

Semantic Segmentation of Tumor from 3D Structural MRI using U-Net Autoencoder

by

Maisha Farzana

16101108

Md. Jahid Hossain Any

16101164

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
Brac University
March 2020

© 2020. Brac University
All rights reserved.

Declaration

It is hereby declared that

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

Student's Full Name & Signature:



Maisha Farzana
11.07.2020

Maisha Farzana
16101108



Jahid
11.07.2020

Md. Jahid Hossain Any
16101164

Approval

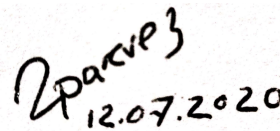
The thesis/project titled “Semantic Segmentation of tumor from 3D structural MRI using U-Net Autoencoder” submitted by

1. Maisha Farzana (16101108)
2. Md. Jahid Hossain Any (16101164)

Of Spring, 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on March 7, 2020.

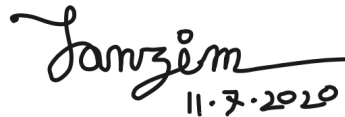
Examining Committee:

Supervisor:
(Member)



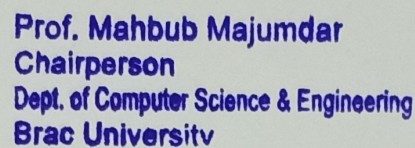
Mohammad Zavid Parvez, PhD
Assistant Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Member)



Tanzim Reza
Lecturer
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)



Prof. Mahbub Majumdar
Chairperson
Dept. of Computer Science & Engineering
Brac University

Mahbubul Alam Majumdar, PhD
Professor and Chairperson
Department of Computer Science and Engineering
Brac University

Abstract

Automated semantic segmentation of brain tumors from 3D MRI images plays a significant role in medical image processing. Early detection of these brain tumors is highly requisite for the treatment, screening, diagnosis and surgical pre-planning of the anomalies. The physicians normally follow the manual way of delineation to process the diagnosis of tumors which is time consuming, requires too much knowledge of anatomy and is too much expensive. To resolve these limitations, convolutional neural network (CNN) based autoencoder model is proposed which performs automated segmentation of brain tumors from 3D MRI brain images. Several algorithms such as image normalization, image augmentation, image binarization are used for data pre-processing. Furthermore, autoencoder based U-Net architecture is developed to extract the key features of the tumor and train the model. Later on, the model is applied to the new 3D MRI brain images to test the accuracy of it by segmenting the tumor region. The proposed model enables enhancing the performance and accuracy of semantic segmentation of brain tumor as compare to the other existing models. Applying the proposed method, the accuracy is obtained upto 96.06% considering the 66 subjects. Finally, this approach is a well-structured model for segmenting the tumor region from MRI brain images which may assist the physicians for providing therapy and better treatment to the patient.

Keywords: Brain Tumor, Semantic Segmentation, MRI, CNN, Pre-processing, Autoencoder, U-Net Architecture.

Dedication

We would like to dedicate our work to all the patients of Tumor Disease who motivated us to work in this field. We also want to dedicate our work to our parents, without whom we could never come so far in life. And a special Thanks to our supervisor who supported us wholeheartedly.

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis have been completed without any major interruptions.

Secondly, to our supervisor Dr Mohammad Zavid Parvez sir for his constant support and encouragement in our work. He helped us whenever we needed help.

Thirdly, to our co-supervisor Tanzim Reza sir for his kind support and advice in our work.

And finally to our parents without their throughout support it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

Declaration	i
Approval	ii
Abstract	iii
Dedication	iv
Acknowledgment	v
Table of Contents	vi
List of Figures	viii
List of Tables	ix
Nomenclature	x
1 Introduction	1
1.1 Motivation	2
1.2 Major Contribution	2
1.3 Thesis Orientation	3
2 Literature Review	4
3 Background Study	7
3.1 Brain	7
3.2 Brain Tumor	8
3.3 MRI	9
3.4 Image Classification	11
3.4.1 Image Segmentation	12
3.5 Machine Learning	12
3.5.1 Supervised Learning	13
3.5.2 Unsupervised Learning	13
3.5.3 Semi-supervised Learning	14
3.5.4 Reinforcement Learning	14
3.6 Artificial Neural Network	14
3.6.1 Activation functions in ANN	15
3.7 Convolutional Neural Network	17
3.7.1 Convolution layer	18

3.7.2	Pooling Layer	18
3.7.3	Fully Connected Layer	18
3.8	Autoencoder	19
3.8.1	U-Net	20
4	Proposed Model	22
4.1	Data Collection	23
4.2	Data Pre-processing	24
4.2.1	NII to PNG conversion	24
4.2.2	Image Normalization	25
4.2.3	Image Augmentation	25
4.2.4	Image Binarization	25
4.2.5	Test-Train Data Split	26
4.3	Model Training	27
4.3.1	Adam Optimizer	27
4.4	Model Test and Validation	28
5	Results and Analysis	29
5.1	Result	29
5.2	Analysis	32
6	Conclusion	35
6.1	Future Work	36
	References	40

List of Figures

3.1	Structure of a Human Brain.	8
3.2	T2-weighted images of high-grade and low-grade gliomas respectively	9
3.3	T1, T2, T2-FLAIR images.	10
3.4	An Image Classification Algorithm Pipeline.	11
3.5	Supervised learning	13
3.6	Unsupervised learning.	13
3.7	Reinforcement learning.	14
3.8	A simple Artificial Neural Network.	15
3.9	Sigmoid function graph.	16
3.10	ReLU activation function graph.	17
3.11	Convolutional Neural Network.	18
3.12	Basic Autoencoder Model.	19
3.13	U-Net Architecture.	20
4.1	Workflow of Proposed Model.	22
4.2	Coronal, Axial, Sagittal 3D MRI images of a brain volume.	24
4.3	Segmented image before and after applying Image Binarization.	26
5.1	Semantic Segmentation example from given brain and segmented tumor images (T1ce).	29
5.2	Semantic Segmentation of tumor region from 3D MRI (T1ce) input brain images.	31
5.3	Accuracy of the model from random six subject's data.	32
5.4	Loss Value of the model from random six subject's data.	33
5.5	Epoch vs. Accuracy graph.	33
5.6	Epoch vs. Loss graph.	34

