Application of Deep Learning in MRI Classification of Schizophrenia

by

Ramisa Fariha Joyee
19301250
Lamia Hasan Rodoshi
19301248
Yasmin Nadia
19301241

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
The School of Data & Sciences
Brac University

23 January 2023

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing the degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Ramisa Fariha Joyee

Ramisa Fariha Joyee

19301250

Lamia Hasan Rodoshi

Lamia Hasan Rodoshi

19301248

Nadia

Yasmin Nadia

19301241

Approval

The thesis/project titled "Application of Deep Learning in MRI Classification of Schizophrenia" submitted by

1. Ramisa Fariha Joyee (ID:19301250, Summer 2019)
2. Lamia Hasan Rodoshi (ID:19301248, Summer 2019)
3. Yasmin Nadia (ID:19301241, Summer 2019)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on January 19, 2023.

Examining Committee:

Supervisor:
(Member)

Faisal Bin Ashraf

Lecturer
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Member)

Md. Shahriar Rahman

Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Md. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Ms. Sadia Hamid Kazi

Chairperson, Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

In today's world, when people are suffering from complex brain diseases, MRI has been playing a very significant part in understanding brain functionalities and its abnormalities. Deep learning has been recently used for the analysis of MRI, fMRI, structural MRI etc. and through this, we have achieved better performance than traditional computer-aided diagnosis for brain disorders. However, similar composition of brain diseases makes it hard to find out and differentiate the accuracy of exact disease from the acquired neuroimaging data. Accordingly, in this paper, a multi channel 2D CNN based architecture was implemented on COBRE dataset [1] which presents a significantly high accuracy over some models. Our modified multichannel 2D CNN architecture achieves around 97% accuracy which improves our classification performance. Furthermore, the paper discusses the boundaries of existing studies, the DL methods and present future possible directions.

Keywords: Schizophrenia; Deep learning (DL); Neuro-image; MRI; Computer-aided diagnosis; Neuro-psychiatric disease; DNN; CNN; SVM; RNN; COBRE; NUS-DAST;

[1] http://cobre.mrn.org/

Dedication

We dedicate the research project to all the people who are going through neurological brain disorders like Schizophrenia, to all those who faced different levels of trauma, especially intensive childhood trauma. Most commonly, Schizophrenia appears in late teens to 20 years old for men and early 20 to 30 years old for women. So we also dedicate to all of those men and women[2]. It is rare to indicate the symptoms of Schizophrenia in children below 12 or over 45 years old. Nowadays, the new variations of schizophrenia arise in teenagers between 16 to 25 years old. So, we also dedicate to those teenagers who can be going through this mental disorder.

[2]https://www.webmd.com/schizophrenia/who-gets-schizophrenia-by-age-sex-more

Acknowledgement

All thanks to the Almighty ALLAH(SWT), The creator and the owner of this universe, the most merciful, beneficent and the most gracious, who provided us guidance, strength and abilities to complete the thesis.
We are especially thankful to Faisal Bin Ashraf sir, our thesis supervisor, and MD. Shahriar Rahman, our co-supervisor, for their help, guidance and support in completion of our project. Moreover, we are immensely grateful to Mahfuzur Rahman Chowdhury for his unconditional help in our project. We also are thankful to the BRAC University Faculty Staffs of the Computer Science & Engineering, who have been a light of guidance for us in the whole study period at BRAC University, particularly in building our base in education and enhancing our knowledge.
Finally, we would like to express our sincere gratefulness to our beloved parents, brothers and sisters for their love, care and support. We are grateful to all of our friends who helped us. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

List of Figures

List of Tables

Abbreviation

SZ – Schizophrenia
MRI - Magnetic Resonance Imaging
DL – Deep Learning
ML - Machine Learning
CNN – Convolutional Neural Network
SVM - Support Vector Machine
DMN - Default Mode Network
AUD – Auditory Cortex
FA - Fractional Anisotropy
DNN - Deep Neural Network
KNN - k-Nearest Neighbor
RNN - Recurrent Neural Network
MD - Mean Diffusivity
STFT - Short-Time-Fourier transform
RBF - Radial Basis Function
VBM - Voxel-Based Morphometry
SBM - Source-Based morphometry
MKCapsnet - Multi Kernel Capsule Network
M2D Model - Multichannel 2D CNN Model
KPCA - Kernel principal component analysis
FLD - Fisher's linear discriminant analysis

Chapter 1

Introduction

1.1 What is Schizophrenia?

Schizophrenia is a serious mental disorder linked to structural and functional brain abnormalities. According to a research approved by WHO, it affects approximately 24 million people worldwide. It may result in hallucinations, delusions, disordered thinking and behavior which is beyond curing and the patient has to live with it for lifetime. Early treatments help to get the symptoms in control before serious development of complications arise.
Schizophrenia is a neurological brain disorder in which people show abnormal behavior. According to recent research [1], Schizophrenia becomes one of the reasons for being aggressive. Nowadays, in a schizophrenic patient, the rate of aggressive behavior is four to six times higher than in healthy people. Among 3,941 schizophrenic patients, there are high symptoms of aggressive behavior which is 15.3% to 53.2%. During the 5 years from 2011 to 2015, the outbreak of schizophrenia provoked from 0.63% to 0.94%. Signs of Schizophrenia may vary but it usually involve delusions, hallucinations, disorganized speech, extremely abnormal behavior, depression, withdrawal from society, suicidal thoughts etc.
Researchers have not yet found any possible reason behind Schizophrenia. It was thought that interaction between genes and some environmental factors may be responsible for this psychiatric disorder.

1.1.1 Problems of Diagnosing Schizophrenia

Diagnosing SZ is very difficult most of the time because this mental disorder is characterised by a wide range of symptoms which can vary person to person. Some main characteristics of SZ are - hallucinations, delusions, chaotic speech, catatonic behaviour, negative contracted emotional expression. But these symptoms also can appear in other mental disorders. So,there are no specific symptoms which can easily be verified as SZ. Besides, many people don't want to believe that they have this neurological brain disorder as there are no specific lab tests for this mental disorder. Nowadays, SZ is determined by psychiatric consultation which is not a reliable identification. That is why MRI classification can be a better option to diagnose SZ. But there are also some barriers as MRI classification is a costly process. However, as there are no specific symptoms and lab tests it is the biggest obstacle to diagnose SZ at an early age.

1.1.2 ML and its Use in Schizophrenia Diagnosis

Diagnosis of schizophrenia using sMRI and fMRI neuroimaging modalities along with conventional method ML is widely used. This is because of limited number of available public dataset and conventional ML do not need powerful hardware elements, therefore high performance is gained for not using complex feature. Practical implementation is hard to be done with limited dataset, therefore, dataset with variety of sz disorder and large number of subjects is very helpful for clinical diagnosis. Implementation of CADs using conventional ML needs vast knowledge in the field of AI. So DL needs to be occupied. DL model GANs is introduced to address MRI data shortage problem. GAN architecture can be used to design CAD system for effective diagnosis of schizophrenia. Lack of free accessibility of sMRI and fMRI neuroimaging modalities for specific class of sz is a great problem which can be resolved using a class of AI technique called zero-shot learning. Moreover, sMRI and fMRI scan can be sent to cloud and appropriate DL model can replace them. Therefore, produced result will be sent out to hospital server and after getting confirmation from the clinician, final report can be presented to the patient.

1.2 Problem Statement

In this paper, we will focus on the research of detecting SZ through MRI classification by DL algorithm methods. Firstly, we will pre-process the dataset of SZ. We will divide the dataset into SZ predicted patients and healthy group. The precision will increase if the group of SZ predicted patients are more in number. Applying quality control and feature engineering, the images will have to be processed. Then we will apply suitable classification such as – logistic regression, k-nearest neighbor, SVM (non-liner mapping), random forest, DNN (gives high accuracy), CNN etc.

CNN, DNN are powerful DL methods which helps to detect SZ with higher accuracy. These methods use multiple hidden layers between input and output data which helps to accurately give the output. CNN's multiple hidden layers are responsible for noise removing. Noise removing technique is an interesting attribute for detecting SZ. In one study [2], researchers used DNN to detect SZ through five multicenter datasets of structural MRI which gave higher accuracy.
Severe psychiatric disorder, SZ deeply affect the patients physically and mentally and drives them to suicide. Because of the failure of early detection of SZ, patients know about the disease in the later stage when they cannot get proper treatment and get better. As SZ is not curable, early detection helps to control the symptoms by treatment. Moreover, MRI data of SZ is quite similar with other brain disorders which makes it harder to detect SZ. Previously it was mentioned, traditional diagnosis of SZ has lower accuracy and cannot be detected early. For this reason, SZ patients along with their family worldwide are suffering greatly. That is why, DL methods instead of traditional ML algorithms give a better liability to detect SZ.

1.3 Paper Orientation

This chapter (chapter 1) introduces readers to Schizophrenia and provides a brief discussion of the research's problem statement and objectives.
The remainder of the paper is organized as follows:
Chapter 2 provides a literature review of some past published researches on the use of DL methods in classifying MRI data of Schizophrenia and major findings and future scope of research.
Chapter 3 describes data analysis and preparation. In this chapter, dataset collection, description of the challenges regarding dataset and pre-processing of dataset to get better outcome are mentioned thoroughly.
Chapter 4 is about methodology where we mentioned that which deep learning network we prefer for MRI image classification to detect SZ. For this research, we used CNN network and six models to check out the result of our dataset by using 2D images.
There is a section for results which is in chapter 5.
And the last chapter is for conclusion and future plan which is in chapter 6.
Finally, there is a bibliography at the end that lists out all the sites, research papers and journals that were referred to in this paper.

1.4 Research Objectives

We intend to detect SZ using MRI classification through various DL methods. To do such, we have come up with a better proposition – feature modification of dataset using various data pre-processing techniques and use the processed dataset to train models for the purpose of detecting SZ. Comparing to the traditional diagnosis accuracy of SZ detection, DL methods give us better accuracy but different DL methods give different accuracy. Our proposition will accomplish the following objectives:

- Research suitable DL method which will give us higher and precise accuracy.
- Train models by using proper modified features of dataset.
- Check the accuracy by using the trained model.
- Use of cost-efficient and time-efficient methods.

Chapter 2

Literature Review

To be able to start the correct treatment earlier, we need to detect SZ as soon as possible. SZ is generally detected through MRI but it is quite difficult to detect because of the other mental diseases with similarities. That is when ML techniques come to rescue as ML methods help to detect SZ accurately in the earlier stage. Some of the past papers that found such appropriate techniques have been reviewed to find possibilities for future research and summarized below.

In this article [3], authors used only age and sex-matched 72 subjects from both healthy people and Schizophrenia patients to ensure a balanced study design.They have used COBRE dataset where subjects were screened and excluded based on their history of abuse or dependence in the scanning period of 12 months.[4] In addition, patients were also screened based on the history of neurological disorders, severe head trauma or intellectual disability. In the time of rs-fMRI scans, SZ patients were screened for hallucinations and delusions. Authors have proposed 3D-CNN based DL classification to separate SZ patients and healthy people. They used rs-fMRI data from COBRE dataset and then pre-processed the data using FMRIB Software Library (FSL version 6.0). They applied standard pre-processing which are included of slice timing, motion correction and denoising, spatial smoothing with 6 mm full width half maximum(FWHM) gaussian kernel, temporal filtering. Finally, the functional data was co-registered to its corresponding structural image and registered with MNI152. These data were also high filtered at 100Hz. The noise and artifacts of the features were normalized and separated by using the FSLNets automated and unsupervised clustering tool[1]. They used FSL Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC), version 3.14 for creating the group ICA to measure connectivity. The classifier had taken every 3D ICA maps as input without any masking and thresholding to avoid bias data. They used ten0-fold cross validation to avoid overfitting of the model. They used a modified version of VGG-Net[5]. In this article, authors applied 3D resting state connectivity networks instead of 1-dimensional time series information to get better performance and they achieved 98.09±1.01% accuracy. They did not apply any other DL methods on this dataset. But in this study, there are some limitations, it can not classify the first episode psychosis patient data. Besides, there are some restrictions on its use in critical practice. Furthermore, it was not specifically known that which ICA components were more responsible for classification because of not

[1] https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/%20FSLNets

using feature selection method.

Accordingly, in another paper[6], writers used the ROI and segmented ventricle images as input for deep neural networks to diagnose SZ. In their works, center slice of MR images with a axial view selected ROI. In this image deep gray matter region is consists of ventricle and caudate where caudate is in left and right sides of brain region. Besides, they observed that there is a shape variation in the ventricle region in SZ patient compared to normal subject. They used Adagrad, SGD and RMSPop for training to minimize the mean square error between the ground truth labels and network output.They used 10 images in per batch and ran 1000 epochs to trained the network. By doing research they are able to find out that the right cerebellar white matter, cortical thickness and enlarged ventricles gave the largest difference between SZ and normal groups. Analysing cortical thickness they obtained 73.6% accuracy to identify SZ patient. Authors got 90% accuracy by using segmented ventricle which is higher than ROI images.

In another paper[7], authors selected and labeled the functionally informative slices and applied 2D-CNN for classification. Motion correction and spatial normalization were applied in data pre-processing. This paper has classified both slice level and subject level. For slice level classification, accuracy was of 72.65% in DMN and 78.34% in AUD. In this work, they were able to enhance the accuracy using 2D-CNN by reducing training parameters compared to 3D CNN.

According to another article[8], authors used COBRE dataset to classify Schizophrenia versus normal subjects using deep learning[3]. They took 74 healthy subjects and 72 SZ patients with age range from 18 to 65 years old of both men and women. They collected data with an echo time TE= 29ms and repetition time TR= 2s. In every brain volume, the size of each slice is 64×64 where total slice is 32 and the voxel size is 3×3×4 mm^3. In this paper, authors used SPM8 (Statistical Parametric Mapping)[2] to pre-process these fMRI data. They eliminated first 5 volumes to allow magnetization to reach the steady state. Besides, for motion correction they used the rest 145 functional volumes which ensures the head motion below 2mm or voxel to voxel according time. So their proposed method was two-stage SAE based architecture for the classification of SZ versus Normal subjects and for training and testing accuracy they used SVM classifier.Moreover, we observed that in this article, they got 95% and 93.33% accuracy in one fold by using proposed method and model. But using 10 fold cross validation they got average accuracy 92%. One thing to notice in this article is their proposed architecture works directly on active voxels' time series instead of converting them into region-wise mean time series.

In another article[9], writers used Extreme Learning Machine(ELM) to classify SZ from MRI image which is a speedy training approach for single layer feed forward neural networks(SLFN). However, there is a criticism that ELM has uncertainty in performance as it generates random hidden layer weights. In order to overcome this problem authors used composition of elementary classifiers into ensembles,like as the Voting ELM (V-ELM) which was trained independently. They used an open source software pipeline which is known as Configurable Pipeline for the Analysis of Connec-

[2]https://www.fil.ion.ucl.ac.uk/spm/software/spm8/

tomes (C-PAC). This software was made upon AFNI [10], FreeSurfer [11] and FSL (the FMRIB Software Library)[12]. Additionally, FSL FAST and AFNI SkullStrip used for tissue segmentation and brain extraction. Also authors discarded first 6 volumes and took 144 volumes for head motion correction. Besides, in this article for data pre- processing they used slice timing, head motion correction,[13][14] nuisance correction (whiter matter, cerebrospinal fluid and principal components regression, first principal component removal and linear detrending)[15] and band-pass temporal filtering between 0.01 and 0.1 Hz[16]. In this article, they used feature selection based filter algorithms which are represented by Pearson's correlation. Moreover, they wrote the result of two classification experiments. Firstly, they accomplished an exploratory evaluation of ELM Parameters to describe the hidden unit output and the number of hidden units.

