Multi Modal Detection of Alzheimer's Disease Using Structural MRI Images

Yusuf Suleiman Tahir

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Electrical and Electronic Engineering

Eastern Mediterranean University
August 2019
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Acting Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Prof. Dr. Hasan Demirel
Chair, Department of Electrical and
Electronic Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Prof. Dr. Hasan Demirel
Supervisor

Examining Committee

1. Prof. Dr. Hasan Demirel

2. Assoc. Prof. Dr. Önsen Toygar

3. Asst. Prof. Dr. Kamil Yurtkan

ABSTRACT

Alzheimer's disease (AD) is the most prevalent case of dementia and a progressive brain disorder. It is an irreversible neurodegenerative disease characterized by a decrease in cognitive and memory functions that are ultimately sufficiently severe to interfere with ordinary operations. The disease has no cure, but the related symptoms are managed by many therapy alternatives. While current treatments cannot prevent Alzheimer's progression, early detection can momentarily assist slow down the deterioration of dementia symptoms and enhance the quality of life for people with Alzheimer's disease and their caregivers. With clinical and neuroimaging data, attempts were produced to use multiple classical machine learning algorithms to automatically diagnose this disease. More recently, in-depth learning methods have been introduced for this purpose due to their superior efficacy. In this thesis, we suggest the use of the brain's structural magnetic resonance image (sMRI), acquired from the ADNI database, to construct a model for detecting Alzheimer's disease based on a profound convolutional neural network (CNN) ensemble. The proposed method relies on initial categorization of the brain data into three different categories: the full brain, the grey matter and the white matter. Three separate CNN models are built for each of these data types based on a pretrained network called VGGNet. After training, the decision (i.e Alzheimer’s patient or Healthy Control) from each of the models are then combined using a simple majority vote to obtain a single, final decision. This approach will give better and more accurate predictions than the use of a single model.

Keywords: Alzheimer’s disease, Structural MRI, CNN, VGGNet, machine learning, Deep learning, ensemble network,

ÖZ

Alzheimer hastalığı (AD), en sık görülen demans ve ilerleyici beyin hastalığı vakasıdır. Normal operasyonlara müdahale edebilecek kadar şiddetli bilişsel ve hafıza fonksiyonlarında azalma ile karakterize geri dönüşümsüz bir nörodejeneratif hastalıktır. Hastalığın tedavisi yoktur, ancak ilgili semptomlar birçok terapi alternatifi tarafından yönetilir. Mevcut tedaviler Alzheimer'ın ilerlemesini engelleyemezken, erken teşhis, demansın bozulmasını yavaşlatmaya hemen yardımcı olabilir ve Alzheimer hastalığı olan kişiler ve bakıcıları için yaşam kalitesini iyileştirebilir. Klinik ve nörogörüntüleme verileriyle, bu hastalığı otomatik olarak teşhis etmek için çoklu klasik makine öğrenme algoritmaları kullanmaya çalışılmaktadır. Daha yakın zamanlarda, bu amaç için üstün etkinliklerinden dolayı derin öğrenme yöntemleri geliştirilmiştir. Bu çalışmada, evrişimli sinir ağı (CNN) tabanlı Alzheimer hastalığını saptamak için bir model oluşturmak üzere ADNI veri tabanından elde edilen beynin yapısal manyetik rezonans görüntülerinin (sMRI) kullanılması önerilmektedir. Önerilen yöntem, beyin verilerinin üç farklı maddede sınıflandırılmasına dayanmaktadır: tam beyin, gri madde ve beyaz madde. Bu çerçevede VGGNet adlı önceden belirlenmiş bir ağa dayanarak bu veri türlerinin her biri için üç ayrı CNN modeli oluşturulmuştur. Eğitimden sonra, modellerin her birinden alınan karar (örneğin AD veya HC), daha sonra tek bir nihai karar almak için basit çoğunluk oyu kullanılarak birleştirilmiştir. Bu yaklaşım, tek bir modelin kullanılmasından daha iyi ve daha doğru tahminlerde bulunmaktadır.

Anahtar Kelimeler: Alzheimer hastalığı, Yapısal MRG, CNN, VGGNet, makine öğrenmesi, Derin öğrenme, topluluk ağı.

DEDICATION

This work is dedicated to Mr. and Mrs. Tahir, my parents, for eternal assistance during my moment at EMU.

ACKNOWLEDGMENT

I wish to express my profound gratitude to Almighty Allah, the most Gracious and the ever Merciful. I would also like to thank my parents, for their unrestrained support in every situation throughout my whole life.

A Special thanks go to my mentor, my teacher, big brother and friend **Mubarak Abdu-Aguye** for watching over me throughout this journey. Words cannot express how grateful I am for his help.

I also wish to thank Prof. Dr. Hasan Demirel, my thesis supervisor for his extensive coaching and guidance, my colleagues and staff of the department of Electrical and Electronic Engineering.

My appreciation goes to the entire members of my family both real and adopted: Idris Tahir, Hafsat Tahir, Maimunatu Tahir Suleiman, Ibrahim Isah, Muhammad Abdu Usman and Sagir Lawan Ahmad for always being supportive and encouraging in all my endeavors.

Finally I would like to thank all my lecturers who have taught me more than Electrical Engineering. While my journey only starts now, I will not forget what I have learnt from you all. Thank you all for your patience and guidance.

God bless you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

2D	Two Dimensional
3D	Three Dimensional
AD	Alzheimer’s disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
AE	Auto Encoders
ACC	Accuracy
BET	Brain Extraction Tool
CAD	Computer aided diagnosis
CDR	Clinical dementia rating
CNN	Convolutional Neural Network
CSF	Cerebrospinal Fluid
CT	X-ray computed Tomography
FMRIB	Functional Magnetic Resonance Imaging of the Brain.
FN	False Negative
FP	False Positive
FSL	FMRIB Library software
GM	Grey Matter
HC	Healthy Control
ILVSRC	Image Large Scale Visual Recognition Challenge
KNN	k-nearest-neighbor
MEG	Magneto Encephalography
MMSE	Mini-mental state exams
MRI	Magnetic resonance imaging

PET	Positron emission tomography
ReLu	Rectifier Linear Unit
ROI	Region of Interest
SEN	Sensitivity
SMRI	Structural Magnetic Resonance Imaging
SPECT	Single Photon Emission Computed Tomography
SPE	Specificity
SPM	Statistical Parameter Mapping
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
VGGNet	Visual Geometry Group Network
WM	White Matter

Chapter 1

INTRODUCTION

1.1 Introduction

Alzheimer’s disease (AD) is a progressive brain disorder and the most common case of dementia. It is an irreversible neurodegenerative disease characterized by the decline of cognitive and memory functions, which are eventually serious enough to interfere with normal activities. Although AD mainly affects elderly individuals over the age of 65, younger people may also be susceptible to it. According to the Alzheimer’s Association, approximately 200,000 Americans under the age of 65 have younger-onset Alzheimer’s disease (also known as early-onset Alzheimer’s). In a study, It is anticipated that in every 85 people, 1 will be afflicted by the AD ailment by 2050 [1].

Currently, Alzheimer’s disease has no cure but there are many treatment options for managing the associated symptoms. While current treatments cannot prevent Alzheimer's progression, early detection can momentarily assist slow down the deterioration of dementia symptoms and enhance the quality of life for people with Alzheimer's disease and their caregivers [2]. MMSE (Mini-mental state exams) and MRI (magnetic resonance imaging) of the brain are some of the popular methods of diagnosing and detecting Alzheimer’s disease, both expressed according to a CDR (clinical dementia rating) standard.

Imaging techniques like x-ray, ultrasound and MRI used in the medical field contains a great deal of information which helps radiologists and other medical professionals to carefully analyze and evaluate the data extensively, for investigative or diagnostic purposes. In most cases the proper interpretation of such images requires long and intensive training of health care professionals, which is time and resource intensive. Additionally, the process of understanding or inferring from such images is still quite prone to human error, due to the complexity of the images themselves. Computer Aided Diagnostics (CAD) have thus been introduced with a view to automating the analysis and interpretation of such images using computers. This has seen serious adoption in the diagnostic processes of many diseases [3] [4] [5] due to the high efficacy of such methods, as well as the savings in cost and resources obtained through their use. Research on MRI scans with machine learning algorithms has been shown to be more precise in the classification of AD patients from HC (Healthy Control) making it a very interesting field of computer-assisted diagnosis. (CAD) [6].

In this work, we build on this paradigm. We propose the use of structural MRI brain images obtained from the public domain ADNI database to build a model for the detection of Alzheimer's disease based on an ensemble of deep convolutional neural networks. The proposed method relies on an initial categorization of the brain data into three different categories: the full brain data, the gray matter and the white matter. Three separate models are built for each of these data types based on a pre-trained CNN called VGGNet [7]. After training, the decisions (i.e AD or Healthy Control) from each of the models are then combined using a simple majority vote to obtain a single, final decision. We believe that this multimodal approach will give better and more accurate predictions than the use of a single model.

1.2 Thesis objectives

Within the boundary of work, the use of structural MRI brain data is proposed for detecting AD. The primary goals are:

- Employing preprocessing techniques like skull tripping in FSL, to remove the bone part of the brain scan and segmentation in Matlab (SPM) to extract the gray matter and white matter separately.
- Adapting a convolutional neural network architecture (VGG16) to build models based on the three types of data; full-brain, gray matter and white matter for AD diagnosis problem through transfer learning.
- Using ensemble method, refers to the practice of combining predictions from multiple statistical models to form one final prediction. In this work, the predictions from the three models are fused to form the final prediction.
- Compare the outcomes produced with the alternative outcomes of the other available state-of-the-art techniques.