List of Tables

3.1	Different color of parts of brain in MRI modalities.	11
5.1	Accuracy, Loss, Validation Accuracy and Validation Loss for the four subjects (after 20 epochs).	30
5.2	Accuracy and Loss value of the random six subjects (after 20 epochs).	32

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

β Beta

ϵ Epsilon

η Eta

σ Sigma

BSE Brain Surface Extractor

CNN Convolutional Neural Network

CNS Central Nervous System

CSF Cerebrospinal Fluid

ECNN En-hanced Convolutional Neural Network

FLAIR Fluid Attenuated Inversion Recovery

HGG High Grade Glioma

LGG Low Grade Glioma

MRI Magnetic Resonance Imaging

MVMN Multivariate Multinomial Distribution

PET Positron-Emission Tomography

ReLU Rectified Linear Unit

ROI Region of Interest

SVM Support Vector Machine

TE Time to Echo

TR Repetition Time

VAE Variational Autoencoder

Chapter 1

Introduction

Brain tumors can be categorized into two types which are primary and secondary tumor. Between these two types, primary brain tumors are originated from the brain cells and the secondary type of tumors are spread into the brain from other organs of the body. Gliomas, being one of the most general forms of primary tumors of the central nervous system which originate from the glial cell of brain. Gliomas contain several types of heterogeneous sub-regions such as edema, enhancing, non-enhancing core etc. This tumor can be categorized into two sub-types which are high-grade gliomas and low-grade gliomas. High-grade gliomas directly originate from the central nervous system and are highly malignant solid tumors[4]. Also, these tumors can induce the development of new tumors by being able to migrate within the CNS. On the contrary, low-grade gliomas normally have an indolent course with longer-term survival comparatively[3].

There are several imaging techniques available such as X-rays, Computed Tomography (CT) scan, Magnetic Resonance Imaging (MRI), Positron-Emission Tomography (PET), EEG etc. Among these, Magnetic Resonance Imaging (MRI) is a key technique that has been widely used for the analysis of brain tumors, monitoring and surgery pre-planning about the tumors. It has advantages of very good spatial resolution and a good temporal resolution and so this technique has been adopted for tumor segmentation task of brain images widely. There are various modalities available for MRI images such as T1-weighted, T2-weighted, T2-FLAIR etc. These images are used to mark several types of properties of tissue, areas of tumor spread and identifying the tumor itself.

Image segmentation is a very essential element for medical imaging processing purposes. Image segmentation technique is considered to be very useful as it can divide an image and extract the region of interest through some semi-automatic or automatic process. The main goal of image segmentation for medical imaging is to make the image simpler by focusing and cropping the main element of the image. Thus, it can be used for identifying brain tumors and other abnormalities in the brain images so that the segmented images can be further analyzed and monitored by the physicians. Recently, for this type of semantic segmentation, deep learning-based algorithms are applied and among those techniques, convolutional neural networks are able to learn from the given examples and can be applied to segment the image based on the learning.

Recently, Autoencoder based neural network has been widely used for segmenting medical images as it can learn efficiently from the given example, compress the code

and then learn how to reconstruct the segmented image from the reduced encoded representation [33]. There are some models of autoencoders and U-Net autoencoder is highly used for segmenting tumor from an input brain image. This model takes an input and reconstructs an output in form of a segmentation map. As it is an autoencoder model and so, it follows unsupervised learning.

In this chapter, we will discuss the motivation behind our work, our contributions towards this research and the thesis orientation which will cover the contents of each chapter in this paper.

1.1 Motivation

Brain tumor is the most common and most lethal of all form of tumors even after decades of research and it is on the rise every day. And early detection of brain tumor might be the major key to survive it. Nowadays many computational tools are used to assist doctors in making decision about the patients' medication. Abundance of study has been done on brain tumor and many had proposed commendable approaches to detect and locate brain tumor. And yet all of it was not enough to change the death ratio caused by brain tumor. So, we decided to make a contribution to this cause with a view to establishing a unique approach of our own.

Our motive is to find an approach to detect brain tumor early with substantially more accuracy. We found numerous research work on this subject. But everyday there are better technology, and with it comes better opportunity to bring the brain tumor detection to perfection. A brain tumor is a growth of abnormal cells in the tissues of the brain. Many fatal brain tumors like Glioblastoma, Astrocytoma, Meningioma, Oligodendroglioma can be treated better and bring a chance of survival with early detection. So, we have come up with the method of detecting the disease at its early stages so that we get a perfect accuracy of the results.

We have chosen MRI data for our research. MRI (Magnetic Resonance Imaging) is a scientific procedure that the radiologist primarily uses to image the human body's internal structure without any surgery. We have collected necessary MRI datasets to test and train our algorithm and bring better results. We have discussed our contributions in this subject in the next section (1.2).

1.2 Major Contribution

Most researches in the field of detecting Brain Tumor disease include MRI image processing techniques as it helps show the cellular structure of brain better than other imaging process. These researches include various segmentation process, deep learning methods and Machine Learning approaches on the MRI dataset with the hope of finding a better detection process.

