Design and Evaluation of Convolutional Neural Network for Detection of Alzheimer's Disease Using MRI Data

by

Riasat Mahbub
17201146
Muhammad Anwarul Azim
18101624
Khondaker Masfiq Reza
18301104
Md Nafiz Ishtiaque Mahee
18101489
MD. Zahidul Islam Sanjid
18101564

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

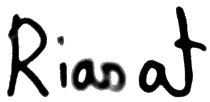
Department of Computer Science and Engineering
Brac University
September 2021

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:

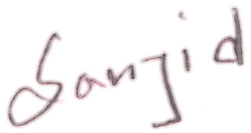


Riasat Mahbub
17201146

Muhammad Anwarul Azim
18101624

Khondaker Masfiq Reza
18301104

Md Nafiz Ishtiaque Mahee
18101489



MD. Zahidul Islam Sanjid
18101564

Approval

The thesis/project titled "Design and Evaluation of Convolutional Neural Network forDetection of Alzheimer's Disease Using MRI Data" is submitted by

1. Riasat Mahbub(17201146)
2. Muhammad Anwarul Azim(18101624)
3. Khondaker Masfiq Reza(18301104)
4. Md Nafiz Ishtiaque Mahee(18101489)
5. MD. Zahidul Islam Sanjid(18101564)

Of Summer, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on September 26, 2021.

Examining Committee:

Supervisor:
(Member)

Zavid Parvez Digitally signed by Zavid Parvez
DN: cn=Zavid Parvez, o=Brac University, ou=CSE, email=zavid.parvez@bracu.ac.bd, c=US
Date: 2021.10.05 11:20:17 +11'00'

Mohammad Zavid Parvez , PhD
Assistant 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

Alzheimer's Disease (AD) is a neurological condition where the decline of brain cells causes acute memory loss and severe loss in cognitive functionalities. Various Neuroimaging techniques have been developed to diagnose AD; among those, Magnetic Resonance Imaging (MRI) is one of the most prominent ones. Recent progress in medical image analysis using deep learning especially has automated this task significantly. Although the state-of-the-art architectures have achieved human-level performance in classifying AD images from Normal Control (NC), they often require predefined Regions of interest as a basis for feature extraction. This condition not only requires specialized domain knowledge of the human brain but also makes the overall design complicated. In this study, we designed a 15 layer Neural network architecture that can facilitate AD diagnosis without being dependent on any such neurological assumption. The network was tested over ADNI-1, a benchmark MRI dataset for AD research, and found an accuracy of 92.41% (AUC = 0.93). This network was further augmented with the help of ensemble learning other well known pre trained models for more accurate and consistent results, resulting in an overall accuracy of 92.44% for the entire system.

Keywords: Alzheimer's Disease, Magnetic Resonance Imaging, Convolutional Neural Network

Dedication

Our work is dedicated to our parents, without whom we could never come this far in our lives. And a special Thanks to our supervisor who provided us their utmost support.

Acknowledgement

Firstly, all praise to the Almighty Allah for whom our thesis have been completed without any major disruption.
Secondly, to our honourable Advisor Mohammad Zavid Parvez sir for his utmost support.
And finally to our beloved parents without their constant support it may not be possible.

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

λ Lambda

AD Alzheimer's Disease

$ADNI$ Alzheimer's Disease Neuroimaging Initiative

AUC Area Under the Curve

CN Cognitively Normal

CNN Convolutional Neural Network

MCI Mild Cognitive Impairment

ML Machine Learning

MRI Magnetic Resonance Imaging

$NIfTI$ Neuroimaging Informatics Technology Initiative

$ReLU$ Rectified Linear Unit

$ResNet$ Residual Networks

ROC Receiver Operating Caracteristic

ROI Region of Interest

$SMRI$ Structural Magnetic resonance imaging

VGG Visual Geometry Group

Chapter 1

Introduction

Alzheimer's Disease (AD) is a neurodegenerative disease that slowly but indeed deteriorates one's memory and other cognitive skills[11]. It is one of the most common causes of dementia for adults older than 65 [18]. People with advanced age, possessing symptoms of Mild Cognitive Impairment (MCI), have the highest risk of developing AD [29]. About 50 million people were affected by various forms of dementia [23]. Alarmingly, by 2050, one new case of AD is expected to develop every 33 seconds, which is nearly 1 million new cases per year [23]. AD treatment is expensive as it alone costs the United States Government approximately $100 billion each year [4]. Making things worse, most of the affected are not in well-developed regions [8], which quickly becomes a reason for many people to take proper treatment only at a late stage of AD. Thereby, early diagnosis of AD becomes a crucial factor for adequate treatment.

MRI has been widely used as a common AD biomarker to accurately detect its onset as it gives a detailed representation of the brain's structure which is necessary for identifying symptoms co-related to AD [12].

Recently, machine learning techniques from a wide range have been developed to automate the diagnosis of AD using MR images [13]. We characterized these techniques into two main groups- linear statistical learning and deep nonlinear learning. Regardless of the group, convolution is seen as a standard method for feature extraction [45]. Since features correlated to AD consist of multiple modalities, nonlinear classifiers like Convolutional Neural Network (CNN) generally outperform linear classifiers like Support Vector Machnie (SVM) or Random Forest [16],[45]. Due to this superior performance in classification, and the availability of sufficiently large dataset like ADNI[44], MIRIAD [14], and OASIS **oasis**, numerous research has been conducted focusing on CNN exclusively [39],[25], [34], [32].

Chapter 2

Background Information

2.1 Human Brain and Neurodegeneration

The human brain, the principal organ of the central nervous system, consists of the cerebrum, brainstem, cerebellum, and spinal cord. Two cerebral hemispheres make up the cerebrum- the most distinguished portion of the human brain. Each hemisphere has a white matter inner core and a grey matter outer surface-also known as the cerebral cortex [3].

The frontal, temporal, parietal, and occipital lobes are the four lobes that make up each hemisphere. Each of the lobes has a strong correlation with a specific function. For example, The frontal lobe is correlated with the cognitive control of human behavior, which includes a large set of tasks- ranging from abstract thought to self-control. The Amygdala, situated in the Temporal lobe, plays a significant role in memory formation and emotion regulation. Another component of the Temporal lobe- The Hippocampus, coordinates the learning process along with memory generation [19],[10]. Regardless of the functionality, all brain segments have Neuron, electrically excitable cells as their fundamental components.

Generally, a Neuron has three major parts- The cell body, a single axon, and several dendrites. The axon is connected to other Neurons dendrites creating a Neural Circuit. Several Neural Circuits consolidate the large-scale brain networks and execute specific functions. However, several factors can cause systematic loss of this Neural structure leading to Neurodegeneration- a process manifested by diseases like multiple sclerosis, Parkinson's disease, and most importantly- Alzheimer's disease[9].

2.2 Disease Mechanism and Stages

Alzheimer's disease has been classified as a proteopathy where specific proteins deviate from their regular structure[6]. As a result, the cell containing those misfolded proteins can not execute their biological functionalities. Most importantly, these abnormal proteins often act like toxic substances, disrupting the whole tissue. For Alzheimer's disease, two misfolded protein occurrences, amyloid-beta [2] and tau proteins [5], have been identified as the most significant biomarkers. Although both biomarkers have a strong correlation with aging, their causal relationship is still

unknown [15].

Generally, the accumulation of misfolded proteins in the human brain is a slow process that makes AD hardly distinguishable in the early stage. This is also because of the subtle differences between AD symptoms and normal aging. Nevertheless, the transition between normal aging and AD is clinically identified, known as Mild Cognitive Impairment (MCI). MCI is categorized into two categories- Amnestic MCI (aMCI) and Nonamnestic MCI (naMCI) [7]. aMCI is the earlier stage where only the person's memory is affected and manifested by Short Time Memory Loos. However, in naMCI, which is the later stage, noticeable cognitive skills other than memory also deteriorate, including language and Visuospatial functions like understanding depth from the visual scene. Gradually, the patient's ability to speak declines and sentences diminished to a single word. In the severe stage of AD, a patient loses most of the cognitive skills and becomes dependent upon the caregiver entirely [1].

