Identifying the Best Machine Learning Techniques for the Prediction of the Course of Alzheimer's Disease

An Assignment
Submitted in fulfillment of the requirements for the
End of Year Project

by

M'Saydez Campbell
Advisor:
Prof. Dr. Mulenbach

17 February 2022

Declaration

UJM St. Etienne
17 February 2022

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I declare that I have properly and accurately acknowledged all sources used in the production of this report. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

M'Saydez Campbell

Abstract

Alzheimer's disease is the most common neurodegenerative disease that causes demantia. It is important that Alzheimer's disease is detected in its early stages as it enhances the effectiveness of the treatments and therapies.The bag of words/ features [1] method was used to classify T1 weighted Magnetic resonance Imaging scans. The model demonstrates that Alzheimer's disease, mild cognitive impairment and cognitive normal can be distinguished using local features via multi representation learning. The highest accuracy achieved was 0.83 using Support Vector Machine. Also, the model demonstrates that the stages of mild cognitive impairment can be distinguished using local features via multiview representation with an accuracy of 0.74 using a Support Vector Machine

Table of Contents

Abstract v

List of Figures ix

1 Introduction 1

2 Related Work 3

2.1 Machine Learning Techniques for Detection of Neurodegenerative Diseases 3

2.2 Multi-view Representation Learning 4

2.3 Contributions 5

3 Model 7

3.1 Data 7

3.1.1 ADNI 7

3.1.2 MIRIAD 8

3.2 Architecture 8

3.2.1 Preprocessing 9

3.2.2 Feature Extraction 10

3.2.3 Disease Classification 10

3.2.4 MCI Classification 11

4 Results and Discussion 13

4.1 Performance Evaluation Measures 13

4.2 Results 13

4.2.1 Disease Classification 14

4.2.2 MCI Classification 14

5 Conclusion and Future Work 17

References 19

List of Figures

1.1 Sagittal, Coronal, Axial Views for the Alzheimer's Disease, Mild Cognitive Impairment and Cognitively Normal 2
1.2 Sagittal, Coronal, Axial Views for Early Mild Cognitive Impairment and Late Mild Cognitive Impairment 2

3.1 Flowchart of the Model 9
3.2 Sagittal, Coronal and Axial (Traversal) Plane [1] 10
3.3 Sample of Feature Extraction using subject classified as Alzheimer's Disease 11
3.4 Sample of Feature Extraction using a subject classified as Late Mild Cognitve Impairment 11

4.1 The Accuracy of the Classifiers 14
4.2 The confusion matrix for SVM 15
4.3 The confusion matrix for Logistic Regression 15
4.4 The confusion matrix for Random Forest 15
4.5 The confusion matrix for Decision Trees 15
4.6 The confusion matrix for SVM 16

Chapter 1

Introduction

Alzheimer's Disease (AD) is an age- related neurodegenerative disease which causes dementia. It is diagnosed by evaluating the psychiatric history, clinical examination, and neurological evaluation of the patient. Mild cognitive impairment (MCI) is seen as the transitional stage between cognitively normal aging and the onset of AD. When detected early, intervention and treatment of AD are most effective. AD can only be definitively diagnosed posthumous by observing the senile plaques and neurofibrillary using neuropathology stains.

The symptoms associated with AD include disorientation, mood, and behaviour changes, deepening confusion, memory loss, eventually advancing to difficulty speaking and swallowing and walking. This is due to the progressive and rapid loss and shrinkage of neurons within the brain resulting in the atrophy of the brain.[2] The areas of the brain most affected are the cerebral cortex and the medial temporal lobe.[3] On a structural magnetic resonance imaging (MRI), the enlargement of ventricles, extreme shrinkage of the hippocampus and increased indentation of the sulci and gyri of the cortex caused by the progression of atrophy can be observed.[2] [4] [5] [6]

This project aims to identify strategies that differentiate mild cognitive impairment, cognitively normal and Alzheimer's disease on the brain and differentiate early and late cognitive mild impairment using magnetic resonance imaging. Consequently, the orthogonal planes of the magnetic resonance imaging was extracted. To detect the stages of Alzheimer's disease, a multi modal and multi view architecture were developed for classification.

