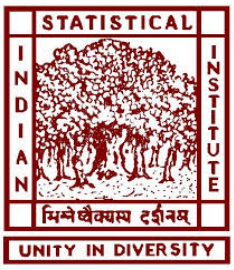


Some Insight Into Alzheimer's Disease Progression

PREPARED BY

Soumyadip Sikdar

Roll: CS2231

SUPERVISED BY

Dr. Kuntal Ghosh

Associate Professor

Machine Intelligence Unit

A thesis submitted in partial fulfillment of the requirements for the degree of MASTER OF TECHNOLOGY IN COMPUTER SCIENCE

Contents

Contents **1**

List of Figures **3**

Acknowledgments **5**

1 Introduction **1**

1.1 DEMENTIA AND ALZHEIMER'S DISEASE 1

1.1.1 DEMENTIA 1

1.1.2 ALZHEIMER DISEASE 2

1.1.3 CHALLENGES WITH EARLY DETECTION 2

1.1.4 DISEASE STAGES 3

1.1.5 BIOMARKERS 4

1.1.6 MOTIVATION 4

1.1.7 PROBLEM STATEMENT 4

1.1.8 THESIS OVERVIEW 5

2 Related Work **6**

3 Data Description **8**

3.1 ADNI DATASET 8

3.2 MRI 8

3.3 COGNITIVE SCORES 9

3.4 DATA DISTRIBUTION 11

3.4.1 SUBJECTS WITH COGNITIVE SCORES 11

3.4.2 SUBJECTS WITH sMRI 12

3.4.3 SUBJECTS WITH sMRI AND COGNITIVE SCORES 13

4 DATA PREPROCESSING **15**

4.1 MRI PREPROCESSING 15

4.2 MRI PADDING 17

5 MODEL **19**

5.0.1 LIGHTWEIGHT 2D-CNN MODEL 19

5.0.2 2D-CNN USED IN STUDY 19

6 Experiment And Results Analysis **22**

6.1 CLASSIFICATION USING COGNITIVE SCORES 22

6.2 CLASSIFICATION USING MRI SCANS 27

6.3 CLASSIFICATION USING MRI AND COGNITIVE SCORES 32

7 Discussion **36**

8 Conclusion and Future Work **37**

Bibliography **38**

List of Figures

1.1 Dementia Classification 1
1.2 Alzheimer and healthy brain 2
1.3 AD Stages 3
1.4 AD Progression,Source: [10] 3

4.1 Rectangular Preprocessed MRI of height 189 units and width 156 units 16
4.2 Square Preprocessed MRI of height and width as 192 units 17
4.3 Preprocessed MRI image of CN subject 18
4.4 Preprocessed MRI image of MCI subject 18
4.5 Preprocessed MRI image of AD subject 18

6.1 Methodology For Disease Stage Classification With Cognitive Assessment Scores 23
6.2 SVM Confusion Matrix 24
6.3 KNN Confusion Matrix 25
6.4 RF Confusion Matrix 26
6.5 2D-CNN Confusion Matrix 27
6.6 Methodology For Disease Stage Classification With MRI 28
6.7 Train-Test Accuracy Curve fold-wise 28
6.8 Train-Loss in fold 1 29
6.9 Train-Loss in fold 2 29
6.10 Train-Loss in fold 3 30
6.11 Train-Loss in fold 4 30
6.12 Train-Loss in fold 5 31
6.13 Methodology For Disease Stage Classification With MRI and Cognitive Scores 32

6.14 SVM Confusion Matrix 33
6.15 KNN Confusion Matrix 34
6.16 RF Confusion Matrix 35

Acknowledgments

I would like to convey my deepest appreciation to my mentor, **Dr. Kuntal Ghosh** (Associate Professor, Machine Intelligence Unit, Indian Statistical Institute, Kolkata), **Anjan Chakraborty** (SRF, Machine Intelligence Unit, Indian Statistical Institute, Kolkata), **Dr Swarup Chattopadhyay**, former SRF, MIU, ISI and **Sandipa Ray**(WOS-B, Indian Statistical Institute, Kolkata), for granting me the opportunity, and support to pursue my thesis. Their support and guidance were indispensable, and without them, achieving this milestone would have been impossible. I extend my heartfelt gratitude to my parents and family for their unwavering support. Finally, I would like to express my thanks to all my friends and batch-mates for their assistance and encouragement throughout this journey.

Abstract

Alzheimer's disease is a neurodegenerative disease that affects a multitude of people globally. It usually affects people of 60 yrs and older causing changes in the anatomy of the brain.Subjects diagnosed with Alzheimer's disease often gets to live for 5-10 yrs.This makes early and accurate detection of the disease of paramount importance not only for the victim but to better understand disease progression in subjects. Over the years with the advancement of machine learning and deep learning a number of studies have come up to perform disease classification basis different biomarkers which acts as indicators for disease progression. This paper presents a multi-modal study that compares performances of machine learning and deep learning models on 2 sets of inputs namely MRI and cognitive scores. It considers dataset from ADNI (Alzheimer Disease Neuroimaging Initiative) which is a longitudinal multi centre study designed to develop clinical, imaging, genetic and biochemical biomarkers for the early detection and tracking of Alzheimer's disease (AD) and performs multi-class classification of subjects into 3 groups of CN(Normal Cognition), MCI(Mild Cognitive Impairment) and AD(Alzheimer Disease). For the deep learning model classification using MRI the study proposes use of a modified 2D-CNN that works with MRI scans. Contrary to Deep Convolutional Neural Networks that outputs better accuracy at the cost of higher execution times 2D-CNN performs faster at the cost of accuracy. In addition the study also considers a mix of both forms as input i.e.features extracted from 2D-CNN and cognitive scores to classify subjects basis machine learning models. This hybrid input captures not only brain anatomical changes but also symptoms that manifest.

Chapter 1

Introduction

1.1 DEMENTIA AND ALZHEIMER'S DISEASE

1.1.1 DEMENTIA

Dementia is an umbrella term that encapsulates symptoms that affects memory, thinking, social abilities and other activities of daily living (ADLs). People with dementia suffers from both cognitive and psychological changes such as trouble planning and organization, communication problems, depression, agitation etc. Dementia is usually classified into 2 types namely PROGRESSIVE/IRREVERSIBLE and REVERSIBLE.[10]

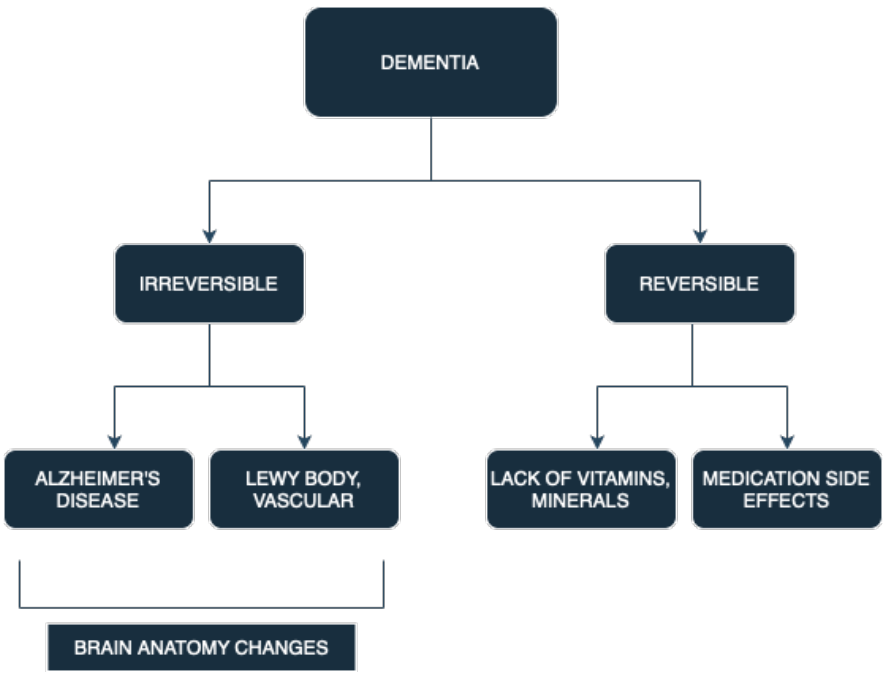


Figure 1.1: Dementia Classification

Figure 1.1 indicates dementia types and causes. Dementia caused by reasons other than changes that occur in the brain is treatable.

1.1.2 ALZHEIMER DISEASE

It is a progressive, neurodegenerative disease that occurs when nerve cells in the brain dies. This is the most common cause of dementia that results in decline of memory, thinking, learning and other activities of daily living. Figure 1.2 indicates how a healthy and Alzheimer's inflicted subject's brain differs.

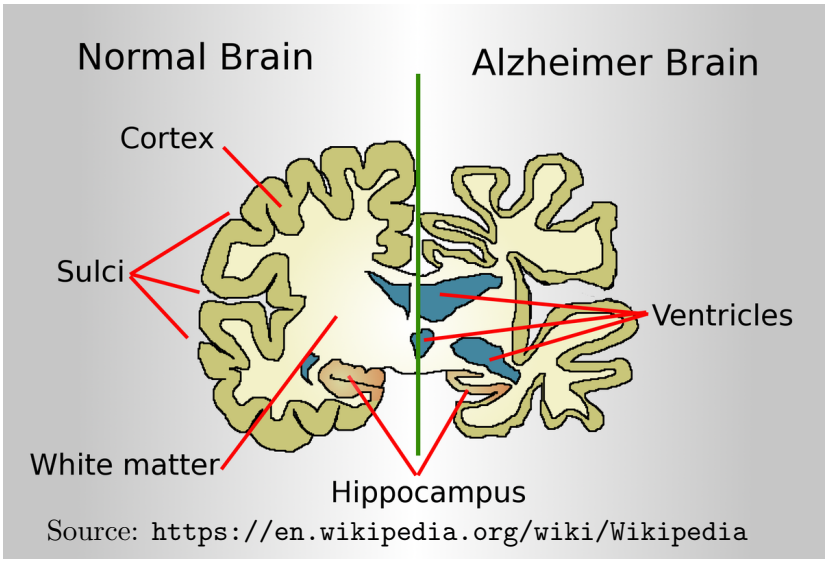


Figure 1.2: Alzheimer and healthy brain

1.1.3 CHALLENGES WITH EARLY DETECTION

With Alzheimer detecting disease progression becomes complex owing to symptoms that is usually attributed to old age and ignored.Such symptoms includes,

- Occasional lapses in short-term memory
- Difficulty completing routine tasks
- Misplacing things etc.