2.3 Literature Review

As our research topic is very specific, a structured literature search process is needed for the literature review. Having this in mind, we used Scopus,(https://www.scopus.com), a large abstract and citation database containing around 78 million literature records from various research discipline.

We started by searching with the string "Alzheimer's Disease" and got a large record list. Then, we narrowed down the list by specifying search string gradually. This search process was conducted in January 2021. Details of this process is demonstrated in fig.3.1.

There was a total of 205 literature records in the narrowed down list which we obtained using search string " ("Alzheimer's disease" AND "MRI" AND "Convolutional Neural Network") ". All of these research work were published between 2014 and 2020.However, we had to exclude 119 records as their access were closed and we included 85 records for literature review finally. Among those, 60 were journal article, 22 were conference paper and 3 were review articles . We reviewed these research works on the basis of the architectures and solutions used during during those studies. Some of the frequently used architectures include-

1. **3D convolutional neural network :** It is similar to conventional CNN however, it uses 3D matrix in both conventional and maxpool layer. We have gotten a total of 8 research where 3D-CNN was used [21], [25], [27], [30], [35], [37], [40], [42], [43].

2. **Convolutional Autoencoder :** It is a neural network architecture which follows an unsupervised approach to regenerate the input. This architecture is effective for dimensionality reduction and feature extraction.We have gotten a total of 9 research where 3D-CNN was used for capturing anatomical

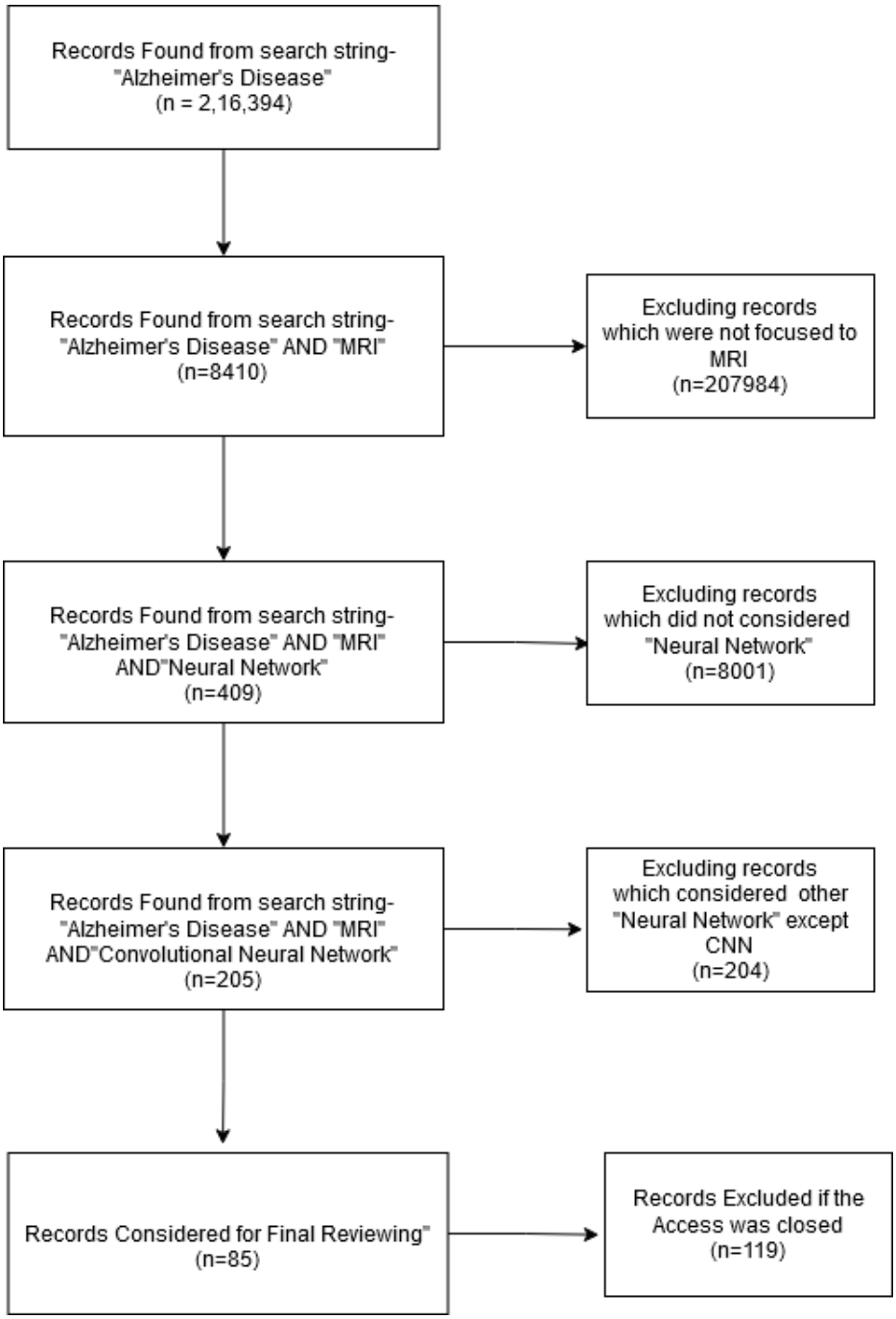


Figure 2.1: Literature Search Process

shape variations in sMRI [21],increasing detection performance [37],classifying progressive mild cognitive impairment (pMCI) and stable mild cognitive impairment (sMCI)[38],and achieving robust performance in small dataset.

3. **Long Short Term Memory :** Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture which has feedback connections between neurons. This feature makes it capable of sequence data and time series. We found LSTM being used for classifying pMCI and sMCI[28] and for detecting Amnestic Mild Cognitive Impairment (aMCI) [41].

From the literature review, we found different CNN architectures providing various advantages towards AD research.Some models provides extraordinary accuracy, some models providing robust classification result (Higher AUC for ROC), and some model ensure satisfactory performance even without the burden of preprocessing or training with large datasets.We are describing some such findings which were disclosed in the literature we reviewed.

Liu et al. designed a landmark-based deep multi-instance learning framework and evaluated the model using both ADNI and MIRIAD dataset. This framework was proved as effective [34]. Lin et al. designed a CNN model for predicting MCI as a early diagnosis for AD patients. Their model, which was trained using 2.5D patches, showed a moderate accuracy of 79.9%. However, the area under Receiver Operating Characteristic Curve was 86.1%, depicting a well balance between sensitivity and specificity [33]. Khan et al. tried to mitigate the dependency on large dataset for CNN training. They integrated transfer learning with the very famous VGG architecture, tested over ADNI dataset, and showed that this approach provide a 4% and a 7% increase in accuracy over the for AD vs. MCI and MCI vs. NC, classification respectively [36].

In 2017, Korolev et al. designed a 21 layers Residual Neural Network and a 17 layers 3D Convolutional Neural Network (3D-CNN) for the binary classification of different AD stages[25]. In an experiment conducted over the ADNI dataset, these two architectures achieved 80% (AUC = .88) and 79% (AUC = .87) classification accuracy, respectively, after running 50 epochs.

In the same year, Li et al. designed a "Y shaped" residual network architecture, where two identical sub-networks with residual blocks extracted features from the Right and Left Hippocampus separately [26]. Later in a fully connected (FC) layer, outputs from these two networks were merged for binary classification. This network was trained on the ADNI I dataset and validated on ADNI Go & ADNI 2 datasets, achieving 0.939 AUC.

In 2018, Khvostikova et al. designed a 3D-CNN that was somewhat similar to the previously mentioned network in the sense that it also leveraged separate identical networks and merged their output in an FC layer [31]. However, unlike relying on two major brain components al a whole (Right and Left Hippocampus) and running two sub-networks, they considered several Regions of Interest (ROI) throughout the

hippocampus and generated that number of sub-networks for feature extraction. They experimented on ADNI several times with varying numbers of RoI- ranging from 28 to 48. This experiment achieved a maximum of 96.7% accuracy while considering 48 RoI.