So, in this paper we have initiated a unique approach of our own with U-Net based Autoencoder architecture which will help us preserve the structural integrity of the image during the segmentation process. In this autoencoder based CNN architecture we apply the supervised approach for the data driven feature learning. This uniquely selected approach gave us the best accuracy in the early detection of Brain Tumor. As we had collected abundance of data from proper source it helped us do the feature extraction for our desired outcome and test and confirm our process accuracy.

1.3 Thesis Orientation

The following parts of the study were structured in the following way. Chapter 2 is the literature review which contains related research works and existing approaches relevant to our proposed model. Chapter 3 includes all the background information related to our work such as Brain tumor, MRI, Image Classification, Machine learning, CNN and how we are using the resources to get the desired output. In chapter 4, we have described the proposed model of our works along with relevant graphs and figures. It includes our overall working methodologies, Dataset information about patients, data split process and preprocessing criteria about dataset will also be included. Algorithms related to our prediction will also be described in this chapter. The predicted results and relevant discussions are showed in chapter 5. Lastly, the summary of the report and conclusion as well as some future work plan is done in chapter 6. In future work part, we will talk about our future ambition on this work.

Chapter 2

Literature Review

In this chapter, to get a proper stance on the topic, a brief survey was carried out among abundance of research paper related to Brain Tumor and MRI to find out about the previous works done in the field and what could be improved.

Amin, Sharif, Yasmin and Fernandes proposed an automated system to differentiate between cancerous and noncancerous MRI brain image[30]. For datasets they collected 39 MRI with healthy brain cells and 46 MRI with tumor from Nashtar hospital Multan. They used Brain Surface Extractor (BSE) method in their local dataset to eliminate the artifacts from MR images. And after mixture of feature extraction they applied SVM classification with Linear, Gaussian and Cubic kernel function tested. They applied three variant kernels of SVM on benchmark dataset to compare presented methodology results. In another paper Myronenko presented an autoencoder decoder regularization which won the 1st place in the BraTS 2018 challenge[44]. They portrayed a segmentation approach with the encoder-decoder structure of CNN with an asymmetrically larger encoder to extract image features. They focused on additional guidance and regularization in this regard. To solve the problem of limited dataset they also added the variational autoencoder (VAE) branch so that it can reconstruct and regularize the shared encoder. They used various techniques such as histogram matching, affine image transforms and random image filtering which didn't indicate any additional improvements. And they found that increasing the network width consistently improves their result, whereas increasing the network depth didn't help with the performance.

Islam and Rishad provided us with some significant insight in their paper regarding CNN filter size and number in brain MRI image classification[41]. Before applying CNN algorithm they preprocessed the dataset with modified tracking algorithm for removal of skull and other artifacts, and Median filtering for noise removal. They confirmed that same size for all images is also a significant concern for image classification field. In CNN architecture combining the input image with some filters they performed rectification, pooling, fully connected layer and classification layer. They found out that square filter is better than rectangular filter because convolution stride is same for horizontal and vertical transverse. Again rectangular filter may create gap in time of convolution. They also stated that large number of filters provides higher classification accuracy in ordinary Neural Network but for CNN it does not bring much of a change. In fact it comes with more complexity. Examining their results, they suggested not to consider stride more than twice as increasing convolution strides results into decreasing CNN accuracy. Gordillo,

Montseny and Sobrevilla elaborated various segmentation methods of MRI images emphasizing both semiautomatic and fully automatic techniques[11]. They clarified that threshold-based techniques, even though is a fast segmentation process, are generally used as a first step in the segmentation process. They also added that semi-automatic approaches such as watershed transformation, which have also been reported to bring very accurate results. They concluded that Pixel classification are the most frequently used technique for brain tumor segmentation even though it's limited to clustering technique.

Vinoth and Venkatesh have proposed an idea of using both deep learning and machine learning algorithm in their approach[45]. In this paper the power scales of MRI datasets are institutionalized by Force standardization. They have completed their extension work using CNN by calculating certain parameter of the image. And they used the parameters to get result applying SVM. SVM classifier helps classify the type of tumor which helps a great deal while deciding treatment procedure. And they added that applying CNN they could find both the HGG and LGG parts of the tumor. But the depth and stage of tumor can be identified by using parameters in the SVM classifier. Isselmou, Zhang and Xu presented a new approach combining threshold segmentation and morphological operation[24]. For pre-processing they used median filter to remove noise and histogram equalization for image enhancement. Then threshold segmentation created binary images from grey level ones by turning pixels below a given threshold to zero and pixels above the threshold to 1. Furthermore, they applied binary dilation and binary erosion to detect the tumor. Mengqiao, Jie, Yilei and Hao also presented an CNN based solution where they use online evaluation platform to test the results[35]. They used 100 MRI to train the algorithm. They use N4ITK in the proposed model to correct the intensity inhomogeneity of the MRI dataset. They added dropout to reduce over fitting and batch normalization technique to speed up the training. Furthermore, they included domain labeling method to rightly assemble the wrongly classified clusters which often tends to linger after applying CNN. Madheswaran and Dhas worked with an adroit Naïve Bayesian based sequence mining approach to predict brain tumor[13]. They used ORNRAD filter to remove the noise from the MRI while pre-processing and Tamura method to characterize the texture by structure and tone. They went through normal, kernel and multi-variable multinomial distributions which are supported by Naïve Bayesian technique and found that MVMN distribution predicts a lot better than the other two approaches. 98 out of 100 samples were accurately predicted with MVMN distribution.