1.1.4 DISEASE STAGES

There exists 5 stages through which the disease progresses:

- Preclinical Stage: During this stage subject may have no noticeable symptoms with minor changes in brain anatomy
- Mild cognitive Impairment (MCI): Subject develops minor symptoms but not significant enough to affect daily activities with further brain anatomical changes.
- Mild Dementia: At this stage subjects may see symptoms manifest to affect activities of daily living such as trouble planning or organizing things, thinking, communication etc.
- Moderate Dementia: At this stage subjects may need assistance performing daily activities, have more severe memory losses,confusion,personality and behavioural changes.
- Severe Dementia: Subjects become entirely dependent with loss of control on bladder and bowel movements and others with eventual death.

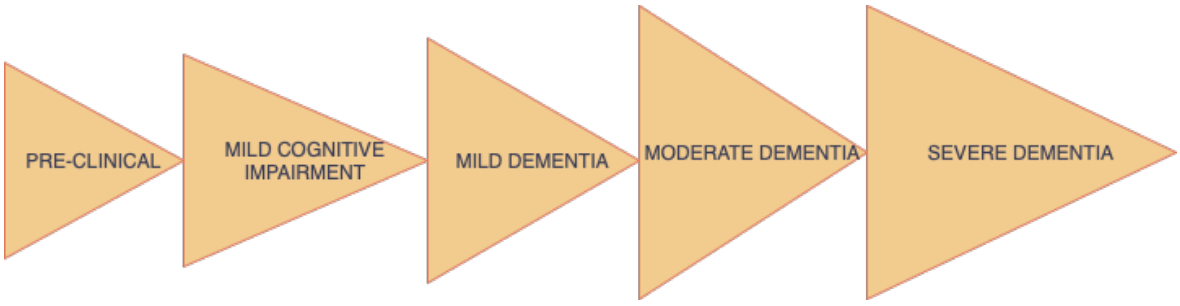


Figure 1.3: AD Stages

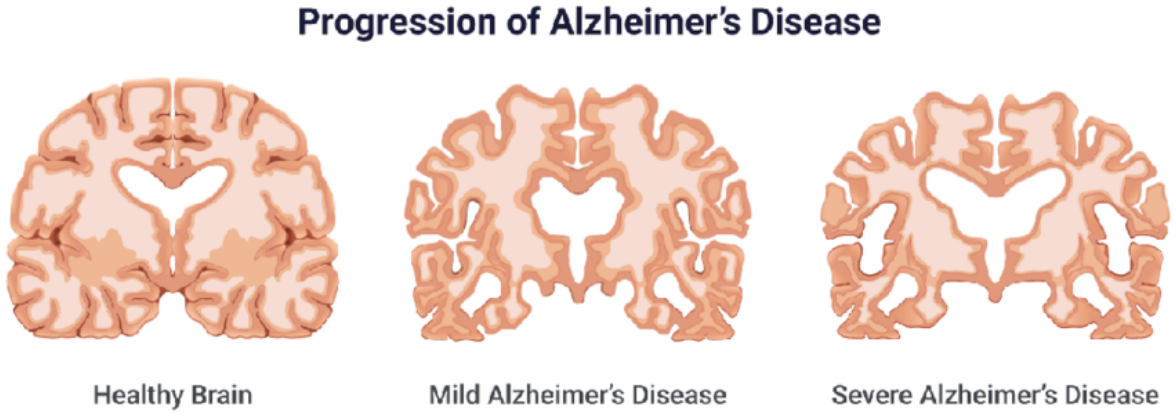


Figure 1.4: AD Progression,Source: [10]

1.1.5 BIOMARKERS

Biomarkers are measurable biological changes that can show if a disease is present or a person is at risk for developing a disease.Scientists and health care professionals screen for certain diseases by detecting and measuring biomarkers.[1].
There exists 2 sets of biomarkers:

- INVASIVE: Biomarkers obtained via methods that invades the body.Ex: Blood tests etc.
- NON-INVASIVE Biomarkers obtained by methods which do not invade the body. Ex: fMRI,MRI etc.

1.1.6 MOTIVATION

Since the rate of progression of Alzheimer disease varies greatly among subjects early detection of the disease is paramount to:

- Improve quality of life for patients and caregivers
- Improve economic impact
- Comprehend disease mechanism across subjects which helps develop treatments to mitigate progression
- Help reduce its ever growing prevalence globally.

1.1.7 PROBLEM STATEMENT

ASSUMPTIONS

The problem statement is based upon 2 primary assumptions namely,

- Changes in symptoms of subjects captured through multiple tests better sheds light on the accurate disease stage.
- Magnetic Resonance Imaging (MRI) has the ability to capture complex brain changes that occurs over time as disease progresses.

OBJECTIVES

The thesis has 3 main objectives:

- Predict the disease stage a subject is in with the help of multiple cognitive assessment scores available in ADNI
- To develop a faster and accurate 2D-CNN architecture to better classify disease stage from structural MRI.
- To consider a mix of both cognitive scores as well as features extracted from MRI as input to improve prediction accuracy on disease stage classification.

1.1.8 THESIS OVERVIEW

This thesis is organised as follows:

- Chapter-3 sheds light on previous works performed towards classification of Alzheimer's disease stage.
- Chapter-4 gives insights into the dataset used for the study along with its distribution and the challenges posed.
- Chapter-5 outlines the data preprocessing carried out for the study.
- Chapter-6 discusses the architectural details of the 2D-CNN model considered to work with MRI.
- Chapter-7 provides details regarding the steps taken for each modality and results obtained.
- Chapter-8 summarizes the results obtained.
- Chapter-9 gives conclusion and future line of work.

Chapter 2

Related Work

In recent years methods of deep learning and machine learning have contributed significantly towards AD diagnosis. Ahmed A et al[4] in 2023 developed a lightweight 2D-CNN to work with 3-channel MRI images taken from Kaggle.The work considers classifying subjects into very mild dementia, mild dementia, moderate dementia and subjects with no dementia. Ruhul Amin Hazarika et al[5] in 2023 used a Deep CNN model combining AlexNet and LeNet architectures with multi-sized filters (1x1,3x3,5x5) at the convolution layers to combine outputs and classify subjects into CN,MCI and AD. Alejandro Mora-Rubio et al[9] in 2023 used 3 different deep CNN architectures to classify subjects into CN,MCI and AD.In 2020 Jong Bin Bae et al[2] used the Inception-v4 architecture fed with MRI scans from AD and CN subjects from different demography and ethnicity towards disease stage classification. Menagadevi et al.[7] developed a computer-aided diagnosis system for detecting AD based on a combination of a deep learning model with traditional classification methods. They first start with preprocessing stages on the input MRI images to enhance the images. After that, they perform segmentation on the preprocessed images to obtain the region of interest. Then, they extract the features using the presented multi scale pooling residual auto encoder model. Later on they made use of separate classifiers such as KNN which provided an overall accuracy of 96.88%.The study only considers MRI and that too with a relatively small dataset. Rizwan Khan et al [6] in 2023 used a transfer learning approach. They used pre-trained models of VGG-16 and VGG-19 to perform the classification of subjects into 4 classes of NC(Normal Cognition), EMCI(Early Mild Cognitive Impairments), LMCI(Late Mild Cognitive Impairments)

and AD(Alzheimer disease) with input as MRI computing the volume of gray matter (GM) in the brain.Mohammad et al[8] in 2021 presented a study that used OASIS MRI images from Kaggle as input to AlexNet and ResNet-50 architectures to help classify subjects into Mild demented, Moderate demented,Very mild demented and non-demented.With the features extracted from the CNN models they used several machine learning classifiers such as SVM(Support Vector Machine), RF(Random Forest), KNN(K-Nearest Neighbour) to classify subjects. K.Swetha et al [11] in 2022 used 2 different sets of data to predict disease category. One set contains data such as MMSE (Mini-Mental State Exam), CDR (Clinical Dementia Rating) etc and the other consisted of information such as age, number of visits to clinic, education, handedness etc. Amrutesh et al [1] in 2022 used both MRI and a combination of CDR (Clinical Dementia Rating) and MMSE (Mini-Mental State Exam) scores seperately from OASIS dataset to classify patients into 4 groups of CN (Normal Cognition), EMCI (Early Mild Cognitive Impairment), LMCI (Late Mild Cognitive Impairment) and AD (Alzheimer's Disease). Many machine learning models were with CDR and MMSE as input of which Random Forest gave the best accuracy of 92.1385%. Transfer learning using InceptionNet v3 on MRI data achieved an accuracy of 90%.

Chapter 3

Data Description

3.1 ADNI DATASET

Alzheimer's Disease Neuroimaging Initiative(ADNI) is a longitudinal multi-centre study carried out in phases(ADNI1,ADNI-GO,ADNI2,ADNI3) [1] in North America.The study objective aimed at creating a multi-modal dataset to help research in the domain of Alzheimer's disease detection and progression with the help of biomarkers.
This study considers dataset from ADNI1,ADNI-GO and ADNI-2 phases and includes 3 different classes namely NC(Normal Cognition), MCI(Mild Cognitive Impairment) and AD(Alzheimer disease).The dataset contains various modalities such as structural MRI (sMRI) and cognitive assessment scores for each subject.