In this article [2], the authors trained 3D-CNN to identify SZ. For this, they used five data sets of structural MRI scans and examined the classification performance of the algorithm. Furthermore, they examined which brain regions were mainly responsible for the DL algorithm by diving the MRI scan into eight regions. This DL algorithm has four 3D convolutional layers with max-pooling-based down-sampling in each convolutional layer. In the training dataset, the accuracy was 97% in which 97% images were classified correctly. Though the accuracy rate decreased in completely new dataset.

In an article [17], the writers used FNC and SBM datasets together and created a full set of feature vectors which was later used to train models. They obtained results from training traditional ML and DL techniques and then made comparisons. They got the accuracy of 82.77% by applying logistic regression, 82.68% by applying SVM, 83.33% by applying random forest and 94.44% by applying DNN approach. For DNN method, depth of three hidden layers was used where the number of nodes in each hidden layer were 512, 256 and 128, and dropout rate was of 50%. Though the DNN approach has some limitations. Exploration of hyper-parameter tuning options was prevented due to the small size of the training dataset whereas for DNN, big set of dataset is usually needed. There was also the risk of overfitting.

In one study [18], authors used MRI data collected from structural MRI scans. MRI scans were based on gray-matter densities of brain using VBM. Authors used the cross validation technique and got the accuracy for trained data of 86% and 83% for the test dataset. Using SVM as classifier, they got high accuracy to detect SZ patients using MRI data. It proves high performance of ML methods for differentiating nonlinear associations between input and output data. It can also efficiently handle high-dimensional data and avoid overfitting.

In this article [19], the authors have developed an in-depth feature approach based on 2D CNN and Naive 3D CNN models. these models were trained previously by spreading 3D structural MRI for identifying SZ. In this research paper, specialists suggested that sMRI evidence reduces extensive grey matter in frontal, temporal, thalamic and stiatal regions in Schizophrenic patients than healthy people. Besides, dMRI studies expressed lower FA and higher MD in numerous white matter spreads to associate frontal-striatal-thalamic nerves in Schizophrenia. Furthermore, 2D and

3D CNN technologies performed well with the help of ML. Nonlinear SVM along with RBF acquired an average accuracy of 70.22%. Deep characteristic pre-trained SVM obtained an accuracy of 72.41%. Naive 3D CNN models exposed a good performance on pre-trained 2D CNN with an accuracy of 79.27%.

As reported by this article [20], a training set is essential to recognize the pattern and to teach a computer to classify instances automatically and accurately. Firstly, an image analysis technology is needed to expose the most pertinent information by fetching the image data. Secondly, a pattern or a model classification method has to be designed to stretch out the information to identify mental disorders. Here, in n-dimensional space, a subspace can be defined as (n-1)-dimensional surface. Thus, SVM, MRI instances are prominent in a severe dimensional space. However, to determine SZ patients' structural and functional brain abnormalities, neither ROI analysis nor VBM or SBM methods enable diagnosis. In this paper, they stated an ML approach in diagnosing SZ by using ROI analysis. On the other hand, SVM technology has some limitations as it requires feature vectors and there are various subjects which could express different features. To conquer this problem, dissimilarity vectors can be applied in lieu of feature vectors. Besides, the alteration process-MRI creates a neuro image to reveal the functional and structural image to determine SZ. This study result suggests the application of ML based on MRI with a tremendous result in diagnosing SZ.

In this article [21], the author firstly talked about simple ML (linear discriminant analysis) and complex ML (DNN). Both of them are used to distinguish SZ based on functional connectivity features. The problem with simple method is that, it disables simultaneous tuning for best feature selection and classifier training. The problem with complex method is that, it needs big size of dataset to train models. To get rid of the mentioned problem, MKCapsnet is used. Kernels are set according to the partition size of brain anatomical structure to capture inter-regional neurological connectivities at various scales. Also, vector dropout strategy was used in the capsule layer to prevent the risk of overfitting of the model. Though capsule neural network is a successful method, it requires improvements for SZ identification. In MKCapsnet, firstly, some participants are excluded due to excessive head movement and unavailable category information after preprocessing the MRI data collected from public dataset. Connectivity strength is estimated and normality is improved of the strength. All values are represented as matrix and fed to multi-kernel capsule network which has three layer. The first layer extracts connectivity information, also represented as vector in the second layer. This vectors are assigned to six channels correspond to six kernels. After that, a capsule dropout strategy is set at the capsule layer and a routing algorithm is used to learn based on the capsule. Since different types of layer settings are used for the proposed model, different settings provided different performance measures. For example, MKCapsnet performed better in terms of sensitivity, capsule dropout strategy performed better than scalar dropout strategy and so, MKCapsnet gives accurate result while identifying SZ. However, MKCapsnet can be improved by replacing vector representation with tensor representation and by adding additional information, this model can be used to detect other brain diseases too.

According to this article [22], deep learning method has been used to assess potential of a biomarker which is the functional connectivity in resting stage. Here, fMRI is used to capture a stable and clear image of brain functional organization at resting stage. Excessive higher dimension of data is an obstacle for fMRI to design and robust classifier. Therefore, two strategies were used to decrease the volume of the problem which are feature selection and classifying these selected features from patient to controls. Feature selection vastly relies on mathematical method and thus, it overlooks physiological necessities which may have prominent role in developing the disease. Moreover, classifying schizophrenic patient from controls only based on the previously selected feature may lead to a one-sided decision. So, deep learning comes to rescue. The idea is that, high dimensional data can be turned into low dimensional code through training a multilayer neural network with small central layer so that high dimensional input vector can be reconstructed. In data preprocessing, firstly, patients were diagnosed and the symptoms were evaluated. Secondly, resting state fMRI was done and result was gathered through 3T MRI scanner. Thirdly, Data preprocessing was done with the help of statistical parametric mapping software. Some volumes were rejected. The enduring rs-fMRI images were rectified using least square approach. rectified images were normalized. Resultant images were filtered and detrended using typical temporal bandpass. Cerebrospinal fluid and white matter were used to reduce influences of head notion and non-neural fluctuation. Mean frame wise displacement was evaluated to differentiate head movement between different groups. Finally, participants maintaining more than 80 percent of the original signal were included for the analysis. In image preprocessing images were smoothed. Average signal of brain region was computed and Pearson's correlation was calculated between each pair of signal and thus, matrices were created. These matrices were splitted into train data and test data. Train data was used to build model and test data was used to evaluate performance. In conclusion, resting state functional connectivity has promising classification capacity and can be used as biomarker.

According to another paper [23], Authors proposed a multi channel 2D model (M2D) to determine task-based functional mri data. Blood oxygenation level dependent (BOLD) signals, which determine how well the brain is functioning, are constantly changing. Therefore, it must categorize 3D voxel wise fmri data. Many papers employed 2D and 3D CNN for this. However, this study used multi-channel 2D CNN with axial, coronal, and sagittal brain data, which represent all three dimensions of the brain. In order to train the 2D CNN network, the authors converted each 3D brain image into 2D image slices and then extracted the highest entropy from calculating the entropy of each slice of brain image. In this model, three 2D CNN portions extract features in simultaneously, after which their outputs are flattened and merged into 1D features in sequence, which are fed to the fully connected neural network for additional learning. In particular, they claimed that M2D CNN model achieved a classification accuracy and precision of around 83%. Overall, M2D CNN performs better than all other models, with M2D CNN coming out ahead of 3D CNN, S2D CNN, 1D CNN, and 3D SepConv models and also PCA + SVM.

Chapter 3

Data Analysis and Preparation

3.1 Dataset

3.1.1 Features

Some public MRI datasets for SZ detection are available at SchizConnect[1] database such as BrainGluSchi, COBRE, MCICShare, NMorphCH, NUSDAST etc. We have used the COBRE dataset for our research which is taken from website[24]. Each control's MRI image is in shape (53, 64, 52, 150). In this dataset, there are 74 Normal Subjects (51 males and 23 females, age range = 18-65 years) and 72 SZ Patients (58 males and 14 females, age range = 18-65 years). Figure 3.1 shows the SZ and normal control's number in percentage.

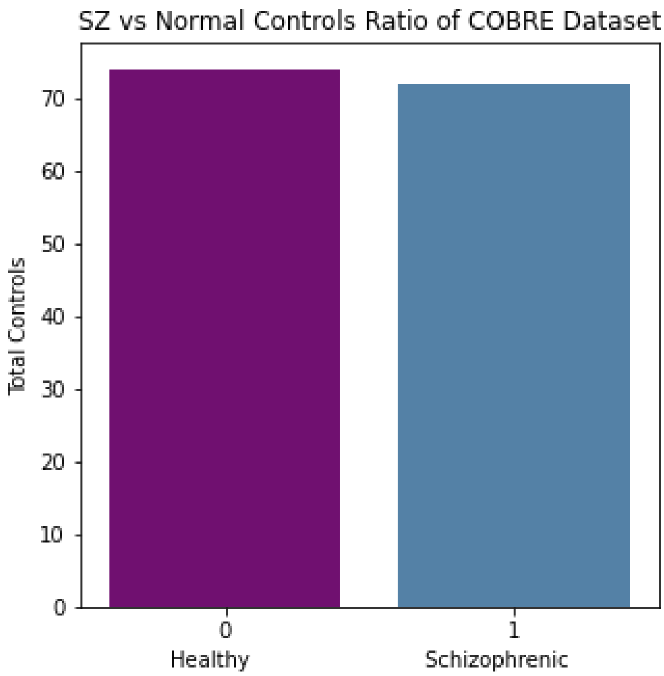


Figure 3.1: SZ vs Normal controls ratio for COBRE dataset

[1] http://schizconnect.org/

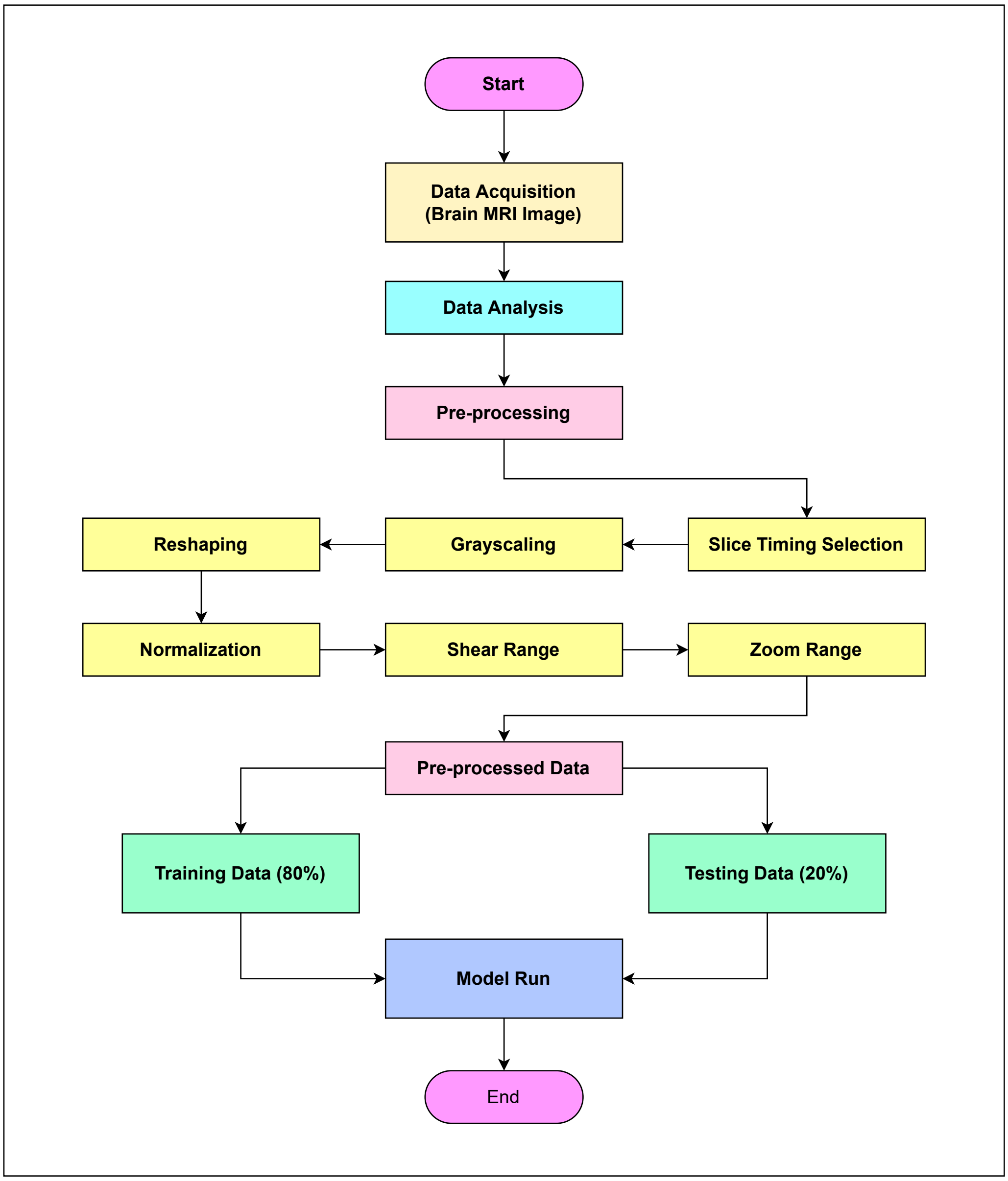


Figure 3.2: Workflow of data preparation

3.1.2 Functional Data Acquisition

The dataset contains resting-state functional MRI images of each subject. This dataset is mainly from the Center for Biomedical Research Excellence (COBRE) dataset. Information on this dataset can be found at website[2]. COBRE is reliable as many professors have used data from here for research works. The fMRI dataset are in Neuroimaging Information Technology Initiative (NIFTI) format (.nii.gz extension). It features 150 EPI (echo planar imaging) blood-oxygenation level dependent (BOLD) volumes and these were obtained in 5mns (TR = 2s, TE = 29 ms, FA = 75°, 32 slices, voxel size = 3x3x4 mm3, matrix size = 64x64, FOV = mm2). The MRI images from all timestamps can be visualized by using MRIcron software though in normal eyes, the differences cannot be seen clearly. Table 3.1 gives an overview of the COBRE dataset.