In the same year, Liu et al. conducted a similar study from a different paradigm. Instead of extracting features from pre-defined RoI, they used a fixed patch landmark detector for identifying landmarks throughout the brain before using those in a pre-trained CNN for binary classification [34].They used three independent datasets (ADNI-1,2 and MIRIAD) and achieved 91%-92% accuracy every time for AD vs. NC classification.

In 2020, Lian et al. designed a Hierarchical Fully Convolutional Network (H-FCN) for the same purpose. Unlike the previously mentioned 2 staged networks, this architecture was three-stage, and the same network was responsible for both the region proposal and classification [32]. As a result, the feature extraction became coupled with the classification process. This H-FCN was trained by the ADNI-1 dataset and tested over ADNI-2 dataset-achieving 90% accuracy (AUC = 0.95) in AD vs. NC classification.

These are some of the existing literature in AD diagnosis using MRI, all of which disclosed significant findings. Nevertheless, we are addressing some of the points which can be explored differently. Most importantly, some of the studies were heavily dependent on the domain knowledge of brain anatomy for feature selection. For example, [26] considered features extracted only from the hippocampus on the assumption of a higher concentration of correlated features. [31] considered 28 to 48 pre-defined RoI based on specialized neuroimaging knowledge. In the same way, [34] identified 1741 landmarks by the statistical measurement of brain anatomy and utilized 50 of them as information regions. Besides the additional specialized knowledge requirement, all but one of the mentioned studies used multi-staged architecture, which can be considered complex from a design perspective. For example, [26], [31] and [34] used two-staged networks, and their primary stages have consisted of a minimum of 2 to a maximum of 50 sub-networks. Moreover, [32] designed 3 staged hierarchical network which was comparatively more complex. Last but not least [25] did use a single-stage architecture, and that was also not dependent on specialized knowledge; however, its sub-optimal performance comparing to other mentioned works creates room for improvement.

This leads to the conclusion that a simple but robust Architecture needs to be designed from the domain of Deep learning exclusively. Addressing this need, we designed a 14 layer Convolutional Neural Network that is able to assist us in the early detection of Alzheimer's Disease by employing various Deep Learning techniques to distinguish between different stages of dementia using structural MRI scans. The rest of the paper is organized as follows. In section II, we have described our proposed methods for work for the whole study. In section III, we have discussed about our experimental results and its implications. Finally, In section IV, we concluded our study with our plans for the future of this study.

Chapter 3

Dataset Preparation and Methodology

This chapter provides a brief idea about our chosen dataset and the proposed ensemble learning based model. All of this information relayed in parts by separating them via section of this chapter. At first, we delve into the various techniques used in data collection and processing in section 3.1. Then, in section 3.2 we describe the architecture of our proposed model.After that, in section 3.3 we describe all our methods for enhancing the performance of our model during training. And at last, at section 3.4 we describe our ensemble learning based solution for early Alzheimer's Detection. The figure 3.1 describes our workflow and methodology for developing our solution from beginning to end. The performance of the entire solution is evaluated based on a multitude of factors such as accuracy, precision, f1 score etc.

3.1 Data Collection and Preprocessing

3.1.1 Data collection

During dataset selection, we decided to go with the collection of MRI images provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI). The ADNI dataset is comprised of several collections of MRI images which are all taken in different locations and different times. For this experiment, we decided to use the 3T baseline image collection.This particular image collection is comprised of the MRI images of 133 subjects. Among these 133 subjects, 45 subjects are tagged as Alzheimer's Disease (AD), 45 with Mild Cognitive Impairment (MCI) and the rest of the 43 subjects are labeled as Control Normal (CN).

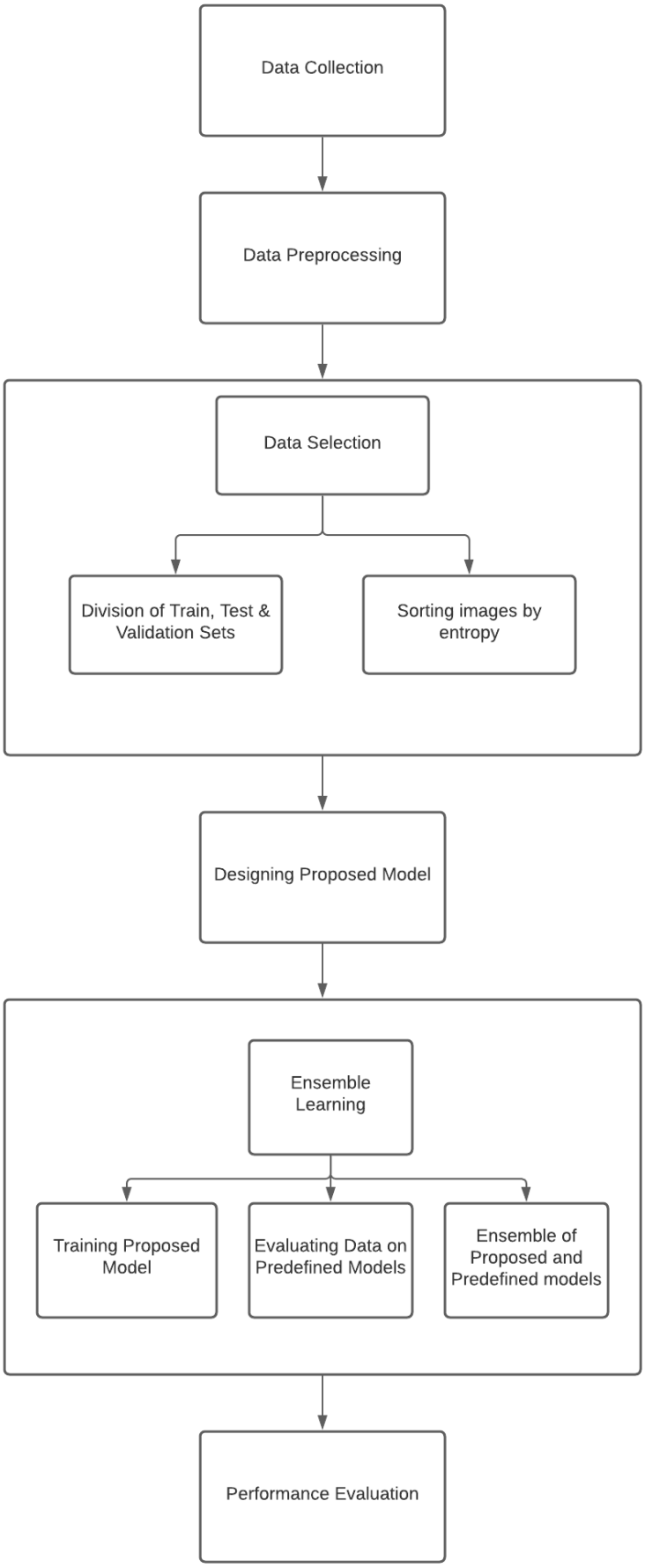


Figure 3.1: Proposed Method for Alzheimer's Detection

3.1.2 Data Prepossessing

In order to take out unwanted details of MRI images that might cause poor training of our classification task, we perform various prepossessing tasks on our data. These processes are:

Motion correction and conform:

This process is responsible for correcting minor motions between multiple sources of a volume by averaging them together

Non Uniform intensity normalization

Also known as N3, this process amends MRI data by mitigating non-uniform intensity. This process is performed with the help of the following equation:

$$I(x) = U(x)f(x) + n(x) \tag{3.1}$$

Where, I represents the given image, U denotes uncorrupted image, f describes the bias field and n is the noise.

Talairach transform computation:

This process converts all the pixel co-ordinates of the image into talairach co-ordinates and applies an affine transformation to the newly obtained co-ordinates.

Intensity normalization:

This step helps to correct for fluctuations on intensity. It does this by scaling intensities of all voxels by taking the mean intensity of white matter as 110.