3.2 MRI

Magnetic Resonance Imaging (MRI) is a non-invasive imaging method that captures high quality images of organs with the help of magnetic and electric fields.
They are of 3 types: structural MRI (sMRI),functional MRI (fMRI) and diffusion MRI (dMRI). Structural MRI (sMRI) is the non-invasive technique to image organs usually employed for pathology detection in the brain. It measures the amount of water in a given location which outputs a detailed picture of the brain anatomy.This leads to easily distinguish between various tissue types such as gray matter (GM) and white matter (WM).
Since we capture features that represents the structural changes that happens in the

brain due to Alzheimer's disease we restrict our scope to sMRI.[3]

3.3 COGNITIVE SCORES

This study makes use of the following cognitive scores:

- **Mini-Mental State Exam(MMSE)**: [1] A basic screening tool to test for cognitive impairment.It contains 11 questions that tests subjects on 5 cognition domains namely,
  - ORIENTATION
  - REGISTRATION
  - ATTENTION/CALCULATION
  - RECALL
  - LANGUAGE

  The score ranges between 0-30. Scores <= 23 indicates cognitive impairment.

- **Clinical Dementia Rating(CDR)**: [1] A semi-structured tool to assess dementia severity.It contains questions that tests 2 domains namely,
  - COGNITION
  - FUNCTION

  The score ranges between 0-3 with the following indications,
  - 0 : No dementia
  - 0.5 : Questionable dementia
  - 1.0 : Mild Dementia
  - 2.0 : Moderate Dementia
  - 3.0 : Severe Dementia

- **Functional Activities Questionnaire(FAQ)**: [1][2] An observer-reported outcome that provides insights into instrumental activities of daily living(IADLs) basis a set of structured questions.Instrumental activities of daily living includes,

  - Making a cup of tea/coffee
  - Managing personal finance etc.

  The score ranges between 0-30 with the following indications against each set of question

  - 0 : Normal
  - 1 : Has difficulty but does by self
  - 2.0 : Needs a bit of assistance
  - 3.0 : Dependent

- **Geriatric Depression Scale(GDS)**: [1][2] Screening test that determines depression in elderly for further evaluation.It has 30 yes/no questions with each scored either 1/0. The score ranges between 0-30 with the following indications,

  - 0-9 : Normal
  - 10-19 : Mild Depression
  - 20-30 : Severe Depression

- **Neuropsychiatric Inventory Questionnaire(NPI-Q)**: [1] Measures severity and impact of manifested symptoms on the caregiver as indicated by the caregiver/informant.It covers 12 domains namely,

  - Delusions
  - Hallucinations
  - Agitation
  - Depression
  - Anxiety
  - Elation
  - Lack of Interest/Concern
  - Dis-inhibition
  - Aimless Wondering without purpose

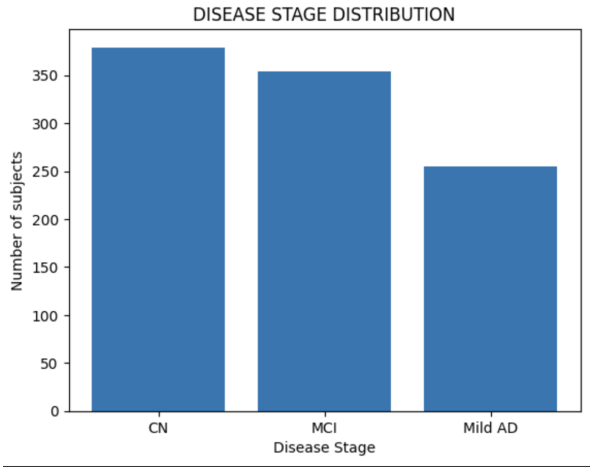
- Nighttime Behaviour
- Appetite
- Irritability

Responses are yes/no.If caregiver responds yes he/she rates severity on a 3 point scale and distress on a 5 point scale.Score ranges between 0-36 with higher score indicative of higher symptoms severity.

3.4 DATA DISTRIBUTION

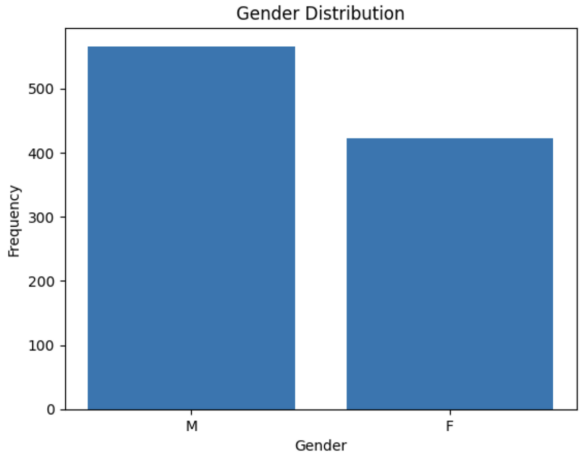
3.4.1 SUBJECTS WITH COGNITIVE SCORES

- Number of subjects with cognitive scores: 988
  Subject distribution across disease stage:
  - CN: 379
  - MCI: 354
  - AD: 255



Subject distribution gender-wise:

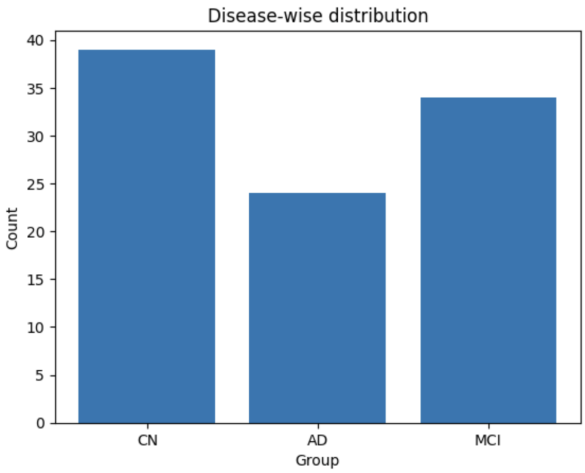
- MALE: 566
- FEMALE: 422



3.4.2 SUBJECTS WITH sMRI

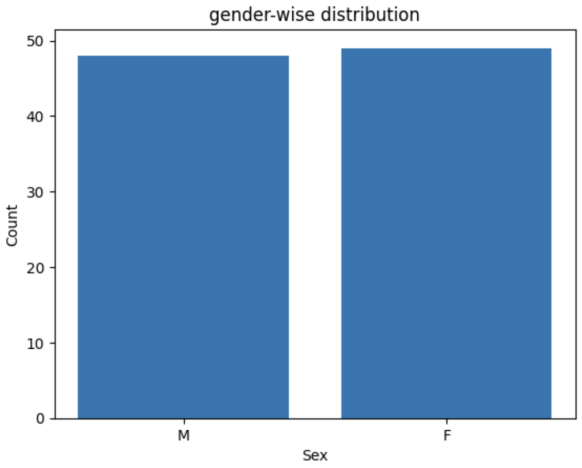
The total number of subjects considered for classification into 3 categories basis MRI scans across disease and gender-wise is as follows:

- Number of subjects with MRIs: 99
  Subject distribution across disease stage:
  - CN: 39
  - MCI: 34
  - AD: 24



Subject distribution gender-wise:

- MALE: 48
- FEMALE: 49



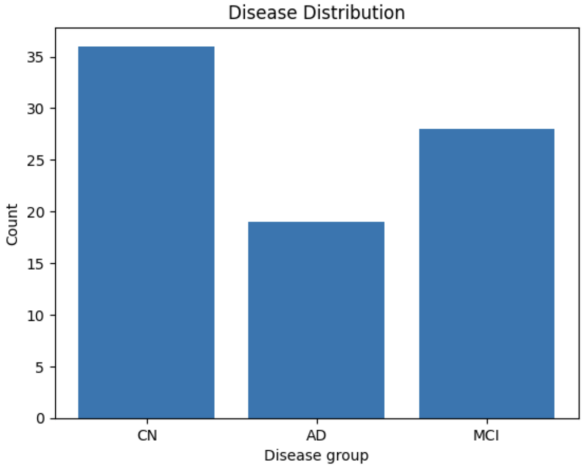
3.4.3 SUBJECTS WITH sMRI AND COGNITIVE SCORES

Number of subjects considered for classification into 3 categories basis features extracted from MRIs along with cognitive scores across disease and gender-wise is as follows:

- Number of subjects with cognitive scores: 85

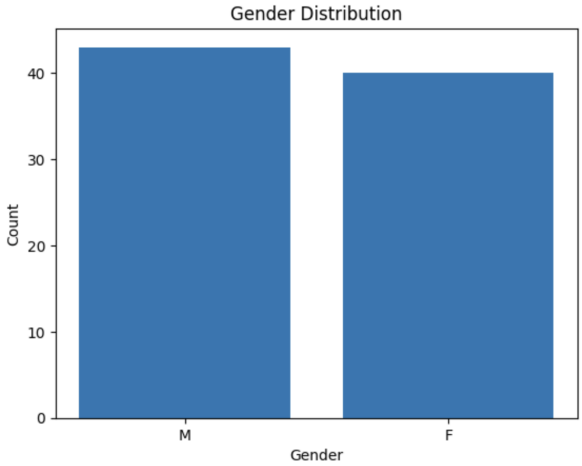
  Subject distribution across disease stage:

  - CN: 36
  - MCI: 28
  - AD: 19



Subject distribution gender-wise:

- MALE: 43
- FEMALE: 40



Chapter 4

DATA PREPROCESSING

4.1 MRI PREPROCESSING

Since the structural magnetic resonance imaging (sMRI) images suffer from the above problems feeding it as is to the model makes it challenging for the model to train.Hence images are first pre-processed using SPM-12 software in MATLAB which involves the following steps:

- **De-oblique**: During MRI acquisition the scan is usually tilted from the horizontal line to cover the whole brain and to avoid MRI induced artifacts. Hence this step is needed to have the brain align with the horizontal line.
- **Re-orientation**: MRI scans must be oriented in a way to ensure smooth registration. Most often the scans are found to be oriented in different directions and hence must be re-oriented. Orientation is dictated by 3 directions namely (1) right to left (2) Anterior to posterior and (3) Inferior to superior.
- **Magnetic Field In-homogeneity Correction**: In general magnetic field as it passes through brain must remain constant which rarely happens since different tissues in the brain absorbs magnetic field to varying extent. This forms abnormally bright and dark areas that makes it difficult to identify tissue boundaries in the brain. This step is necessary to be performed prior to skull-stripping and tissue segmentation.