Class	Control	Male	Female	Age
SZ	72	58	14	18-65
Normal	74	51	23	18-65

Table 3.1: Number of Controls in each Class and their Age

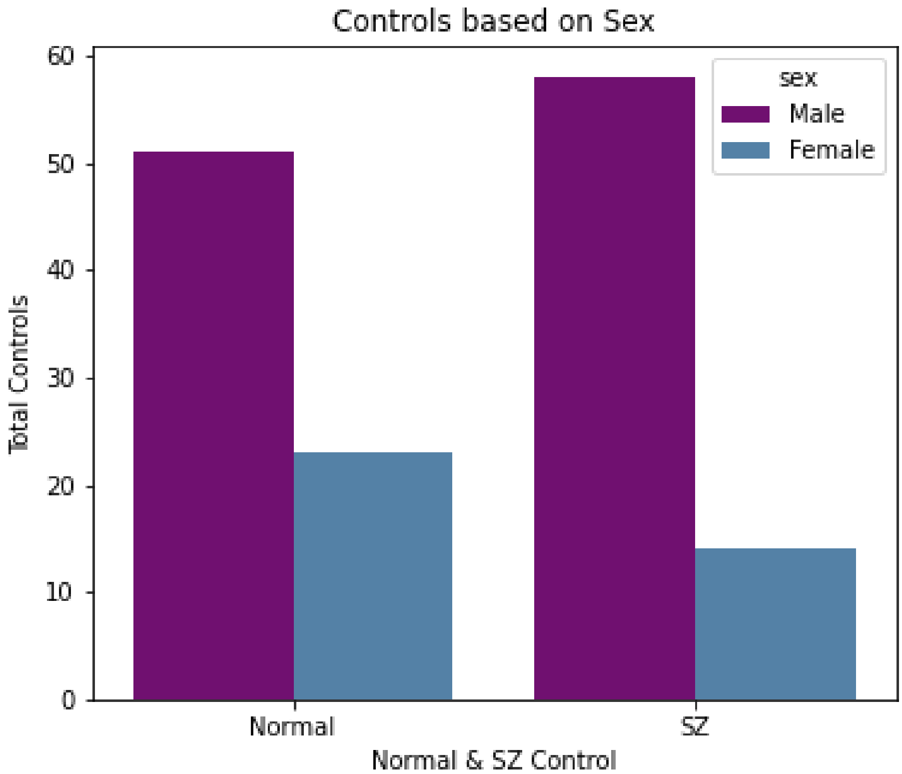


Figure 3.3: Controls based on sex

Figure 3.3 shows the male-female ratio in the mentioned dataset. In our graph, it is showed that there is a large difference in SZ patients' gender classification. According to our graph, men are facing SZ disorder more than 50% than women. A study

[2]http://cobre.mrn.org/

[25] reported that brain-derived neurotrophic factor (BDNF) may be engaged with patho-physiology of SZ for which male are getting more effected by SZ than female.

According to the figure 3.4, we can see that, the SZ disorder also varies age to age along with gender. From the graph, it is shown that, at lower age(18-30) and higher age(48-65), the SZ discloses in higher range in men than women. Moreover, at the age of 30 to 36, we can see the range of SZ is approximately equal in both men and women.

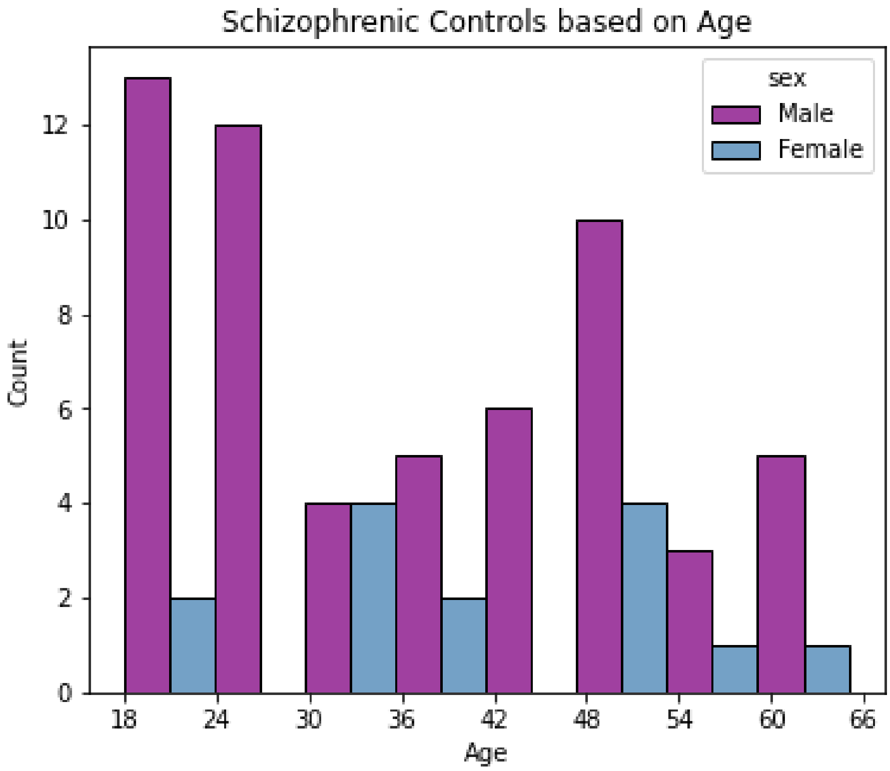
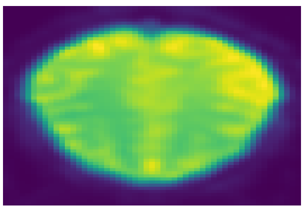
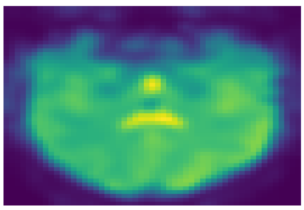


Figure 3.4: SZ Controls based on age

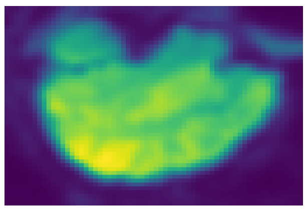
Figure 3.5 refers to the 2D images based on signal intensity from different angle (Axial, Coronal, Sagittal) of a SZ patient at 10 volume.



(a) Above angle

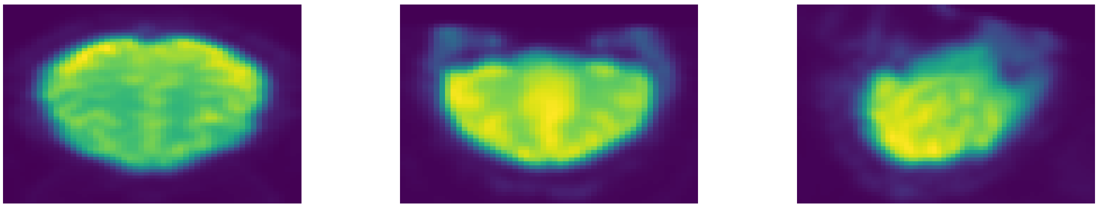


(b) Rear angle



(c) Side angle

Figure 3.5: SZ control's MRI image from different angles at volume 10 according to signal intensity



(a) Above angle (b) Rear angle (c) Side angle

Figure 3.6: Normal control's MRI image from different angles at volume 40 according to signal intensity

Figure 3.6 refers to the 2D images based on signal intensity from different angles of a normal control at 40 volume. These images are generated by using Matplotlib library. We can see from two different classes' images that SZ patient's brain image is more swollen than normal control's brain image. The reason behind swelling is neuroinflammation caused by Microglia, stress etc [26].

We have plotted first 95 volumes' 2D MRI image of SZ patient in figure 3.7 and normal control in figure 3.8 to see the differences in timestamps. Activity fluctuations over time in fMRI are very small. In normal eyes, it looks blurry but the changes are there.

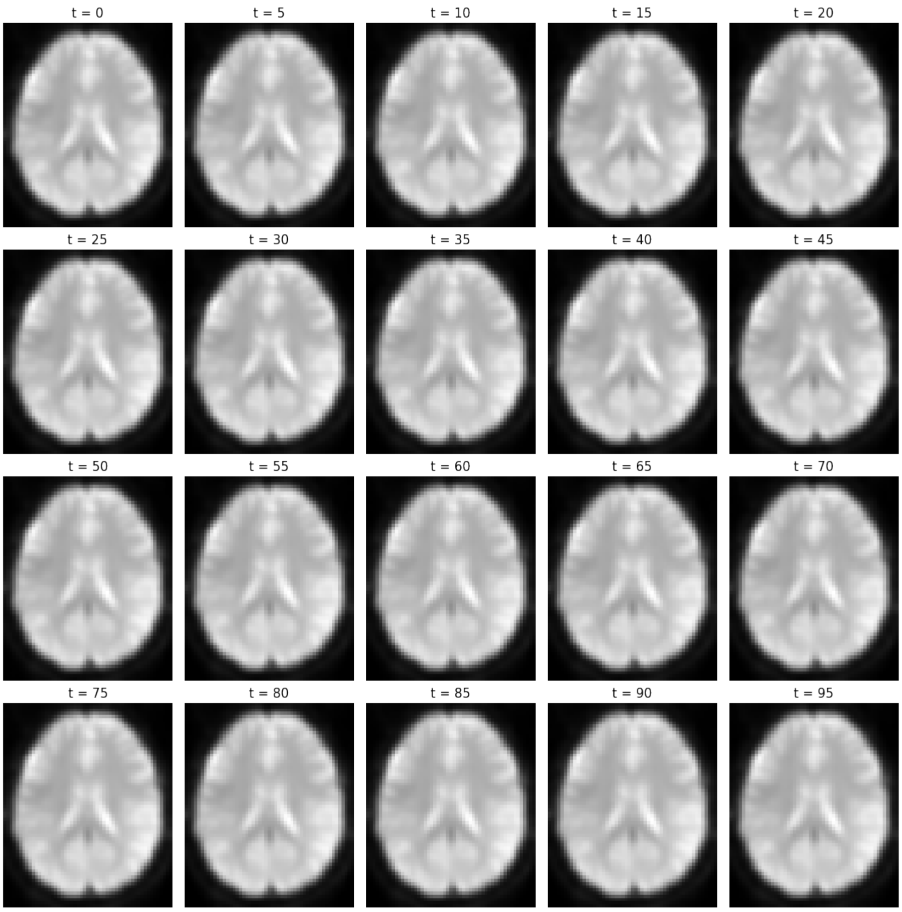


Figure 3.7: SZ patient's first 95 volumes' 2D MRI images (Axial region)

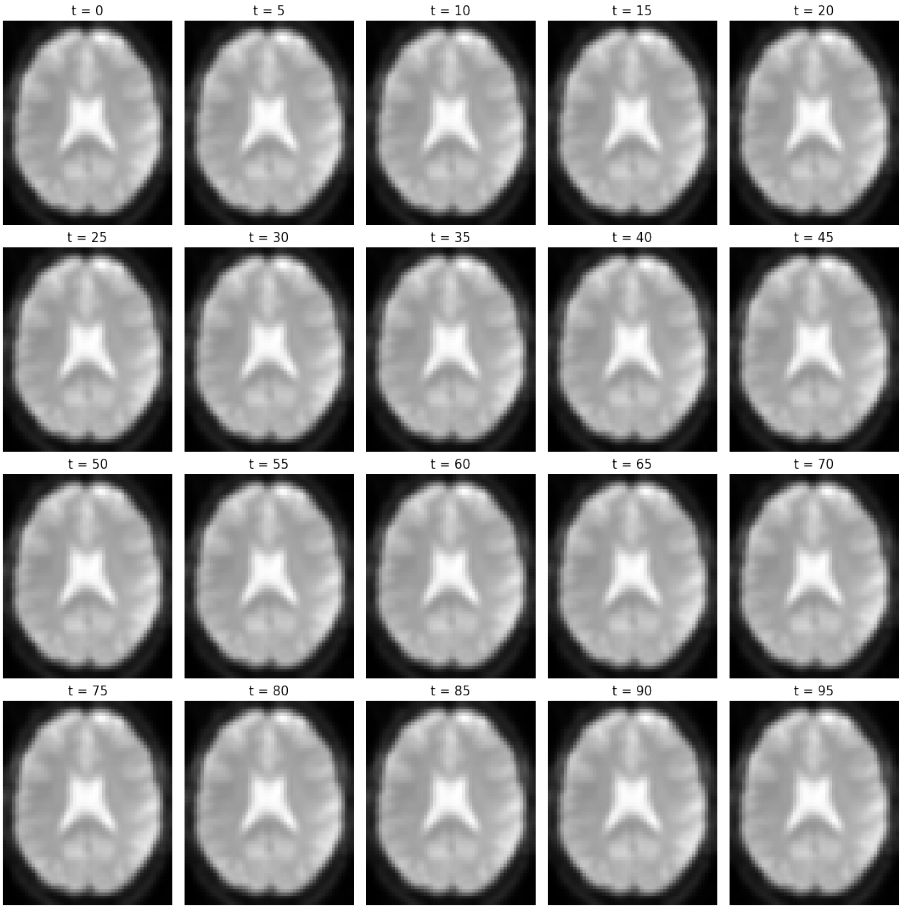


Figure 3.8: Normal control's first 95 volumes' 2D MRI images (Axial region)

As from the first 95 volumes' images with interval 5, we cannot identify the differences, we plotted graph of all timestamps' signal intensity to identify the differences from one timestamp to another timestamp. Figure 3.9 and 3.10 refers to the fluctuations of SZ and normal control's MRI image in signal intensity. From the graph, it is clear that the fluctuations are abnormal in SZ patient's MRI image in signal intensity [27].

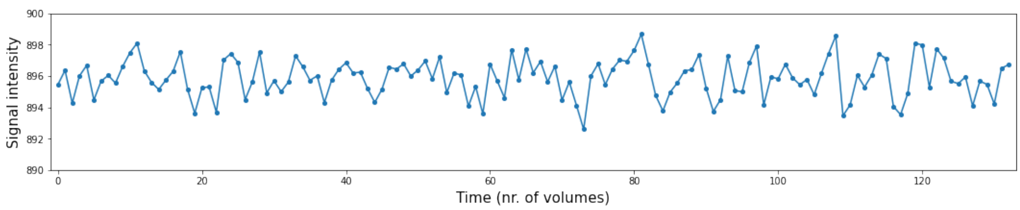


Figure 3.9: SZ patient's all volumes' MRI images' differences according to signal intensity

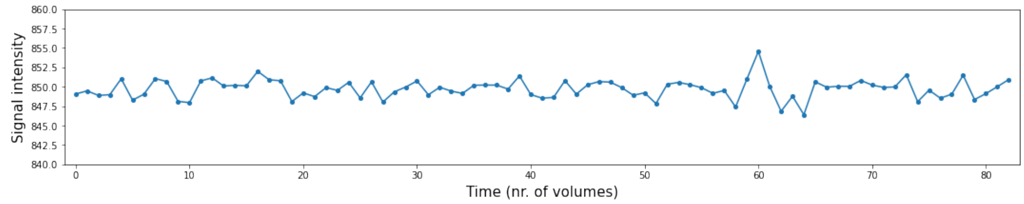


Figure 3.10: Normal control's all volumes' MRI images' differences according to signal intensity

3.2 Challenges regarding Dataset

We have collected COBRE dataset[3] after searching many sites. Even though this dataset is available at the official COBRE site, but before getting the dataset, we need to go through many official procedures which makes it problematic to acquire the data. As our research is a part of the biomedical sector, datasets are not publicly available that much for security reasons. So it creates hindrance to explore more in this sector.
The dataset has only 146 control's MRI data. Less data gives us less variation. Model cannot learn properly with less data. Moreover, we cannot easily get medical data. These are the challenges we had to face in our research.
Although only 146 controls' MRI data are present in the dataset. But the size of data is too huge. As maximum of 150 timestamps' and minimum of 40 timestamps' 3D images are there in controls, we get in total of 207,524,400 images from every angle from the whole dataset. We got the total number of images by calculating the following process (3.1):

$$\sum_{i=1,2,...,146} (((c_{i_x} * c_{i_y}) + (c_{i_x} * c_{i_z}) + (c_{i_z} * c_{i_y})) * c_{i_t}) \tag{3.1}$$

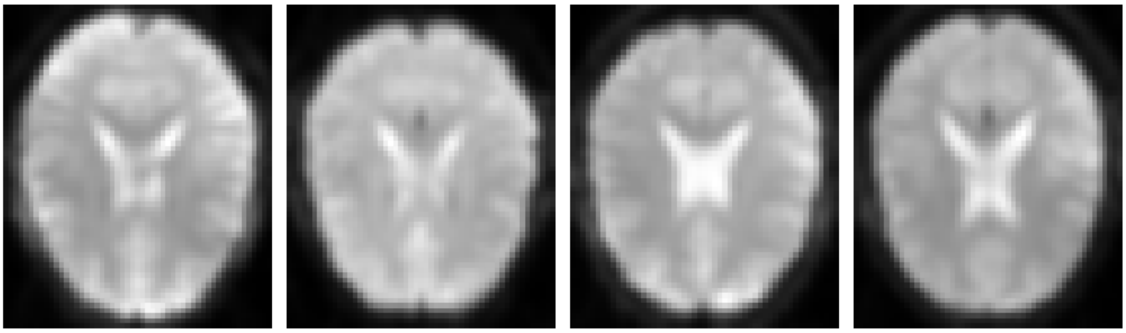
Here,
(x, y, z, t) = Shape of each MRI image
i = i-th control
c = controls
x = 1st axis of shape of i-th c
y = 2nd axis of shape of i-th c
z = 3rd axis of shape of i-th c
t = Time axis of shape of i-th c

Handling large data is tough as it increases the processing time and complexity. As we have raw MRI images for our research, we have to manually process each image so that our model can give better performance but large data tends to increase the complexity and creates data storage issues.