Figure 1.1: Sagittal, Coronal, Axial Views for the Alzheimer's Disease, Mild Cognitive Impairment and Cognitively Normal

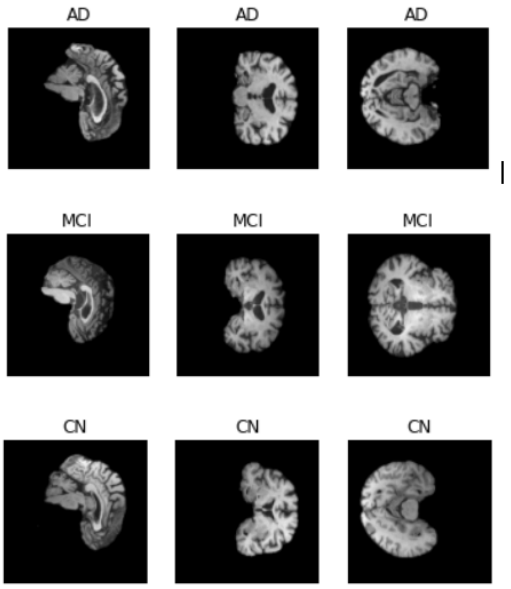
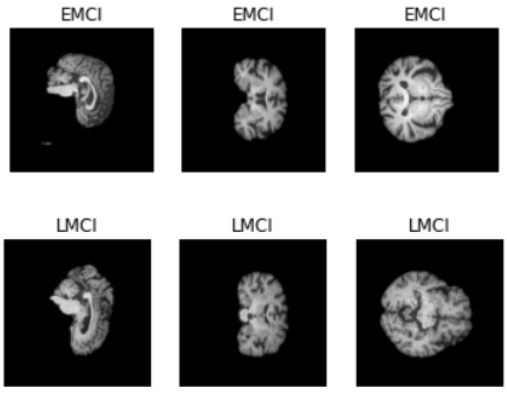


Figure 1.2: Sagittal, Coronal, Axial Views for Early Mild Cognitive Impairment and Late Mild Cognitive Impairment



Chapter 2

Related Work

2.1 Machine Learning Techniques for Detection of Neurodegenerative Diseases

Neurodegenerative diseases are a list of conditions that cause the cells of the nervous system (neurons) to shrink or die. They are incurable because neurons cannot reproduce or regenerate. Machine learning techniques are used for the diagnosis [7] [8] [9] [10] [11]and prognosis [12] [13] of neurodegenerative diseases such as Alzheimer's, Parkinson, and Schizophrenia to enhance their treatment. Researchers have used magnetic resonance imaging [10] [11] [12], psychological tests, proteins [14] [15], genetic material [16] [17] [18] and electroencephalogram [7] [8]to identify biomarkers that can correctly detect neurodegenerative diseases in subjects with high accuracy.

The use of classifiers such as Logistic Regression, Support Vector Machines (SVM), Random Forest, Decision Trees, Naïve Bayes, K Nearest Neighbours (KNN) and Ada Boost are mentioned in multiple papers [9] [10] [11]. Notable differences in previous research are the pre-processing, feature extraction, data representation and dimensionality reduction measures.

Predictability of AD at an early stage has been linked to the use of various methods by several researchers. Rohni et al [9] proposed a method that can predict AD in the early stages and categorize it by combining classifiers such as KNN, SVM and Gaussian Naïve Bayes classifier into one single model using MRI scans. Lahmiri et al [10] proposed a method for early detection of AD using Alzheimer's disease assessment scale (ADAS) scores and SMRI by extracting the cerebral cortex, cortical thickness, gyrification index. They suggested that a SVM trained with cortical thickness, gyrification index and ADAS scores distinguish between AD and cognitively normal subjects had the best results compared to other classifiers and features. Singh et al [11] proposed a method using

the Kohonen self-organizing map (KSOM) and least squares SVM to diagnose patients in different stages of Parkinson's disease using T1 weighted MRI scans. They suggested that their system can be translated to other neurodegenerative diseases.

Another method to predict AD at an early stage is deep learning. It is a branch of machine learning that has been developed to identify high level features from raw input data. The main architectures used are convolutional neural network [19] [20] [21], recurrent neural network [22], deep neural network [23] [24] [25] and autoencoder [24].

2.2 Multi-view Representation Learning

Multi-view representation learning deals with learning features of multi-view data to extract valuable information. The first studies of multi-representation learning are canonical correlation analysis [26] and its kernel expansions [27] [28]. Multi-view representation is important because machine learning models are reliant on the expressiveness of the data. The applications of multi-view representation include the combination of different data types such as video and text or images and text as well as the combination of the same data type but from multiple different angles. It can be split into two categories: multi-view representation fusion and multi-view representation alignment.

Multi-view representation fusion is the combination of multi-view inputs into one representation. Types of multi-view representation fusion learning techniques are multi modal latent Dirichlet allocation [29], multi-view sparse coding [30] [31], multi-view latent space Markov networks [32] and multi-modal deep Boltzmann machine [33]. On the other hand, multi-view representation alignment is the alliance between representations learned from multiple different views. These multi-view representation alignments can be divided into correlation-based alignment and distance and similarity-based alignment.