Thaha, Kumar, Murugan, Dhanasekeran, Vijayakarhick and Selvi proposed an Enhanced Convolutional Neural Networks (ECNN) for auto segmentation of brain tumor MRI[48]. ECNN performed a pixel-wise segmentation. And the new Novel Bat Optimization Algorithm (NBOA) is applied to reduce errors. They proposed optimization-based MRI segmentation with small kernels for deep architecture which positively affects overfitting provided the lesser weights. They used both skull stripping and image enhancement algorithm. They compared their ECNN results with the existing CNN method with regards to precision, recall and accuracy and it shows noticeably more efficient performance in each field. Priya and Shobarani presented a Contextual Clustering based segmentation technique in their paper which focuses on accuracy by reducing false segmentation, and as it takes least computation it takes less time to compute[27]. Keeping accuracy in mind they preprocessed the image in

multiple layers like resizing, gray scale conversion, noise removal, median filtering and morphological opening operation. After the vigorous amount of preprocessing they applied the contextual clustering algorithm which improved the segmentation accuracy.

Meena, Pavitra, Nishanthi and Nivetha also proposed an admirable method using CNN in their paper[49]. They went through bias field correction, intensity and patch normalization for pre-processing. They showed the importance of using intensity normalization to solve heterogeneity caused by multisite multi-scanner acquisition of MRI. They also applied data augmentation and deep architecture through small kernels. They concluded that to effectively train CNN activation function LReLU is more important than ReLU. Natarajan, Krishna, Kenkre, Nancy and Singh focused on proper and accurate threshold operation for efficient brain tumor detection with MRI[10]. They preprocessed the MRI dataset with gray scale imaging and histogram equalization. Then they applied High pass filter to sharpen the images by enhancing contrast and Median filter for noise reduction. Next after the threshold segmentation they applied Morphological operation to remove as much imperfection as possible for the final image subtraction.

Benson and Lajish also proposed a Morphological operation based MR image enhancement and skull stripping technique[12]. Their proposed method worked on all T1, T2 and FLAIR axial images. They addressed the problem of low contrast MR images. To solve that they used erosion and dilation fundamental morphological operation which smoothens the contours of an object, breaks narrow isthmuses and eliminates thin protrusions. Pereira, Pinto, Alves and Silva proposed another CNN based model which is built over convolutional layers with small 3*3 kernels to allow deeper architecture[26]. They prepared the dataset by bias field correction, intensity and patch normalization. They addressed the heterogeneity caused by multisite multi scanner of MRI using intensity normalization. They evaluated their method with the BRATS 2013 and won the 1st position of the online evaluation.

Chapter 3

Background Study

In this chapter, amplitude of papers, books and articles were observed and considered to gather as much information about human brain, brain tumor, MRI, Image Classification, Machine learning, Artificial Neural Network, Convolutional Neural Network and Autoencoder as possible to gather sufficient information before moving forward in the paper.

Brain tumors are categorized into primary and secondary tumor types where primary brain tumors originate from brain cells and secondary tumors metastasize into the brain from other organs[44]. Glial cells are the origin point for the most common type of primary brain tumor. Gliomas are the most common primary central nervous system malignancies which exhibit highly variable clinical prognosis, usually contain various heterogeneous sub-regions (i.e., edema, enhancing and non-enhancing core) with variable histologic and genomic phenotypes[31]. The gliomas can be divided into two parts and these two are High-grade gliomas and Low-grade gliomas. Low-grade gliomas are a diverse group of primary brain tumors that often arise in young, otherwise healthy patients and generally have an indolent course and the patient who gets this type of tumor has a comparatively long survival rate and the High-grade gliomas directly originates from the central nervous system which shows malignant behavior and grows very fast.

Magnetic Resonance Imaging (MRI) is an essential tool for brain tumor analysis as it has very good spatial resolution images and good temporal images. Usually, there are some MRI modalities such as: T1-weighted, T1 with contrast agent(T1c), T2-weighted and Fluid Attenuation Inversion Recover (FLAIR) to identify different tissue properties and regions of tumor.

Automated segmentation of 3D brain tumors can save time and produce an accurate reproducible solution for further tumor analysis, monitoring and classification. Deep learning based segmentation techniques surpassed traditional computer vision methods for dense semantic segmentation as convolutional neural networks (CNN) are able to learn from the examples and develop state-of-the-art segmentation accuracy both in 2D natural images[39] and in 3D medical image modalities[25].

3.1 Brain

The human brain is the largest brain of all vertebrates compared to the weight of approximately 3.3 lbs of the body. The average male is 1,274 cubic centimeters in brain volume while the female brain is 1,131 cubic centimeters in volume. It

contains approximately 86 billion nerve cells (neurons) called ‘gray matter’[6]. It also contains billions of nerve fibers (axons and dendrites) called ‘white matter’. According to the Mayfield Clinic, the main part of the human brain is the cerebrum separated into two hemispheres. The brainstem lies below, and the cerebellum sits behind it[18]. The cerebral cortex is the outermost layer of the cerebrum, composed of four lobes: the frontal, parietal, temporal and occipital lobes. The human brain grows, like all vertebrate brains, from three parts known as the forebrain, midbrain, and hindbrain. Each of these includes ventricles called fluid-filled cavities. The forebrain grows into the cerebrum and underlying structures; the midbrain is part of the brainstem; and the hindbrain causes brainstem and cerebellum regions to grow[8]. The cerebral cortex is considered the center of complex thinking and is significantly expanded in human brains. Vision perception takes place in the back of the skull in the occipital lobe. The frontal lobes are responsible for problem solving and judgement related tasks and motor function. The temporal lobes processes sound and expression, including the hippocampus and amygdala, respectively, which play roles in memory and emotion. For spatial orientation and movement, the parietal lobes combine information from different senses. The brain is surrounded by a layer of tissue called the meninges. The skull (cranium) helps protect the brain from injury. The labels of different parts of human brain is show in the Figure 3.1.