Figure 3.11 and 3.14 show 4 different normal controls' and SZ controls' axial region's brain images in same slice and volume, figure 3.12 and 3.15 show 4 different normal controls' and SZ controls' sagittal region's brain images in same slice and volume

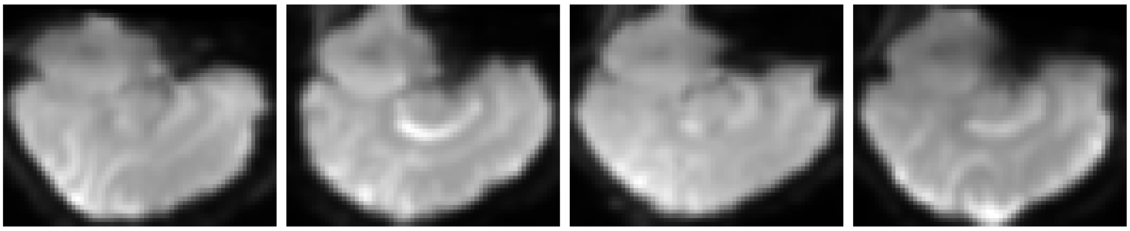
[3]http://cobre.mrn.org/

and figure 3.13 and 3.16 show 4 different normal controls' and SZ controls' coronal region's brain images in same slice and volume. We can see from the figures that the differences between normal and SZ control's axial region's brain images are subtle which is why it is challenging to classify SZ MR images accurately. For example, figure 3.11b and 3.14d MR images are almost same and so in human eyes, it is difficult to say which is SZ patient's brain image and which is normal control's brain image. For this reason, in early stages, SZ detection is hard.



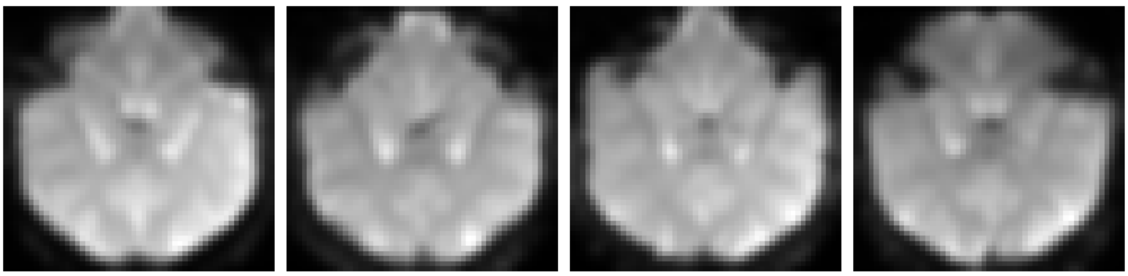
(a) (b) (c) (d)

Figure 3.11: 4 different Normal control's brain image (Axial region)



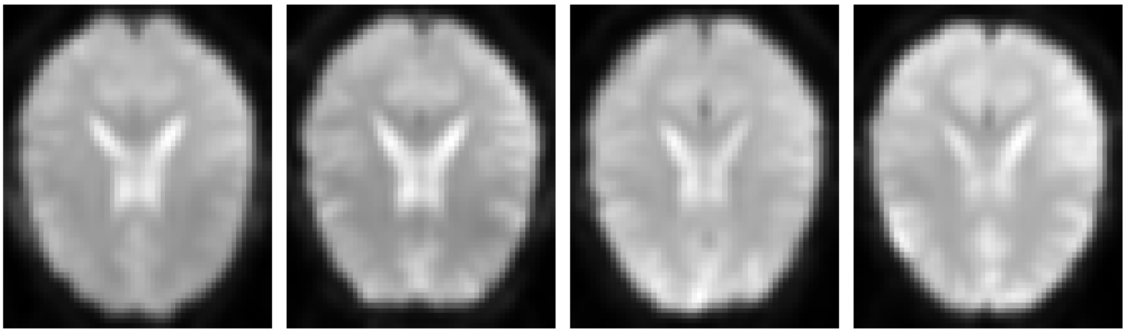
(a) (b) (c) (d)

Figure 3.12: 4 different Normal control's brain image (Sagittal region)



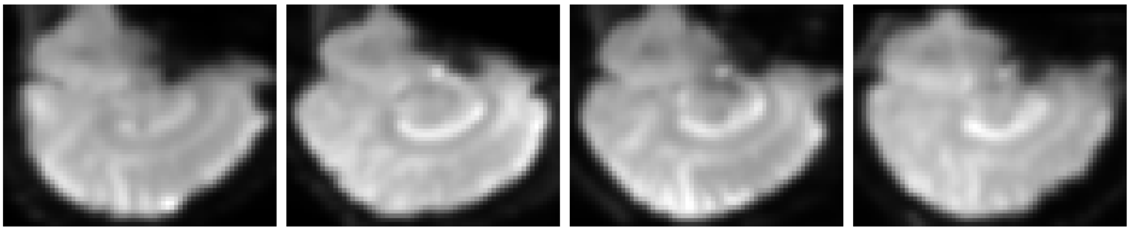
(a) (b) (c) (d)

Figure 3.13: 4 different Normal control's brain image (Coronal region)



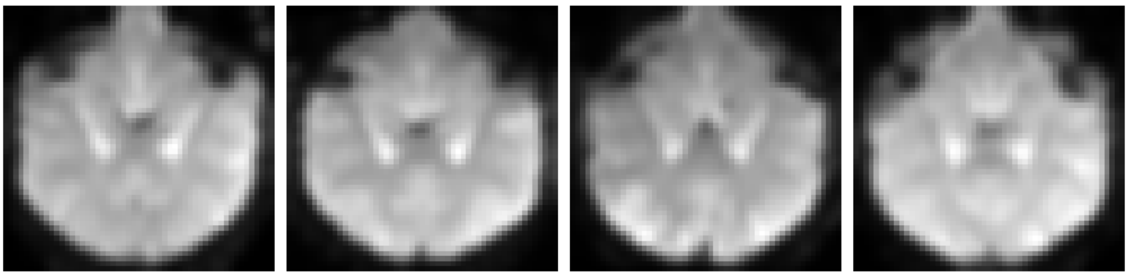
(a) (b) (c) (d)

Figure 3.14: 4 different SZ control's brain image (Axial region)



(a) (b) (c) (d)

Figure 3.15: 4 different SZ control's brain image (Sagittal region)



(a) (b) (c) (d)

Figure 3.16: 4 different SZ control's brain image (Coronal region)

3.3 Data Pre-processing

Data pre-processing is an important step for model training as model's performance depends in it significantly. We took pre-processed MRI data from this paper[28] for our research, though we had to do more pre-processing of the multi-dimensional data to fit our model. This paper[28] analysed the dataset using the NeuroImaging Analysis Kit[4] version 0.12.14.. Firstly, we loaded our nii data by importing nibabel package. The size of the data is (53, 64, 52, 150). Here, 150 is the timestamps in which 3D data (53, 64, 52) are volumized. We get various angles such as axial, coronal and sagittal 2D pictures from these $(3D + t)$ data.

3.3.1 Scrubbing

Using scrubbing method, the excessive motioned frames were removed from the dataset. Basically, frames with frame displacement greater than 0.5mm were identified with excessive motion and removed for further pre-processing.

3.3.2 Spatial Smoothing

To improve the signal to noise ratio and spatial normalization, the fMRI volumes of the data were smoothed with a 6 mm isotropic Gaussian blurring kernel. The improvement of signal to noise ratio increases sensitivity. But there is a drawback which is, it reduces spatial resolution which may deduct any important feature.

3.3.3 Brain Extraction

Brain extraction is the first and most important step in MRI analysis. Removing non-brain tissues to improve accuracy and analysis speed of data is essential and that is why brain extraction is needed. Brain extraction tool of FMRIB Software Library (FSL) can be used in this case. For our dataset, the data already have extracted brain MR images, so we didn't need to do this pre processing part.

3.3.4 Slice Timing Selection

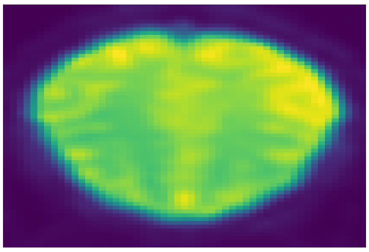
The overall size of the dataset is huge. As the changes over time is very subtle, so, for our research, we only took 2D images from timestamp 10 to 12 and 8 informative slices from all three regions as in these slices in each region, the differences between normal control and SZ patient is clear. We have manually analysed each slice from each subject with the help of MRIcron software.

3.3.5 Grayscaling

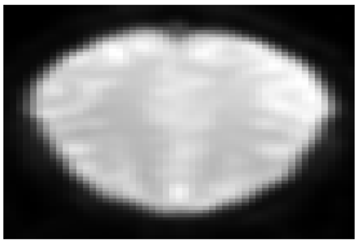
We have kept our original 2D MRI data in grayscaled image for reducing complexity. An RGB image is the combination of three different color images (Red, Green, Blue) stacked on top of each other. RGB color is a m X n X 3 array of color pixel.

[4]NIAK https://github.com/SIMEXP/niak

The range of values a colour component plane can have is (0 – 255). So, 256 X 256 X 256 = 16777216 combination of colour can be represented in an RGB image. On the other hand, grayscale image has only one channel which can have (0-255) or 256 values. Grayscale image's total possible values is one third of the RGB image's total possible values. Color increases the compution time and complexity of the model. In some model, there is no need for color images such as ours. We only need the intensity points to classify the images. That is why, we have kept our images into grayscale images. Figure 3.17a and figure 3.17b shows the conversion from colored MRI image to grayscale image.



(a) Colored image

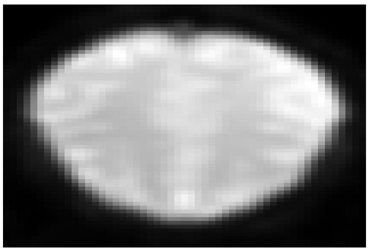


(b) Grayscale image

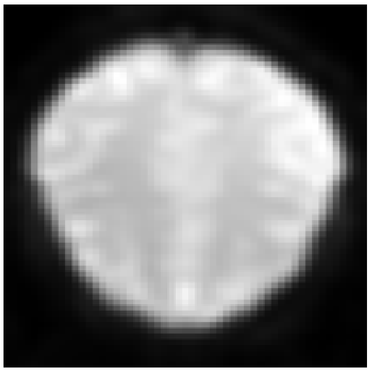
Figure 3.17: colored vs grayscaled MR image

3.3.6 Reshaping

We have resized our grayscale images to [224, 224] pixel as our models take [224, 224] or [256, 256] pixel images as input data. Figure 3.18 shows the resizing procedure of an image.



(a) Before reshaping



(b) After reshaping

Figure 3.18: Before vs after reshaping to 224 X 224 image

3.3.7 Data Augmentation

To maximize the diversity of images, we set horizontal flip to true. To normalize the data, we also rescaled our images to 1/255. Moreover, we set shear range and zoom range to 0.2 to increase the variety of training images so that we can get better accuracy.

3.4 Dataset Split

We have split the entire dataset[5] into training and testing in 80:20 ratio for Inception-V3, VGG16 and VGG19 models. Among of total 146 controls, for training part, healthy controls are in number of 60 and SZ patients are in number of 60 and the remaining are for testing part. Then we have perfomred image pre-processing and after image pre-processing, among total of 1314 2D MR images of all 146 of controls, 1080 images are of training part and 234 images are of testing part. Table 3.2 shows the size of dataset details in training and testing for Inception-V3, VGG16, VGG19 models.

Class	Normal	SZ	Total Images
Training	540	540	1080
Testing	126	108	234

Table 3.2: Number of images in training and testing set

We have split the entire dataset into training and testing in 75:25 ratio for our modified multichannel 2D CNN model. Among of total 146 controls, for training part, healthy controls are in number of 55 and SZ patients are in number of 54 and the remaining are for testing part. Then we have perfomred image pre-processing and after image pre-processing, among total of 2336 2D MR images of all 146 of controls in each region, 1744 images are of training part and 592 images are of testing part. Table 3.3 shows the size of dataset details in training and testing for modified multichannel 2D CNN model.

Class	Region	Normal	SZ	Total Images
Training	Axial	880	864	1744
Training	Coronal	880	864	1744
Training	Sagittal	880	864	1744
Testing	Axial	304	288	592
Testing	Coronal	304	288	592
Testing	Sagittal	304	288	592

Table 3.3: Number of images in training and testing set for M2D model

[5] http://cobre.mrn.org/

Chapter 4

Methodology

4.1 Convolutional Neural Network

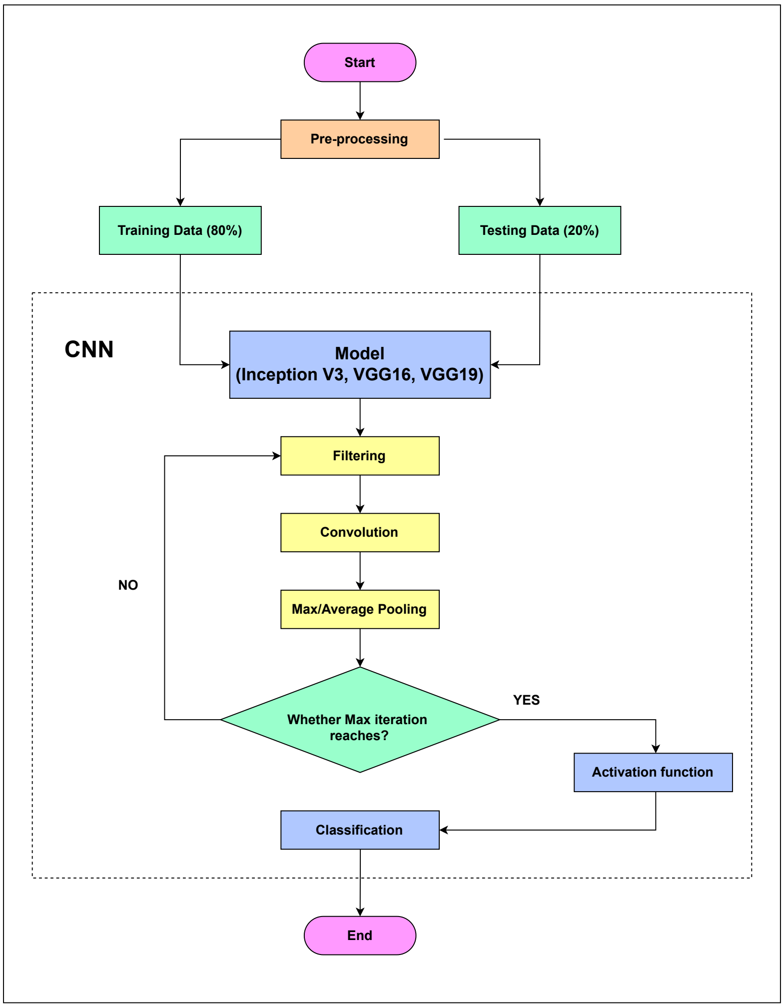


Figure 4.1: Workflow of CNN

Convolutional Neural Network(CNN) is a well known method in computer vision application which is a part of Deep Neural Network. This neural network analyze visual images and classify them, identifies texts, videos, documentation by recognising patterns[29]. In CNNs, artificial neurons take inputs from objects, then process them and send the result as output. CNNs use multiple layers which perform feature extraction. During convolutional layer, the network extracts the significant features from the image or dedicated objects and remove the irrelevant noise. Convolutional Neural Network(CNNs) have four steps to process an image or data. Figure 4.1 shows the workflow of CNN algorithm. The steps of the CNN algorithm is described below in details.