1.3 Thesis contributions

Structural MRI images of the brain are used to build a multi-model network for AD diagnosis. This involves the adaptation of a popular CNN architecture (VGG16) through the use of transfer learning. This involves three main steps. The first step is the image preprocessing, which involves removing the bone part of the brain from the scan (skull tripping) using a software developed by Analysis Group of FMRIB, Oxford, UK Called FSL [8], segment of the gray matter and white matter in the brain image using the SPM software in MATLAB [9], converting all the 3D brain scans to 2D slices. The second step involves selecting the most informative slices from a scan for training, this is done by calculating the entropy of the slices in each scan. The last step involves independently fine-tuning a pre-trained CNN architecture called

VGG16 on three different types of image data; the full-brain slices, the gray-matter slices and the white-matter slices to get optimal results for AD detection problem from each data type. Subsequently, the decisions from the three models are combined to form a more accurate and robust multi-model network. We evaluate the proposed technique on a publicly-available dataset from the ADNI [2] and Show that the findings obtained are comparable to the state-of - the-art literature.

1.4 Thesis overview

Chapter 2 presents theoretical background and literature review of latest AD diagnosis research. Chapter 3 provides the methodology adopted in this work, which includes the acquisition of images and pre-processing. It also includes information of the techniques used to evaluate the efficiency of the classification. In this job, Chapter 4 offers the suggested structure in detail. Chapter 5 analyzes the efficiency of the suggested technique and discusses the comparison between the outcomes and other methods. On the grounds of assessment and debate, Chapter 6 provides the thesis findings and also discuss the contributions made. It also involves scope for study enhancement and future research.

Chapter 2

BACKGROUND AND LITERATURE REVIEW

2.1 Introduction

Early diagnosis of AD is very important for improving the effectiveness of any treatment and can also assist to understand the root of the illness system (biomarkers) for detection and tracking purposes [5]. Much effort has been directed towards the design of automatic computer-aided diagnostic techniques to distinguish AD subjects from healthy ones based on Neuro-imaging data. this is because traditional methods of diagnostic procedures are very time-consuming (i.e., completing a full evaluation usually takes a few weeks) [6] and are highly dependent on expertise of the physician, which can lead to erroneous diagnosis. In the following sections, we discuss the standard pipeline employed in detecting Alzheimer's disease with image based computer aided diagnosis.

The pipeline starts with preprocessing techniques which are performed on the image data. these involves removing artifacts introduced by imaging devices in the process of acquiring them, correcting pixel intensities , removing parts of scan that are not relevant to the problem, segmentation of different brain tissue classes which are used analyzing the disease. The preprocessing step is necessary to achieving the overall goal. Machine learning algorithms, both classical techniques and the recent deep learning methods are then used on the preprocessed data. in the classical methods, features from these images are first extracted, which are then used as inputs

to a classifier (like SVM, Random forest, logistic regression etc.) while deep learning methods learn these features directly from the data using successive layers of increasingly meaningful representations.

2.2 Pre-processing

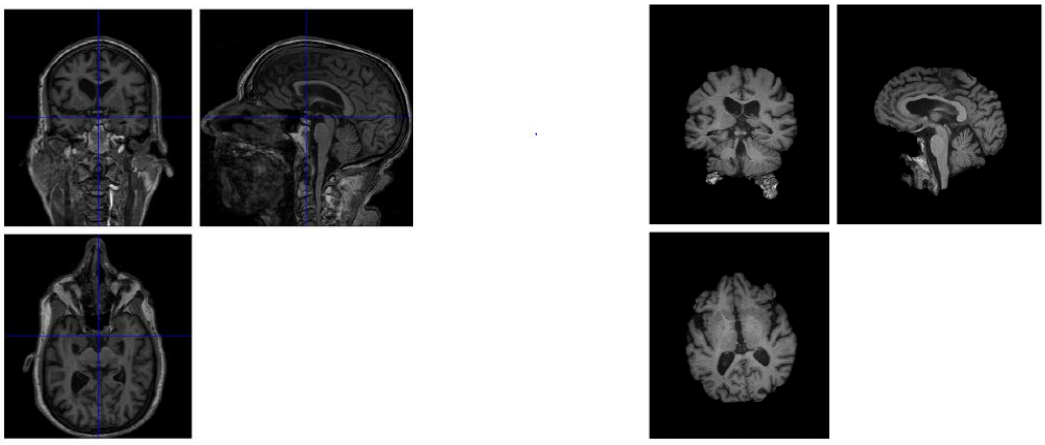
Data preprocessing is a required step to achieve meaningful results in machine learning. Neuroimaging data are typically volumetric images (3D) and due to the different nature of acquisition, some modalities requires special pre-processing techniques (e.g. intensity normalization, de-noising, bias-field correction, modulation, etc.) to remove introduced artifacts that are not originally part of these images.

In this thesis, our preprocessing steps are particular to those required for MR images. Specifically, we perform skull tripping to remove the bone part of the scans, segmentation to segment the different tissues of the brain (gray matter, white matter, CSF) and then flattening to convert these scans to 2D slices to make them suitable for use with the deep CNN architecture used in this work.

2.2.1 Skull tripping

Generally in brain images, images of brain itself are of the most importance and not the tissues surrounding it e.g the skull, fat or skin surrounding the brain. This is because these external tissues can complicate the learning procedure and negatively affect classification, segmentation, or regression tasks. To eliminate these tissues from the brain images used, a tool called Brain Extraction Tool (BET) in the FSL software was used.

In [10], the series of process BET tool used to achieve the overall brain extraction is explained. In the first step, the intensity histogram is processed to get the upper and lower intensity values for the image and a rough brain/non-brain threshold. Next a rough size of the entire head and the Centre of the head in the image is found. Then collocation of a sphere's surface in a triangular manner is initialized inside the brain which is allowed to slowly deform, attempting to move toward the brain's edge while following forces that keep the surface well-spaced and smooth.



(a) (b)

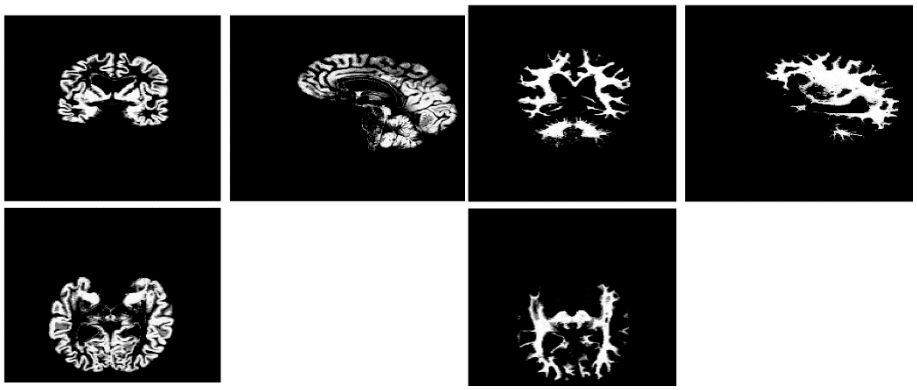
Figure 2.1: MRI brain scan (a) with skull (b) without skull.

2.2.2 Segmentation

Significant changes occur in the brain tissues (gray-matter, white-matter and Cerebrospinal fluid) of patients suffering from Alzheimer's disease. Separating these tissues and investigating each separately can improve the accuracy of detecting the disease. Segmentation or tissue classification is the process employed in separating these tissues. It can be achieved mainly in two ways: one is a tissue classification approach where voxels (volume element representing a value in the three dimensional space, corresponding to a pixel for a given slice thickness) are assigned

to a tissue class according to their intensities and by registration to a template where a brain template is warped to match the brain volume to be segmented [11].

The segmentation is achieved using ‘segment’ function in statistical parameter mapping (SPM) software version. This function is a one click method that encompasses spatial normalization, bias field correction preprocessing techniques and tissue segmentation together. The algorithms of the function as explained in [11] unifies the two forms of brain segmentation, tissue classification approach and registration with a template into a single probabilistic framework.



(a) (b)

Figure 2.2: (a) Grey matter segment (b) White matter segment of the brain.

2.2.3 Flattening / conversion of 3d scan to 2d slices

Brain MR Images are captured in a 3D volumes of either a DICOM or NIFTI file format. These volumes can be directly used with a 3D based deep learning network, but to make use of the readily available popular architectures of deep learning that proves more efficient and accurate in learning process, these volumes have to be converted to 2D image slices.

Med2image is one of the several utilities built for this purpose, it is a simple Python3 command line based program that converts medical image formatted files to 2D image format (JPEG, TIFF, PNG, etc.).

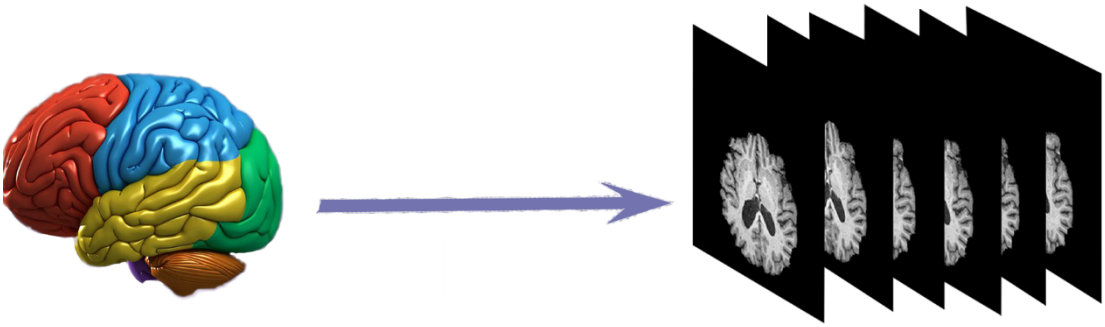


Figure 2.3: Conversion of 3D to 2D image slices.