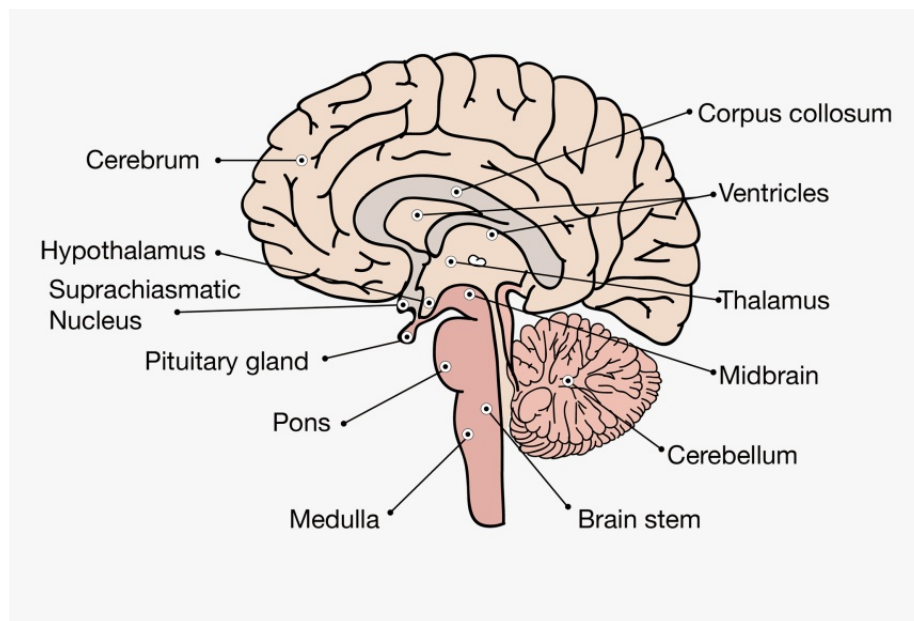


Figure 3.1: Structure of a Human Brain.

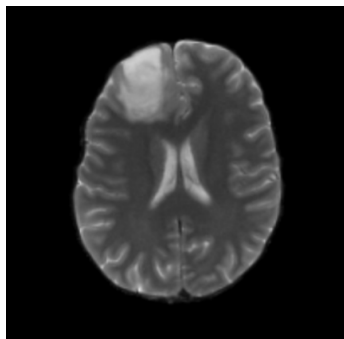
3.2 Brain Tumor

Brain tumors are common, requiring general medical providers to have a basic understanding of their diagnosis and medical management and the most prevalent brain tumors are intracranial metastases from systemic cancers, meningiomas, and gliomas and specifically, glioblastoma[43]. Meningiomas are tumors of the meninges, mostly benign and often managed by surgical resection, with radiation therapy and

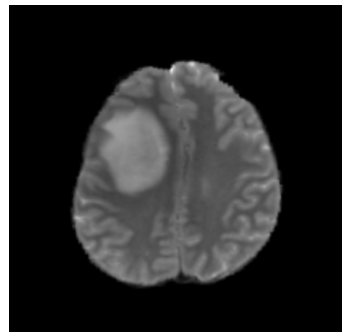
chemotherapy reserved for high-risk or refractory disease and glioblastoma is the most common and aggressive malignant primary brain tumor, with a limited response to standard-of-care concurrent chemo radiation. This paper will discuss about two types of gliomas and those are low-grade glioma and high-grade glioma as the dataset is based on these two types of tumors.

Low-grade gliomas (LGGs) are a diverse group of primary brain tumors that often arise in young, otherwise healthy patients and generally have an indolent course with longer-term survival in comparison with high-grade gliomas[3]. There is not much to worry regarding this type as it can be cured simply because of the treatment options including observation, surgery, radiation, chemotherapy, or a combined approach and the particular treatment will be based on tumor location, histology, molecular profile, and patient characteristics too.

On the contrary, high-grade gliomas, also known as primary central nervous system (CNS) tumors as they directly originate from the CNS, are highly malignant solid tumors arising from transformed cells of the brain and the spinal cord[4]. The rate of HGG is comparatively low among the children and the teenagers whereas the adults are at high risk regarding the tumor. However, they often annihilate the healthy brain tissues as they show considerably malignant behavior and usually, they grow very fast. Additionally, HGG can induce the development of new tumors by being able to migrate within the central nervous system for some centimeters. The treatment for HGG is not simple and easy as similar to LGG because the HGG normally grows very rapidly and thus the treatment is difficult to control the situation. The T2-weighted images of the high-grade and low-grade gliomas are shown in the Figure 3.2.



(a) T2-weighted Low grade glioma.



(b) T2-wighted High grade glioma.

Figure 3.2: T2-weighted images of high-grade and low-grade gliomas respectively

3.3 MRI

Magnetic Resonance Imaging (MRI) is a radiology imaging technology which uses powerful magnets to create a strong magnetic field, radio frequency waves and a computer to produce three dimensional detailed anatomical images of the internal organs and tissues and structures[2], [1]. MRI technique is widely used in scientific experiments as a patient will not have any pain during an MRI test and also because of its invasive quality which means there will not be any additional harmful side effects during and after the test. Due to its very good spatial resolution, this