4.1.1 Convolutional Layer

The main motive of this layer is to stretch out features of the objects or images. In this layer the network targets to resize the images to reduce some pixels or unnecessary weights which are predicted as noise to minimize the computational time. Convolution is an element-wise multiplication which is also called matrix multiplication that identifies a dimension of 3X3 part of the image. Then it multiplies this part with a filter. Until all of the images processed these steps repeat.

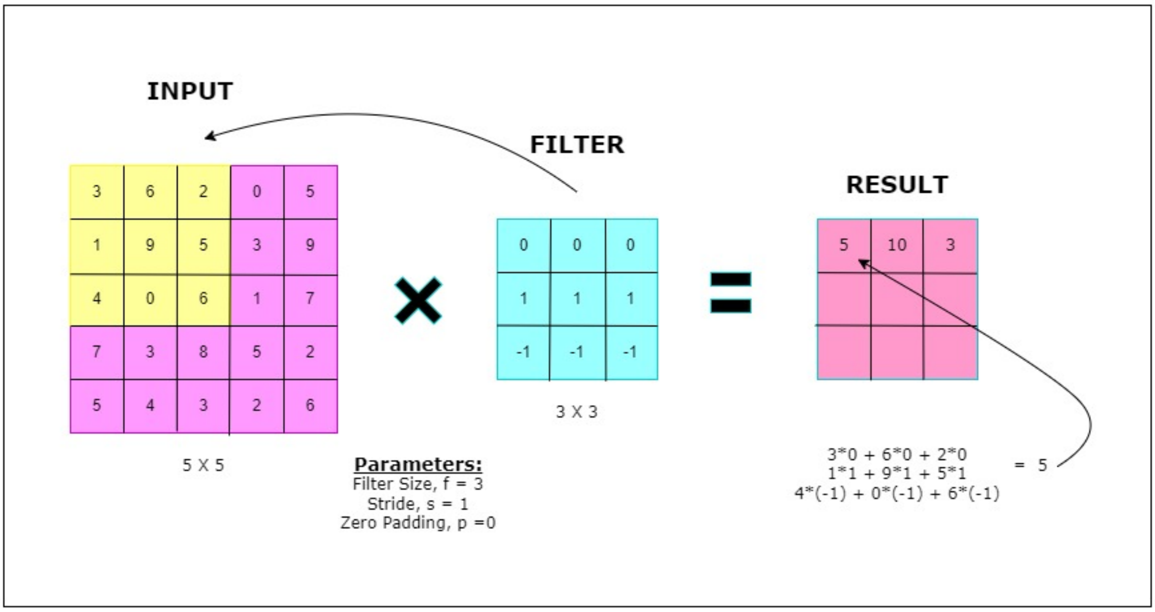


Figure 4.2: Pictorial representation of Convolutional Layer

In the Convolutional Neural Network, we calculate the volume of image. Here in the above 4.2 we showed an input image with a shape of 6 × 6 and a filter of shape 3 × 3. Then doing matrix multiplication we got 4 × 4 shape output.
From an input volume of Dimension $H_{in} \times W_{in} \times$ D (where, H= Height, W=Width, D=Depth) we can use the following formula to compute the result volume:

$$W_{out} = \frac{W_{in} - F + 2P}{S} + 1 \tag{4.1}$$

Here,

W_{out} = Updated Width
W_{in} = Input Width
F = Spatial Dimension
S = Stride
P = Padding

From this equation (4.1), the network generate the output volume with the dimensions ($H_{out} \times W_{out} \times D$).
Here,
$H_{out} = W_{out}$
$D(depth)$ = The number of filters Kernel (K).

4.1.2 Activation Layer - Non Linearity

After processing the each convolution layer, the activation layer comes and it represents the non linear function. In this layer all the pixel with negative value is replaced by zero.According to a research, we found that activation function does not hamper the gradients during back-propagation [30]. Mainly, four types of non linear functions are used in CNNs which are Rectified Linear Unit(ReLU), sigmoid, Tanh and Softmax. Activation Layer succeed the following formulas:

For ReLU Activation Function:

$$Output = Max(zero, Input) => \phi(x) = Max(0, x) \tag{4.2}$$

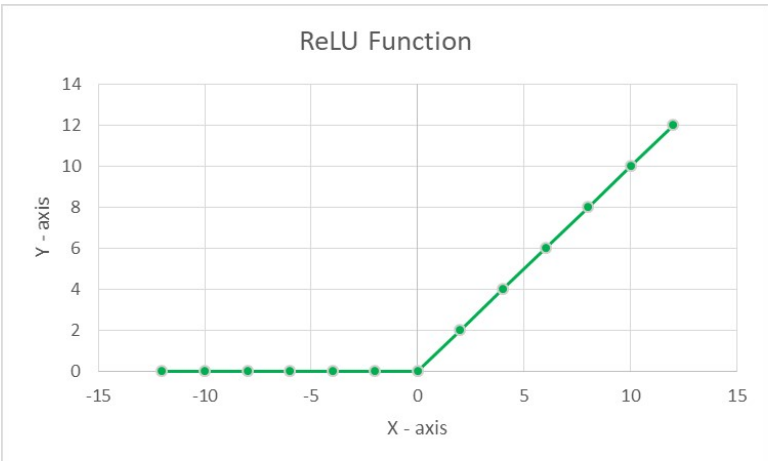


Figure 4.3: Pictorial representation of activation layer ReLU

From this equation (4.2), we can say that this function returns the input value of x back when the value is positive, but it returns 0 when the input value is negative.

So, basically it gives the output rang of 0 to infinity.

For Sigmoid Activation Function:

$$f(x) = \frac{1}{(1 + e^{-x})} \tag{4.3}$$

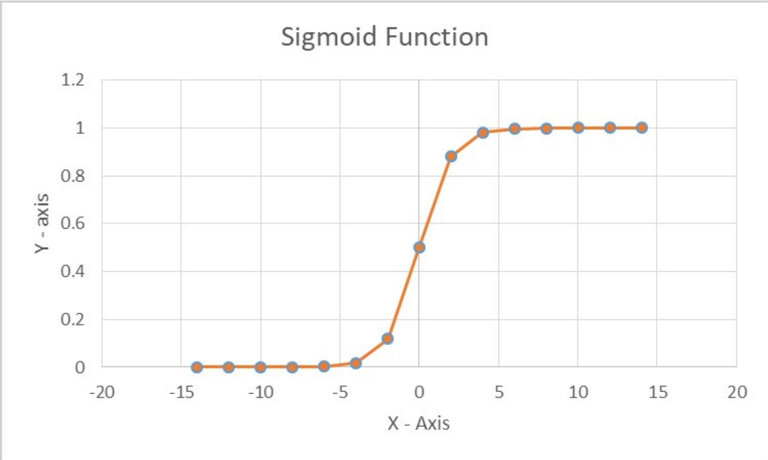


Figure 4.4: Pictorial representation of activation layer Sigmoid Function

Sigmoid Activation Function transform all the input values in range of 0 to 1. It transform all the values into 0 which are below than 0 and the above value of 1 turned into 1 by this function. In the equation (4.3), x is the input value and f(x) is the output value which always gives the output in range of 0 to 1.

For Tanh Activation Function:

$$tan(h) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{4.4}$$

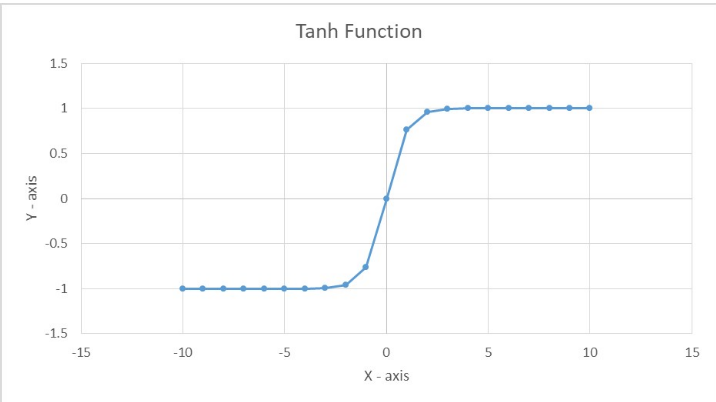


Figure 4.5: Pictorial representation of activation layer Tanh Function

For Softmax Activation Function:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \tag{4.5}$$

Here,

σ = softmax function
$\vec{z}$ = input vector
e^{z_i} = standard exponential function for input vector
K = number of classes in the multi-class classifier
e^{z_j} = standard exponential function for output vector
e^{z_j} = standard exponential function for output vector

The ideal activation functions should have the following features [31]:

- Non-linearity: It allows the following layers to elaborate on each other. This assure that non-linear multi-layer network will not degrade into single layer linear network.
- Differentiability: As more layers are stacked so gradient should be smaller and optimized.
- Simplicity : As a complex activation function takes more calculation time. So it should be simple.
- Saturation: For the saturation the gradient is close to zero in particular intervals to update the parameters impossible.
- Fewer Parameters: Most activation functions should have few parameters.

4.1.3 Pooling

Basic Convolution Network architecture have alternating conv layers and pooling layers and some functions to reduce the dimensionality of each feature map to figure out important information. There are many types of Spatial Pooling. Average Pooling, Max Pooling, Sum Pooling etc. are the some of pooling types [32].

In case of Average Pooling, it takes the average value from the rectified feature map. For instance, If the window is 2 $\times$ 2 Then after average pooling we will get 1 value from the average of 4 values. So it reduces the computational time of the network. Besides, The Max Pooling takes the maximum value and then work like as Average Pooling. Also we could use sum of all elements of the allocated window.

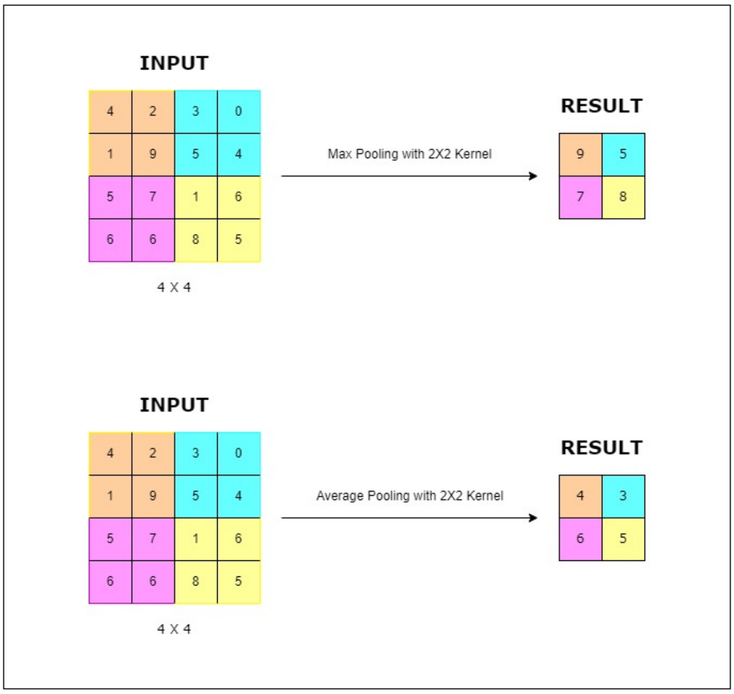


Figure 4.6: Pictorial representation of max pooling and average pooling

Advantages of Pooling:

- Reduce feature dimension and make data more manageable.
- This layer shorten the computational cost and reduce number of parameters which also control overfitting.
- As we use max pooling or average pooling in a local neighborhood it transform the large data into small data without loosing any important data. But these small data help a network to work properly.

4.1.4 Normalization

Convolutional Neural Network the part of Deep Learning have developed the state-of-art to perceive the images such as image segmentation, image classification, scene description etc. To make the network train faster and give better accuracy result normalization plays an important role which is a most popular method.In this layer data is being scaled in a range of 0 to 1. There are many types of normalization[33]. They are:

- Batch Normalization
- Weight Normalization
- Layer Normalization
- Group Normalization
- Weight Standardization

One important thing is Batch Normalization performs its functionality in different ways during training and interface [30]. Moving mean and variance are determined for each batch separately during the training phase. Instead, the network that we employed at the moment the interface was happening saw the moving mean and variance of all the photos concurrently [30]. So lots of research and experimentation of the state-of-art of network shows that batch normalization not only gives us better performance and accuracy but also it is really required for training part to get a proper output. Without batch normalization, sometimes deep learning fail to train the model[30]. Basically normalization is a pre-processing technique to standardize input data to get a modified value which can be trained by the model[1].

As there are some disadvantages of batch normalization, T.saliman and P. Kingma proposed Weight Normalization[33]. The idea was separate the length of the weight from the direction and change the parameter of the network to make fast training[33]. Though it speed up the training but Weight Normalization is not stable as Batch Normalization. That is why it is not widely used in neural network.

4.1.5 Fully connected (FC) Layer

In general, Convolutional Neural Network is one or more fully connected layers which is a multi layer perceptron. In this network, all the neurons of previous layer are connected to every single neurons of current layer. So, all the information of data is transformed through this connection. The last fully connected layer of this CNN is known as output layer. For tasks classification, Softmax activation is used to give better performance probability of the outputs[34]. Furthermore, SVM is also used to solve different classification of tasks[35].

4.2 CNN Model Architectures

We have trained in total six CNN model architectures for SZ MRI image classification. But in paper, we have described best three models. The descriptions of the models are given below.

4.2.1 Inception-V3 Model

Inception-V3 is an image recognition model of CNN which is an optimized and advanced version of Inception-V1. The accuracy of this model on the image net dataset is 78.1%. This model is based on the paper: "Rethinking the Inception Architecture for Computer Vision" by Szegedy, et. al [36].

Architecture

The inception model employs numerous filters of various sizes at the same level. Therefore, we have parallel layers rather than deep layers. We applied batch normalization and softmax activation function at each convolutional layer. The loss is calculated using the Softmax activation function. This model employs four methods.