Skull Stripping:

In this process we remove the skull and any other visible organs other than the brain in the MRI so that the final image only contains the necessary features for classification.

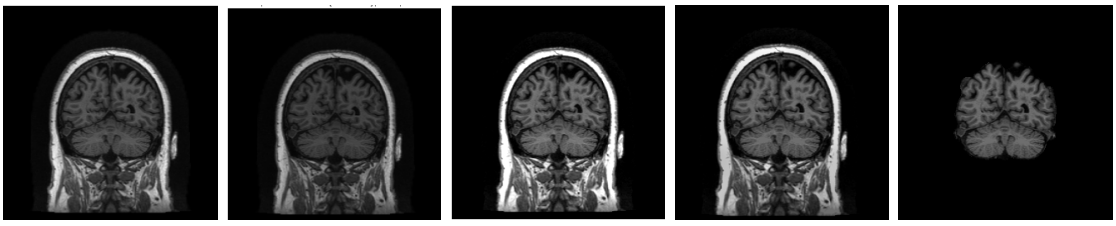


Figure 3.2: Transformation of a slice after going through all pre processing steps.

3.1.3 Data selection:

After prepossessing, we have images of 133 patients, with each subject having 256 slices of images. Among these 256 images, only a few are needed for classification as a limited sample size would raise the chances of success of the model. So, in order to choose the best possible slices, we only choose to only include the first 32 slices with the highest entropy values for each subject. After the necessary slices were chosen we decided to segregate the processed dataset into separate training, test and validation sets.

3.2 Model Architecture

In this section, we discuss our proposed CNN model that we use to detect Alzheimer's Disease using MRI images. We go through all the steps of our model construction below.

1. **Convolutional Layer:** The Convolutional is used to perform convolution operation on images using various filters.It contains various parameters and hyper-parameters such as filters, kernels, K etc. Convolutional layers extract features using the aforementioned filters, which are then used to compare images part by part to differentiate the similarities and differences among them. The main goal of a convolutional layer is to extract and identify high level features. We perform this operation in order to preserve the relationship between each pixel by learning image aspect using small samples of input data.

2. **Pooling Layer:** The pooling layer is used to reduce the dimensions of our feature maps. The Pooling layer is used in conjunction with the convolutional layer to reduce the size of the volume of the image when it is too large. This layer is responsible for making computation relatively fast, prevents overfitting and reduces strain on memory. For our model, we decided to go with a Max Pooling layer as we are mostly interested in the lighter portions of the MRI images.

3. **Flatten Layer:** This layer is used to convert the pooled feature map into a single column. We must flatten the pooled outputs in order to pass it to the fully connected layers

4. **Fully Connected Layer:** After flattening the inputs we must pass them onto a fully connected layer. These layers act as a kind of feed forward neural network that passes on the data to other layers, assigning weights to each part of the data on the way.

5. **Activation Functions:** In our CNN we have mainly used two activation functions. These are:

(a) Softmax: The softmax function is used in the neural network to normalize the output of a network to a probability distribution over the specified output classes. The equation 3.2 surmises the softmax activation function.

$$f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{k} e^{x_j}} \tag{3.2}$$

(b) ReLU: The Rectifier Linear Unit (ReLU) function works by taking only the positive parts of its arguments. Unlike the sigmoid function, ReLU does not have any problems with vanishing gradient. The equation 5.2 summarises the ReLu function

$$f(x) = max(0, x) \tag{3.3}$$

3.3 Learning Enhancements

This section describes the learning enhancement techniques used during training of our model. Each of these processes are described below-

1. **Image Augmentation** Data Augmentation is the process by which we can insert more variation into our already existing dataset by the help of various transformations. Data augmentation helps us to increase the amount of relevant data in our dataset by applying transformations to already existing data. Image augmentation is one of the most popular forms of data augmentations as most image datasets do not contain enough images to sufficiently train a neural network. As we only have the MRI images of 133 patients, we must properly augment our data in order to get reliable results. To this end, we have opted for some popular image augmentation techniques such as: randomly zooming on parts of the subject within a given range, increasing the brightness on an image by a random amount, suddenly flipping an image horizontally and vertically etc. We hope that by using such techniques of Image Augmentation we will be able to introduce much needed variation in our dataset, thereby making it more robust.

2. **Regularization** In certain cases a model might try to take extreme values for its coefficients during training in order to get predictions, which might result in overfitting. To avoid this scenario we introduce additional penalty terms in the training function so that the model does not overfit on the training set. Such techniques are known as regularization, which are able to reduce the error coefficients in a network to ensure that the model does not stray too far from the training set. And although there are multiple regularization methods to choose from, the l1 and l2 regularizers remain the most popular. For our particular model, we shall use l1 regularization as it can eliminate features from being detected while l2 just shrinks their impact on the dataset. The equation for L1 regularization can be written as-

$$Loss = Error(y, \hat{y}) + \lambda \sum_{i=1}^{n} |w_i| \tag{3.4}$$

Here, the lambda is called the regularization parameter which is manually tuned, which prevents the error function from blowing up by either deducting from or summing lambda to w. This prevents our model from suddenly overfitting on the training set due to rapid changes in weights of our network.

3. **Optimization** In Deep learning algorithms, optimization refers to the act of tuning hyper parameters of a network by adjusting by minimizing the cost function of the network. For our particular case we will be using the Adam optimizer which combines various approaches to optimization such as AdaGrad and RMSprop and combining them together to get both their benefits. It works by calculating an individual adaptive learning rate for each parameter from its estimates of first and second moments of the gradients. This makes Adam work much like a heavy ball with friction, preferring flat parts of error surface.

4. **Transfer Learning** Transfer learning is a technique used in deep learning problems where a previously developed model used for one particular is reused for a different one. Transfer learning is used in cases where instead of starting from scratch, we leverage the already defined weights of a network to make training faster and more accurate. For our study, we have used transfer learning to take advantage of various pre trained models such as Xception [24], InceptionV3 [22], ResNet50 [20], Vgg16 [17] etc. making them fit into our dataset.

3.4 Ensemble Learning

Convolutional Neural Networks and other neural network models all have the same drawback over fitting, also known as variance. This is because neural networks learn via a stochastic approach and can be sensitive to the training set. As a result, after running a substantial number of epochs, the final model may still not be generalized. Consequently, the model will perform poorly in classifying inputs that are new to it.

Intriguingly, running fewer epochs barely can help because it will increase the bias, and the model will perform poorly even in the train set. This trade-off between variance and bias can be reduced by running multiple models and combining their outputs for a stable and accurate prediction. This approach is also known as ensemble learning

In ensemble learning multiple neural network models with different hyper parameters are run in parallel. Conventionally, three, five or seven models are used in an ensemble. Most noticeably, these models are chosen in such way so that they are

highly accurate but diverse in terms of their predictions. Various methods such as Averaging, Max Voting etc are used for determining the final output from these multiple models output. This final output is guaranteed to be better than any of the individual model.
For our proposed solution we decided to use our proposed model along with other pre trained models for the highest accuracy. Specifically, the pre trained models VGG16, ResNet50, InceptionV3 and Xception are being run in parallel with our own proposed model in order to get diverse yet accurate results. In order to get the final result from such an arrangement, we opted for a Max voting strategy which is very popular for classifications problems such as ours. This method takes the predictions from each model and then bases the final result on the label predicted my a majority of all the models predictions. The figure 3.3 shows our ensemble learning strategy in detail.