- **Non-brain tissue removal/Skull-stripping**: The ROI(region of interest) lies within the brain and hence non-brain parts such as skull,neck,eyes,nose and mouth are omitted.

- **Registration**: Process of aligning images from different geometric spaces (native spaces) to a common or standard space. There exists 3 main components to it. At first a geometric transformation in space needs to be specified. The 3D transformation parameters are translation, rotation, scaling, and shearing in the x-, y-, and z-directions. Rigid-body transformation consists of six DOF(Degrees Of Freedom), involving three translations and three rotations, while affine transformation consists of 12 DOFs involving three scaling and three shearing factors in addition to the rigid-body parameters.

- **Segmentation**: This step segregates the tissues lying in GM(Gray Matter), WM(White Matter) and CSF(Cerebrospinal Fluid) region of brain.

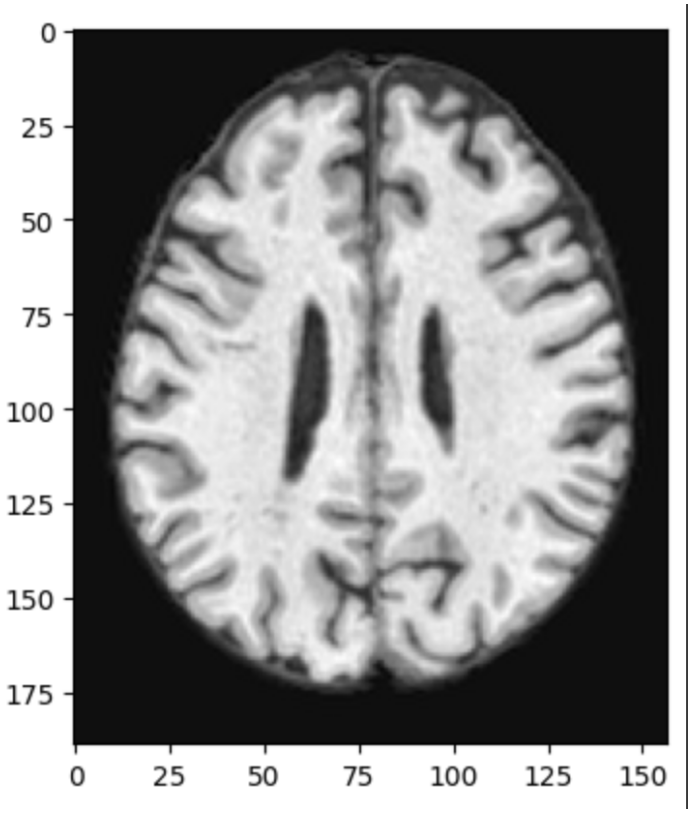


Figure 4.1: Rectangular Preprocessed MRI of height 189 units and width 156 units

4.2 MRI PADDING

Since MRI scans are fed into a 2D-CNN which considers images with equal height and width we pad the original MRI with dimensions 156 x 189 x 157 where the number of slices is 156,height is 189 and width is 157 respectively. After padding the dimension becomes 156 x 192 x 192.

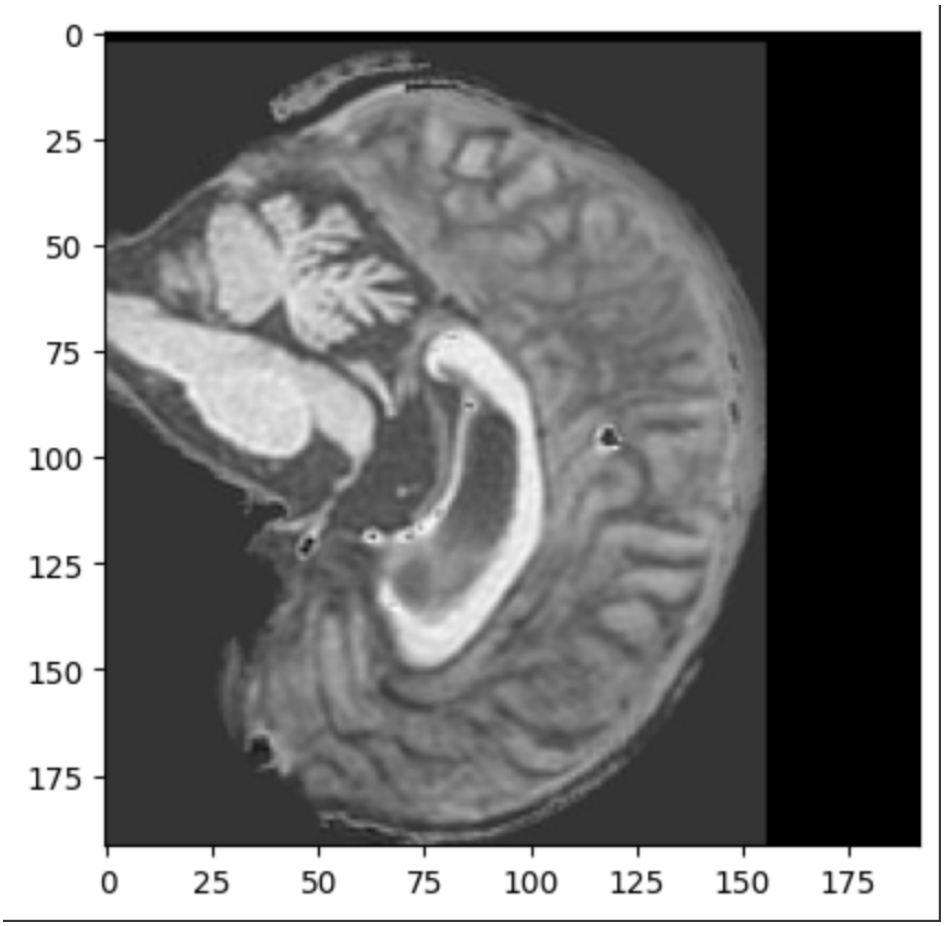
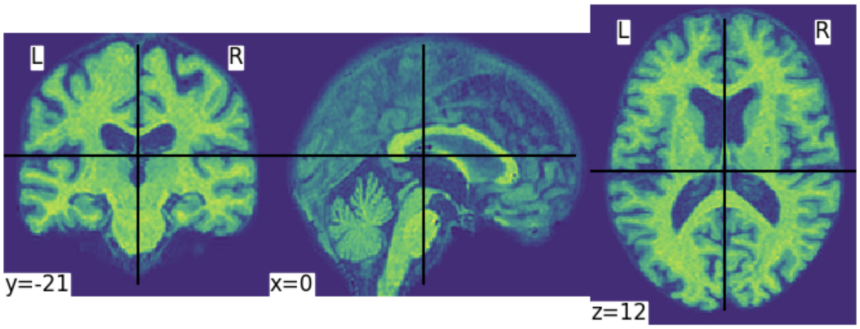


Figure 4.2: Square Preprocessed MRI of height and width as 192 units



-

Figure 4.3: Preprocessed MRI image of CN subject

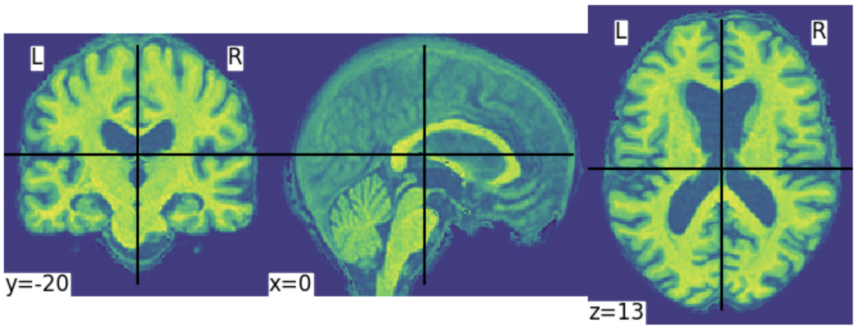


Figure 4.4: Preprocessed MRI image of MCI subject

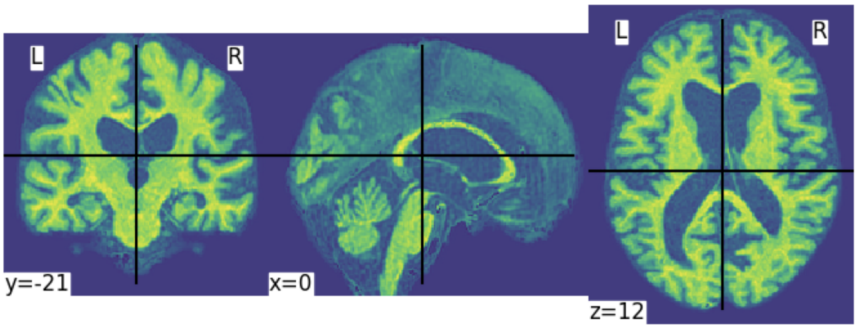


Figure 4.5: Preprocessed MRI image of AD subject

Chapter 5

MODEL

5.0.1 LIGHTWEIGHT 2D-CNN MODEL

The 2D-CNN model considered in the study is a modified version of the model used by Ahmed A et al[4] that classifies subjects into 4 classes of non-demented,very mild dementia,mild dementia and moderate dementia using dataset from Kaggle that uses 10 layers with ReLU activation in hidden layers.
The architecture of the lightweight 2D-CNN used by Ahmed at al in his study contains 2 Conv2D layers followed by 2 Maxpool2D layers,flatten layer and 2 dense layers.