2.2.4 Slices selection based on entropy

There is a large number of 2D slices in a 3D MRI scan. Most of these slices do not contain enough information about the disease and training a network on all the slices will lead to poor generalizability of the resulting model. In most recent studies, the slices for training are extracted randomly (cite, cite, cite). We adapt a method used by [12] in our work to extract the most informative slices in each scan for training the network. To achieve this the image entropy for each slice is calculated.

In general, the entropy or average information of an image can be determined approximately from the histogram of the image. As set forth in [11], Taking into account a group of m symbols with p1, p2, ……, Pm, probabilities, the following formula is used to get each slice's value for entropy:

$$H = - \sum_{i=0}^{m} P_i log_2 P_i \quad (1)$$

This can also be gotten likewise with the use of histogram plot for a single slice [11]. Entropy is a measure of a slice's variety. Therefore, by sorting the image slices in descending order with respect to their entropy value, those with the highest values are considered to be the ones with most information.

2.3 Feature extraction

Feature Extraction is a technique used to extract distinctive features from an image in such a way that those features represents the information in that image in a highly-descriptive and usually lower-dimensional form. These features can be local and or global characteristics of the image such as edges, entropy, color, shapes, regions with similar properties and any combination of these [12]. Feature fusion is another method that combines distinct important characteristics to create a more robust description / descriptor of a picture.

In image processing, there are different methods and algorithms for extraction and fusion of features. The usual pipeline used for brain images involves skull tripping and segmentation, then selection of region of interest (ROIs) and computation of features from each of these ROIs [13] [14] [5] .

These features then serve as inputs to various type of machine learning algorithms (e.g. SVM, Naïve Bayes, Logistic Regression, kNN, Random Forest, etc.) for classification or diagnostic purpose.

2.4 Machine learning

Machine learning is a subfield in artificial intelligence in which algorithms are built to enable systems learn on their own. The basic principle of machine learning is to build algorithms that can receive input data, learn patterns and features from the data and predict an output while updating outputs as new data becomes available without being explicitly programmed. Machine learning are basically classified into three types of algorithms: supervised learning, unsupervised learning and Reinforcement learning [15].

In supervised learning, the algorithm train on data set containing training examples associated with correct labels which then facilitate its ability to reach an accurate conclusion when given a set of new data. In contrast, data are not labeled in unsupervised learning. The Algorithm must find patterns and relationships given unlabeled and uncategorized data without any prior training. In reinforcement learning the system is exposed to an environment where it trains itself continually using trial and error to make specific decisions.it learns from the past experience and try to make the best possible predictions [13] [14] [16][16].

Some of the various application of Machine learning in our daily activities include image recognition, speech recognition, prediction, medical diagnosis etc.

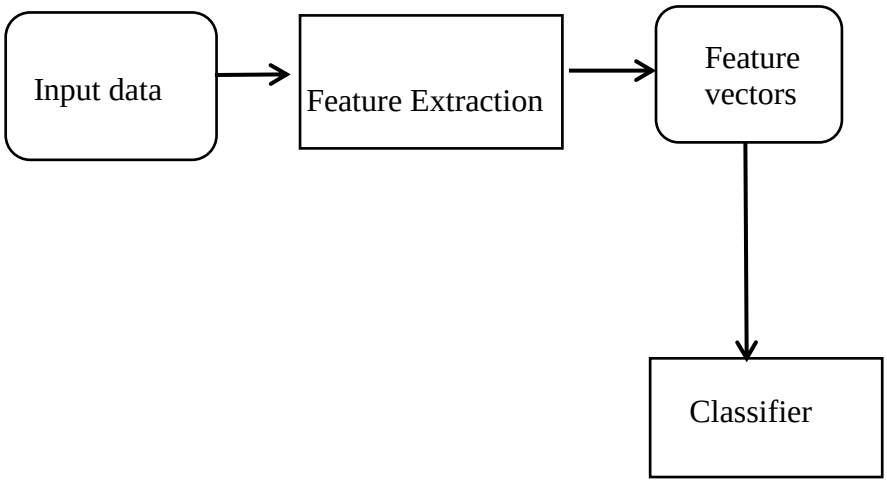


Figure 2.4: An overview of machine learning process.

2.5 Deep learning

In classical machine learning methods, a good representation of input information is essential, i.e. an excellent set of characteristics which will generate a feature vector that can be used by algorithms as a bias for learning. The primary concept of deep learning is to fix this issue by learning it straight from the information [17]. Deep learning is a field in machine learning that involves learning representations hierarchically from input data using successive layers of increasingly meaningful representations. Modern deep learning method often contains a lot of layers of representation which all learned automatically from the exposure to training data. These layered representations are always learned via models called neural networks, which are structured in layers stacked on top of each other [18].

Deep learning methods steadily learn more about images as it goes through each layer of the neural network. Early layers learn low-level features like curves, lines and edges, subsequent layers combine features from earlier layers into a more meaningful representation. While deep learning algorithms prove more powerful in

dealing with images with their ability to learn features automatically from these images, it can lack the ability to generalize on a new data for less complex problems because they require a large amount of data to be effective.

While the most popular among deep learning methods are the convolutional neural networks, others such as Generative Adversarial Networks (GANs), AutoEncoders (AE) are also very effective in solving various types of machines learning problems.

2.5.1 Convolutional neural networks

Convolutional Neural Networks (CNNs) are popular deep learning models used mostly to solve computer vision problems like image classification and object detection. The main construction blocks of CNNs are layers of convolution, layers of pooling (downsampling), features of activation and fully connected layers [17]. These network input data can be 1-dimensional, 2- dimensional or even 3-dimensional data.

The convolution layer consists of m x m filters to be applied to the entire input picture matrix in a way called convolution. Convolution in this context, refer to a linear operation that involves multiplication of the filter weights with image matrix in a special manner, each weight is a model parameter to learn. After the convolution operation; each filter produce an affine transformation of the input known as feature map which serves as input to the next layer.

Activation functions are very helpful characteristics of neural artificial networks. These features determine whether or not to activate a neuron. Whether the data received by the neuron is applicable to the data provided or should be overlooked.

$$Y = \text{Activation}\left(\sum(\text{weight} * \text{input}) + \text{bias}\right) \quad (2)$$

The activation function is the nonlinear transformation that we do over the input pixels. This transformed output is then sent to the next layer of neurons as input. Some of the popular activation function used in convolutional neural networks includes: the sigmoid function; the rectified linear unit (ReLu) function and the softmax function.

Pooling layer is a layer usually added after the convolutional layer. The purpose of the layer is to down sample each of the resulting feature maps from the previous convolutional layer to create a new set of the same number of pooled feature maps their by reducing the spatial dimensionality of the data. Average pooling and max pooling are the two main functions used in the pooling operations [18]. The reduction in the spatial dimensionality reduces the amount of training data needed, and also gives the network some measure of translation invariance.

Fully- connected layers comes after many convolutional and pooling layers, in this layer the feature maps from the previous layer are flattened into a single vector with many neurons. This flattened vector goes to fully connected layers for classification.

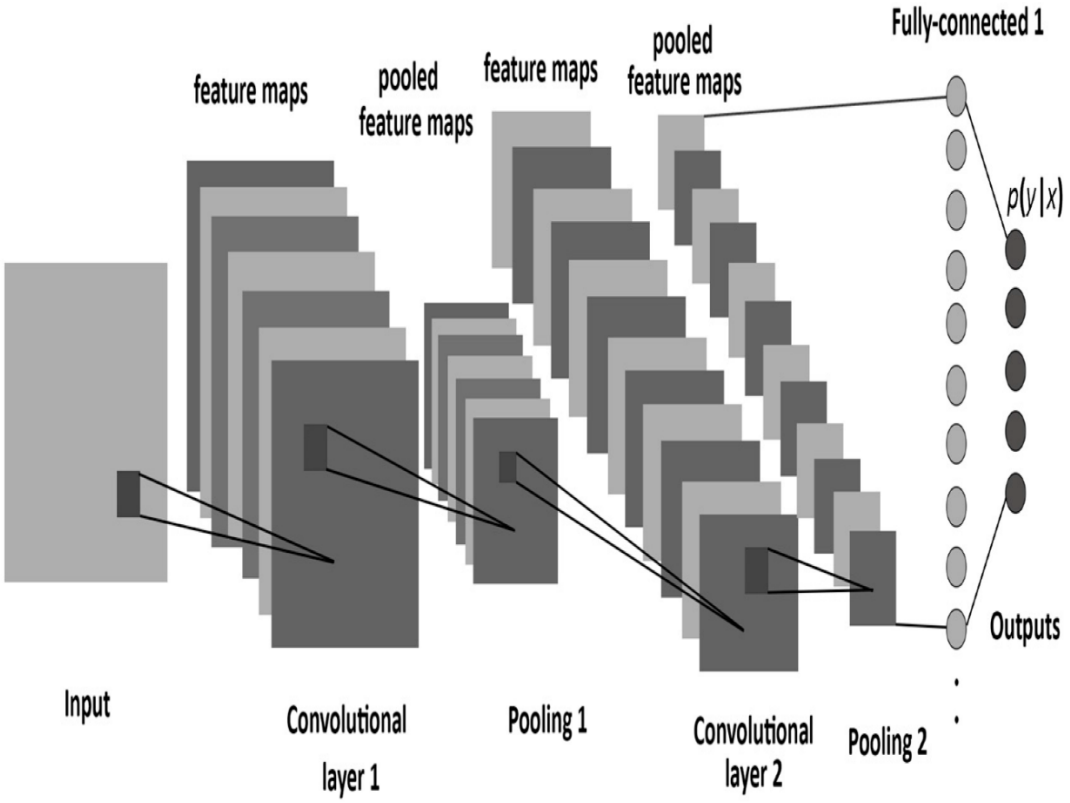


Figure 2.5: An overview of CNN architecture [19].