Over time/In recent time, several researchers have contributed to the strategies and technologies utilized for predicting and diagnosing AD. Zhang et al [34] proposed a multi-view clustering model (Consensus Multiview Clustering) for the prediction of AD based on non-negative matrix factorization. Lui et al [35] proposed a structure based multi-view learning method for AD/MCI classification. They retrieved multi-view representation of feature data for each subject, clustered the subjects in one class into different subclasses and applied a multitask feature selection model. The final decision was based on learning each view separately using Support Vector Machines and fusing their result.

Xu et al [36] used genetic variations and MRI features to propose a new sparse Bayesian approach to identify a link between genetic variations and MRI features and for AD diagnosis. Lui et al [37] also used sparse representation however it was based on

hypergraph construction. They developed a view aligned hypergraph learning method to obtain the consistency among views and a multi view label fusion method to make the final decision.

2.3 Contributions

This project proposes to use multi-view representation learning using the bag of features based on the bag of words method [1] for disease and MCI classification using sagittal, coronal, and axial slices of MRI images.The bag of features method is an adaptation of the bag of words method used in text classification. In the bag of words method, a histogram of feature occurrences for each image are created by extracting the keypoints and clustering the features. The histograms are then used to train the classifier.

Chapter 3

Model

3.1 Data

The data in this project used was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [38] and the Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD) [39] databases.

3.1.1 ADNI

ADNI was established in 2004 as a private- public partnership. The purpose of ADNI is to collect information such as medical, clinical, cognitive, functional, biochemical, and behavioural data from MCI, CN and AD individuals to detect AD at the earliest phase, identify biomarkers and assist researchers to develop effective treatments and therapeutics. This initiative can be split into four cycles ADNI 1, ADNI Go, ADNI 2 and ADNI 3. ADNI 1 focused on amnestic mild cognitive impairment (aMCI). ADNI Go focused on early MCI (EMCI) and ADNI 2 focused on the EMCI patients and late MCI (LMCI). ADNI 3 includes rollover participants from ADNI2 and newly enrolled subjects. The data used in this project was collected under the ADNI 3 cycle.

The ADNI classifies subjects using the following criteria:

- CN - subjects that have no signs of depression, mild cognitive impairment, or dementia.
- AD - subjects that have been diagnosed with Alzheimer's disease by a physician.
- SMC or Significant Memory Concern - subjects that have memory concerns considered to be the gap between CN and MCI.

- MCI - subjects that have reported memory concern autonomously or by an informant or clinician with continued regular activities and no sign of dementia.
- EMCI - subjects with MCI; however they are classified in an early stage of MCI using the Wechsler Memory Scale Logical Memory II
- LMCI - subjects with MCI; however they are classified in a late stage of MCI using the Wechsler Memory Scale Logical Memory II.

ADNI 3 consists of 2583 subjects categorized by patient, CN, MCI, LMCI, EMCI and SMC. The number of patients without a research group are 20, CN individuals are 813, MCI individuals are 663, LMCI individuals are 185 and EMCI are 340, AD 447 and SMC are 115. There are 1217 females, 1361 males and 5 where the gender was unknown. The age groups are 2 subjects under 2, 1 subject between the ages of 19-29, 100 between the ages of 50-59, 744 between the ages of 60-69, 1250 between the ages of 70-79, 454 between the ages of 80-89, 29 above the age of 89 and 3 where the gender was unknown.

The MRI scans retrieved were T1- weighted magnetization-prepared rapid gradient-echo (MP-RAGE). This led to 500 AD, 500 CN, 500 EMCI and 500 LMCI scans.

3.1.2 MIRIAD

MIRIAD is a dataset of longitudinal T1 MRI scans. The purpose of MIRIAD is to determine the smallest period where it would be possible to have clinical trials in AD using atrophy. It consists of 69 subjects categorized by AD and healthy control (HC). The number of subjects in the AD research group are 46 and 23 in the HC research group. There are 31 males and 38 females. The subjects were required to attend seven imaging visits up to 52 weeks from the first scan. On three visits, two back-to-back scans were taken.

MIRIAD classifies subjects using the following criteria:

- HC - subjects that have no signs of neurodegenerative diseases.
- AD - subjects that have been diagnosed with Alzheimer's disease by a physician.

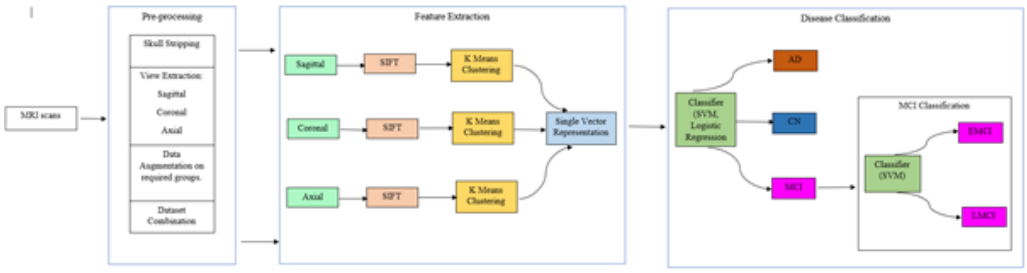
The MRI scans were T1 weighted inversion recovery prepared fast spoiled gradient recalled (IRPFSG) sequence. This led to 667 scans with 424 AD and 243 HC.