5.0.2 2D-CNN USED IN STUDY

The input consists of 156 x 192 x 192 dimensional images which the classifier considers to classify subjects into 3 groups of CN, MCI and AD respectively.
The model is a stack of 3 2D convolution layers, each followed by a batch normalization layer, a maxpool 2D layer with fully connected layer having 5760 neurons and a final softmax layer.All hidden layers pass through ReLU activation.
Performance was gauged through several hyper-parameter tuning such as increasing the number of neurons in convolution layers, adjusting the learning rate, modifying the optimizer and changing the number of iterations but found this 14 layer 2D-CNN performs better with epochs set to 25, learning rate of 0.001 with Adam as optimizer.

LAYER	OUTPUT SHAPE
INPUT	(150,150,3)
CONV2D	(150,150,32)
ReLU	(150,150,32)
MAXPOOL2D	(75,75,32)
CONV2D	(75,75,32)
ReLU	(75,75,32)
MAXPOOL2D	(37,37,32)
FLATTEN	(43808)
DENSE	(150)
DENSE	(4)

Table 5.1: Lightweight 2D-CNN

LAYER	OUTPUT SHAPE
INPUT	(156,192,192)
CONV2D	(192,192,10)
BATCHNORM2D	(192,192,10)
ReLU	(192,192,10)
MAXPOOL2D	(96,96,10)
CONV2D	(96,96,10)
BATCHNORM2D	(96,96,10)
ReLU	(96,96,10)
MAXPOOL2D	(48,48,10)
CONV2D	(48,48,10)
BATCHNORM2D	(48,48,10)
ReLU	(48,48,10)
MAXPOOL2D	(24,24,10)
FLATTEN	(5760)

Table 5.2: 2D-CNN our study uses

Chapter 6

Experiment And Results Analysis

6.1 CLASSIFICATION USING COGNITIVE SCORES

The following depicts steps for classification using cognitive assessment scores of subjects with a total of 14843 records.

1. Obtain cognitive scores of subjects.
2. Since cognitive scores do not lie within the same range each of them is normalised to have similar distribution.
3. The normalized scores of each of the subject along with the subject disease group(CN/MCI/AD) is fed to the models
4. Perform 5-fold cross validation to assess model performance.

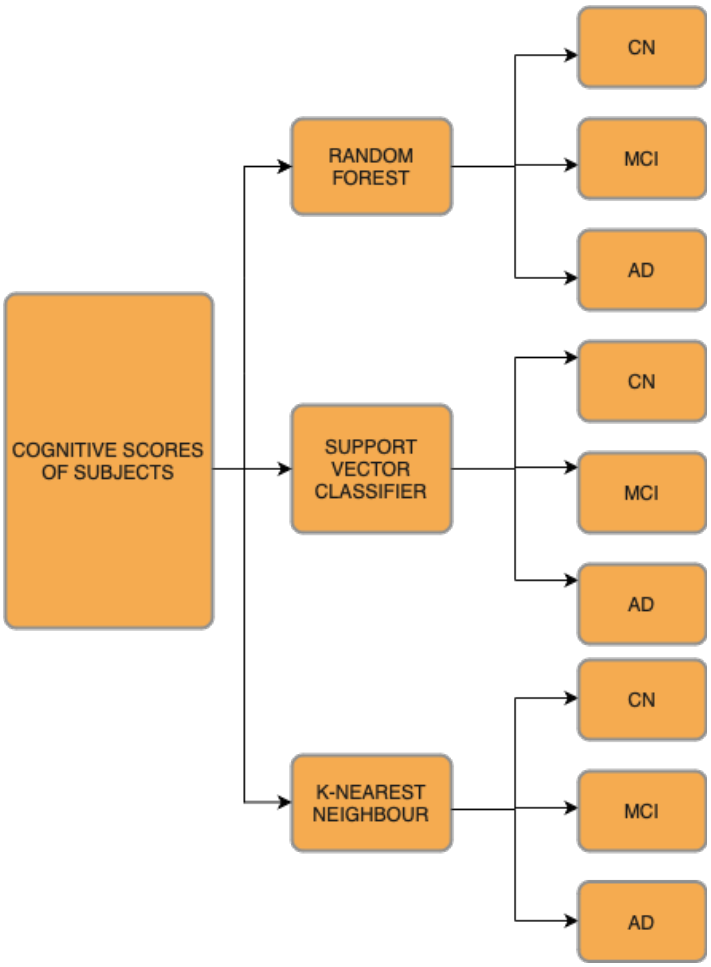


Figure 6.1: Methodology For Disease Stage Classification With Cognitive Assessment Scores

We obtain the following results:

- SUPPORT VECTOR CLASSIFIER

CLASS	PRECISION	RECALL	F1-SCORE	ACCURACY
CN	0.93	0.88	0.90	0.8
MCI	0.74	0.85	0.79	
AD	0.66	0.46	0.54	

Table 6.1: SVM Classification Report

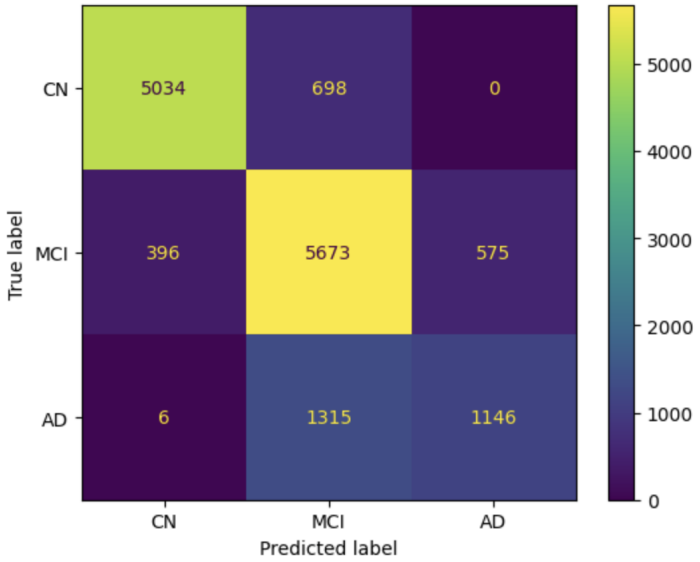


Figure 6.2: SVM Confusion Matrix

- K-NEAREST NEIGHBOUR CLASSIFIER

CLASS	PRECISION	RECALL	F1-SCORE	ACCURACY
CN	0.93	0.96	0.95	0.95
MCI	0.96	0.93	0.95	
AD	0.99	0.99	0.99	

Table 6.2: KNN Classification Report

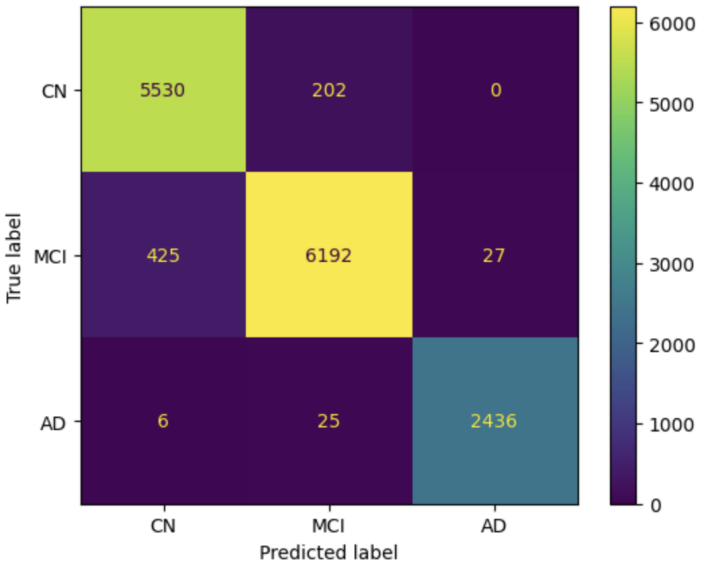


Figure 6.3: KNN Confusion Matrix

- RANDOM FOREST CLASSIFIER

CLASS	PRECISION	RECALL	F1-SCORE	ACCURACY
CN	0.93	0.88	0.90	0.80
MCI	0.75	0.85	0.80	
AD	0.67	0.51	0.58	

Table 6.3: RF Classification Report

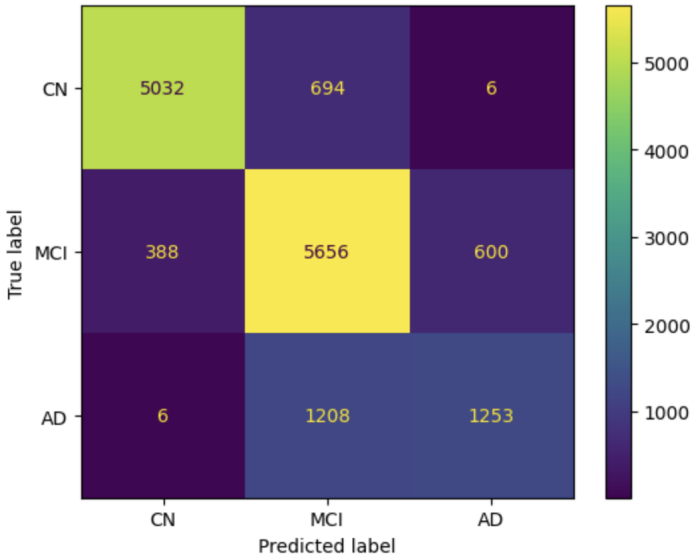


Figure 6.4: RF Confusion Matrix

6.2 CLASSIFICATION USING MRI SCANS

The following depicts steps for classification using MRI images of subjects:

1. Gather raw MRI of subjects.
2. Raw MRIs are preprocessed.
3. Pixels with values as NAN/nan are set to zero for every MRI.
4. The pre-processed MRIs with unequal height and width per slice is padded with zeros to make them have equal height and width before and used as input to 2D-CNN.
5. Perform 5-fold cross validation to assess model performance.