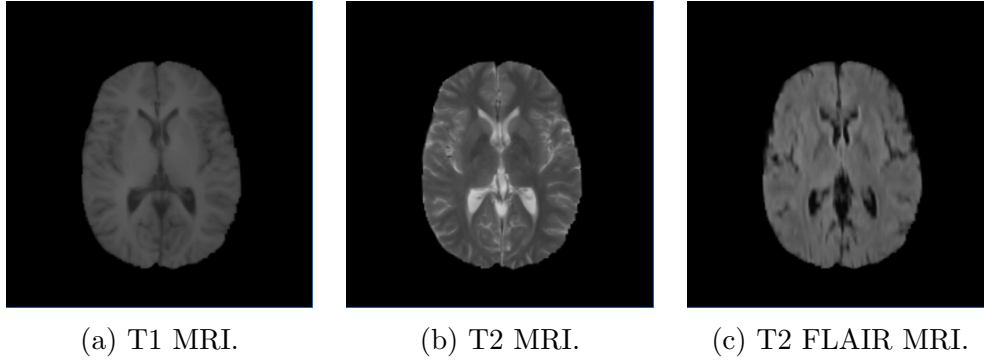


Figure 3.3: T1, T2, T2-FLAIR images.

technique is being used to detect abnormality of brain and spinal cord, detect tumor, cysts etc. in various part of the body, injuries or abnormalities of the joints, certain types of heart problem and also in other scopes as well[7].

Up to 60 percent of an adult human body is water and water is consisted of Hydrogen and Oxygen. The only one proton of each Hydrogen molecule always spin through their own random axis and thus creating a small individual magnetic field of their own but as a total the net value of that magnetic field is close to zero because each proton is spinning through a random axis, not through the same axis with same alignment. In the MRI machine a strong magnetic field is created with the powerful magnets around the machine and then all the protons get excited and aligned with that magnetic field[9]. Then a radio-frequency pulse is applied to the protons and then the protons again get excited and realigned with either 90 degree or 180 degree against the magnetic field. So, when the process of applying radio-frequency pulse is stopped then the protons release electromagnetic energy and realign with the static magnetic field. Then the computer is used to convert the analog data sent by the machine into digital data and then with the help of fourier transformation and image processing the actual MRI is obtained.

The most commonly used MRI modalities are T1-weighted and T2-weighted images and there is also another commonly used modality which is T2-Fluid Attenuated Inversion Recovery (FLAIR). T1-weighted images are produced by using short repetition time (TR) and short echo time (TE) whereas T2-weighted images are produced by using longer TR and TE times. The T2-Flair modality is similar to a T2-weighted image but the difference between these two is the repetition time and echo time is very long in Flair modality. T1-weighted imaging can also be performed while infusing Gadolinium (Gad). Gad is a non-toxic paramagnetic contrast enhancement agent. When injected during the scan, Gad changes signal intensities by shortening T1. Thus, Gad is very bright on T1-weighted images. We can differentiate T1 and T2-weighted images easily with the help of the color of Cerebrospinal Fluid (CSF) which is dark on T1-weighted images and bright on T2-weighted images. The images of the T1-weighted, T2-weighted and T2-Flair modalities are shown in the Figure 3.3 [14], [31], [38], [32], [36].

Several parts of the human brain show different color in each modality and by looking at those we can identify the particular modality. Cerebrospinal Fluid (CSF), White matter, Cortex, Fat (within bone marrow), infection etc. are some parts to be mentioned in Table 3.1.

Table 3.1: Different color of parts of brain in MRI modalities.

Tissue	T1-Weighted	T2-Weighted	FLAIR
CSF	Dark	Bright	Dark
White Matter	Light	Dark Grey	Dark Gray
Cortex	Gray	Light Grey	Light Grey
Fat	Bright	Light	Light
Inflammation	Dark	Bright	Bright

As the goal of our research is to detect tumor from human brain region and so we need clear and detailed brain images so that our algorithms and methods work best on those. So, we will use MRI images for our research work because MRI images give extremely clear, detailed images of soft-tissue structures that other imaging techniques cannot achieve. Thus, we can accomplish our goal with higher accuracy with MRI images.

3.4 Image Classification

Image classification refers to a process in computer vision task which can classify an image into its visual content class[19]. For example, an image classification algorithm can be designed so that it can classify whether an image contains a human figure or not. In image classification algorithm, features are the values that have been extracted from the image. Figure 3.4 shows a high-level algorithm for image classification.

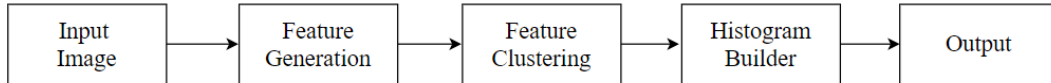


Figure 3.4: An Image Classification Algorithm Pipeline.

There are various image classification algorithms which are used for classification of data and among those algorithms, some are rule based and some are learning based algorithms[42]. There may be a good image classification but some pixels are always misclassified or even unclassified. The reason behind this type of problem is mixed pixel of the input data. In general, a classification system is designed based on the user’s need, spatial resolution of selected remotely sensed data, compatibility with previous work, image-processing and classification algorithms available, and time constraints[5]. So, any particular type of input data needs a classification algorithm which can be applied into that type. For example, a image classification algorithm which is designed to detect whether the input is an image of dog or not, will not be able to classify the image of a cat.

3.4.1 Image Segmentation

One of the significant aspects of image processing is image segmentation that aims to extract the ROI (Region of Interest) for image analyzing, object representation, visualization and further research. Deep learning-based image segmentation has been widely applied in detecting any objects in images, handwritten digit recognition, categorizing whole images etc.[22], [15]. An image is a set of pixels which has a range of 0 to 255 and the pixels which have the similar attributes can be batched together with the help of image segmentation. By dividing the image into several segments, the important features can be classified for further modification and analyzing whereas the other regions of the image which do not contain any information can be neglected. There are several types of segmentation and two of those are semantic segmentation and instance segmentation.