3.2 Architecture

The model is designed to be multi-view in its feature representation and multi modal in its architecture. Thereby, its architecture can be split into four parts. The first component

is data cleaning and data augmentation. The second component is feature extraction and the final component is disease classification and MCI classification.

Figure 3.1: Flowchart of the Model



3.2.1 Preprocessing

Magnetic Resonance Imaging (MRI) uses radiology to form pictures of the anatomy and the physiological processes of the body. To clean the data, deepbrain python library [40] was used to strip the non-brain elements. The MRI images were skull stripped using a custom U-Net trained on multiple datasets including the ADNI dataset.

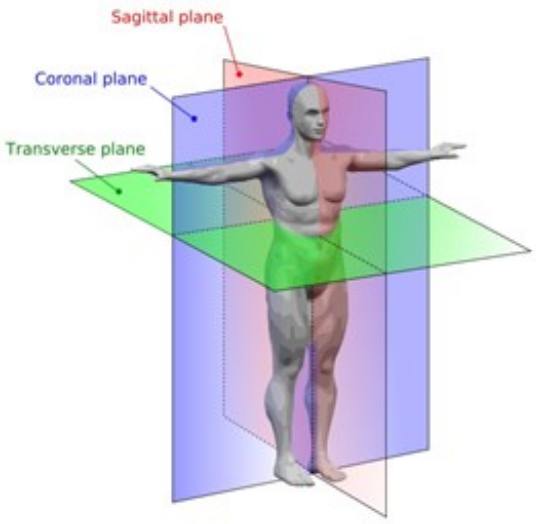
They are stored in the form.nii or .nii.gz or as a dicom file. To retrieve a representative slice of the sagittal, coronal, and axial planes, the nibabel library was used to obtain the 3D volume of the images and the slices were manually extracted.

MRI images use the anatomical coordinate system to define the planes that describe the anatomical system of a human. The planes are defined as the sagittal, coronal, and axial planes. The sagittal plane/ y-z plane separates left from right, the coronal plane or x-z plane separates the anterior from the posterior / front from the back and the axial plane or x-y-z plane separates the head from the feet. This can be seen in Figure 3.2.

The MRI images from the dataset were extracted as 3D volumes in the form of [x,y,z] as the distances from the y-z plane/ anterior from the posterior, x-z plane/left from right and x-y plane/ head from feet respectively. Using the principles highlighted, three slices representing the sagittal, axial, and coronal view were extracted.

Oversampling using data augmentation was performed on the CN category of the MIRIAD dataset increasing the number of subjects to 486. The ADNI and MIRIAD data were combined to train the model on MRI images captured under different conditions. The MCI category is a combination of the LMCI and EMCI datasets. The final dataset for disease classification is 925 AD, 986 CN and 1000 MCI and for MCI classification is 500 LMCI and 500 EMCI.

Figure 3.2: Sagittal, Coronal and Axial (Traversal) Plane [1]



3.2.2 Feature Extraction

Using Scale Invariant Feature Transform (SIFT) algorithm [41], the main features of the images were extracted. The SIFT algorithm can identify local features of an image regardless of transformation, scaling and rotation and noise. The steps of the SIFT algorithm are:

- Scale Space Extrema Detection
- Keypoint Localization
- Orientation Assignment
- Keypoint Descriptor Generation

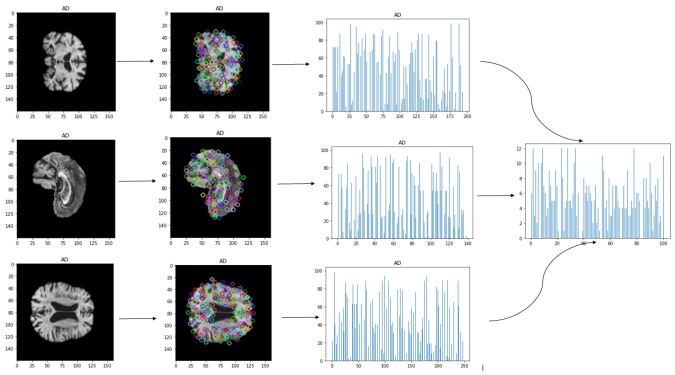
The k- means clustering algorithm [42] [43] was used to cluster the keypoints to determine similar features in each view of the MRI scans. It clusters similar data points to identify underlying relationships. After clustering each view, they were combined to form a single feature vector representation [44] of that subject.