CLASS	PRECISION	RECALL	F1-SCORE	ACCURACY
CN	0.0	0.0	0.0	0.35
MCI	0.35	1.0	0.52	
AD	0.0	0.0	0.0	

Table 6.4: 2D-CNN MRI Classification Report

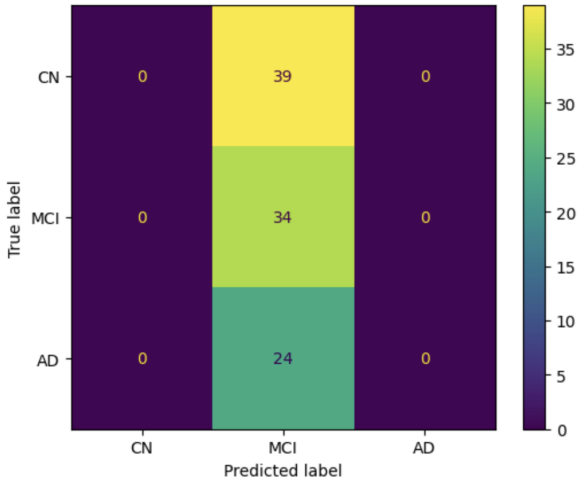


Figure 6.5: 2D-CNN Confusion Matrix

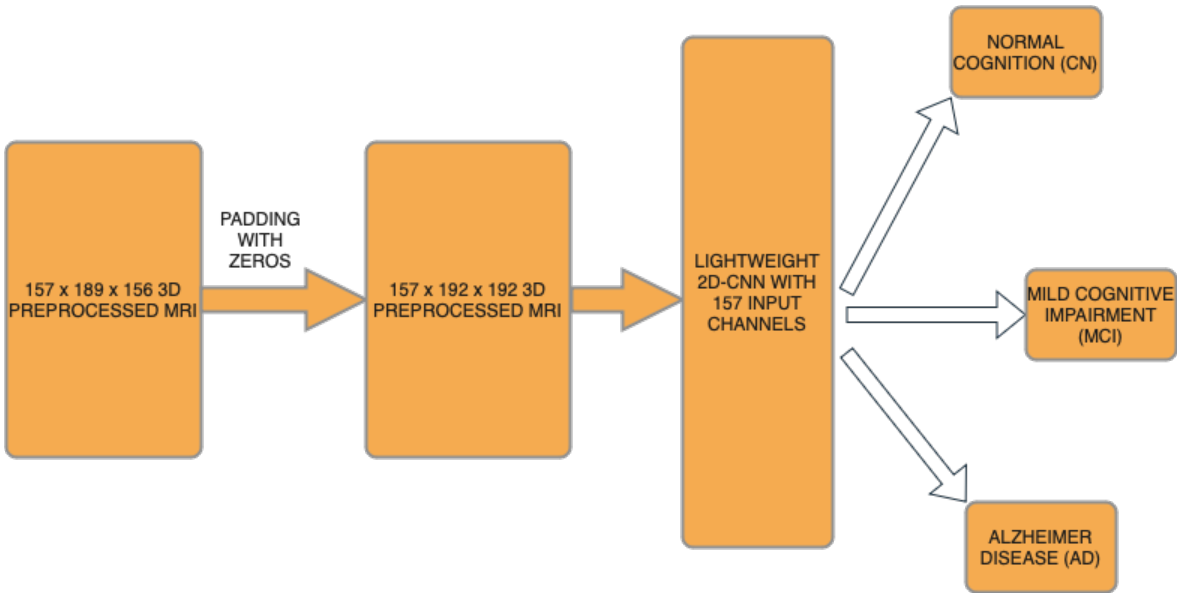


Figure 6.6: Methodology For Disease Stage Classification With MRI

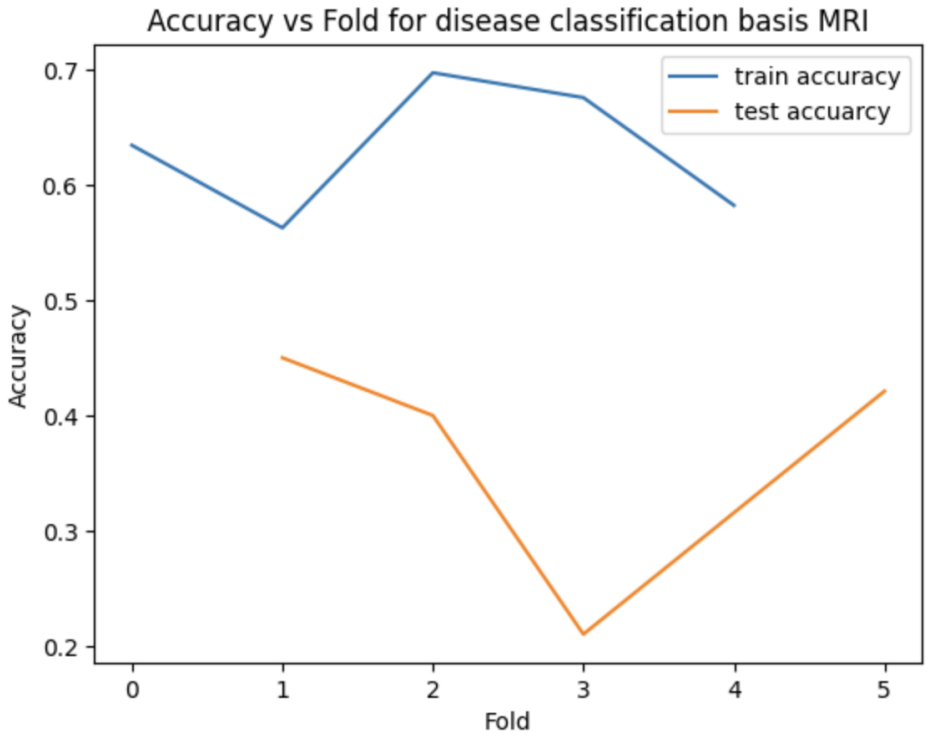


Figure 6.7: Train-Test Accuracy Curve fold-wise

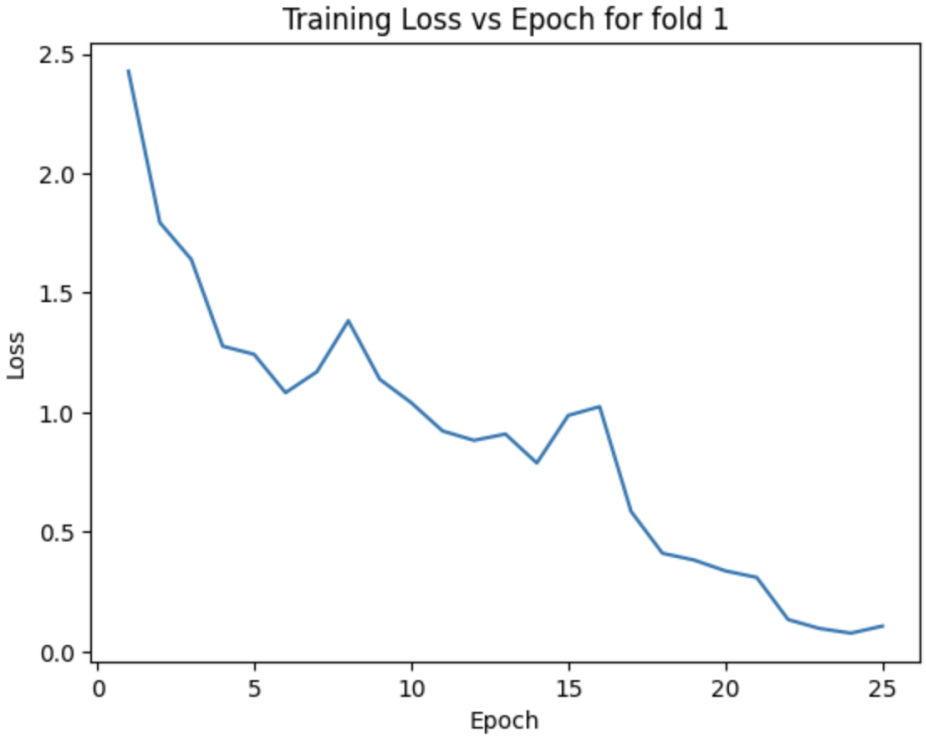


Figure 6.8: Train-Loss in fold 1

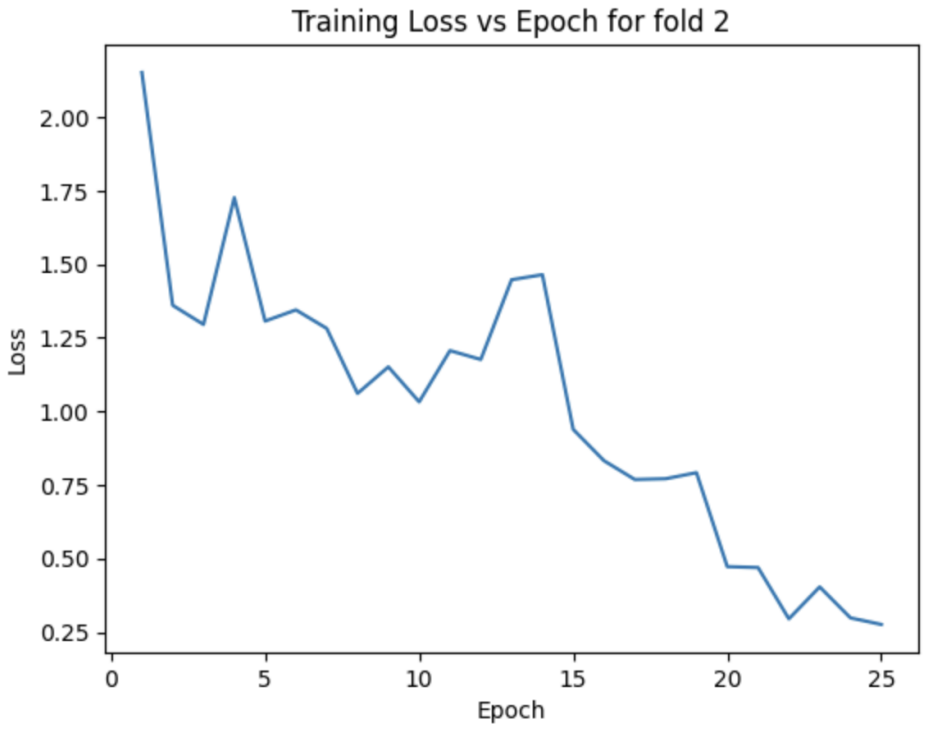


Figure 6.9: Train-Loss in fold 2

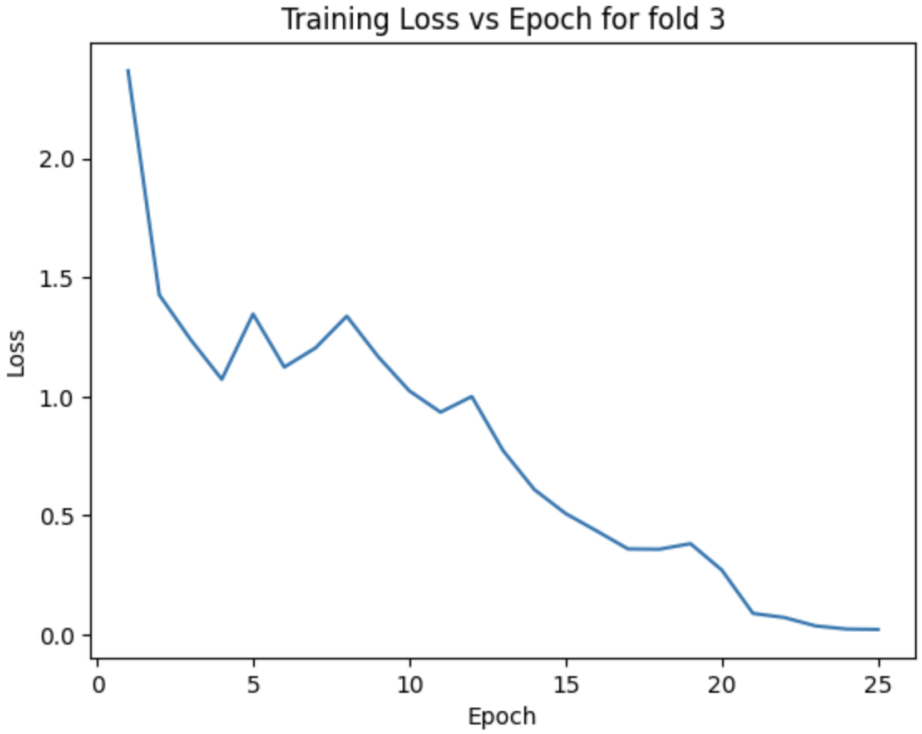


Figure 6.10: Train-Loss in fold 3

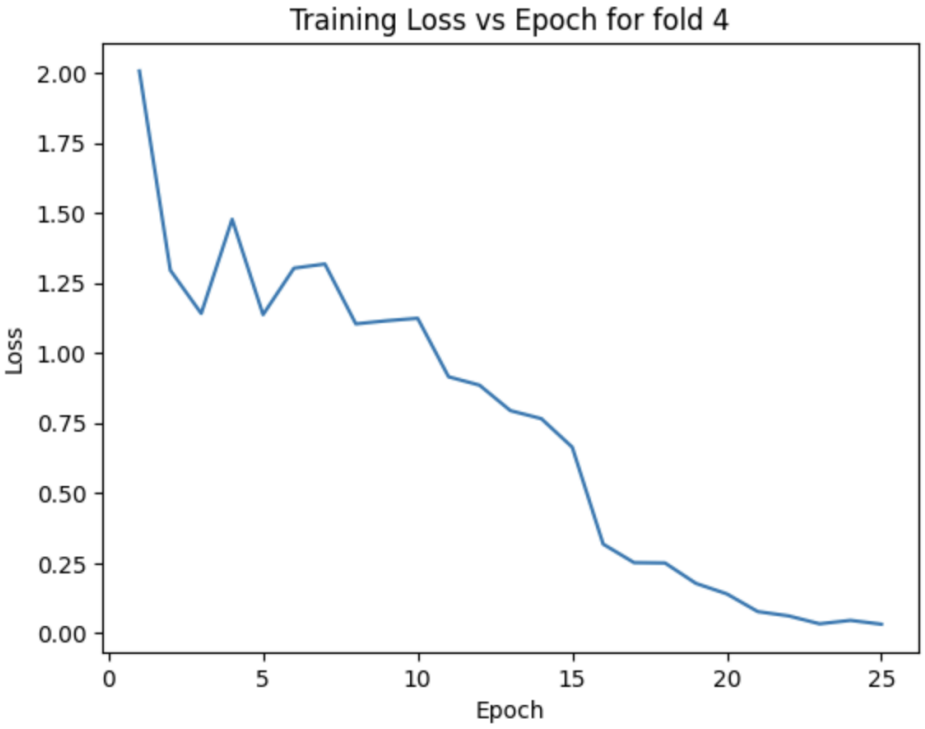


Figure 6.11: Train-Loss in fold 4

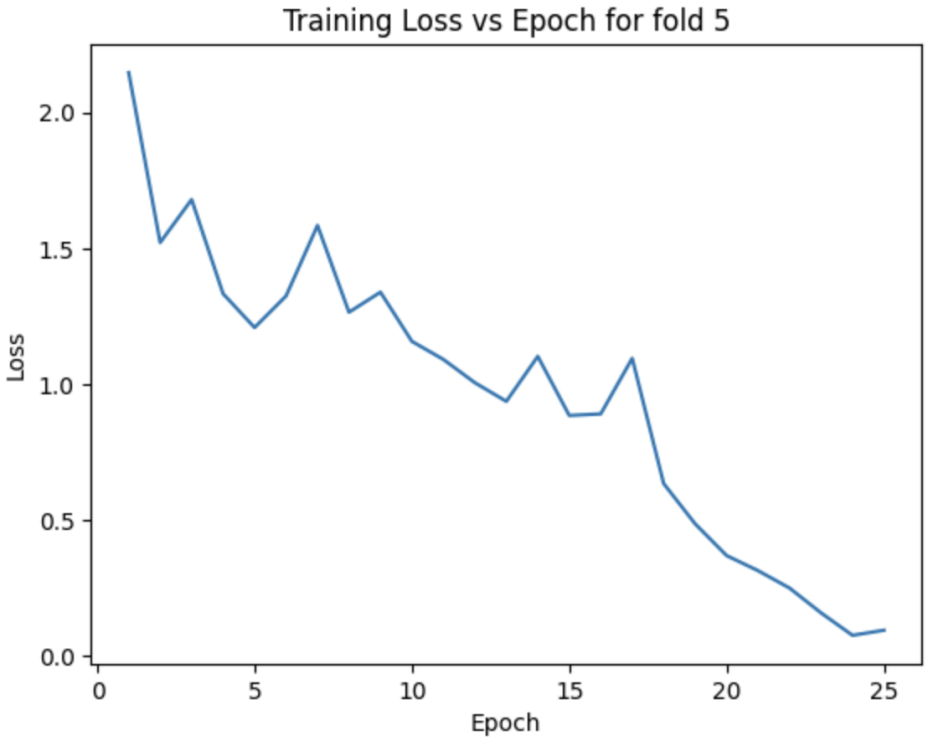