2.5.2 CNN Architectures

Among the many CNN architectures used in image classification and recognition are: AlexNet (2012), VGG-net, Residual Networks (ResNet), GoogLeNet, etc.

AlexNet was first presented in [20]. It has 60 million parameters, 650,000 neurons and consists of 5 convolution layers, some of which have been followed by max-pooling layers, and 3 fully linked layers with a final softmax of 1000-way. To avoid overfitting, Dropout Regularization is also included. It attained a winning top-5 error rate of 15.3% compared to 26.2% achieved by the second-best ILSVRC-2012 entry in the ILSVRC-2012 competition.

In VGG-net, the impact of the convolutional network depth on its precision is studied in the large-scale image recognition environment while keeping a very tiny (3/3) network-wide convolution filter. Pushing the depth on the prior art setting to 16-19 weight layers indicates a substantial rise in precision. In the 2014 imageNet challenge, the VGG-net secured the first and second places in the localization and classification tracks respectively [21].

Residual Networks (ResNet) [22] were proposed to ease the training process of very deep networks, even deeper than those prior art like the VGG-net. Instead of the traditional way, layers are reformulated to learn residual functions with reference to the layer inputs. This method achieved 1st position on the ILSVRC 2015 classification task and COCO 2015 competitions.

Another very popular architecture are the GoogLeNet [23] and Inception [25], the GoogLeNet presented its architecture based on modules called inception modules, although VGG-net in the 2014 ImageNet challenge have very similar performance to the GoogLeNet, GoogLeNet received more attention due this inception modules and its relatively less computational expensive. The inception modules utilize the idea of very small convolution filters of different size to learn representations which are later concatenated at the end of each module.

2.6 Transfer learning

In most real life applications of machine learning, training data is limited. Training a deep network from scratch can be difficult and quite data intensive. This is due to the large number of model parameters and complexity. However, the models described in the above section which are trained using large dataset like the ImageNet, can be

re-employed for problems different from the original classification problem. This is possible because weights already learned by those models can be useful to other datasets [18]. The process is termed as Transfer Learning.

There are basically two ways to use a pre-trained model: feature extraction and fine-tuning. In feature extraction, representations learned by a network trained on some data previously are used to extract features from a new dataset. These features are then passed to a new classifier, which is trained from scratch. On the other hand, fine-tuning consists of freezing and unfreezing some top layers in the pre-trained network and jointly training these layers with the added layers (classifiers) [24].

2.7 Ensemble learning technique

Models of the neural network are a nonlinear method. They are flexible and are capable of learning in data complex nonlinear relationships. A downside of this flexibility is their sensitivity to original circumstances, both in terms of original random weights and statistical noise in the training dataset. Each time a network is trained, it learns a (slightly) different version of mapping functions from input to network output. This will lead to different predictive performance even on the same dataset. Training multiple models on the same problem and combining the vote of their prediction to get a single strong and accurate prediction is termed ensemble learning [26].

There are different types of ensemble network: voting method, bagging method, the boosting method and the stacking method.as explained in [25], In bagging method, learning occurs independently from each other in parallel and then combined using an averaging process. Learning in boosting method occurs differently, the models

learn sequentially in a very adaptive way, i.e. they are dependent on each other and then combined in a deterministic strategy. In the stacking method, heterogeneous learners are considered, representations are learned in parallel and then combined by training a base model to output predictions based on the different models.

The methods described above operates based on some combination rules which they have built-in, such rules include the majority voting, weighted majority voting, mean rule, sum rule, product rule etc. however , ensemble classifiers built on different on models trained with different subset of training data can be combined using one of the many rules.

2.8 Literature review

2.8.1 Classical methods

Classical machine learning techniques like the Support Vector Machine and feedforward neural networks were effectively implemented to detect AD using various neuroimaging biomarkers such as CT, PET, SPECT, MRI images [26] [14] [15], Magneto encephalography (MEG) and Electroencephalography (EEG). One such approach is discussed in [26], where a multi-modality classification frame-work was proposed and implemented to efficiently exploit the complementary in the multi modal data. First, for each modality, pairwise similarities are separately calculated, then similarities from various modalities are combined into a nonlinear graph fusion method that generates a unified graph for final classification. The technique has accomplished 91.8 percent precision. Also, [14] proposed a multi-modal method that combined three modalities; MRI, FDG-PET and CSF to discriminate between AD and HC (healthy controls), using kernel combination method. The combined kernel is

made by a fusion of features extracted from all the above mentioned modalities and SVM is then used to evaluate the classification accuracy.

In another research, [27] Used a stacked Deep Polynomial Network(S-DPN) algorithm to further improve feature representation and furthermore proposed a two-stage multi-modality S-DPN algorithm together with SVM classifier to generate more discriminate, robust features and final classification accuracy, the proposed algorithm achieved 97.2% accuracy. Similarly [5] in their work used a dual-tree complex wavelet transform for extracting features from an image, these features are then sent to a feed forward neural network to classify AD and HC.

2.8.2 Deep learning methods

Subsequently, deep learning techniques have outperformed the traditional classical methods with a large margin. [28] In their paper proposed a simple approach which achieved a high performance with high classification accuracy, the research uses a key technique of cross domain feature to represent MRI data. Sparse auto-encoders are used to learn a set of basis from natural images and then convolutions is applied to extract features from the Alzheimer's disease Neuroimaging Initiative (ADNI) dataset. This method was further improved by [29], in which a 3D convolutional neural network was used to predict the disease status of a patient using the MRI scan of the brain. This method also achieved a higher accuracy in detecting patients with the disease.

In [30], a deep learning architecture based on stacked auto-encoders was proposed. The network included a softmax layer as the output, the choice of which was to overcome the bottleneck problems that was shown in most of the previous research and aid a better diagnosis of AD and its prodromal stage. Prominent CNN

architectures like LeNet and the first model of Inception were adopted in [30], both structural and resting state functional MRI scans are used to distinguish brains affected by the AD disease from normal healthy brain in older adults. The study presents two robust pipeline, including extensive preprocessing modules and deep learning based classifiers. Features are extracted from the whole brain data using CNN architectures resulting in a highly accurate predictive model.

Using a very small amount of structural MRI images, transfer learning is used in another work [12], to identify AD. Initialized with pre-trained weights from the ImageNet dataset, two common architectures: VGG16 and Inception V4 are fine tuned to achieve more precise outcomes. The authors used a smart entropy-based method in the preprocessing steps to pick the most informative slices for training from a brain scan.

Chapter 3

METHODOLOGY

3.1 Introduction

The chapter deals with the methodology adopted regarding the acquisition of the sMRI data, steps involved in fine-tuning a pre-trained network and model testing methodology.

3.2 Image acquisition

Structural MRI data from the Alzheimer's disease Neuroimaging Initiative (ADNI) database [2] are used in this work. Based on the acquisition plane, ADNI database provides three types of scans: AXIAL, SAGITTAL and CORONAL. Since our target was to simply differentiate between AD and HC patients through the images, we used the AXIAL data. Briefly, the scans are obtained from a 3 Tesla, T1-weighted scanner (Siemens) with 3D acquisition.

3.3 Fine-tuning a pre-trained CNN model

In all the three models, a technique widely adopted when reusing models called fine-tuning is employed. Fine-tuning allows the use of a small dataset on a network with a very large number of parameters and it also requires less computing power during training. To perform fine-tuning in this work, we perform the following steps:

1. In the first step, the original model classifier is replaced with a custom built classifier which is more appropriate to the task at hand.

2. The newly-added, randomly-initialized classifier is then trained for a few epochs on feature vectors from the inputs as extracted by the pre-trained model. In this step it is necessary to freeze the convolutional base of the VGG16 network to prevent the pre-trained weights from updating while training the classifier. This is done to avoid the network-propagating error signal being too large during training and the representations previously learned to be destroyed by the layers being fine-tuned.

3. In this step, the top convolutional block of the pre-trained network is unfreeze and jointly re-train along with the classifier on the input images. The reason for unfreezing only the top block is to avoid overfitting while training on a small dataset and also because the top convolutional blocks encode more specialized features which we need to repurpose to match the problem at hand.

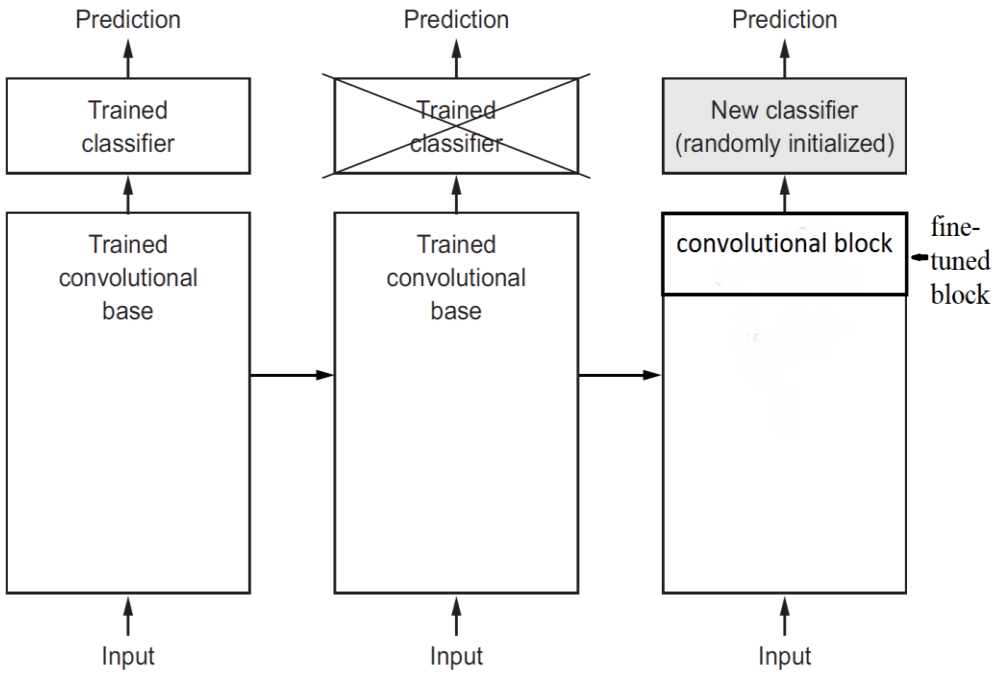


Figure 3.1: Fine tuning a pre-trained CNN network.