[1] https://www.baeldung.com/cs/batch-normalization-cnn

Facorial convolution comes first. Therefore, the model becomes computationally effective. Moreover, the network generates fewer parameters. Secondly, the model swaps out larger convolutions for smaller ones. The training is expedited as a result. For instance, two 3X3 filters can be used in place of a 5X5 filter and so the parameters are decreased from 25 to 18. Asymetric convolution is the third method. 1X3 and 3X1 filters are used in place of 3X3 filters. This further narrows the parameters. Auxiliary classifier comes in fourth. This has been added inside the layer. Then, the main network loss is increased by the consequent loss. Lastly, grid size reduction which is done by pooling method.

Implemented Model

For our data training, firstly, input layer of shape (None, 224, 224, 3) is used. Then, 2D convolution layer of shape (None, 111, 111, 32) is applied. After that, Batch Normalization is done which is of (None, 111, 111, 32) shape. Consequently, activation function is implemented of size (None, 111, 111, 32). Again, 2D convolution layer of shape (None, 109, 109, 32) is used. Next, Batch Normalization is done which is of (None, 109, 109, 32) shape. Then, again, activation function is implemented of size (None, 109, 109, 32). Another Conv2D layer is implemented of shape (None, 109, 109, 64).Next Batch Normalization is done of shape (None, 109, 109, 64). Activation function was again implemented of size (None, 109, 109, 64). Later, 2D Max Pooling layer is done of shape (None, 54, 54, 64). Then, Conv2D layer is implemented of size (None, 54, 54, 80). Again, Batch Normalization is done of shape (None, 54, 54, 80). Repeating these processes multiple times with different shape, we finally concatenated the layers and flattened them. After flattening, we get a dense layer of shape (None, 2). While applying the model, total trainable parameters were of 102,402. We have used Softmax activation function in our dense layer. We have also used Categorical crossentropy as loss function, Adam as optimizer and Accuracy as metrics. As for the weights' values, we have used ImageNet dataset. We trained our model in batch size of 64 for 20 epochs.

Formula for categorical cross entropy is,

$$L_i = -\sum_j y_{i,j} log(\hat{y}_{i.j}) \tag{4.6}$$

Here,
L_i= Sample loss value
i = ith sample in a set
j = label/output index
y = target values
$\hat{y}$ = predicted values

4.2.2 VGG16 Model

We have used 2D CNN architecture - VGG16 model to classify our MRI image data. In their publication "Very Deep Convolutional Networks for Large-Scale Image Recognition", K. Simonyan and A. Zisserman from the University of Oxford introduced the convolutional neural network model known as VGG16[37]. In the top five tests, the model performs 92.7% accurately in ImageNet, a dataset of over 14 million images divided into 1000 classes. It was a well-known model that was submitted to ILSVRC-2014. By sequentially substituting several 33 kernel-sized filters for AlexNet's big kernel-sized filters (11 and 5, respectively, in the first and second convolutional layers), it improves upon AlexNet. Weeks of training were put into VGG16 using NVIDIA Titan Black GPUs.

Architecture

The RGB image with a fixed size of 224X224 is the input to the convolution layer. Convolutional filters were applied with a very small receptive field, 33 (which is the smallest size to capture the notion of left/right, up/down, and center). The picture is then passed through the stack of convolutional layers. It also uses 11 convolution filters in one of the setups, which may be thought of as a linear transformation of the input channels (followed by non-linearity). Convolution layer input's spatial padding retains the spatial resolution.

Implemented Model

Initially, input layer of shape (None, 224, 224, 3) is used. Then, two 2D convolution layer is used of shape (None, 224, 224, 64). After that, A 2D Max Pooing layer is applied of shape (None, 112, 112, 64). Consequently, two 2D convolution layer layers are used of shape(None, 112, 112, 128). Again, A 2D Max Pooing layer is applied of shape (None, 56, 56, 128). Later on, three 2D convolution layer layers are used of shape (None, 56, 56, 256). 2D Max Pooing layer is applied of shape (None, 28, 28, 256). Next, three 2D convolution layer layers are used of shape (None, 28, 28, 512). After this, 2D Max Pooing layer is applied of shape (None, 14, 14, 512). Next, three 2D convolution layer layers are used of shape (None, 14, 14, 512). Then, 2D Max Pooing layer is applied of shape (None, 7, 7, 512). Finally, after flattening, we get the dense layer of size (None, 2). In our applied model, we had 50,178 trainable parameters. The model had 13 2D convolution layers along with 5 max pooling layers. We have used Softmax activation function in our dense layer. We have also used Categorical crossentropy as loss function, Adam as optimizer and Accuracy as metrics. As for the weights' values, we have used ImageNet dataset. We trained our model in batch size of 64 for 20 epochs.

4.2.3 VGG19 Model

A convolutional neural network with 19 layers is called VGG-19. A variation of the VGG model called VGG19 has 19 layers in total (16 convolution layers, 3 Fully connected layer, 5 MaxPool layers and 1 SoftMax layer). There are further VGG variations, including VGG11, VGG16, and others. 19.6 billion FLOPs make up VGG19.

Architecture

VGG19 model receives a fixed-size (224X224) RGB picture as input, indicating that the matrix was shaped (224, 224, 3). The mean RGB value of each pixel, calculated throughout the whole training set, was the only preprocessing that was carried out. They were able to cover the entirety of the image by using kernels that were (3X3) in size with a stride size of 1 pixel. To maintain the image's spatial resolution, spatial padding was applied. Stride 2 was used to conduct max pooling over a 2X2 pixel window. This was followed by the Rectified Linear Unit (ReLu), which introduced non-linearity to increase the model's ability to categorize data and shorten computation time. Rectified linear unit (ReLu) was used after this to add non-linearity to the model in order to enhance classification accuracy and computation time. As opposed to earlier models that used tanh or sigmoid functions, this one performed far better. Implementing three fully linked layers, the first two of which had a size of 4096, followed by a layer with 1000 channels for classification using the 1000-way ILSVRC, and the third layer being a softmax function.

Implemented Model

We have used modified 2D CNN architecture - VGG19 model to classify our MRI image data. Initially, input layer of shape (None, 224, 224, 3) is used. Then, two 2D convolution layers are used of shape (None, 224, 224, 64). After that, a 2D Max Pooing layer is applied of shape (None, 112, 112, 64). Consequently, two 2D convolution layers are used of shape (None, 112, 112, 128). Again, A 2D Max Pooing layer is applied of shape (None, 56, 56, 128). Later on, four 2D convolution layers are used of shape (None, 56, 56, 256). 2D Max Pooing layer is applied of shape (None, 28, 28, 256). Next, four 2D convolution layers are used of shape (None, 28, 28, 512). After this, 2D Max Pooing layer is applied of shape (None, 14, 14, 512). Next, four 2D convolution layers are used of shape (None, 14, 14, 512). Then, 2D Max Pooing layer is applied of shape (None, 7, 7, 512). Finally, after flattening, we get the dense layer of size (None, 2). In our applied model, we had 50,178 trainable parameters. The model had 16 2D convolution layers along with 5 max pooling layers. We have used Softmax activation function instead of ReLu function in our dense layer. We have also used Categorical cross entropy as loss function, Adam as optimizer and Accuracy as metrics. As for the weights' values, we have used ImageNet dataset. We trained our model in batch size of 64 for 20 epochs.
Mathematically RelU function is expressed as,

$$f(x) = max(0, x) \tag{4.7}$$

4.3 Multichannel 2D CNN Model

CNN, a subclass of DNN, has been considered to be particularly effective at extracting features. To categorize 3D fMRI data, Multichannel 2D CNN model[23] can be used. The model incorporates multichannel information gained from 2D CNN networks and works with 2D fMRI data slices. In task-evoked fMRI scanning [23], candidates are assigned with few task stimuli and simultaneously engaged in particular actions that produce distinct BOLD signals which outputs time series of a 3D volume of the brain. This volume consists of millimeter level spatial resolution for each task block. Task evoked fMRI data or resting state fMRI data can be presented as 3D data with coronal, sagittal, and axial axes of the brain. Comparing several CNN models, we observe that 3D CNN can perform well in extracting information from 3D fMRI data, but it is computationally expensive and requires exponentially more model parameters as each additional dimension is added. Additionally, supervised learning can suffer from data overfitting due to the use of too many parameters. In the meantime, to facilitate computing, models with separable 3D CNN has been developed in computer vision research. However, 3D SepConv model can be less efficient due to usage of less parameters as compared to the general 3D CNN model. Contrarily, 2D CNN models lose few spatial information of 3D fMRI data for using fewer parameters and less processing power. After addressing issues of various CNN models, we are proposed with M2D CNN. This model has two phases. Firstly, 3D fMRI images are sliced into 2D fMRI images along with one dimension. The 2D CNN model receives multichannel 2D images as input. This model is used to extract feature from these three brain anatomical planes- coronal as x-plane, sagittal as y-plane, and axial as z-plane. Secondly, fully connected hidden layer is used to concatenate the information from M2D CNN learning network. Each 2D CNN performs convolution computation independently using a single type of multichannel 2D picture as an input. All three 2D CNN segments' outputs are flattened and joined together to create 1D feature and passed to fully connected neural network which outputs the categorization outcome. The speciality of this model is that this model considers the 3D spatial information of the brain as the features include the characteristics derived from all three orthogonal planes.

4.3.1 M2D CNN Model Architecture

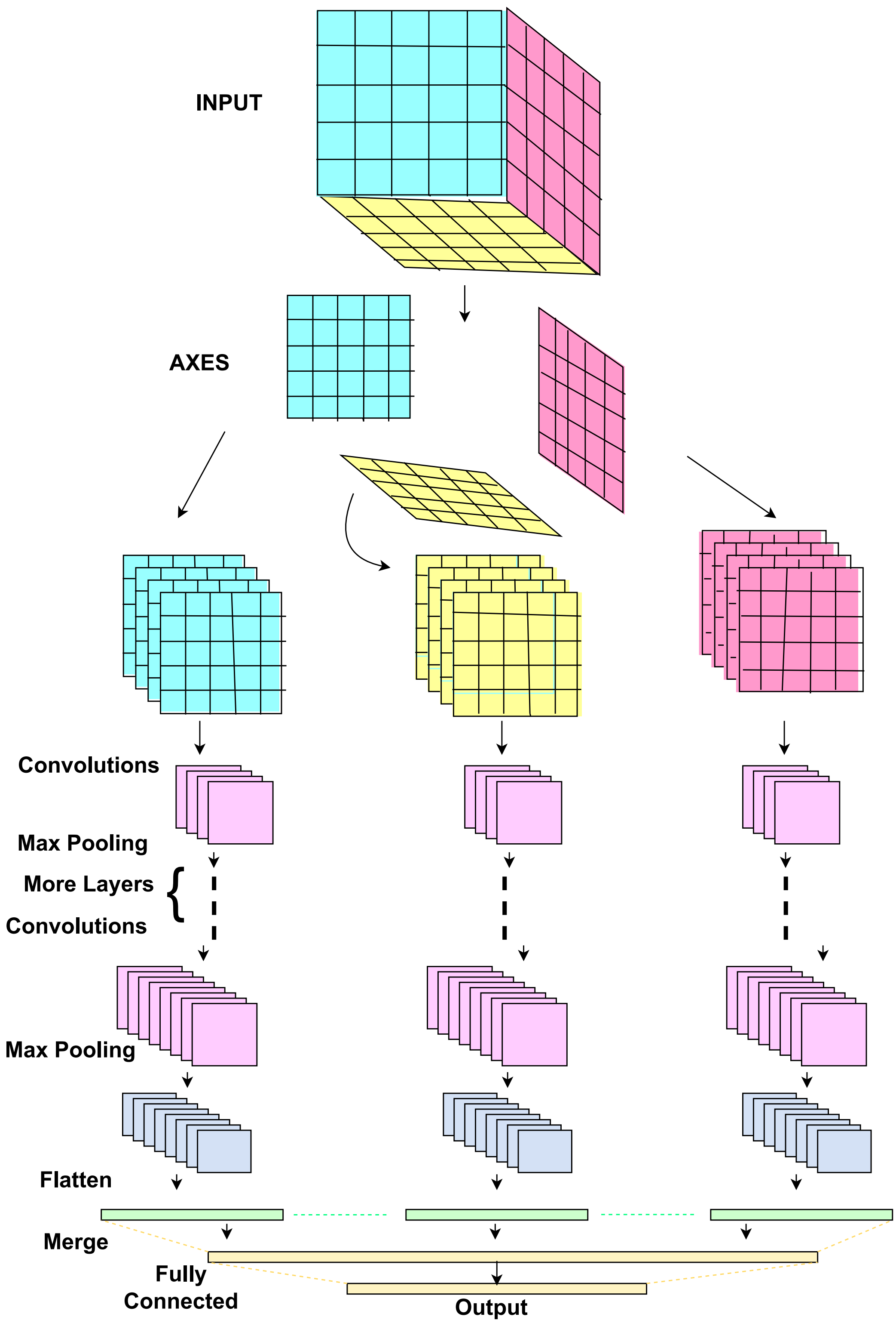


Figure 4.7: Architecture of M2D CNN

This architecture consists of input layers, convolutional layers, pooling layers, merge layers, fully connected layers, and output layers.

4.3.2 Input Layer

Each 3D fMRI brain imaging sample has dimensions of dim X, dim Y, and dim Z. By slicing per unit length along the y axis to obtain 2D images of size dim X x dim Z on the sagittal plane, along the z axis to obtain 2D photos of size dim X dim Y, and along the x axis to obtain 2D images of size dim Y x dim Z, we obtain three groups of images in separate planes. As a result, 2D pictures with dim X, Y, and Z channels are produced. These channels are similar to the stereostatic brain space.

4.3.3 Convolutional Layer

The convolution computation is processed by all the 2D CNNs, and features are extracted on each plane. From previous layer each convolutional kernel is convolved across the width and height of the 2D input and the dot product between the kernel and the input is computed. The results are summed up and 2D activation map is produced for each kernel.

4.3.4 Pooling Layer

Pooling layer filters out the results of convolutional layer and performs feature selection.

4.3.5 Merge Layer

Result of all the 2D CNNs are then concatenated. We get 1D vector as output. In other words, this output represents the merged spatial feature of brain data.

4.3.6 Fully Connected Layer

This layer takes merge layer as input. Also, feature combination and complex non linear relationship modelling are performed in this layer.

4.3.7 Output Layer

Softmax function (4.5) is used for output computation of each category that the sample belongs to.