Figure 6.12: Train-Loss in fold 5

6.3 CLASSIFICATION USING MRI AND COGNITIVE SCORES

The following steps depicts classification of subjects basis a combination of MRIs and cognitive scores of subjects.

1. Gather features per subject extracted by 2D-CNN from respective MRIs.
2. The extracted features from MRIs are combined with cognitive scores subject wise.
3. The concatenated cognitive scores and MRI features are normalized and used as input.
4. The input is fed to machine learning classifiers such as SVC(Support Vector Classifier),RF(Random Forest) and KNN(K-Nearest Neighbour).
5. Perform 5-fold cross validation to assess model performance.

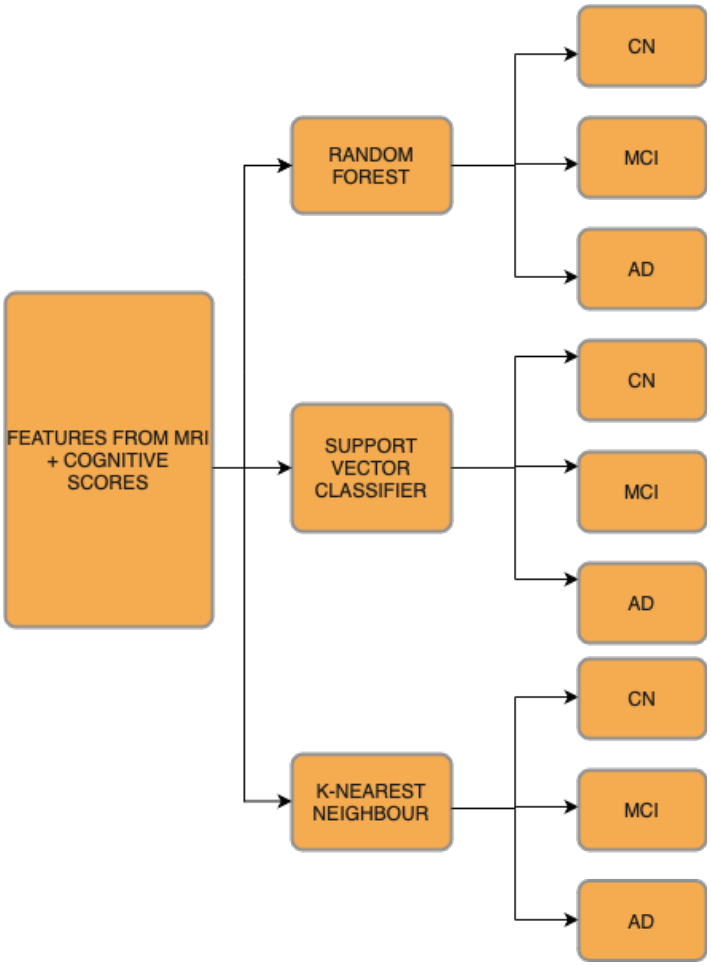


Figure 6.13: Methodology For Disease Stage Classification With MRI and Cognitive Scores