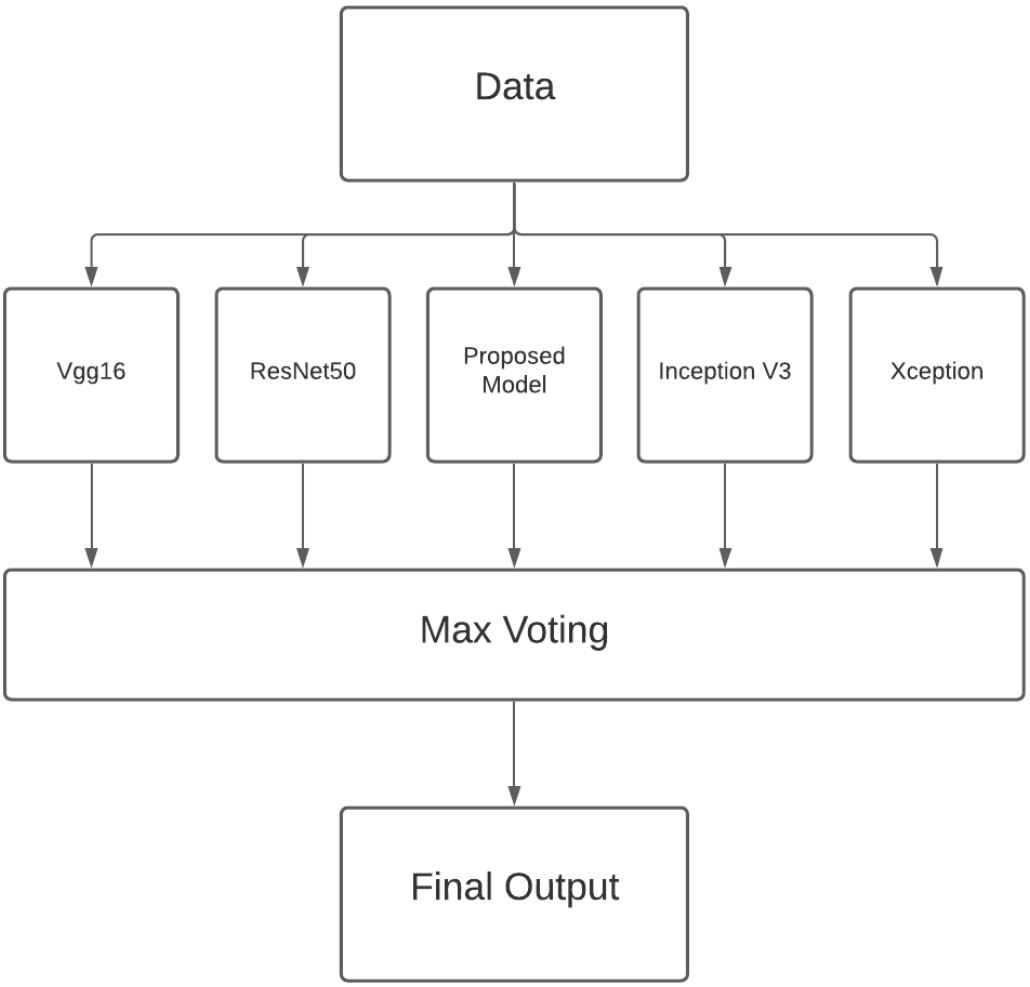


Figure 3.3: Proposed Ensemble Learning Solution

Chapter 4

Implementation and Result Analysis

Previously, we discussed the process of data collection and data preprocessing along with a brief overview of our proposed model and our ensemble learning based solution .In this section, we talk about actual construction of our network as well as our own implementation of Ensemble Learning with other pre trained networks. Figure 4.1 contains a brief overview of our fully constructed model

4.1 Implementation of baseline Model

This section describes our implementation of our proposed model in python using libraries such as keras and tensorflow. This model is then saved after training for use in our ensemble learning solution. Details of this implementation is described below.

1. **Convolutional Layer Selection:** In our own CNN model we used the Conv2D layer in keras. Specifically we use four Conv2D layers in our model.

2. **Pooling Layer Selection:** For pooling, we decided to go with the MaxPooling 2d in keras. Since, every Conv2D layer requires pooling, we decided to use four Maxpooling layers.

3. **Flatten Layer:** Alter the final pooling layer, we use the Flatten layer in keras to flatten our outputs from the pooling layers.

4. **Dense Layer:** After flattening out inputs, we use six Dense layers from keras as the hidden layers of our CNN. These Dense layers in keras act as fully connected networks in the CNN.

5. **Dropout Layer:** The Dropout layer randomly sets a portion of its inputs to 0 with a defined frequency during training, which helps to prevent overfitting.

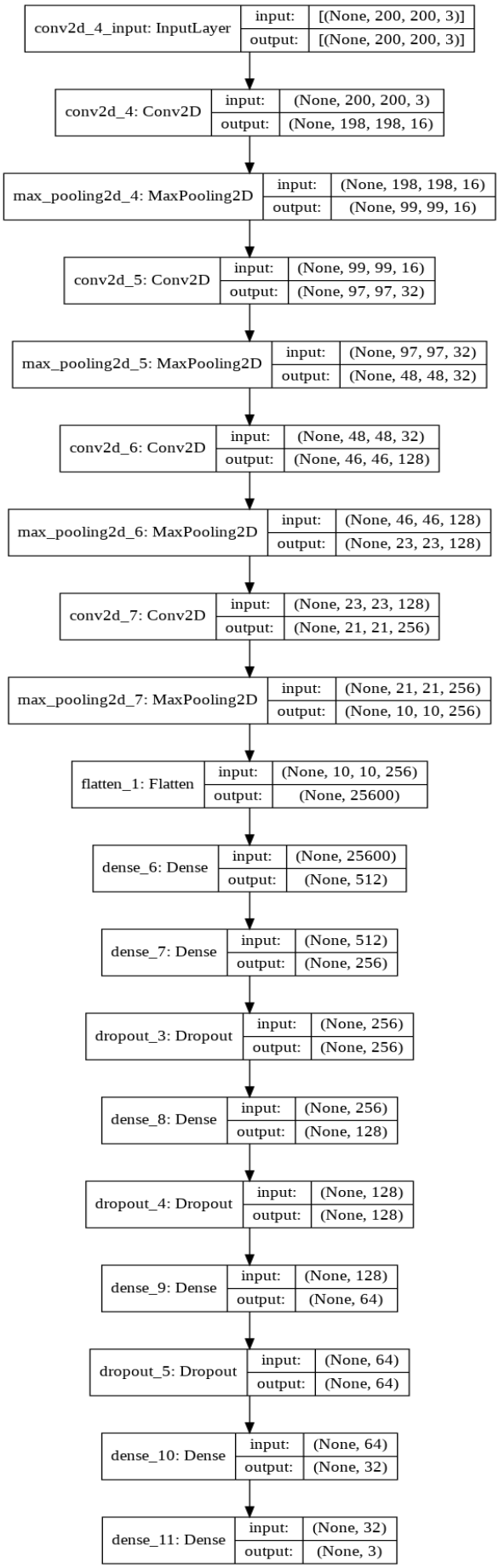


Figure 4.1: Architecture for the proposed model.

We used three dropout layers in prevent overfitting in our network.

4.2 Implementation of Ensemble Learning

In this section we describe how we implemented our ensemble learning based solution by importing various pre trained models from keras and then running them in tandem with our own model.The final result is then acquired by the use of a max voting strategy. A detailed overview of the entire process is described below.

1. **Loading Proposed model:** At first we load our own proposed after it has been trained by using keras's load_model function.

2. **Importing Pre Trained models:** From the Application Module of keras we import the following models: VGG16, Xception, ResNet50 and InceptionV3 by taking their weights from the imagenet dataset. After excluding the fully connected networks at the top of the model, we adapt these models to fit our data.

3. **Max Voting :** After successfully importing all pre trained models, we then run all of them alongside our own proposed model on a validation set of the entire dataset. The final output of is then taken by a majority vote of all these models.

4.2.1 Individual Model performance

This section includes the experimental performance and performance of all of our models. In evaluating the model we have used accuracy as the primary evaluation metric in the compilation. After both the proposed model and the pretrained models have been fitted to our dataset, we evaluate each of them based on their validation accuracy. We have three separate evaluations consisting of AD vs CN vs MCI, AD vs CN and MCI vs CN.