3.2.3 Disease Classification

The classification of AD, CN and MCI categories depends on the clustering of similar features and the combination of each view into a single feature vector using local features. The local features extracted by the SIFT algorithm were differentiated by texture, colour or intensity of the image. They can be an edge, image patch or point.

For disease classification, Logistic Regression, Support Vector Machine, Decision Trees, Random Forest were used.

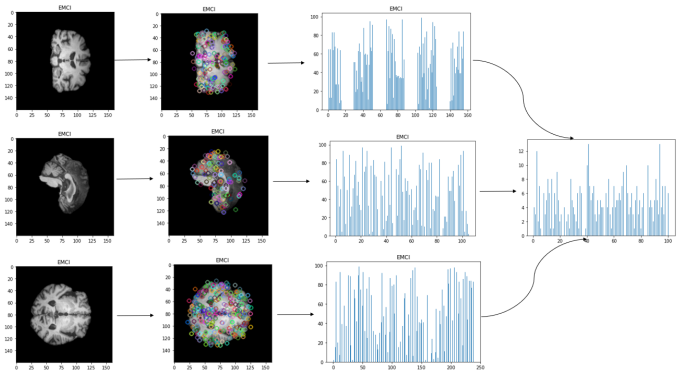
Figure 3.3: Sample of Feature Extraction using subject classified as Alzheimer's Disease



3.2.4 MCI Classification

The classification of LMCI and EMCI categories depends on the clustering of similar features and the combination of each view into a single feature vector using local features. The local features extracted by the SIFT algorithm were differentiated by texture, colour or intensity of the image. They can be an edge, image patch or point.

Figure 3.4: Sample of Feature Extraction using a subject classified as Late Mild Cognitve Impairment



For MCI classification, Support Vector Machine was used.

Chapter 4

Results and Discussion

4.1 Performance Evaluation Measures

The classification accuracy and recall (sensitivity) of each classifier were used to evaluate the performance of the model. The accuracy was calculated by the number of classified instances compared to the correct labels in the test set. The recall was calculated by the number of classified instances of a particular class by the number of correct labels that belonged to that class. The classifiers were trained using 5-fold cross validation to determine the best hyperparameters for each classifier.

4.2 Results

The model created a feature descriptor/ bag of words made up of 8370 items each containing features extracted by the SIFT algorithm. It was used to train the K means algorithm to determine to cluster the features of each image. The K or number of clusters was 100 for disease classification and 100 for MCI classification.

The SIFT algorithm was chosen over other feature extraction techniques because the MRI images are from two different sources captured by two different machines and by different methods. Due to this, the MRI scans vary in terms of rotation, position and noise. During pre-processing, the MRI images were not aligned or normalized hence using the SIFT algorithm helped in correcting these issues. The K means algorithm was chosen over other clustering algorithms because it has low computational cost and is fast.

The results prove that the model generalizes well on MRI images captured on different machines under different conditions. The SVM classifier had the best results. This is because the classifier used a rbf kernel function to project feature vectors in the input space to a feature space of higher dimension. The SVM classifier utilized a low penalty

for misclassified points and low margin. Therefore, maximizing the margin between data points allows for data points that are not close together to be considered in the same class. This permitted more data points to be grouped together.

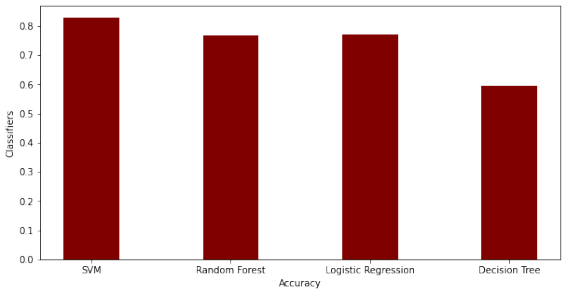
4.2.1 Disease Classification

The accuracies of the classifiers are (as seen in Figure 4.1):

- SVM - 0.83
- Random Forest - 0.77
- Logistic Regression - 0.77
- Decision Trees - 0.60

As seen in Figure 4.2, 4.3, 4.4, 4.5, the classifiers predicted mild cognitive impairment better than AD and CN categories. This is due to the AD and CN categories being a combination of MRI scans taken by two different machines using different methods. Despite this, the recall or sensitivity for the AD and CN categories in all classifiers besides Random Forest are approximately 0.70.