Semantic segmentation

Semantic segmentation is a type of image classification at a pixel level which means it classifies the same kind of objects from an image into a single object. Semantic segmentation has a wide array of applications ranging from scene understanding, inferring support-relationships among objects to autonomous driving[20]. It can only detect a single type of objects in an input image. For example, semantic segmentation will classify an image that has some cars in it, into one single object which is car.

Instance segmentation

Instance segmentation is the process of detecting and delineating each distinct object of interest appearing in an image[28]. It will label each foreground pixel with each object and instances and so, the separate objects in an image can be identified easily. Instance segmentation can be said as object detection + semantic segmentation. This kind of segmentation is applied for the tasks related to counting the number of objects.

3.5 Machine Learning

Machine Learning is a subset of Artificial Intelligence where a machine or program learns from some trained data or experience and makes predictions or results based on the test data or experience. It mainly focuses on making data-driven decisions rather than being explicitly programmed for performing a certain task [10]. These algorithms work in such a way that they learn from their mistakes and improve over time slowly when they are being tested on new data and gradually reach to good accuracy. Moreover, data are separated into training set and test set by which algorithms are trained by training data and predicted data are compared to test data for checking accuracy [11], [12]. Machine learning algorithms are trained and tested repeatedly until satisfactory accuracy results are obtained.

Types of Machine Learning: Machine Learning algorithms can be categorized into four different types[12].

3.5.1 Supervised Learning

Supervised machine learning algorithms can be applied upon what has been learned in the past training data to new test data using labeled examples to predict the desired outcome. The learning algorithm produces an inferred function to make predictions about the output data by starting from the analysis of a given and known training dataset. After sufficient training applied on the data, the system or program can be able to provide particular targets or results for any new data given to that. In order to modify the model and to get more accurate result, the system can compare its output with the given correct, desired output and then find errors and gradually reduces the error on each step. See Figure 3.5.

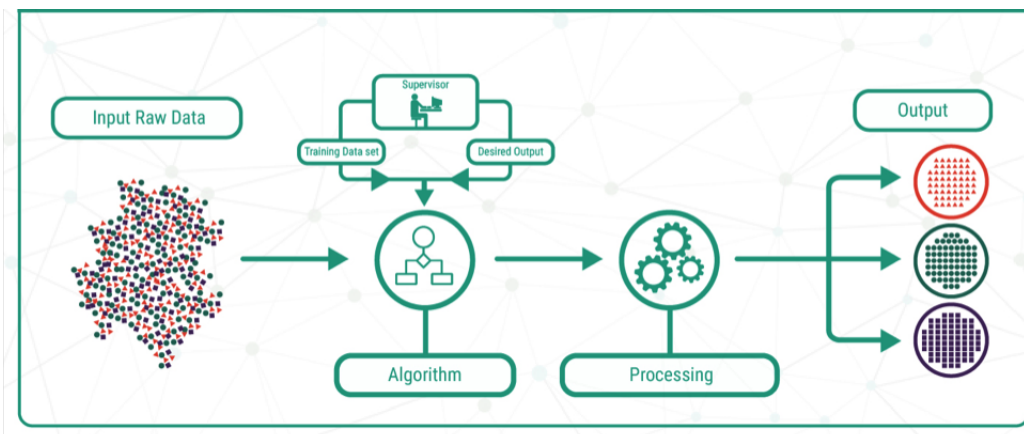


Figure 3.5: Supervised learning

3.5.2 Unsupervised Learning

Unsupervised machine learning algorithms are applied when the information or data used to train is neither classified nor labeled. It is applied by the system to find out a function that shows a hidden structure from unlabeled data. The main focus is not about to get the right, particular output rather its main goal is to explore the data and draw inferences from the datasets to describe hidden structures from unlabeled data. See Figure 3.6.

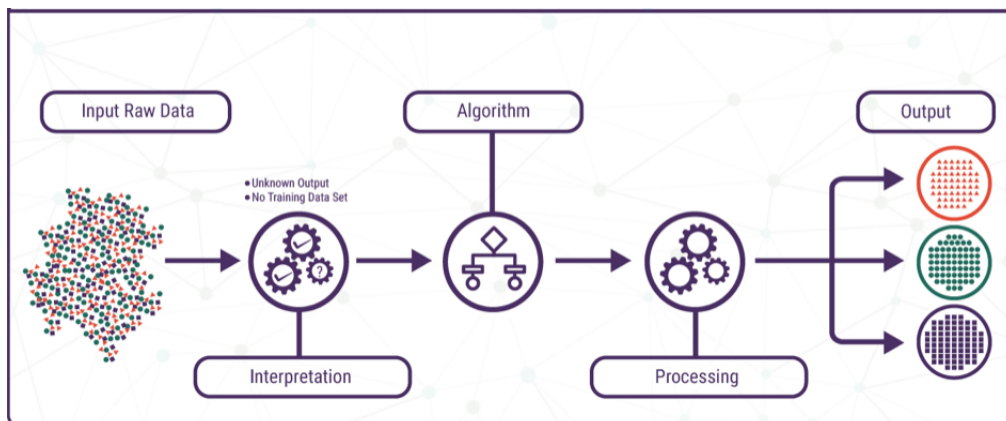


Figure 3.6: Unsupervised learning.

3.5.3 Semi-supervised Learning

Semi supervised machine learning algorithms use both labeled and unlabeled data for training and thus this type falls somewhere between supervised and unsupervised machine learning. Normally, it takes a small amount of labeled data and a large amount of unlabeled data to train the dataset. In order to improve high learning accuracy, this method is useful. When the acquired labeled data requires skilled and relevant resources in order to train it or learn from it, then semi-supervised learning method is used. In other case, acquiring unlabeled data generally does not require additional resources to learn.