4.4 Modified M2D Model Structure

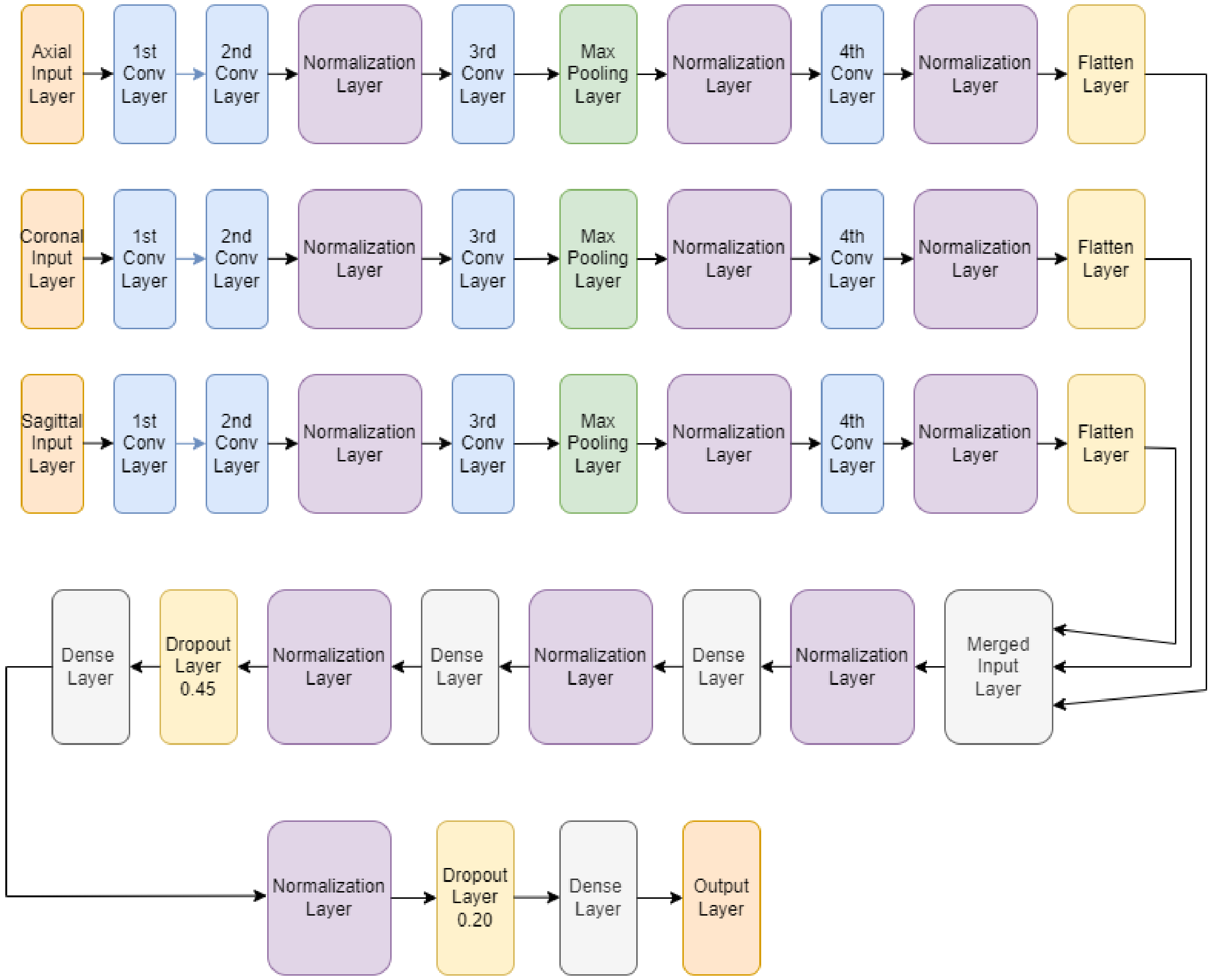


Figure 4.8: Structure of modified M2D CNN

We have passed our inputs through 3 different regions - axial input layer, coronal input layer and sagittal input layer. Then, our model has two consecutive convolution layers with 16 and 32 filters, (3x3) kernel and 'reLu' activation function. After that, we have one normalization layer. We put one convolution layer with 32 filters, (3x3) kernel and 'reLu' activation function after the normalization layer. Then, we put one (2x2) max pooling layer and another normalization layer. We have added another convolution layer with the previous setup. Then, we added normalization layer and finally flattened the 2D matrices through the flatten layer. 3 different regions have gone through these layers in parallel and finally merged into one merged layer and then we passed this merged matrix to a normalization layer and a dense layer with 256 units and 'sigmoid' activation function. After that, we have added another normalization layer and another dense layer with 128 units and 'sigmoid' activation function. We also used another normalization layer with a dropout layer of 0.45. Then, after another set of dense layer, normalization layer, dropout layer of 0.20 and dense layer, we finally get the output layer.

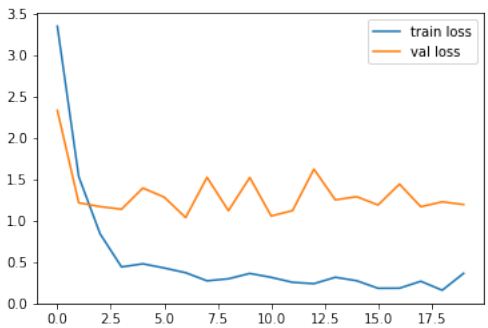
Chapter 5

Result and Analysis

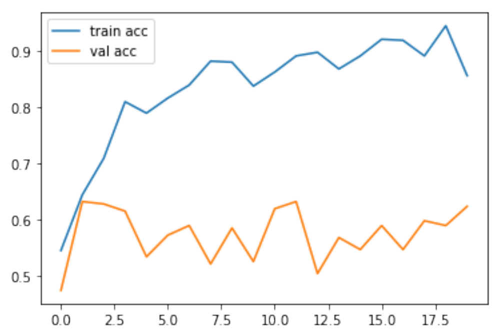
This chapter describes the implementation of some built in models that we used initially to check the performance of our taken dataset for detecting SZ and also narrates our proposed model that we have built to get better perfomance. Initially We used six models - InceptionV3, VGG16, VGG19, ResNet50, ResNet101 and ResNet152. Then we proposed multi channel 2D CNN (M2D CNN) model. We have used two dimensional data for three channels axial(xy), coronal(xz) and sagittal(yz). The models were implemented and tested using Colab notebook. The implementation of the models consist of four stages; input data pre-processing, feature extraction, classification and testing. Firstly, We have equipped our dataset[24] in two part - testing set and training set. In both training and testing part, we have created two classes for normal people and SZ patient where we have stored all MR images of both normal people and SZ patient.

5.1 Initially used Models

Inception V3 is an image classification model of CNN which is a built in model of Keras. From Inception-V3 model, in training, we got 94.54% accuracy and in testing, we got 63.25% acccuracy.



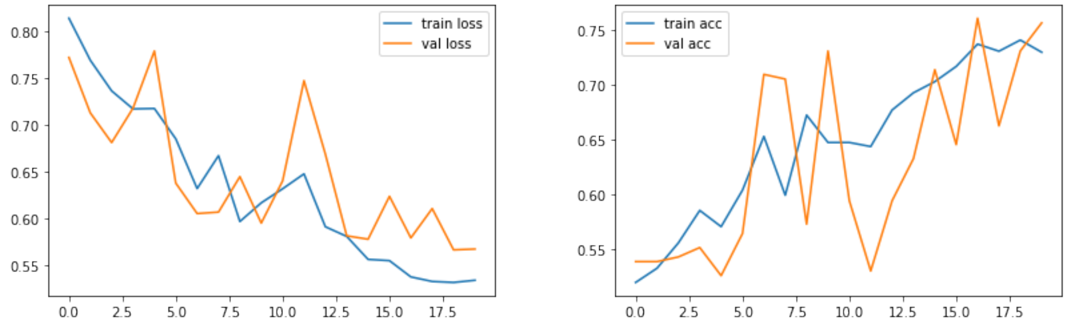
(a) Training loss vs validation loss



(b) Training accuracy vs validation accuracy

Figure 5.1: InceptionV3 model for 20 epochs

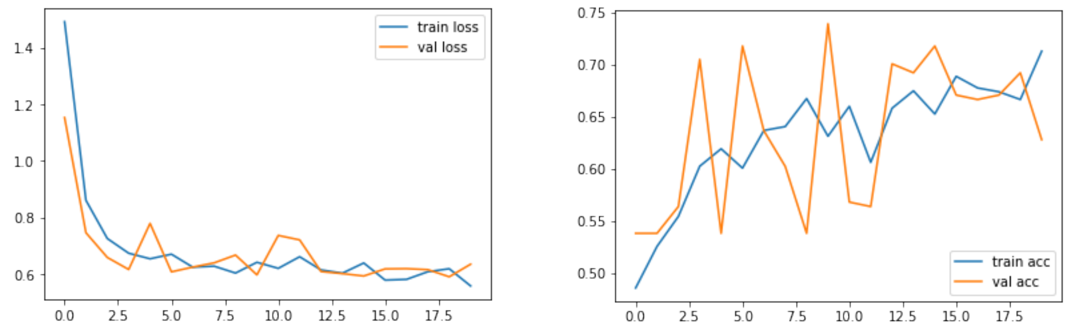
VGG16 is a method which identify the objects and it can reach 92.7% accuracy rate when classifying more than thousands photos into thousands different classes. It is a well-known technique which can classifying images and employ with transfer learning easily. But for our dataset?? From VGG16 model, in training, we got 74.07% accuracy and in testing, we got 76.07% accuracy.



(a) Training loss vs validation loss (b) Training accuracy vs validation accuracy

Figure 5.2: VGG16 model for 20 epochs

By using another built in model VGG19 , in training, we got 71.30% accuracy and in testing, we got 73.93% accuracy.



(a) Training loss vs validation loss (b) Training accuracy vs validation accuracy

Figure 5.3: VGG19 model for 20 epochs

We also trained our data in ResNet50, ResNet101 and ResNet152 model. From ResNet50 model, in training, we got 57.87% accuracy and in testing, we got 67.95% acccuracy. From ResNet101 model, in training, we got 60.37% accuracy and in testing, we got 66.67% accuracy. From ResNet152 model, in training, we got 58.06% accuracy and in testing, we got 58.97% accuracy. Table 5.1 shows the training and testing accuracy of our presented models in tabular form.

Model	Architecture	Training Accuracy	Testing Accuracy
2D-CNN	InceptionV3	94.54%	63.25%
2D-CNN	VGG16	74.07%	76.07%
2D-CNN	VGG19	71.30%	73.93%
2D-CNN	ResNet50	57.87%	67.95%
2D-CNN	ResNet101	60.37%	66.67%
2D-CNN	ResNet152	58.06%	58.97%

Table 5.1: Results of different model architectures

According to these six models we did not get a good performance to detect SZ. The highest accuracy of these six models is 76.07% which is not up to the mark. So we have designed an architecture of M2D CNN model which gives us far better result than these models.

5.2 Proposed Model M2D CNN

Our proposed model is multi channel 2D CNN. As previously, we have tried one channel 2D CNN where we ran the built in models on axial portion of brain. But we did not get a good accuracy. Because the functional MR images of brain basically carry 4D data which are x, y and z axis with time series. It is impossible for MRI scanners to provide a 3D image of the brain using 2D pixels. Instead, the MRI scanner used a 3D equivalent voxel of a pixel. A voxel resembles a small cube with 1-3 mm-wide edges. Then, structural MRI explore extremely high-resolution 3D image of each slice of the brain and fMRI explore low resolution 3D image. The blood oxygenation level dependent (BOLD) signal is measured by the MRI scanner using a powerful magnet.[38] For these reason we took multi channel 2D CNN where we used the data of three axes. We also figured out some specific slices of three axes which ca be differentiated easily with other slices. In our proposed model architecture we had to use the minimum number of max-pooling to get better result as our images were very small. Besides, we had used Kfold process to get better accuracy. For that we divided our dataset into 4 folds.

5.2.1 Accuracy of Our Proposed Model

This section will illustrate the accuracy of our proposed model. For training part we have taken 1744 images and 592 images for testing part. Then We used colab notebook to run our experiments. As we have used K-fold process, so we got 4 training accuracy and 4 test accuracy for 4 folds. Table 5.2 shows the training and testing accuracy of our proposed model architecture in each fold in tabular form.

Fold No	Fold 1	Fold 2	Fold 3	Fold 4	Best
Training Accuracy	93.63%	93.98%	96.04%	97.53%	97.53%
Testing Accuracy	78.12%	89.86%	92.57%	97.13%	97.13%

Table 5.2: Results of 4 folds of M2D model architectures

5.3 Comparison with Related Research Works

Many research works already have done on this dataset ??. We have studied some research papers where different models were used on this dataset to classify SZ and normal people. We have shown the comparisons of some of those papers in 5.3 which worked on this dataset.

Articles	Published Year	Models	Dataset	Accuracy
Our proposed model architecture	-	multi channel 2D CNN	COBRE	97.13%
Qureshi et al. [39]	2019	3D-CNN	COBRE	98.09%
Kim et al. [40]	2016	DNN	COBRE	95.4%
Patel et al. [8]	2016	SAE	COBRE	92%
Du et al. [41]	2012	Feature extraction method	COBRE	98%

Table 5.3: Related Works on COBRE Dataset

So, according to this table 5.3 we can see that one paper [39] used 3D-CNN to detect SZ from COBRE dataset. Another one [40] used DNN and the last paper [41] used feature extraction model where they basically assembled a two-level feature identification scheme with KPCA and FLD. Moreover, other research works had done with different models where our model architecture is less complex and the test accuracy is higher than some works. So it will help to detect SZ patient more easily. Most of the papers used full dataset whereas we have used a small portion and got a good accuracy by using some techniques such as used more specific slices and time stamps of data.
Furthermore, We have studied some papers where authors worked on detecting SZ and normal people by using both machine learning and deep learning on different datasets. According to those papers, Kardy et al. [42] constructed VGG16 model and used MRI slices to detect SZ. Yan et el. [43] suggested a multi-scale RNN model to classify SZ and normal people. Zeng et al. [44] proposed SAE model to optimized a multi site rs-fMRI dataset. Qureshi et al. [39] constructed 3D CNN based on ICA to detect the discriminitive features of SZ. We have compared those papers in table

5.4. Additionally, compared to these research works we can see that the accuracy of most of the works are between 83% to 95%. Only a few work's accuracy is 98.09%. Among all these works our model architecture's accuracy is better than most of the works.

Articles	Published Year	Subjects	Models	Training Set/ Testing Set	Accuracy
Our proposed model architecture	-	74 N, 72 SZ	multi channel 2D CNN	4 fold cross validation	97.13%
Kardy et al. [42]	2021	500 N slices, 500 SZ slices	2D-CNN	-	94.33%
Jo et al. [45]	2020	24 N, 48 SZ	SVM, RF, NB	10 fold cross validation	68.6%
Yan et al. [43]	2019	542 N, 151 SZ	RNN	5 fold cross validation	83%
Qureshi et al. [39]	2019	74 N, 72 SZ	3D-CNN	10 fold cross validation	98.09%
Bae et al. [46]	2018	54 N, 21 SZ	SVM	10 fold cross validation	92.1%
Zeng et al. [44]	2018	377 N, 357 SZ	SAE	5 fold cross validation	85%
Kim et al. [40]	2016	50 N, 50 SZ	DNN	5 fold cross validation	95.4%

Table 5.4: Related works on SZ detection

5.4 Result Analysis

Figure 5.4 represents the training and testing accuracy among all 2D CNN models those we have implemented. From the figure, we can glance that our constructed model architecture of multi channel 2D CNN gives the best testing accuracy among all the models. According to our model we got highest 97.53% training accuracy and 97.13% testing accuracy. Besides, our initial testing models VGG16 and VGG19 gives almost similar testing accuracy which is 76.07% and 73.93% respectively. Moreover, ResNet50 and ResNet101 gives similar accuracy which is 67.95% and 66.67% respectively. InceptionV3 and ResNet152 model gives lowest testing accuracy out of all the models. In terms of training accuracy, InceptionV3 model gives the highest of training accuracy which is 94.54%. And for the other models, training accuracy is lower than testing accuracy which implies that the number of our input data is not sufficient enough to train the models accurately which leads to under fitting. So keep it in our mind we have built a model architecture which has no under fitting issues. In this case, we also used only two time stamp of MR images instead of 150 time stamp as our dataset has 150 time stamp. According to our dataset[1], our pre-processed dataset is size of 8 GB where we have used a small part with only some slices of three axes, axial, coronal and sagittal. Though we have used small data, our own model architecture gives a high training and testing accuracy. It will reduce the both time and hardware complexity. So our proposed model architecture is more machine friendly.