We obtain the following results:

- SUPPORT VECTOR CLASSIFIER

CLASS	PRECISION	RECALL	F1-SCORE	ACCURACY
CN	0.40	0.47	0.44	0.37
MCI	0.36	0.43	0.39	
AD	0.25	0.11	0.15	

Table 6.5: SVM Classification Report

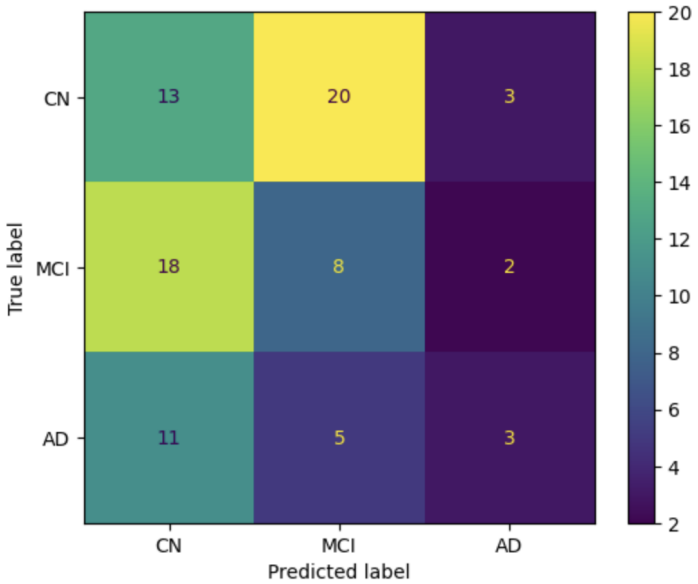


Figure 6.14: SVM Confusion Matrix

- K-NEAREST NEIGHBOUR CLASSIFIER

CLASS	PRECISION	RECALL	F1-SCORE	ACCURACY
CN	0.46	0.61	0.52	0.40
MCI	0.33	0.36	0.34	
AD	0.2	0.05	0.08	

Table 6.6: KNN Classification Report

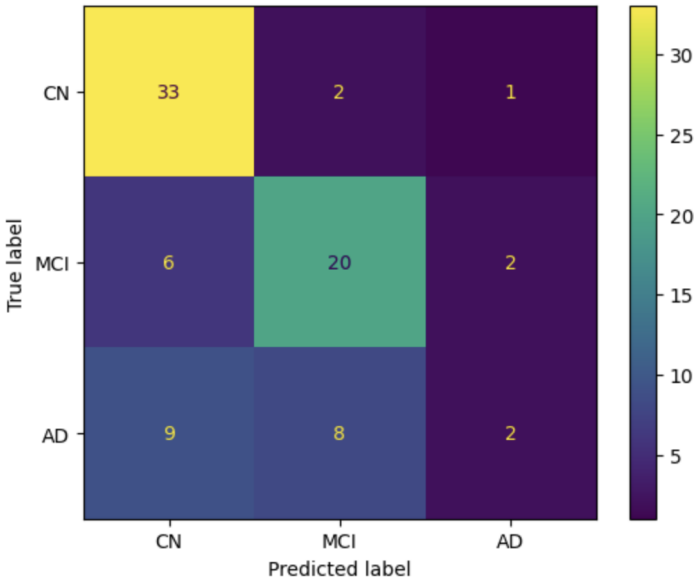


Figure 6.15: KNN Confusion Matrix

- RANDOM FOREST CLASSIFIER

CLASS	PRECISION	RECALL	F1-SCORE	ACCURACY
CN	0.41	0.64	0.50	0.41
MCI	0.44	0.25	0.32	
AD	0.36	0.21	0.27	

Table 6.7: RF Classification Report

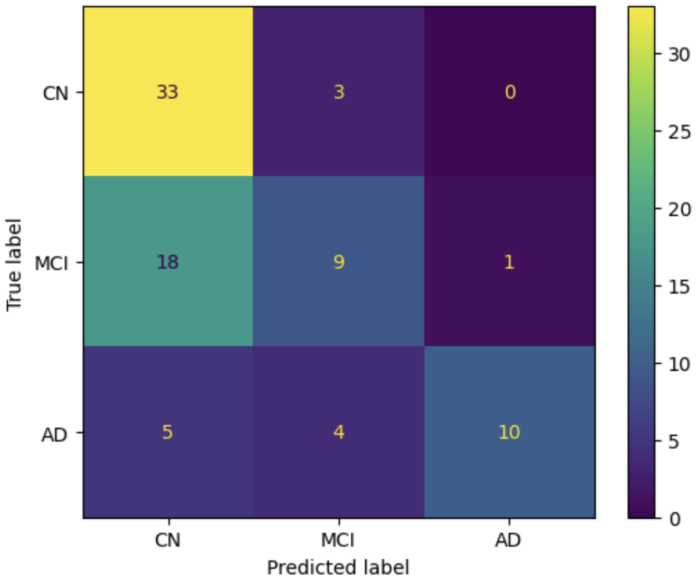


Figure 6.16: RF Confusion Matrix

Chapter 7

Discussion

From the results obtained we see that machine learning models with cognitive assessment scores outperforms achieving 95% accuracy.This highlights the importance of not one but several cognitive scores combined towards prediction of Alzheimer's disease stage.

For 2D-CNN fed with MRI figure 7.7 indicates a case of over-fitting which could be attributed to lack of sufficient data since we use only 100 MRI scans for training and testing purpose. Also figure 7.8 to 7.12 indicates a lot of variation which could be due to the presence of noise in the image. 2D-CNN achieves an average accuracy of 35% with a standard deviation of 0.086.

The above factors have led to the poor performance of machine learning classifiers that considers input as a mix of features extracted from MRI and cognitive scores with a slight improvement over 2D-CNN model achieving an accuracy of 41% by Random Forest (RF).

Chapter 8

Conclusion and Future Work

This is a novel study that looks at how different inputs can help with disease stage prediction in addition to develop a 2D-CNN that not only executes faster but gives decent result towards disease classification. Convolutional Neural Networks (CNNs) with Structural MRI scans as input has shown great result but due to lack of sufficient data our model could not show promising results while models with cognitive scores outperformed.

In future work with perhaps a larger sized dataset hopefully the model would perform better which will definitely improve the accuracy of the machine learning models that considers both MRI and cognitive scores as input.

Hybrid inputs both from imaging modalities and others can help track disease progression better which was limited by our small sized dataset.In future i would apply more advanced CNN models with inputs that includes other imaging modalities such as fMRI etc.

Bibliography

[1] Aniverthy Amrutesh, Gowtham Bhat C G, Amruthamsh A, Asha Rani KP, and Gowrishankar S. Alzheimer's disease prediction using machine learning and transfer learning models. In *2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, pages 1–6, 2022.

[2] Jong Bin Bae, Subin Lee, Wonmo Jung, Sejin Park, Weonjin Kim, Hyunwoo Oh, Ji Won Han, Grace Eun Kim, Jun Sung Kim, Jae Hyoung Kim, et al. Identification of alzheimer's disease using a convolutional neural network model based on t1-weighted magnetic resonance imaging. *Scientific reports*, 10(1):22252, 2020.

[3] Sumana Basu. *Analyzing Alzheimer's disease progression from sequential magnetic resonance imaging scans using deep convolutional neural networks*. McGill University (Canada), 2019.

[4] Ahmed A Abd El-Latif, Samia Allaoua Chelloug, Maali Alabdulhafith, and Mohamed Hammad. Accurate detection of alzheimer's disease using lightweight deep learning model on mri data. *Diagnostics*, 13(7):1216, 2023.

[5] Ruhul Amin Hazarika, Arnab Kumar Maji, Debdatta Kandar, Elzbieta Jasinska, Petr Krejci, Zbigniew Leonowicz, and Michal Jasinski. An approach for classification of alzheimer's disease using deep neural network and brain magnetic resonance imaging (mri). *Electronics*, 12(3):676, 2023.

[6] Rizwan Khan, Saeed Akbar, Atif Mehmood, Farah Shahid, Khushboo Munir, Naveed Ilyas, M Asif, and Zhonglong Zheng. A transfer learning approach for multiclass classification of alzheimer's disease using mri images. *Frontiers in Neuroscience*, 16:1050777, 2023.

[7] M Menagadevi, S Mangai, Nirmala Madian, and D Thiyagarajan. Automated prediction system for alzheimer detection based on deep residual autoencoder and support vector machine. *Optik*, 272:170212, 2023.

[8] Badiea Abdulkarem Mohammed, Ebrahim Mohammed Senan, Taha H Rassem, Nasrin M Makbol, Adwan Alownie Alanazi, Zeyad Ghaleb Al-Mekhlafi, Tariq S Almurayziq, and Fuad A Ghaleb. Multi-method analysis of medical records and mri images for early diagnosis of dementia and alzheimer's disease based on deep learning and hybrid methods. *Electronics*, 10(22):2860, 2021.

[9] Alejandro Mora-Rubio, Mario Alejandro Bravo-Ortíz, Sebastián Quiñones Arredondo, Jose Manuel Saborit Torres, Gonzalo A Ruz, and Reinel Tabares-Soto. Classification of alzheimer's disease stages from magnetic resonance images using deep learning. *PeerJ Computer Science*, 9:e1490, 2023.

[10] Ibrahim Sadek Shafiq Ul Rahman, Lamees M Mahmoud. Ai-based tool for early detection of alzheimer's disease, 2023.

[11] K. Swetha, E. N. Vijaya Kumari, Ajmeera Kiran, and Keerthana Sree Arrola. Alzheimer's disease diagnosis from mri using siamese convolutional neural network. In *2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*, pages 1–6, 2022.