After running all models on our dataset for 20 epochs, we managed to obtain satisfactory accuracy on all our models. For our own proposed model, we managed to obtain an accuracy of 92.41% for 3 way classification between AD vs CN vs MCI and 90.48% and 88.28% accuracy for 2 classification between AD vs CN and MCI vs CN respectively.

As for the other pretrained models, For the 3 way classification the Xception model managed to score 87.02%, with the InceptionV3 model being the closest to it with its 83.07% accuracy followed closely by both ResNet50 and VGG16 models with an accuracy of 83.52% and 83.15% respectively.

For the 2 way classification tasks, all models managed to score pretty well on the AD vs CN tak with Xception model in the lead with its 86.22% accuracy, followed

Model Name	Accuracy AD vs CN vs MCI	Accuracy AD vs CN	Accuracy MCI vs CN
VGG16	83.15%	81.03%	82.73%
ResNet50	83.52%	83.36%	83.23%
Inceptionv3	84.07%	83.95%	83.76%
Xception	87.02%	86.22%	88.28%
Proposed Model	92.41%	90.48%	87.28%
Ensembled Solution	92.44%	91.27%	89.36%

Table 4.1: Comparison of Accuracy of all models

closely by the InceptionV3 model with its own score of 83.95% accuracy. These reults are followed up by both the ResNet50 and VGG16 models with their scores of 83.36% and 81.03% respectively.

As for the other 2 way classification task on MCI vs CN, the Xception model scored a respectable 87.28% with the InceptionV3 model being right behind it with its own score of 83.76%. The other two models, namely ResNet50 and VGG16 also managed to do well with their own scores of 83.28% and 82.73% on the dataset.

A detailed comparison of all models is given at the Table 4.1

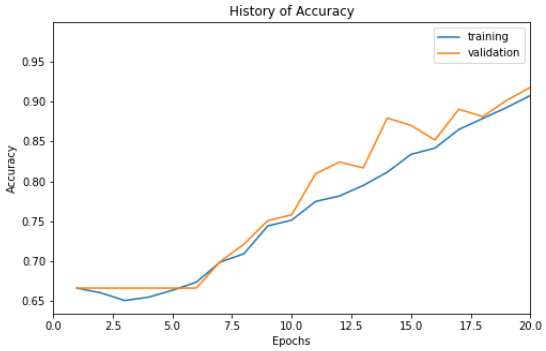
4.2.2 Ensembling Results

This section includes all the results of our ensemble learning solution. After the proposed model and the pre trained models have been fitted on our existing dataset, we run all of them in parallel on a separate validation dataset. After that, we use a max voting strategy on all of these models in order to get the final result. We evaluate the final results on a multitude of factors such as Accuracy, AUC, F1 score, Cross-entropy Loss etc.

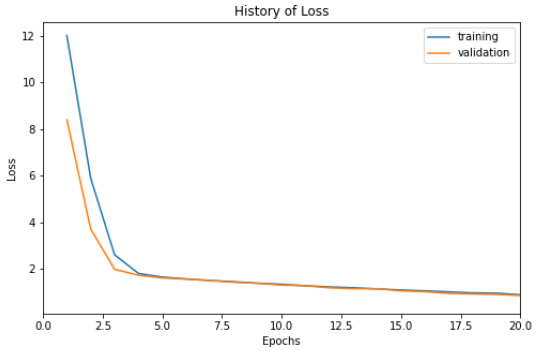
After ensembling, we get an overall accuracy of 92.48% for 3 way classification of AD vs CN vs MCI, which is an improvement on the results of most of our models. Our ensemble learning based solution also managed to get a relatively high precision and recall scores of 87.23% and 91.11% respectively. From this, we get an overall F1 score of 0.8913 from both our precision and recall scores.

For 2 way classification between AD and CN, our solution achieves an overall accuracy of 90.23%. It also manages to have a high precision and recall score of 88.99% and 83.33% respectively. With such results, the model is able to achieve a respectable F1 score of 0.8602.

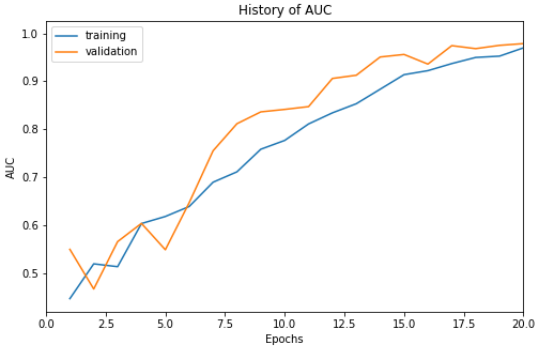
Finally, the ensemble learning solution is attain a respectable accuracy of 89.32% for the 2 way classification between MCI and CN. Its is able to gain precision and recall scores of 81.25% and 88.64% respectively, which it uses to gain a decent F1 score of 0.8478.