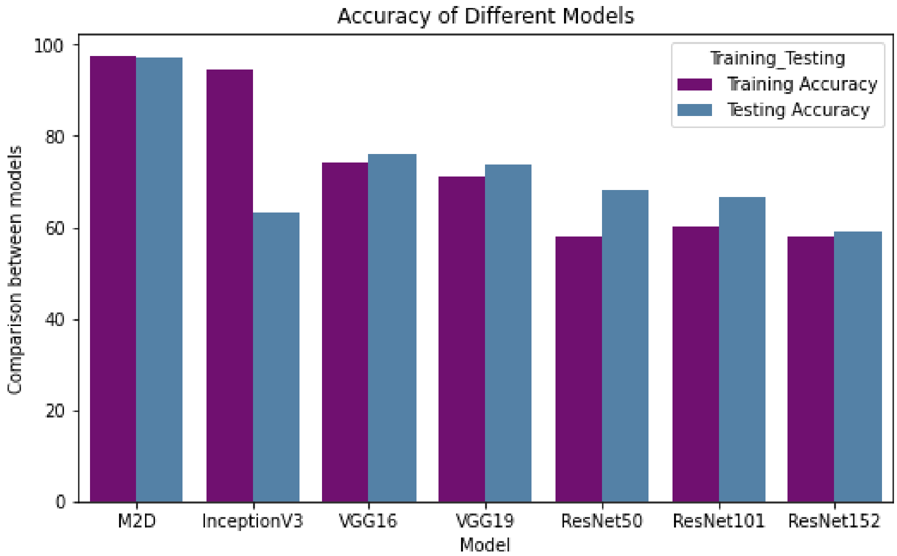


Figure 5.4: Comparison between different 2D CNN models' accuracy

According to the histogram of 5.5 we can see some differences of four folds. Basically we have divided our data in four parts and used different unknown test data to see where we could get better accuracy. As we have used colab notebook for running the code, we could not use large dataset for GPU limitations. Moreover our dataset contains functional MR images which has very low resolution. So it is difficult to detect data properly. Further we have to face some over-fitting issues for using this small data. So to reduce some over-fitting issues and increasing the test accuracy we used kfold methods. From this bar chart 5.5 we got highest accuracy from fourth fold.

[1]http://cobre.mrn.org/

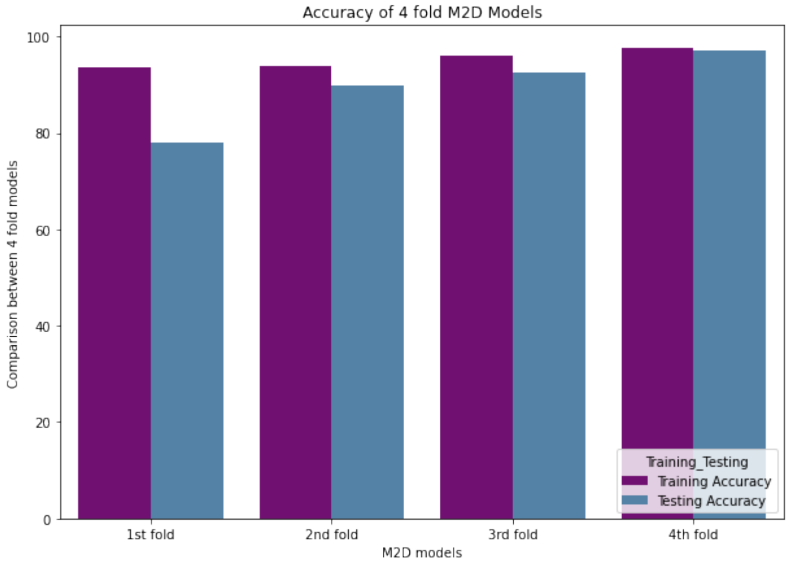


Figure 5.5: Comparison between 4 fold M2D CNN models' accuracy

Chapter 6

Conclusion

The proposed multichannel 2D CNN model outperforms all the other previously used built in models because of using sliced images of selective timestamps , taking into account the 3D information of whole brain contained in fMRI data and changing the architecture by reducing the number of maxpooling layers. The voxel feature of 3D images is preserved since we have used three 2D multichannel CNN in parallel and integrated them using fully connected hidden layer. The limitations that we have faced is overfitting in both validation loss and validation accuracy. This is beacuse we have taken a small part of the entire dataset due to hardware limitations. Moreover, by doing this work we came to understand that our dataset?? is not as much good as it contains images with low resolution which cannot not be figured out by using pre-processing. So it became difficult to detect SZ from MRI images. As our dataset[1] couldn't show a clear MR image, initially we have tried to collect data from hospital. But they denied us the permission for security reasons. Besides, we faced another limitation which is Hardware limitation. To process this large dataset, we need high configured GPU. In case of the shortage of GPU, we could not take all the images of the COBRE dataset to check accuracy of those models we have used.

6.1 Future Plan

By using the six model architectures, we have got moderate accuracy but not satisfactory. So we have tried to solve these problems as given below to get better accuracy.

- We have built a model architecture using the concept of multichannel 2D CNN in order to increase the accuracy.
- We had added necessary activation functions, dropout layers.
- We have increased the number of input data such as- we have taken 3 regions' 8 image slices from one region's 3 slices' to avoid underfitting.

Some of the planned things given below we have not tried yet which we plan to do in future. The points below may stable and increase the accuracy of SZ detection.

- We will modify more specified slices of brain MR images to detect SZ precisely.

[1]http://cobre.mrn.org/

- We have tried 2D CNN with 2D MR images. In future, we will try 3D CNN with 3D MR images.
- We will try to use another dataset such as NUSDAST alongside COBRE dataset as our dataset is kind of blurry and the features are not quite clear. We will again try to collect clinical data from hospitals.
- We will apply further pre-processing of our data so that we can extract as much features possible from the images.
- We will try to collect clinical data and make our own dataset.
- We will select images from the COBRE dataset, containing clearer features and make a different dataset with those specific images and work on that dataset to get rid of overfitting.

To conclude, we hope that 3D structural or functional MR images will give us higher accuracy to detect SZ. Moreover, we hope that by implementing these future plans mentioned above, we can achieve our goal which is to detect sz at an early stage with greater accuracy and make the process cost effective and hassle-free.

Bibliography

[1] Yeqing Wu et al. “Epidemiology of schizophrenia and risk factors of schizophrenia-associated aggression from 2011 to 2015”. In: *Journal of International Medical Research* 46 (Aug. 2018), p. 030006051878663. DOI: 10.1177/0300060518786634.

[2] Jihoon Oh et al. “Identifying Schizophrenia Using Structural MRI With a Deep Learning Algorithm”. In: *Frontiers in Psychiatry* 11 (Feb. 2020). DOI: 10.3389/fpsyt.2020.00016.

[3] Muhammad Naveed Iqbal Qureshi, Jooyoung Oh, and Boreom Lee. “3D-CNN Based Discrimination of Schizophrenia Using Resting-State FMRI”. In: *Artif. Intell. Med.* 98.C (July 2019), pp. 10–17. ISSN: 0933-3657. DOI: 10.1016/j.artmed.2019.06.003. URL: https://doi.org/10.1016/j.artmed.2019.06.003.

[4] Muhammad Naveed Iqbal Qureshi et al. “Multimodal discrimination of schizophrenia using hybrid weighted feature concatenation of brain functional connectivity and anatomical features with an extreme learning machine”. In: *Frontiers in neuroinformatics* 11 (2017), p. 59.

[5] Karen Simonyan and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition”. In: *arXiv preprint arXiv:1409.1556* (2014).

[6] Manohar Latha and Ganesan Kavitha. “Detection of Schizophrenia in brain MR images based on segmented ventricle region and deep belief networks”. In: *Neural Computing and Applications* 31.9 (2019), pp. 5195–5206.

[7] Yue Qiu et al. “Classification of Schizophrenia Patients and Healthy Controls Using ICA of Complex-Valued fMRI Data and Convolutional Neural Networks”. In: June 2019, pp. 540–547. ISBN: 978-3-030-22807-1. DOI: 10.1007/978-3-030-22808-8_53.

[8] Pinkal Patel, Priya Aggarwal, and Anubha Gupta. “Classification of schizophrenia versus normal subjects using deep learning”. In: *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing.* 2016, pp. 1–6.

[9] Darya Chyzhyk, Alexandre Savio, and Manuel Graña. “Computer aided diagnosis of schizophrenia on resting state fMRI data by ensembles of ELM”. In: *Neural Networks* 68 (2015), pp. 23–33.

[10] Robert W Cox. “AFNI: what a long strange trip it’s been”. In: *Neuroimage* 62.2 (2012), pp. 743–747.

[11] Bruce Fischl. “FreeSurfer”. In: *Neuroimage* 62.2 (2012), pp. 774–781.

[12] Mark Jenkinson et al. “Fsl”. In: *Neuroimage* 62.2 (2012), pp. 782–790.

[13] Karl J Friston et al. "Movement-related effects in fMRI time-series". In: *Magnetic resonance in medicine* 35.3 (1996), pp. 346–355.

[14] Theodore D Satterthwaite et al. "Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth". In: *Neuroimage* 60.1 (2012), pp. 623–632.

[15] Yashar Behzadi et al. "A component based noise correction method (CompCor) for BOLD and perfusion based fMRI". In: *Neuroimage* 37.1 (2007), pp. 90–101.

[16] F Gregory Ashby. *Statistical analysis of fMRI data*. MIT press, 2019.

[17] Srivathsan Srinivasagopalan et al. "A deep learning approach for diagnosing schizophrenic patients". In: *Journal of Experimental & Theoretical Artificial Intelligence* 31.6 (2019), pp. 803–816. DOI: 10.1080/0952813X.2018.1563636. eprint: https://doi.org/10.1080/0952813X.2018.1563636. URL: https://doi.org/10.1080/0952813X.2018.1563636.

[18] Rowena Chin et al. "Recognition of Schizophrenia with Regularized Support Vector Machine and Sequential Region of Interest Selection using Structural Magnetic Resonance Imaging". In: *Scientific Reports* 8 (Sept. 2018). DOI: 10.1038/s41598-018-32290-9.

[19] Mengjiao Hu et al. "Structural and diffusion MRI based schizophrenia classification using 2D pretrained and 3D naive Convolutional Neural Networks". In: *Schizophrenia Research* (2021). ISSN: 0920-9964. DOI: https://doi.org/10.1016/j.schres.2021.06.011. URL: https://www.sciencedirect.com/science/article/pii/S0920996421002231.

[20] Elisa Veronese et al. "Machine Learning Approaches: From Theory to Application in Schizophrenia". In: *Computational and mathematical methods in medicine* 2013 (Jan. 2013), p. 867924. DOI: 10.1155/2013/867924.

[21] Tian Wang et al. "Multi-Kernel Capsule Network for Schizophrenia Identification". In: *IEEE transactions on cybernetics* PP (2020).

[22] Han Shaoqiang et al. "Recognition of early-onset schizophrenia using deep-learning method". In: *Applied Informatics* 4 (Dec. 2017). DOI: 10.1186/s40535-017-0044-3.

[23] Jinlong Hu et al. "A multichannel 2D convolutional neural network model for task-evoked fMRI data classification". In: *Computational intelligence and neuroscience* 2019 (2019).

[24] Pierre Bellec. "COBRE preprocessed with NIAK 0". In: *12* 4. (2015).

[25] Xiang Yang Zhang et al. "Gender difference in association of cognition with BDNF in chronic schizophrenia". In: *Psychoneuroendocrinology* 48 (2014), pp. 136–146. ISSN: 0306-4530. DOI: https://doi.org/10.1016/j.psyneuen.2014.06.004. URL: https://www.sciencedirect.com/science/article/pii/S030645301400211X.

[26] Norbert Müller. "Inflammation in Schizophrenia: Pathogenetic Aspects and Therapeutic Considerations". In: *Schizophrenia bulletin* 44 (Apr. 2018). DOI: 10.1093/schbul/sby024.

[27] Rei Wake et al. “Abnormalities in MRI Signal Intensity in Schizophrenia Associated with Idiopathic Unconjugated Hyperbilirubinemia”. In: *Australian & New Zealand Journal of Psychiatry* 43.11 (2009). PMID: 20001401, pp. 1057–1069. DOI: 10.3109/00048670903107526. eprint: https://doi.org/10.3109/00048670903107526. URL: https://doi.org/10.3109/00048670903107526.

[28] Pierre Bellec. “COBRE preprocessed with NIAK 0.12.4”. In: (Jan. 2015). DOI: 10.6084/m9.figshare.1160600.v15. URL: https://figshare.com/articles/dataset/COBRE_preprocessed_with_NIAK_0_12_4/1160600.

[29] Christian Szegedy et al. “Going deeper with convolutions”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1–9.

[30] Vignesh Thakkar, Suman Tewary, and Chandan Chakraborty. “Batch Normalization in Convolutional Neural Networks—A comparative study with CIFAR-10 data”. In: *2018 fifth international conference on emerging applications of information technology (EAIT)*. IEEE. 2018, pp. 1–5.

[31] Wang Hao et al. “The role of activation function in cnn”. In: *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*. IEEE. 2020, pp. 429–432.

[32] Neena Aloysius and M Geetha. “A review on deep convolutional neural networks”. In: *2017 international conference on communication and signal processing (ICCSP)*. IEEE. 2017, pp. 0588–0592.

[33] N Vijayrania. “Different Normalization Layers in Deep Learning”. In: *Towards Data Science* 10 (2020).

[34] Tianmei Guo et al. “Simple convolutional neural network on image classification”. In: *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*. IEEE. 2017, pp. 721–724.

[35] Yichuan Tang. “Deep learning using linear support vector machines”. In: *arXiv preprint arXiv:1306.0239* (2013).

[36] Christian Szegedy et al. “Rethinking the Inception Architecture for Computer Vision”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 2818–2826. DOI: 10.1109/CVPR.2016.308.

[37] Karen Simonyan and Andrew Zisserman. “Very Deep Convolutional Networks for Large-Scale Image Recognition”. In: *arXiv 1409.1556* (Sept. 2014).

[38] P Hoyos, N Kim, and Sabine Kastner. “How is magnetic resonance imaging used to learn about the brain”. In: *Front. Young Minds* 7 (2019), p. 86.

[39] Muhammad Naveed Iqbal Qureshi, Jooyoung Oh, and Boreom Lee. “3D-CNN based discrimination of schizophrenia using resting-state fMRI”. In: *Artificial intelligence in medicine* 98 (2019), pp. 10–17.

[40] Junghoe Kim et al. “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”. In: *Neuroimage* 124 (2016), pp. 127–146.

[41] Wei Du et al. “High classification accuracy for schizophrenia with rest and task fMRI data”. In: *Frontiers in human neuroscience* 6 (2012), p. 145.

[42] Seifedine Kadry et al. “Automated detection of schizophrenia from brain MRI slices using optimized deep-features”. In: *2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*. IEEE. 2021, pp. 1–5.

[43] Weizheng Yan et al. “Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site FMRI data”. In: *EBioMedicine* 47 (2019), pp. 543–552.

[44] Ling-Li Zeng et al. “Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI”. In: *EBioMedicine* 30 (2018), pp. 74–85.

[45] Young Tak Jo et al. “Diagnosing schizophrenia with network analysis and a machine learning method”. In: *International journal of methods in psychiatric research* 29.1 (2020), e1818.

[46] Youngoh Bae et al. “Differences between schizophrenic and normal subjects using network properties from fMRI”. In: *Journal of digital imaging* 31.2 (2018), pp. 252–261.