339, no. 16, pp. 1130–1143, 1998.

[4] T. Vos, C. Allen, M. Arora, R. M. Barber, Z. A. Bhutta, A. Brown, A. Carter, D. C. Casey, F. J. Charlson, A. Z. Chen, *et al.*, "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the global burden of disease study 2015," *The lancet*, vol. 388, no. 10053, pp. 1545–1602, 2016.

[5] B. Peng, S. Wang, Z. Zhou, Y. Liu, B. Tong, T. Zhang, and Y. Dai, "A multilevel-roi-features-based machine learning method for detection of morphometric biomarkers in parkinson's disease," *Neuroscience letters*, vol. 651, pp. 88–94, 2017.

[6] H.-C. Cheng, C. M. Ulane, and R. E. Burke, "Clinical progression in parkinson disease and the neurobiology of axons," *Annals of neurology*, vol. 67, no. 6, pp. 715–725, 2010.

[7] K. R. Chaudhuri, D. G. Healy, and A. H. Schapira, "Non-motor symptoms of parkinson's disease: Diagnosis and management," *The Lancet Neurology*, vol. 5, no. 3, pp. 235–245, 2006.

[8] W. Poewe, "Non-motor symptoms in parkinson's disease," *European journal of neurology*, vol. 15, pp. 14–20, 2008.

[9] K. R. Chaudhuri and A. H. Schapira, "Non-motor symptoms of parkinson's disease: Dopaminergic pathophysiology and treatment," *The Lancet Neurology*, vol. 8, no. 5, pp. 464–474, 2009.

[10] R. F. Pfeiffer, "Non-motor symptoms in parkinson's disease," *Parkinsonism & related disorders*, vol. 22, S119–S122, 2016.

[11] R. B. Postuma and D. Berg, "Advances in markers of prodromal parkinson disease," *Nature Reviews Neurology*, vol. 12, no. 11, p. 622, 2016.

[12] D. Jennings, A. Siderowf, M. Stern, J. Seibyl, S. Eberly, D. Oakes, K. Marek, P. Investigators, *et al.*, "Imaging prodromal parkinson disease: The parkinson associated risk syndrome study," *Neurology*, vol. 83, no. 19, pp. 1739–1746, 2014.

[13] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[14] L. Jeancolas, D. Petrovska-Delacrétaz, G. Mangone, B.-E. Benkelfat, J.-C. Corvol, M. Vidailhet, S. Lehéricy, and H. Benali, "X-vectors: New quantitative biomarkers for early parkinson's disease detection from speech," *Frontiers in Neuroinformatics*, vol. 15, p. 4, 2021.

[15] J. M. Tracy, Y. Özkanca, D. C. Atkins, and R. H. Ghomi, "Investigating voice as a biomarker: Deep phenotyping methods for early detection of parkinson's disease," *Journal of biomedical informatics*, vol. 104, p. 103 362, 2020.

[16] A. Rahman, A. Khan, and A. A. Raza, "Parkinson's disease detection based on signal processing algorithms and machine learning," *CRPASE: Transactions of Electrical, Electronic and Computer Engineering*, vol. 6, pp. 141–145, 2020.

[17] D. Iakovakis, R. E. Mastoras, S. Hadjidimitriou, V. Charisis, S. Bostanjopoulou, Z. Katsarou, L. Klingelhoefer, H. Reichmann, D. Trivedi, R. K. Chaudhuri, *et al.*, "Smartwatch-based activity analysis during sleep for early parkinson's disease detection," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 2020, pp. 4326–4329.

[18] L. Ali, C. Zhu, N. A. Golilarz, A. Javeed, M. Zhou, and Y. Liu, "Reliable parkinson's disease detection by analyzing handwritten drawings: Construction of an unbiased cascaded learning system based on feature selection and adaptive boosting model," *Ieee Access*, vol. 7, pp. 116 480–116 489, 2019.

[19] N. Amoroso, M. La Rocca, A. Monaco, R. Bellotti, and S. Tangaro, "Complex networks reveal early mri markers of parkinson's disease," *Medical image analysis*, vol. 48, pp. 12–24, 2018.

[20] S. Shinde, S. Prasad, Y. Saboo, R. Kaushick, J. Saini, P. K. Pal, and M. Ingalhalikar, "Predictive markers for parkinson's disease using deep neural nets on neuromelanin sensitive mri," *NeuroImage: Clinical*, vol. 22, p. 101 748, 2019.

[21] O. Cigdem, I. Beheshti, and H. Demirel, "Effects of different covariates and contrasts on classification of parkinson's disease using structural mri," *Computers in biology and medicine*, vol. 99, pp. 173–181, 2018.

[22] P. R, S. Dutta Roy, P. Mandal, and S. Ghosh, "High-accuracy detection of early parkinson's disease through multimodal features and machine learning," *International Journal of Medical Informatics*, vol. 90, Mar. 2016. DOI: 10.1016/j.ijmedinf.2016.03.001.

[23] R. Prashanth and S. D. Roy, "Early detection of parkinson's disease through patient questionnaire and predictive modelling," *International journal of medical informatics*, vol. 119, pp. 75–87, 2018.

[24] T. A. Mostafa and I. Cheng, "Parkinson's disease detection with ensemble architectures based on ilsvrc models," *arXiv preprint arXiv:2007.12496*, 2020.

[25] P. Kamencay, M. Benčo, T. Mižďoš, and R. Radil, "A new method for face recognition using convolutional neural network," 2017.

[26] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, Ieee, 2017, pp. 1–6.

[27] K. Basterretxea, J. M. Tarela, and I. Del Campo, "Approximation of sigmoid function and the derivative for hardware implementation of artificial neurons," *IEE Proceedings-Circuits, Devices and Systems*, vol. 151, no. 1, pp. 18–24, 2004.

[28] M. Straat, "On-line learning in neural networks with relu activations," Ph.D. dissertation, 2018.

[29] A. Dertat, "Applied deep learning-part 4: Convolutional neural networks," *Toward DataScience,[Online]. Available: https://towardsdatascience. com/applied-deeplearning-part-4-convolutional-neural-networks-584bc134c1e2*, 2017.

[30] P. Baldi and P. J. Sadowski, "Understanding dropout," *Advances in neural information processing systems*, vol. 26, pp. 2814–2822, 2013.

[31] S. Santurkar, D. Tsipras, A. Ilyas, and A. Madry, "How does batch normalization help optimization?" *arXiv preprint arXiv:1805.11604*, 2018.

[32] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[33] Y. Zheng, C. Yang, and A. Merkulov, "Breast cancer screening using convolutional neural network and follow-up digital mammography," in *Computational Imaging III*, International Society for Optics and Photonics, vol. 10669, 2018, p. 1 066 905.

[34] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[35] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, 2017.

[36] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[37] D. L. Bliwise, "Sleep disorders in alzheimer's disease and other dementias," *Clinical cornerstone*, vol. 6, no. 1, S16–S28, 2004.

[38] S. Chouinard, J. Poulin, E. Stip, and R. Godbout, "Sleep in untreated patients with schizophrenia: A meta-analysis.," *Schizophrenia bulletin*, vol. 30, no. 4, p. 957, 2004.