Figure 4.1: The Accuracy of the Classifiers



4.2.2 MCI Classification

The accuracy of the classifier are :

- SVM - 0.74

As seen in Figure 4.6, the classifier predicted EMCI than LMCI. The recall for LMCI was 0.57 exhibiting that the model had a difficult time distinguishing the EMCI from LMCI.

Figure 4.2: The confusion matrix for SVM

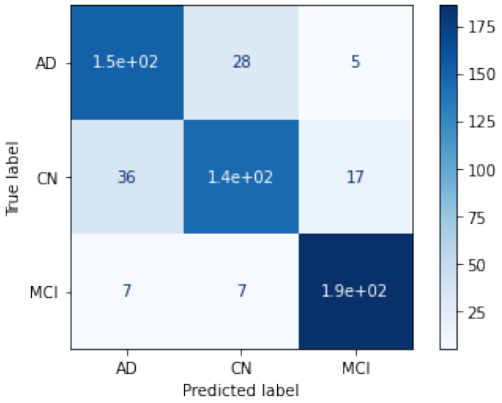


Figure 4.3: The confusion matrix for Logistic Regression

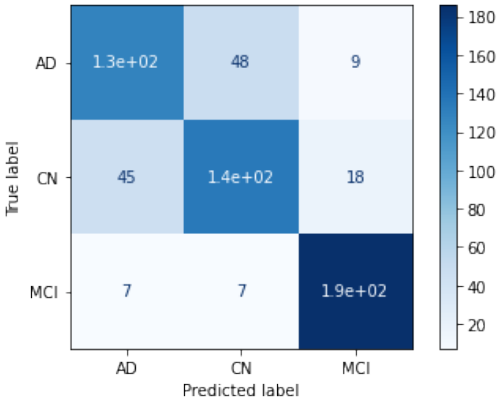


Figure 4.4: The confusion matrix for Random Forest

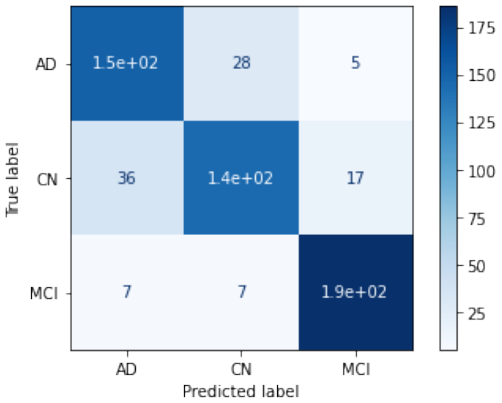


Figure 4.5: The confusion matrix for Decision Trees

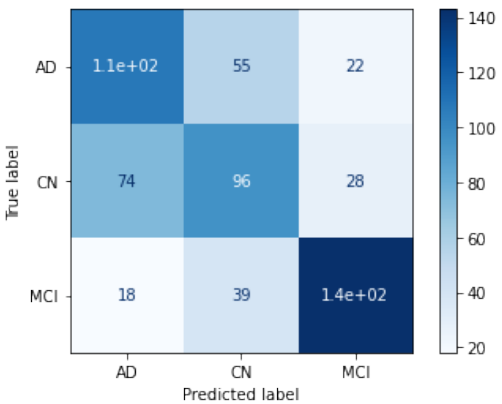
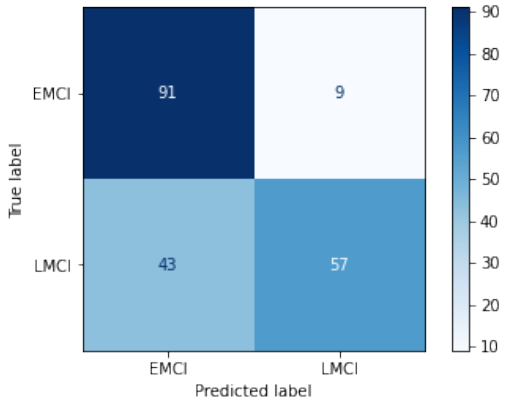


Figure 4.6: The confusion matrix for SVM



Chapter 5

Conclusion and Future Work

The results show that multi-view representation learning using a single vector representation can detect the stages of AD using local features. It suggests that a single vector representation of the three slices representing the sagittal, axial and coronal planes can determine MCI (LMCI and EMCI), AD and CN with high accuracy and limited pre-processing. Also, it implies that SIFT for feature extraction and K Means for clustering are practical tools for classifying MRI scans. This model can be enhanced by including more MRI slices and pre-processing the MRI scans by image registration, normalization. A great enhancement in this research would be to extract high level features such as cortex thickness and hippocampus size. Then applying the model. These options could increase the accuracy and make the model robust.

References

[1] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints." *ECCV Workshop on Statistical Learning in Computer Vision*, pp. 59–74, 2004.