3.4 Decision fusion via majority voting

The voting method is an ensemble technique utilized in this work. It consist of models where each is used to make predictions on each test subject. These outputs are then aggregated together to produce an improved single prediction using a majority voting method called hard voting where the final prediction is taken to be as the mode of all the predictions.

$$\hat{y} = \text{mode}\{c_1(x), c_2(x), \ldots\ldots\ldots c_m(x)\} \quad (3)$$

$\hat{y}$ = Final prediction

$c_1(x), c_2(x), \ldots\ldots c_3(x)$ = predictions from classifier

3.5 Model testing methodology

A reliable measurement is achieved by obtaining all performance results using the 5-fold cross validation where all the data for a model are randomly divided into 5 groups, each unique group get to be used as test data while the other groups for training. The training-testing experiments are repeated five times and in each of the experiments, a different set of samples/slides are used. The predictions from each model are probability values between 0 and 1. For this evaluation, model output values between 0-0.5 are considered indicative of AD and values between 0.51-1.0 are considered indicative of HC). These predictions from each model are then fused together using an ensemble technique called majority voting as explained in section 3.5 to produce a more accurate and robust result.

This method is adopted as we believe that our system can benefit from the fusion of the decisions from different modalities (i.e. whole brain, white matter, grey matter) in order to produce a more accurate result than a single (unimodal) model.

3.5.1 Evaluation metrics

In order to evaluate the performance of the system, the performance metrics of interest are derived over the image slices. An image slice is considered to be in a particular class if the prediction output is between some thresholds discussed in previous section. The metrics of interest are the accuracy (ACC), sensitivity (SEN) and specificity (SPE) of the system. These are calculated from a confusion matrix plot. The metrics are defined mathematically as follows:

$$\mathrm{ACC} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{4}$$

$$\mathrm{SEN} = \frac{TP}{TP + FN} \tag{5}$$

$$\mathrm{SPE} = \frac{TN}{TN + FP} \tag{6}$$

Where: TP - true positives, TN - true negatives, FN - false negatives, FP - false positives.

These values are determined by:

TP: number of correctly predicted AD images.

TN: number of correctly predicted HC images.

FN: number of Incorrectly predicted HC images.

FP: number of Incorrectly predicted AD images.

Chapter 4

PROPOSED CNN BASED MULTI-MODAL FRAMEWORK

4.1 Introduction

The chapter deals with the details of the proposed multi-modal framework and how it is implemented. It is divided into five section: which starts with the general system overview, where the preprocessing pipeline, the model training and the decision fusion adopted are introduced. The preprocessing phase where the data preprocessing is explained in details, the dataset obtained after the steps, the model training and the decision fusion phase.

4.2 System overview

The general structure of the proposed methodology is presented in figures below. Figure 4.1 presents the steps employed in preprocessing the images, which involves skull tripping, segmentation of the grey and white matter, conversion of the volumes to 2D image slices and selecting the few most informative slices for training. Figure 4.2 presents the network structure used for the three models each of which relies on fine tuning a pre trained network (taken from the VGG16 model [7]). Each model naturally contains a classification portion made up of a number of feed forward layers. Each model outputs a prediction score between 0 and 1, indicating the probability of the supplied input image belonging to an AD patient. Note that the models operates on images preprocessed using the pipeline shown in Figure 4.1.

Figure 4.3 shows how the decision fusion level is achieved, mode is taken on set of predictions made by each of the three models on a particular image slice.

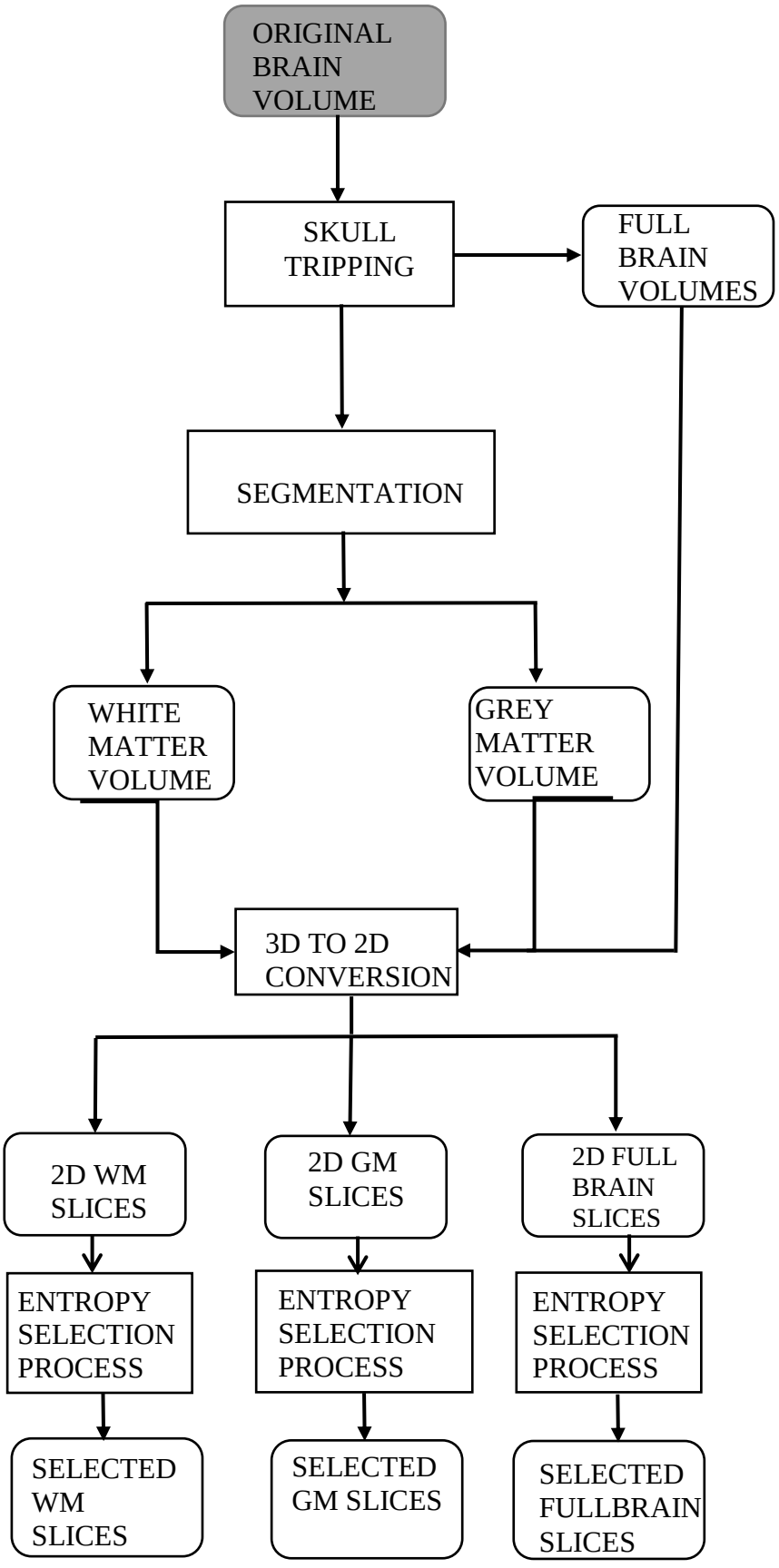


Figure 4.1: The image preprocessing pipeline.