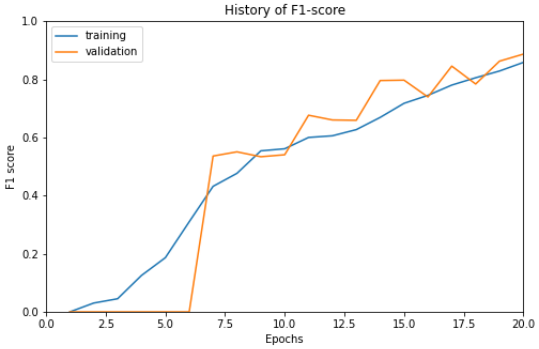
(a) Accuracy



(b) Loss

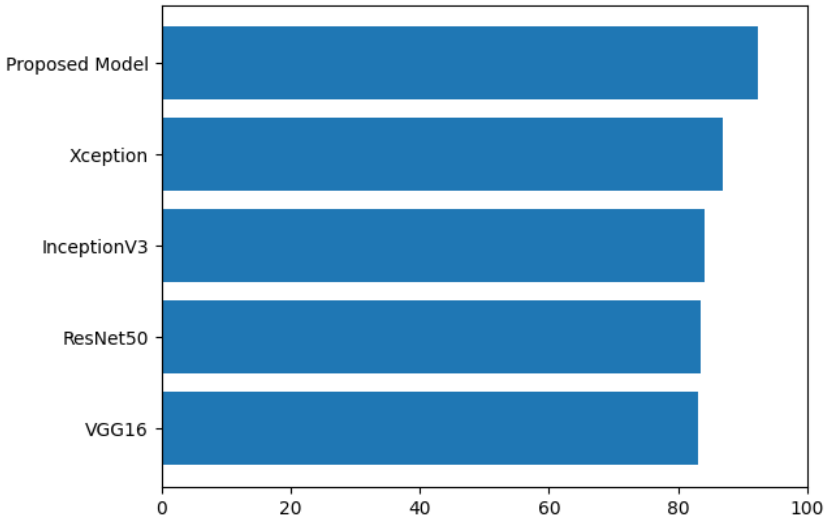


(c) AUC curve

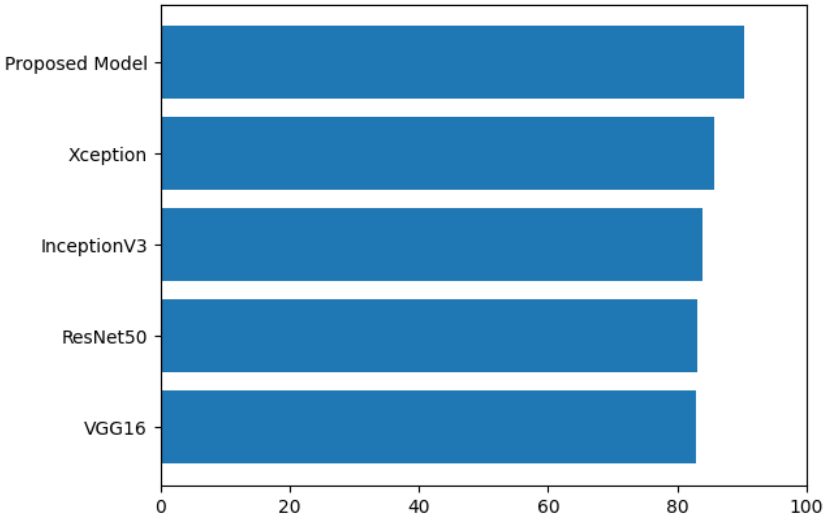


(d) F1 score

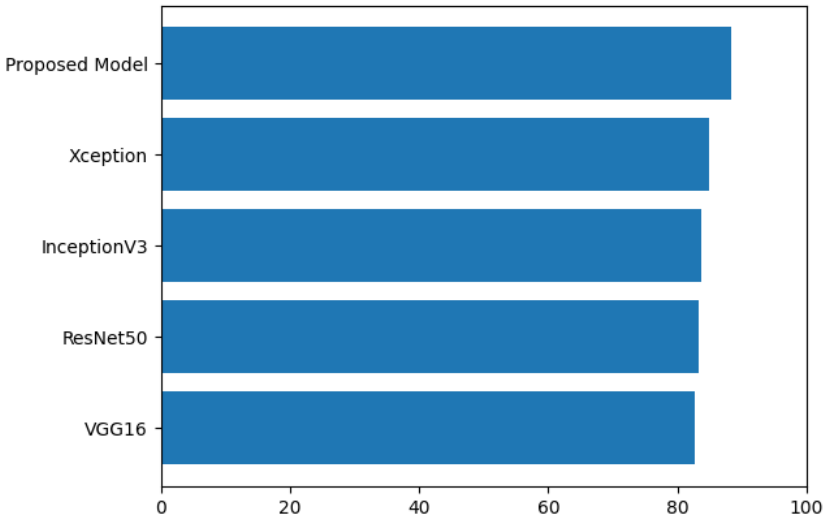
Figure 4.2: Graphs describing (a) Accuracy, (b) Loss, (c) AUC curve and (d) F1 score of the 3 way classification of proposed model



AD vs CN vs MCI



AD vs CN



MCI vs CN

Figure 4.3: Confusion matrix of Ensemble Learning Solution

Study	Sample Size	Feature	Accuracy
Korolev et al.	231	Whole Brain Image	80%
Khovostikol et al.	214	48 predefined ROI from hippocampus	97%
Liu et al	1457	50 pre calculated landmarks	91%
Lian et al	1457	120 anatomically proposed locations	90%
Lin et al	818	35 calculated features	87.06%
Proposed Solution	133	Whole Brain Image	92.44%

Table 4.2: Comparison of Different methods of Alzheimer's Disease detection

4.2.3 Results comparison

This section show a detailed comparison between our solution and other solutions described elsewhere. Here, we see that our proposed solution has a relatively high accuracy, beating other complex solutions while using relatively less resources.A detailed comparison of the various solutions are described in the Table 4.2

4.2.4 Discussion

From our overall results we can see that our proposed solution was able to achieve a higher accuracy than other models with more complex feature extraction and training methods. This means that our proposed model is able to train faster with less data while still being able to achieve surprisingly accurate results in a short amount of time. Our proposed ensemble learning model is able to achieve a total accuracy of 92.44% which is higher than that of most other models that segment the brain into multiple parts using similar unsupervised learning models for a more complete feature extraction. And while models with Pre defined Regions of Interest are able to beat out our solution in terms of accuracy, It is still harder and more time consuming to manually set up Regions of Interest in an MRI image than just using the whole brain MRI image seen in our solution. Overall, our solution provides a faster and more accurate alternative method to existing Alzheimer's Detection systems.

Chapter 5

Conclusion

Neurodegenerative diseases like Alzheimer's Disease (AD) are a common sight on elderly and senior populations. Not only do they take a toll on the patients physical and mental health, but they may also cause financial insolvency to treat its symptoms. So, correct and accurate diagnosis of AD critical to ensure proper medication for its patients. The advent of modern medical imaging technologies have taken great strides to make accurate diagnosis easier to achieve than ever before. And although various state-of-the-art automated solutions have achieved human-level performance in classifying AD images from Healthy patients, they often require predefined Regions of interest as a basis for feature extraction. This not only makes specialized domain knowledge of the human brain a requirement but is also responsible for making the overall design complicated than it needs to be. This research represents our efforts to automate the detection of AD in susceptible patients and provide them with accurate results by developing our own 15 layer CNN which runs in parallel with other pre defined CNN networks to provide us an accurate and fast solution to early detection of Alzheimer's Disease that does not require any domain specific knowledge to operate

5.1 Future Work

In the future, we hope to explore this idea further with more data. As we have only used a certain small subset of the entire ADNI dataset for this project, we hope to expand our model to be trained in the entire ADNI dataset. We also hope to integrate other data sources such as OASIS, MIRIAD etc to get a through understanding Alzheimer's Disease. And although we already have used many state-of-the-art pre trained models in parallel with our own, we wish to one day leverage other sophisticated models such as InceptionV4, MobileNet to be used in our solution. We are also looking forward to exploring other new methods of ensemble learning such as Bagging, Boosting, Stacking etc. in order to get more performance out of combing multiple models together.

Bibliography

[1] B. Reisberg, J. Borenstein, S. P. Salob, S. H. Ferris, *et al.*, "Behavioral symptoms in alzheimer's disease: Phenomenology and treatment.," *The Journal of clinical psychiatry*, 1987.

[2] M. Shoji, T. E. Golde, J. Ghiso, T. T. Cheung, S. Estus, L. M. Shaffer, X.-D. Cai, D. M. McKay, R. Tintner, B. Frangione, *et al.*, "Production of the alzheimer amyloid beta protein by normal proteolytic processing," *Science*, vol. 258, no. 5079, pp. 126–129, 1992.

[3] J. Nolte, *The human brain*. Mosby/Elsevier, 1993.

[4] P. D. Meek, E. K. McKeithan, and G. T. Schumock, "Economic considerations in alzheimer's disease," *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, vol. 18, no. 2, pp. 68–73, 1998.

[5] M. G. Spillantini and M. Goedert, "Tau protein pathology in neurodegenerative diseases," *Trends in neurosciences*, vol. 21, no. 10, pp. 428–433, 1998.

[6] L. C. Walker and H. LeVine, "The cerebral proteopathies," *Molecular neurobiology*, vol. 21, no. 1, pp. 83–95, 2000.

[7] R. C. Petersen, R. Doody, A. Kurz, R. C. Mohs, J. C. Morris, P. V. Rabins, K. Ritchie, M. Rossor, L. Thal, and B. Winblad, "Current concepts in mild cognitive impairment," *Archives of neurology*, vol. 58, no. 12, pp. 1985–1992, 2001.

[8] A. Wimo, B. Winblad, H. Aguero-Torres, and E. von Strauss, "The magnitude of dementia occurrence in the world," *Alzheimer Disease & Associated Disorders*, vol. 17, no. 2, pp. 63–67, 2003.

[9] D. C. Rubinsztein, "The roles of intracellular protein-degradation pathways in neurodegeneration," *Nature*, vol. 443, no. 7113, pp. 780–786, 2006.

[10] J. Sepulcre, J. C. Masdeu, J. Sastre-Garriga, J. Goñi, N. Vélez-de-Mendizábal, B. Duque, M. A. Pastor, B. Bejarano, and P. Villoslada, "Mapping the brain pathways of declarative verbal memory: Evidence from white matter lesions in the living human brain," *Neuroimage*, vol. 42, no. 3, pp. 1237–1243, 2008.

[11] J. Brown, G. Pengas, K. Dawson, L. A. Brown, and P. Clatworthy, "Self administered cognitive screening test (tym) for detection of alzheimer's disease:cross sectional study," *Bmj*, vol. 338, 2009.

[12] P. Vemuri and C. R. Jack, "Role of structural mri in alzheimer's disease," *Alzheimer's research & therapy*, vol. 2, p. 4, 2010.

[13] L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, P. A. Kanda, and R. Anghinah, "Andr," *Nitrini, "Improving alzheimer's disease diagnosis with machine learning techniques," Clinical EEG and neuroscience*, vol. 42, no. 3, pp. 160–165, 2011.