[2] K. Johnson, N. Fox, R. Sperling, and W. Klunk, "Brain imaging in alzheimer disease," *Cold Spring Harb Perspect Med.*, vol. 13, 2012.

[3] X. Wang, M. L. Michaelis, and E. K. Michaelis, "Functional genomics of brain aging and alzheimer's disease: focus on selective neuronal vulnerability," *Current genomics*, vol. 13, 2010.

[4] A. Chandra, D. G., M. Politis, and et al, "Magnetic resonance imaging in alzheimer's disease and mild cognitive impairment," *J Neurol*, vol. 13, pp. 1293–1302, 2019.

[5] P. C. and et al, "Hippocampus and entorhinal cortex in mild cognitive impairment and early ad," *Neurobiol Aging*, vol. 25, pp. 303–310, 2004.

[6] X. Li, D. Coyle, L. Maguire, D. Watson, and T. McGinnity, "Gray matter concentration and effective connectivity changes in alzheimer's disease: a longitudinal structural mri study," *Neuroradiology*, vol. 53, pp. 733–748, 2012.

[7] V. Podgorelec, "Analyzing eeg signals with machine learning for diagnosing alzheimer's disease," *Elektronika Ir Elektrotechnika*, vol. 18, pp. 61–64, 2012.

[8] L. Trambaiolli, A. Lorena, F. Fraga, P. Kanda, R. Anghinah, and R. Nitrini, "Improving alzheimer's disease diagnosis with machine learning techniques," *Clinical EEG and Neuroscience*, vol. 18, pp. 160–165, 2011.

[9] M. Rohini and D. Surendran, "Classification of neurodegenerative disease stages using ensemble machine learning classifiers," *Procedia Computer Science,*, vol. 165, pp. 63–73, 2019.

[10] S. Lahmiri and A. Shmuel, "Performance of machine learning methods applied to structural mri and adas cognitive scores in diagnosing alzheimer's disease," *Biomedical Signal Processing and Control*, vol. 52, pp. 414–419, 2019.

[11] G. Singh and L. Samavedham, "Unsupervised learning based feature extraction for differential diagnosis of neurodegenerative diseases: A case study on early-stage diagnosis of parkinson disease," *Journal of Neuroscience Methods*, vol. 256, pp. 30–40, 2015.

[12] M. Miguel, V. C. J. Francisco, T.-E. Amadeo, and M.-B. Luis, "Imaging biomarkers for the diagnosis and prognosis of neurodegenerative diseases. the example of amyotrophic lateral sclerosis," *Frontiers in Neuroscience*, vol. 12, p. 784, 2018. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2018.00784, 10.3389/fnins.2018.00784,1662-453X

[13] M. ming Xu, M. tian Zhou, S. wei Li, X. chu Zhen, and S. Yang, "Glycoproteins as diagnostic and prognostic biomarkers for neurodegenerative diseases: A glycoproteomic approach," *Journal of NeuroScience Research*, vol. 99, p. 784, 2021.

[14] C. Lin, S. Chiu, T. Chen, J. Jang, and M. Chiu, "Classifications of neurodegenerative disorders using a multiplex blood biomarkers-based machine learning model," *Int J Mol Sc*, vol. 18, 2020.

[15] C.S.Eke, E. Jammeh, X. Li, C. Carroll, S. Pearson, and E. Ifeachor, "Identification of optimum panel of blood-based biomarkers for alzheimer's disease diagnosis using machine learning,"," *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC*, vol. 2020, pp. 3991–3994.

[16] J. Ren, B. Zhang, D. Wei, and Z. Zhang, "Identification of methylated gene biomarkers in patients with alzheimer's disease based on machine learning," *BioMed Research International*, vol. 2020, 2020.

[17] L. Scheubert, M. Luštrek, Schmidt, and R. et al, "Tissue-based alzheimer gene expression markers–comparison of multiple machine learning approaches and investigation of redundancy in small biomarker sets." *BMC Bioinformatics*, vol. 13, p. 266, 2012.

[18] D. Castillo-Barnes, L. Su, J. Ramírez, D. Salas-Gonzalez, F. J. Martinez-Murcia, I. A. Illan, F. Segovia, C. C. Andres Ortiz, M. R. Farlow, C. Xiong, N. R. Graff-Radford, P. R. Schofield, C. L. Masters, S. Salloway, M. Jucker, H. Mori, J. Levin,

and J. M. Gorriz, "Dominantly inherited alzheimer network (dian), autosomal dominantly inherited alzheimer disease: Analysis of genetic subgroups by machine learning, information fusion," *BMC Bioinformatics*, vol. 58, pp. 153–167, 2020.