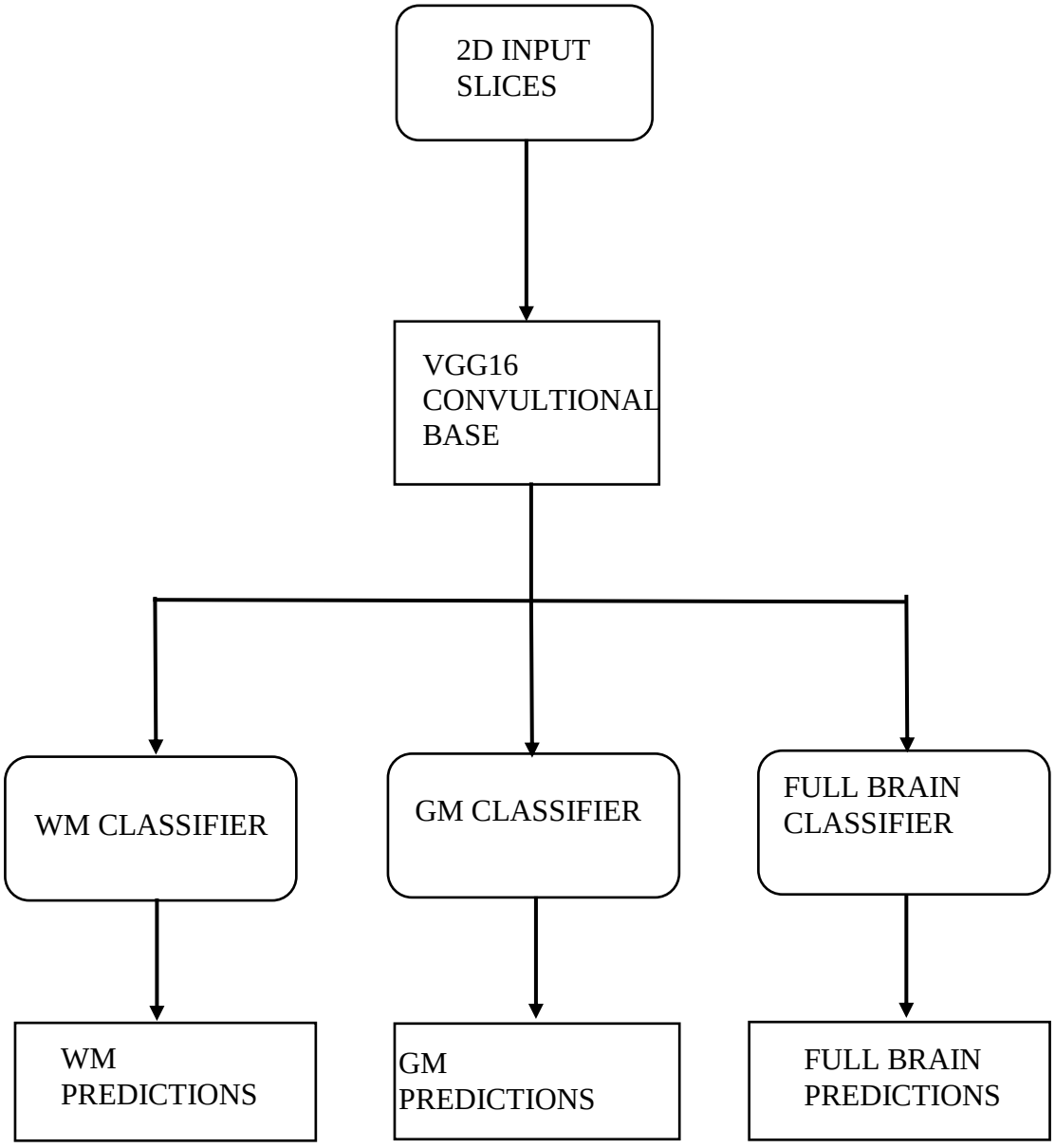


Figure 4.2: Proposed Network Architecture.

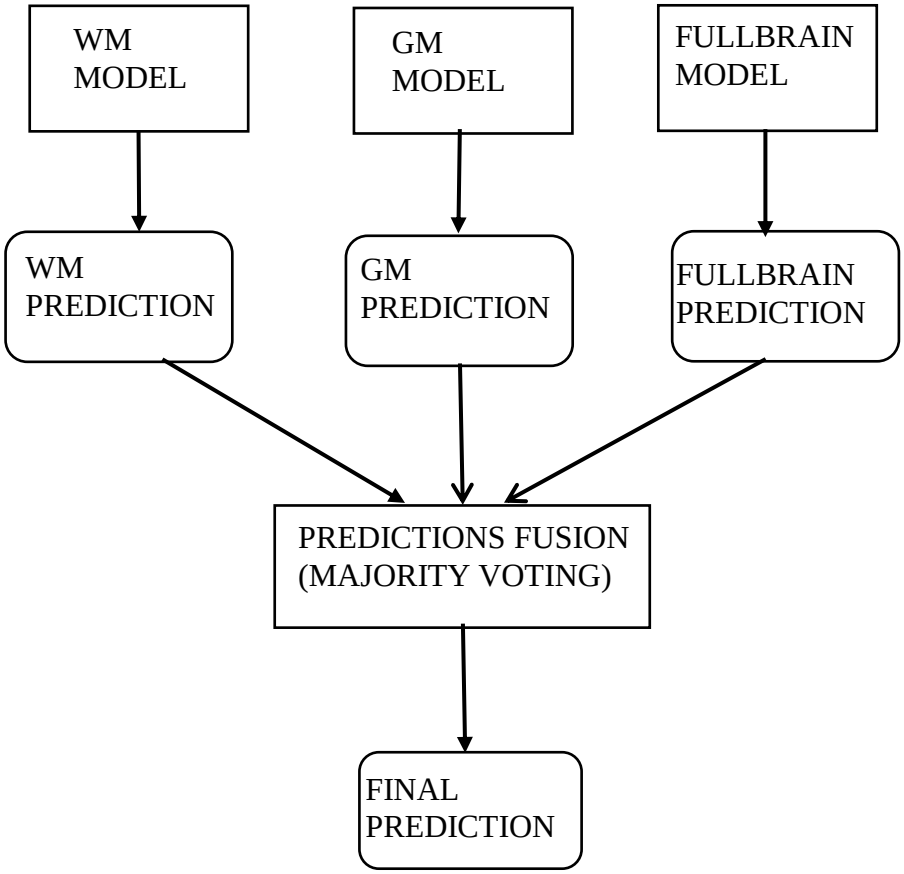


Figure 4.3: Proposed Decision Fusion pipeline.

4.3 Preprocessing phase

The preprocessing of the brain volumes starts with Skull-tripping, it is used to separate the brain from the skull. The BET in the FSL [8] suite was used for this purpose. BET deletes non-brain tissues from the image of the head. The skull tripped images are then duplicated which are all needed in further steps.

Brain tissue segmentation is then carried out on one of the duplicates. For this step, statistical parameter mapping (SPM) is used to separate the grey-matter (GM) and the white-matter (WM) which mostly make up the whole brain. After the

segmentation process, 3 different data points are now available: GM volumes, WM volumes and the other duplicate of the full-brain volumes. These volumes in the next step are then converted to 2D slices using a python utility called Med2Image [31] which is specially written for the task. We converted all the volumes to image slices so as to be able to utilize a pre-trained model used in this work which requires 2D images as input.

In a brain scan, there are many 2D slices and a lot of these slices do not contain information about the disease. As such a selection process is employed in all the slices of the full brain volumes, GM volumes and WM volumes to extract only those useful in this context. To achieve this, the entropy of each slice is calculated and out of the 256 slices in a scan, only 32 are selected from each which are considered to be the most informative. This is done to prevent the network from over fitting and improve its generalization ability.

4.4 Dataset

Our dataset were randomly picked from the ADNI database [2] as described in section 3.2. this consists of 200 scans, 100 of them which are selected from the AD category, the other 100 which are in the HC category all within the age range of 74 ± 8.3 After all the necessary preprocessing steps have been carried out, a balanced dataset is obtained. The details of the dataset are summarized in table 3.1.

Table 4.1: Dataset obtained after the preprocessing steps.

MODALITY	FULL-BRAIN	GREY-MATTER	WHITE-MATTER
AD	3200	3200	3200
HC	3200	3200	3200
TOTAL	6400	6400	6400

4.5 Model training phase

The pre-trained VGG16 model [7] is employed in building all the three proposed networks. It has 16 convolutional layers with very small filters of 3x3 size, 5 max-pooling layers of size 2x2, followed by three fully-connected layers, with the soft-max classifier as the final layer. A rectifier activation function is applied to all hidden layers. The original network also uses dropout regularization in the fully-connected layers to reduce overfitting and improve generalization error. The structure of the VGG16 model used in the ImageNet database is shown in figure 4.3.

The original VGG16 classifier is replaced with new custom built which consists of a flatten layer (which converts the feature maps from the VGG16 model into a y-element feature vector) and two dense layers: the first has 256 neurons with ReLu activation function and the last contains a single neuron with a sigmoid activation function. The new model is then fine-tuned as explained in section 3.3. The structure is also shown in figure 4.4.

Keras Machine Learning library [32] with a Tensorflow backend is used in implementing our system. The pre-trained VGG16 model prepackaged in the Keras

library is adopted. 5-fold cross validation method is used in training all the three models with an 80%-20% splits between training and testing data. The experimentally decided hyper-parameter settings used in all the three models is summarized table 3.2.

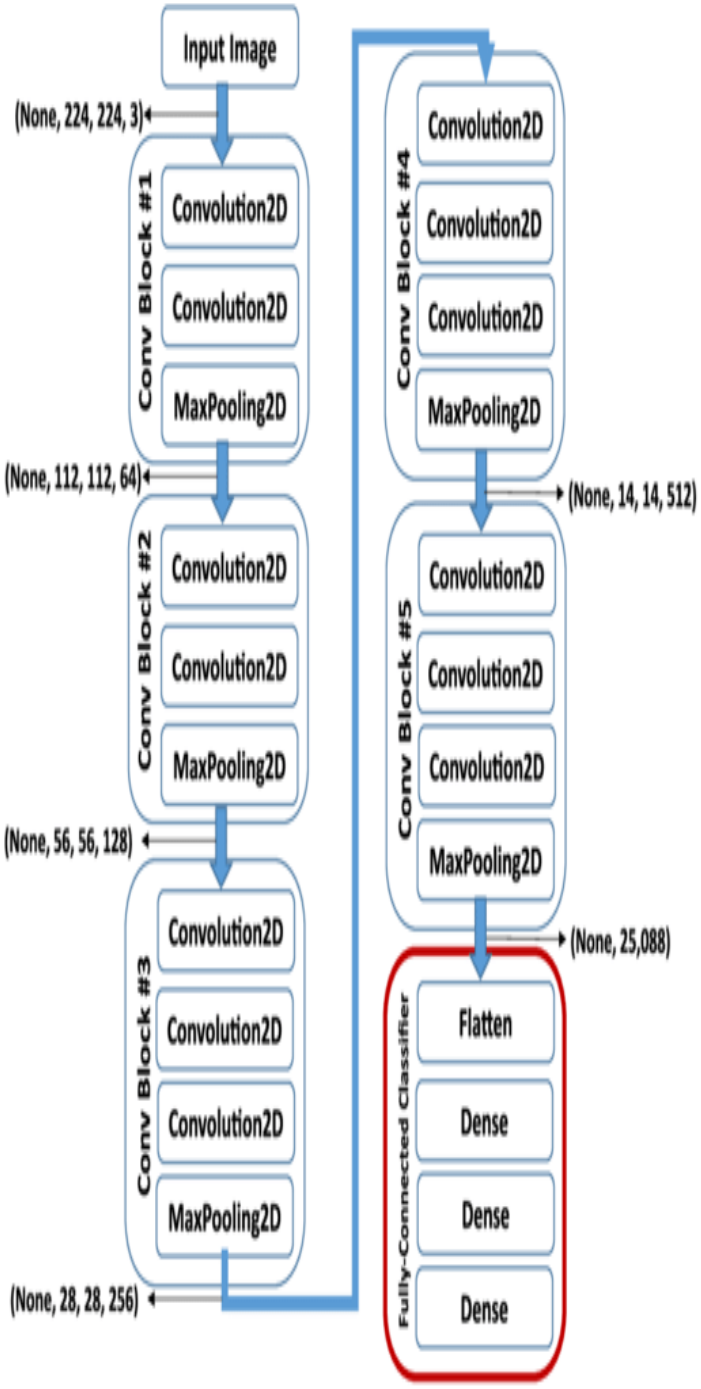


Figure 4.3: VGG16 structure [33].