[14] I. B. Malone, D. Cash, G. R. Ridgway, D. G. MacManus, S. Ourselin, N. C. Fox, and J. M. Schott, "Miriad—public release of a multiple time point alzheimer's mr imaging dataset," *NeuroImage*, vol. 70, pp. 33–36, 2013.

[15] D. A. Drachman, "The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of alzheimer's disease," *Alzheimer's & Dementia*, vol. 10, no. 3, pp. 372–380, 2014.

[16] R. Li, W. Zhang, H. I. Suk, L. Wang, J. Li, D. Shen, and S. Ji, "Deep learning based imaging data completion for improved brain disease diagnosis," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2014, pp. 305–312.

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

[18] R. Guerreiro and J. Bras, "The age factor in alzheimer's disease," *Genomemedicine*, vol. 7, no. 1, pp. 1–3, 2015.

[19] L. Witter and C. I. De Zeeuw, "Regional functionality of the cerebellum," *Current opinion in neurobiology*, vol. 33, pp. 150–155, 2015.

[20] 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*, 2016, pp. 770–778.

[21] E. Hosseini-Asl, R. Keynton, and A. El-Baz, "Alzheimer's disease diagnostics by adaptation of 3d convolutional network," in *Proceedings - International Conference on Image Processing, ICIP*, Cited By :97, vol. 2016-August, 2016, pp. 126–130. [Online]. Available: www.scopus.com.

[22] 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.

[23] W. H. O. al., *Dementia fact sheet*. World Health Organization: Geneva, Switzerland, 2017.

[24] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[25] S. Korolev, A. Safiullin, M. Belyaev, and Y. Dodonova, "Residual and plain convolutional neural networks for 3d brain mri classification," in *Proceedings - International Symposium on Biomedical Imaging*, Cited By :73, 2017, pp. 835–838. [Online]. Available: www.scopus.com.

[26] H. Li, M. Habes, and Y. Fan, "Deep ordinal ranking for multicategory diagnosis of alzheimer's disease using hippocampal mri data," arXiv preprint, 2017.

[27] S. Sarraf and G. Tofighi, "Deep learning-based pipeline to recognize alzheimer's disease using fmri data," in *FTC 2016 - Proceedings of Future Technologies Conference*, Cited By :47, 2017, pp. 816–820. [Online]. Available: www.scopus.com.

[28] S. Esmaeilzadeh, D. I. Belivanis, K. M. Pohl, and E. Adeli, *End-to-end alzheimer's disease diagnosis and biomarker identification*, ser. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2018, vol. 11046 LNCS, pp. 337–345, Cited By :17. [Online]. Available: www.scopus.com.

[29] D. Gallagher, A. Kiss, K. Lanctot, and N. Herrmann, "Depression and risk of alzheimer dementia: A longitudinal analysis to determine predictors of increased risk among older adults with depression," *The American Journal of Geriatric Psychiatry*, vol. 26, no. 8, pp. 819–827, 2018.

[30] M. Khosla, K. Jamison, A. Kuceyeski, and M. R. Sabuncu, *3d convolutional neural networks for classification of functional connectomes*, ser. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2018, vol. 11045 LNCS, pp. 137–145, Cited By :12. [Online]. Available: www.scopus.com.

[31] A. Khvostikov, K. Aderghal, J. BenoisPineau, A. Krylov, and ". G. Catheline, "Cnn-based classification using smri and md-dti images for alzheimer disease studies," arxiv," preprint, 2018.

[32] C. Lian, M. Liu, J. Zhang, and D. Shen, "Hierarchical fully convolutional-network for joint atrophy localization and alzheimer's disease diagnosis using structural mri," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 4, pp. 880–893, 2018.

[33] W. Lin, T. Tong, Q. Gao, D. Guo, X. Du, Y. Yang, G. Guo, M. Xiao, M. Du, and X. Qu, "Convolutional neural networks-based mri image analysis for the alzheimer's disease prediction from mild cognitive impairment," *Frontiers in Neuroscience*, vol. 12, no. NOV, 2018, Cited By :57. [Online]. Available: www.scopus.com.

[34] M. Liu, J. Zhang, E. Adeli, and D. Shen, "Landmark-based deep multi-instance learning for brain disease diagnosis," *Medical image analysis*, vol. 43, pp. 157–168, 2018, Cited By :91. [Online]. Available: www.scopus.com.

[35] C. Feng, A. Elazab, P. Yang, T. Wang, F. Zhou, H. Hu, X. Xiao, and B. Lei, "Deep learning framework for alzheimer's disease diagnosis via 3d-cnn and fsbi-lstm," *IEEE Access*, vol. 7, pp. 63 605–63 618, 2019, Cited By :16. [Online]. Available: www.scopus.com.

[36] N. M. Khan, N. Abraham, and M. Hon, "Transfer learning with intelligent training data selection for prediction of alzheimer's disease," *IEEE Access*, vol. 7, pp. 72 726–72 735, 2019, Cited By :15. [Online]. Available: www.scopus.com.

[37] K. R. Kruthika, Rajeswari, and H. D. Maheshappa, "Cbir system using capsule networks and 3d cnn for alzheimer's disease diagnosis," *Informatics in Medicine Unlocked*, vol. 14, pp. 59–68, 2019, Cited By :18. [Online]. Available: www.scopus.com.

[38] K. Oh, Y. .-. Chung, K. W. Kim, W. .-. Kim, and I. .-. Oh, "Classification and visualization of alzheimer's disease using volumetric convolutional neural network and transfer learning," *Scientific Reports*, vol. 9, no. 1, 2019, Cited By :14. [Online]. Available: www.scopus.com.

[39] E. Hussain, M. Hasan, S. Z. Hassan, T. H. Azmi, M. A. Rahman, and M. Z. Parvez, "Deep learning based binary classification for alzheimer's disease detection using brain mri images," in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, IEEE, 2020, pp. 1115–1120.

[40] L. Nanni, M. Interlenghi, S. Brahnam, C. Salvatore, S. Papa, R. Nemni, and I. Castiglioni, "Comparison of transfer learning and conventional machine learning applied to structural brain mri for the early diagnosis and prognosis of alzheimer's disease," *Frontiers in Neurology*, vol. 11, 2020. [Online]. Available: www.scopus.com.

[41] O. Pelka, C. M. Friedrich, F. Nensa, C. Mönninghoff, L. Bloch, K. .-. Jöckel, S. Schramm, S. S. Hoffmann, A. Winkler, C. Weimar, and M. Jokisch, "Sociodemographic data and apoe- 4 augmentation for mri-based detection of amnestic mild cognitive impairment using deep learning systems," *PLoS ONE*, vol. 15, no. 9 September, 2020. [Online]. Available: www.scopus.com.

[42] J. Wen, E. Thibeau-Sutre, M. Diaz-Melo, J. Samper-González, A. Routier, S. Bottani, D. Dormont, S. Durrleman, N. Burgos, and O. Colliot, "Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation," *Medical image analysis*, vol. 63, 2020, Cited By :4. [Online]. Available: www.scopus.com.

[43] E. Yagis, L. Citi, S. Diciotti, C. Marzi, S. Workalemahu Atnafu, and A. G. S. De Herrera, "3d convolutional neural networks for diagnosis of alzheimer's disease via structural mri," in *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, vol. 2020-July, 2020, pp. 65–70. [Online]. Available: www.scopus.com.

[44] C. R. J. Jr, M. A. Bernstein, N. C. Fox, P. Thompson, D. H. G. Alexander, B. Borowski, P. J. Britson, J. L. Whitwell, and C. W. al., "The alzheimer's disease neuroimaging initiative (adni): Mri methods," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 27, no. 4, p. 2008,

[45] J. Wen, E. ThibeauSutre, M. DiazMelo, J. S.´alez, S. B. A. Routier, D. Dormont, S. Durrleman, N. Burgos, and O. C. al., "Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation," *Medical image analysis*, vol. 63,