[19] S. Basaia, F. Agosta, L. Wagner, E. Canu, R. S. Giuseppe Magnani, and M. Filippi, "Automated classification of alzheimer's disease and mild cognitive impairment using a single mri and deep neural networks," *NeuroImage: Clinical*, vol. 21, 2019.

[20] K. Bäckström, M. Nazari, I. Y. Gu, and A. S. Jakola, "An efficient 3d deep convolutional network for alzheimer's disease diagnosis using mr images," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, vol. 302, pp. 149–153, 2018.

[21] S. Esmaeilzadeh, Y. Yang, and E. Adeli, "End-to-end parkinson disease diagnosis using brain mr-images by 3d-cnn," 2018.

[22] J. Dakka, P. Bashivan, M. Gheiratmand, I. Rish, S. Jha, and R. Greiner, "Learning neural markers of schizophrenia disorder using recurrent neural networks," 2017.

[23] D. D. Nicola Amoroso, A. Fanizzi, M. L. Rocca, A. Monaco, A. Lombardi, C. Guaragnella, R. Bellotti, and S. Tangaro, "Deep learning reveals alzheimer's disease onset in mci subjects: Results from an international challenge," *Journal of Neuroscience Methods*, vol. 302, pp. 3–9, 2018.

[24] S. Han, W. Huang, Y. Zhang, and et al., "Recognition of early-onset schizophrenia using deep-learning method," *Appl Inform*, vol. 4, p. 16, 2017.

[25] J. Kim, V. D. Calhoun, E. Shim, and J.-H. Lee, "Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia," *NeuroImage*, vol. 124, pp. 127–146, 2016.

[26] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.

[27] F. R. Bach and M. I. Jordan, "Kernel independent component analysis," *Journal of Machine Learning Research*, vol. 3, pp. 1–48, 2002.

[28] D. R. Hardoon, S. R. Szedmak, and J. R. Shawe-taylor, "Canonical correlation analysis: An overview with application to learning methods," *Neural Comput*, vol. 16, pp. 2639–2664, 2004.

[29] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[30] Y. Jia, M. Salzmann, and T. Darrell, "Factorized latent spaces with structured sparsity," *NIPS*, pp. 982–990, 2010.

[31] T. Cao, V. Jojic, S. Modla, K. C. D. Powell, and M. Niethammer, "Robust multimodal dictionary learning,," *MICCAI*, vol. 70, pp. 259–266, 2013.

[32] E. P. Xing, R. Yan, and A. G. Hauptmann, "Mining associated text and images with dual-wing harmoniums," *UAI*, vol. 70, pp. 633–641.

[33] L. Pang and C. Ngo, "Mutlimodal learning with deep boltzmann machine for emotion prediction in user generated videos," *ACM ICMR*, vol. 70, pp. 619–622, 2015.

[34] X. Zhang, Y. Yang, T. Li, Y. Zhang, H. Wang, and H. Fujita, "Cmc: A consensus multi-view clustering model for predicting alzheimer's disease progression," *Computer Methods and Programs in Biomedicine*, vol. 199, 2021.

[35] M. Liu, D. Zhang, E. Adeli, and D. Shen, "Inherent structure-based multiview learning with multitemplate feature representation for alzheimer's disease diagnosis," *IEEE Transactions on Biomedical Engineering*, vol. 63, pp. 1473–1482.

[36] Z. Xu, S. Zhe, and P. Y. Yuan Qi, "Association discovery and diagnosis of alzheimer's disease with bayesian multiview learning," *Journal of Artificial Intelligence*, vol. 56, 2016.

[37] M. Liu, J. Zhang, P.-T. Yap, and D. Shen, "View-aligned hypergraph learning for alzheimer's disease diagnosis with incomplete multi-modality data," *Journal of Artificial Intelligence*, vol. 36, 2017.

[38] P. S. Aisen and et al, "Clinical core of the alzheimer's disease neuroimaging initiative: progress and plans."," *Alzheimer's dementia : the journal of the Alzheimer's Association*, vol. 6, pp. 239–246, 2010.

[39] I. B. Malone and et al., "Miriad–public release of a multiple time point alzheimer's mr imaging dataset," *NeuroImage*, vol. 70, pp. 33–36, 2013.

[40] I. Itzcovich, "deepbrain 0.1," Accessed on 25 Apr 2021. [Online]. Available: https://pypi.org/project/deepbrain/

[41] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *ACM ICMR*, 2004.

[42] J. MacQueen, *Some methods for classification and analysis of multivariate observations*, 1st ed. Univ. of Calif. Press, 1967.

[43] Steinhaus, "Sur la division des corps mat eriels en parties," *Bulletin de l'Académie Polonaise des Sciences*, vol. 5, pp. 801–804, 1956.

[44] S. TAFASCA, "Multi-view image classification," *Hackathon hosted by Daimler-Benz*.