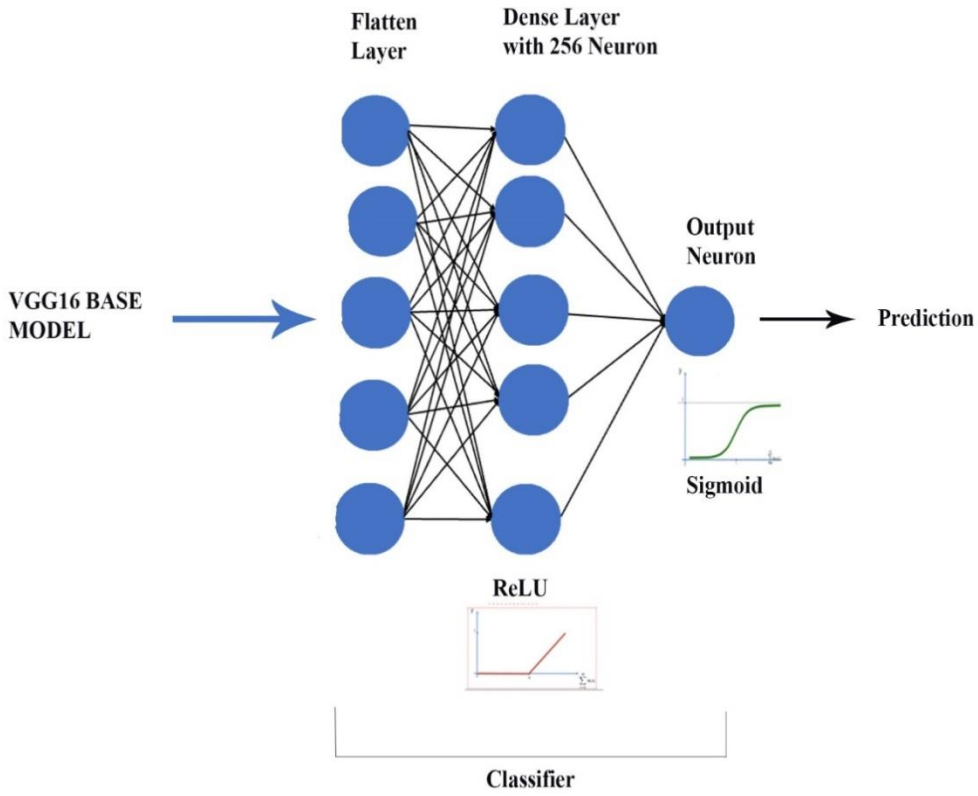


Figure 4.4: The proposed classifier.

Table 4.2: Experimentally decided hyper-parameter settings for all the three models

Hyper-parameter	Full brain model	Grey-matter model	white-matter model
epochs	15	20	28
optimizer	Adam(lr=1e-5)	Adam(lr=1e-5)	Adam(lr=1e-5)
batch size	16	16	16
callbacks	None	None	None

4.6 Decision fusion phase

Predictions made on the test data by each of the three models on a particular image slice pass through the decision fusion method, where the majority vote is taken as the final output for that slice. As explained in section 3.4, The voting method is an ensemble technique. It consist of models where each is used to make predictions on each test subject. These outputs are then aggregated together to produce an improved single prediction using a majority voting method called hard voting where the final prediction is taken to be as the mode of all the predictions.

Chapter 5

RESULTS AND DISCUSSIONS

5.1 Experimental results and discussion

As mentioned in section 3.5, the proposed frame-work is evaluated based on its predictions on test data after training for each fold. This is to comprehensively investigate the effectiveness of the frame-work and get a good estimate of its performance in real-world settings, where the system will be used in diagnosing previously-unseen individuals.

For each model, the testing accuracies obtained after training/testing on all the folds are averaged to give an overall accuracy. A confusion matrix is calculated after each fold which are then summed up to obtain the overall confusion matrix for the model.

Furthermore, the majority voting is done over all the predictions from the three models, considering all testing images, the prediction from each relevant model using the relevant data point are then are then aggregated as explained in section 3.4 to yield a single and more accurate prediction. Confusion matrix for the ensemble method is also calculated. Table 4.1 shows the detailed results for the proposed framework.

Table 5.1: Detailed Results obtained with the proposed framework.

Model	SEN (%)	SPE (%)	ACC (%)
Full brain	98.47	98.96	98.72
Grey matter	96.90	96.43	96.66
White matter	95.50	95.12	95.31
Multi-modal	99.71	99.52	99.62

We further show the confusion matrix based on the performance of each individual model and also the performance of the ensemble network. These are used in calculating the metrics of interest in all the models.

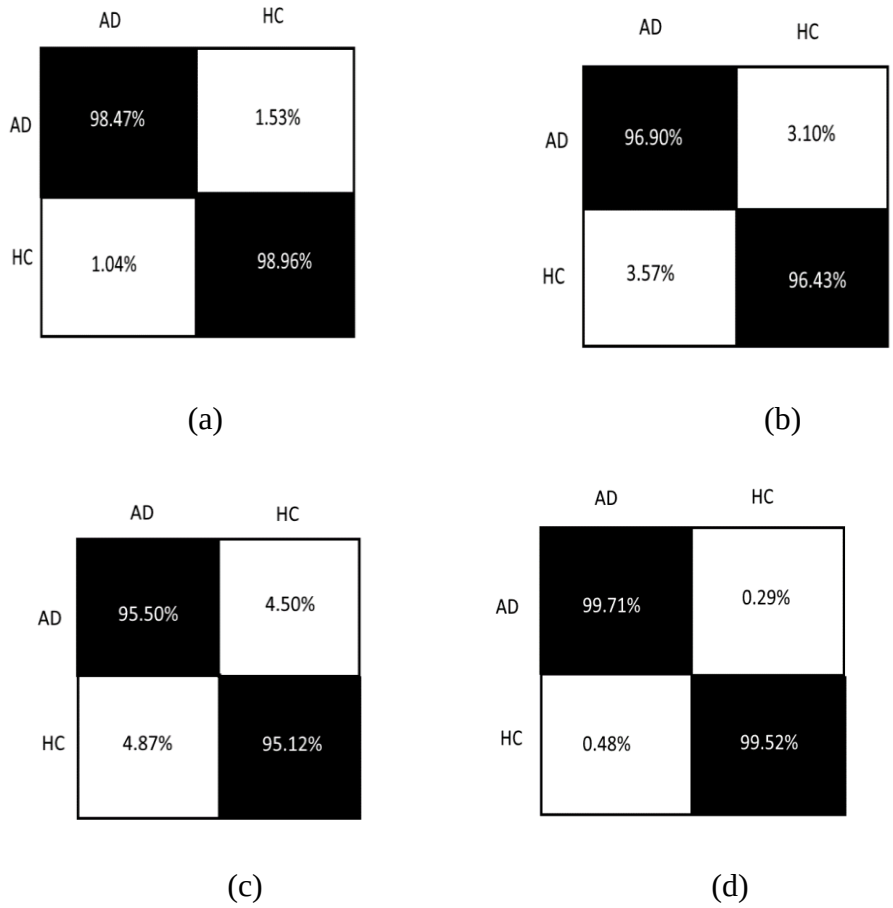


Figure 5.1: Confusion matrix for (a) Full brain model (b) Grey-matter model (c) White-matter model (d) Ensemble network

5.2 Comparison with other methods

A comparison of results obtained in relation to other recent approaches is discussed in this section; it is done both in terms of accuracy, the size of training data and validation type.Although some of the methods might not have used the same datasets and/or experimental configurations as in this work, the results can still be compared, as with all the techniques discussed also used structural MRI images in implementing, Which have a high level of resemblance across datasets, particularly when the images have been preprocessed and the brain has already been recorded and segmented in the published datasets. We compare our findings with the five techniques of deep learning outlined in the chapter 2.

Table: comparison with state-of-the-art methods

Model	Average Accuracies(%)	Training data size	Validation type	Dataset
Auto-encoder + 3D CNN[28]	95.39	2265 volumes	Hold-out	Oasis
3D CNN [29]	97.60	210 volumes	10-fold	ADNI
Stacked Auto-encoders[30]	87.76	21,726 slices	10-fold	ADNI
LeNet,Inception model [31]	94.52	12,675 slices	Hold-out	ADNI
Inception V4 [12]	96.31	5,120 slices	5-fold	Oasis
VGG16[12]	92.30	5,120 slices	5-fold	Oasis
Proposed (white matter model)	95.20	5,120 slices	5-fold	ADNI
Proposed (grey matter model)	96.66	5,120 slices	5-fold	ADNI
Proposed (full matter model)	98.72	5,120 slices	5-fold	ADNI
Proposed multi-model)	99.63	5,120 slices	5-fold	ADNI

As can be seen from Table above, our proposed method using full brain data outperforms every method with respect to the accuracy and data size used for training. It is not just because of transfer learning only, but also because of the

entropy based technique used in selecting the slice for training which is based on calculating and considering slices with highest entropy as more informative is adopted from the research by [12] . The mixture of these two methods leads to very excellent performance with a relatively lower training dataset. Using transfer learning with a limited dataset is important because it not only increases training time and computational costs but also considerably frees the technique from reliance on big, tediously annotated data on training.

5.3 Multi-modal performance

The proposed Multi-modal method performs better than all the methods listed in the table 5.2 achieving an accuracy of 99.63%. The result of majority voting is considered more generally robust and precise which is a very important aspect in diagnosis.

Chapter 6

CONCLUSIONS

6.1 Conclusions

In this work, we proposed the use of structural MRI scans to classify Alzheimer's patient from healthy control. Three data points (skull-tripped full brain data, grey-matter data and white-matter data) obtained from each scan through the preprocessing pipeline discussed in this work. Each of these data point are used to build a model using the VGG16 networks through ImageNet pre-trained weights and fine tuning. Furthermore, we fused the predictions on testing data from the three models using an ensemble technique (majority voting) to obtain more robust and accurate results.

Compared to other methods a small training dataset is used to get optimal results. The slice selection technique explained in this work, which is based on entropy calculation adopted from the work of played a key role in selecting the most informative ones. We evaluated our models on images from ADNI database; where 6400 images for each modality are extracted from sMRI scans of 200 subjects (100-AD, 100-HC) used to train the models. The proposed method provides high performance which is comparable to the state-of-art. We also investigate the use of white-matter as a modality for detecting Alzheimer's disease which proves also very efficient.

6.2 Future work

In the future we intend to explore the utility of including a trainable decision fusion layer, which will allow for the adaptive mixing of decisions from the constituent models. This will involve the training of the system as a whole, but only adjusting the weights of the decision fusion layer during training. We believe that this will yield performance gains over the current approach as a trainable fusion layer would learn the appropriate way to combine the discrete outputs from each model to maximize predictive performance.

s

REFERENCES

[1] Alzheimer's Association. [Online]. https://www.alz.org/alzheimers-dementia/what-is-alzheimers, 2017.

[2] Alzheimer's neuro-imaging initiative. [Online]. http://www.adni.loni.usc.edu/, 2017.

[3] Wenqing Suna, "Computer aided lung cancer diagnosis with deep learning algorithms," , 2016.

[4] Alessandro Giust Dan Ciresan, "Mitosis Detection in Breast Cancer Histology Images with Deep learning Neural Network," pp.. 411–418, 2013.

[5] Ji-In Kim, Goo-Rak Kwon Debesh Jha, "Diagnosis of Alzheimer's Disease Using Dual-Tree Complex Wavelet Transform, PCA, and Feed-Forward Neural Network," Journal of Healthcare Engineering, p. 13, June 2017.

[6] Annina & Singh Deo, Mahima & Selvam, Venkatesan & Babu, Ramesh. Simon, "an overview of machine learning and its applications," international journal of electrical sciences and engineering, vol. 1, no. 1, pp. 22-24, 2015.

[7] (2019) GitHUB. [Online]. https://gist.github.com/baraldilorenzo/07d7802847aaad0a35d3

[8] FMRIB, Oxford, UK Analysis Group. FSL. [Online]. https://fsl.fmrib.ox.ac.uk/fsl/fslwiki

[9] Matlab. [Online]. https://www.mathworks.com/products/matlab.html

[10] S.M. Smith, "Fast robust automated brain extraction.," Human Brain Mapping, vol. 17, no. 3, pp. 143-155, November 2002.

[11] J. Ashburner and K.J. Friston., "Unified segmentation.," NeuroImage, vol. 26, pp. 839-851, 2005.

[12] Marcia Hon and Naimul Mefraz Khan, "Towards Alzheimer's disease classification through transfer learning," in 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), November 2017.

[13] Iman Beheshti, "Structural MRI-based Classification of Alzheimer's Disease," Gazimağusa, North Cyprus, 2016.

[14] Yaping Wang, Daoqiang Zhang, "Multimodal Classification of Alzheimer's Disease and Mild Cognitive Impairment," NeuroImaging, vol. 55, no. 3, pp. 856-867, april,2011.

[15] Asad. Abdi, "Three types of Machine Learning Algorithms.," 2016.

[16] Anitha J. Jude Hemanth D., "Image Pre-processing and Feature Extraction Techniques for Magnetic Resonance Brain Image Analysis.," in Computer Applications for Communication, Networking, and Digital Contents., Berlin, Heidelberg, 2012.

[17] Leonardo Sampaio Ferraz Ribeiro,Tiago Santana Nazare,Tu Bui,John Collomosse Moacir Antonelli Ponti, "Everything You Wanted to Know about Deep Learning for Computer Vision but Were Afraid to Ask," in SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), Niterói, Brazil, 2017.

[18] François Chollet, "fundamentals of deep learning," in Deep learning with python.: manning, 2018, p. 4.

[19] alex yu. towardsdatascience. [Online]. https://towardsdatascience.com/how-to-teach-a-computer-to-see-with-convolutional-neural-networks-96c120827cd1

[20] Ilya Sutskever, Geoffrey E. Hinton Alex Krizhevsky, "ImageNet Classification with Deep Convolutional Neural Networks," in Neural Information Processing Systems (NIPS), usa, 2012.

[21] Andrew Zisserman Karen Simonyan, "Very Deep Convolutional Networks for Large-Scale Image Recognition," CoRR, , vol. abs/1409.1556. , 2014.

[22] Kaiming, He Xiangyu, Ren Jian Sun, "Deep Residual Learning for Image Recognition," in IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[23] Vincent Vanhoucke, Christian Szegedy, "Rethinking the Inception Architecture for Computer Vision," , 2015.

[24] jason brownlee. (2018, december) machine learning mastery. [Online]. https://machinelearningmastery.com/ensemble-methods-for-deep-learning-neural-networks/

[25] Richard Maclin David Opitz, "Popular Ensemble Methods: An Empirical Study," Journal of Artificial Intelligence Research, pp. 169-198, 1999.

[26] Katherine Graya, Qinquan Gaob, Liang Chena, Daniel Rueckerta, The Alzheimer'sDisease Neuroimaging Initiative Tong Tonga, "Multi-modal classification of Alzheimer's disease using nonlinear graphfusion," elsevier, pp. 171-181, march ,2017.

[27] Xiao Zheng, Jun Shi, Yan Li, and Xiao Liu, "Multi-modality stacked deep polynomial network based feature learning for Alzheimer's disease diagnosis," in IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016.

[28] Murat seckin Ayhan, Anthony s. Maida Ashish Gupta, "Natural Image Bases to Represent Neuroimaging Data," in ICML'13 Proceedings of the 30th

International Conference on International Conference on Machine Learning, Atlanta, 2013.

[29] Adrien & Montana, Giovanni Payan, "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks.," in ICPRAM 2015 - 4th International Conference on Pattern Recognition Applications and Methods, Proceedings. 2. , 2015.

[30] Danielle D DeSouza, John Anderson, Ghassem Tofighi Saman Sarraf, "DeepAD: Alzheimer′ s Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI," bioRxiv, p. 070441, 2016/1/1.

[31] med2image. GITHUB. [Online]. https://github.com/FNNDSC/med2image

[32] Francios chollet. Keras: The Python Deep Learning library. [Online]. https://keras.io/

[33] kasthurirangen gopalakrishnan, "deep convolutionalneural network with tarnsfer learning for computer vision based data driven pavement distress detection," elsevier, vol. 157, pp. 322-330, 2017.

[34] V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Szegedy, "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[35] Cynthia M. Stonnington, Josephine Barnes, Frederick Chen, Carlton Chu, Stefan Kloppel, "Accuracy of dementia diagnosis - A direct comparison between radiologists and a computerized method," Brain, vol. 131, no. 11, pp. 2969-2974, november, 2008.

[36] Elizabeth Johnsona, Kathryn Ziegler-Grahamb, H. Michael Arrighic Ron Brookmeyer, "Forecasting the global burden of Alzheimer's disease," alzhemer's & dementia, pp. 186-191, july,2007.

[37] Siqi Liu et al., "Early diagnosis of Alzheimer's disease with deep learning," in 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing,china, 2014.

[38] Christian Igela, Akshay Paia, Ioana Balasa, Cecilie Ankerb, Lauge Sørensena, "Differential diagnosis of mild cognitive impairment and Alzheimer's disease using structural mri cortical thickness, hippocampal shape,hippocampal texture and volumetry," neuroimage: clinical, vol. 13, pp. 470-482, 2017.

[39] Jason Brownlee. (2019) machine learni g mysteries. [Online]. https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/

[40] Hilll and hawkes studholme, "An overlap invarient entropy measure of 3D medical images alignment," pattern recognition, vol. 32, pp. 71-86, 1999.

[41] D. Louis Collins. Holmes. Peters Evans, "Automatic 3-D model-based neuroanatomical segmentation," human brain mapping, vol. 3, no. 3, pp. 190-208, 1995.

[42] Coleman, and Doraiswamy Petrella, "Neuroimaging and Early Diagnosis of Alzheimer Disease : A Look to the Future," pp. 315–336, 2003.