3.5.4 Reinforcement Learning

Reinforcement machine learning algorithms interacts with its environment by producing actions and thus discovers errors or rewards. In an unpredictable, potentially complex environment, the agent learns to reach a goal by this learning. A simple reward feedback is required for the agent in order to learn which action is best and which is not and this is known as the reinforcement signal. This method allows machines and systems to automatically learn the ideal behavior within a specific context and thus it can produce maximum performance. See Figure 3.7.

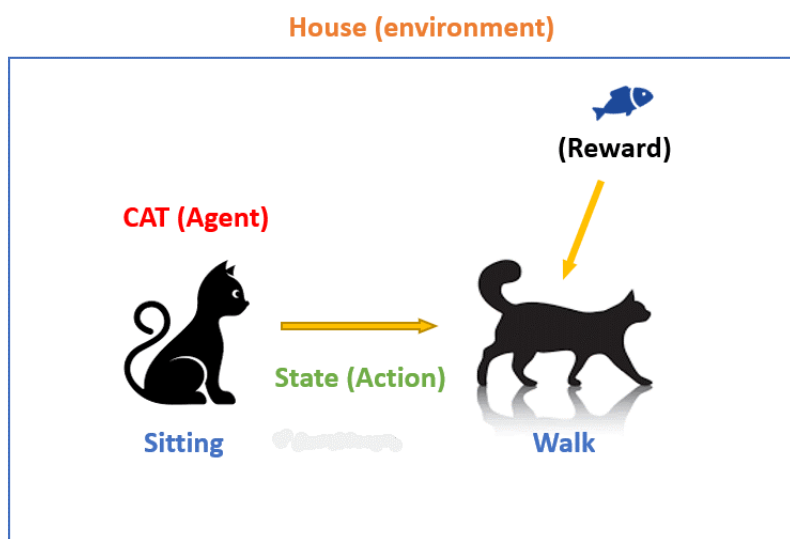


Figure 3.7: Reinforcement learning.

3.6 Artificial Neural Network

An artificial neural network (ANN) is a computational model which is focused on the structure of biological neural networks and their functions. ANN changes or learns based on the input and output data and that is why the information which flows through the neural network affects the structure of it[29]. ANNs are known to be nonlinear statistical data modeling methods where modeling or patterns are found for the complex relationships between inputs and outputs. This network is made up of multiple nodes that represent human brain's biological neurons[21]. The neurons

are connected with each other by links and they also interact with each other as shown in Figure 3.8. Input neurons get activated through sensors when they perceive the environment and the other neurons get activated through weighted connections that those have with the previously activated neurons. The output of each node is called the activation or node value for that node. A multitude of examples are given to an artificial neural network, and then it attempts to get the same response as the given example. If it's incorrect, an error is measured and the values at each neuron and synapse propagated by the ANN for the next time. The value of the weights and biases of the neurons get updated on the basis of the error at the output. This process is known as back-propagation. Artificial neural network works with three types of layers where layers are just sets of neurons.

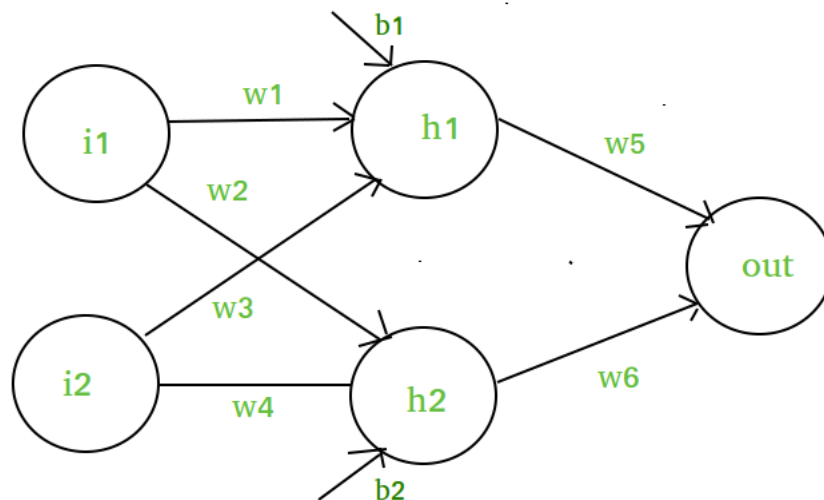


Figure 3.8: A simple Artificial Neural Network.

As visualized above, i_1 , i_2 are the nodes of input layers and w_1 , w_2 , w_3 , w_4 are the weighted values and b_1 , b_2 are the additional bias values. h_1 , h_2 are the nodes of hidden layers and out is the output of the network. In the input layers, the input data are stored and in the finished computations of the network are stored to the output layers. The hidden layers transform the input data into something meaningful that the output layer can apply. Each of the hidden layers is individually connected to the neurons in its input and output layer. This allows for a full learning process, and that takes place to the max limit as the weights inside the artificial neural network are updates after each iteration.

3.6.1 Activation functions in ANN

The neurons calculate the weighted sum of its input, adds a bias value and then an activation function decides whether the neuron should be activated or not. So, an activation function normally introduces non-linearity into the output of a neuron. An activation function's objective in an artificial neural network is to make the input data non-linear so that the input will be easier to learn and perform other complex activities. There are several activation functions applied for different kinds of approaches and algorithms. Two of those activation functions are sigmoid function and ReLU (Rectified Linear Unit) activation function.

1. Sigmoid function

Sigmoid activation function is a non-linear function that has a range of value 0 to 1 [37]. This function is plotted as ‘S’ shaped graph. Usually, sigmoid function is used in the output layer of a neural network for a binary classification where the results are in between 0 to 1. The sigmoid activation function is given in equation (3.1)

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (3.1)$$

Graphical representation of a sigmoid function is shown in Figure 3.9.

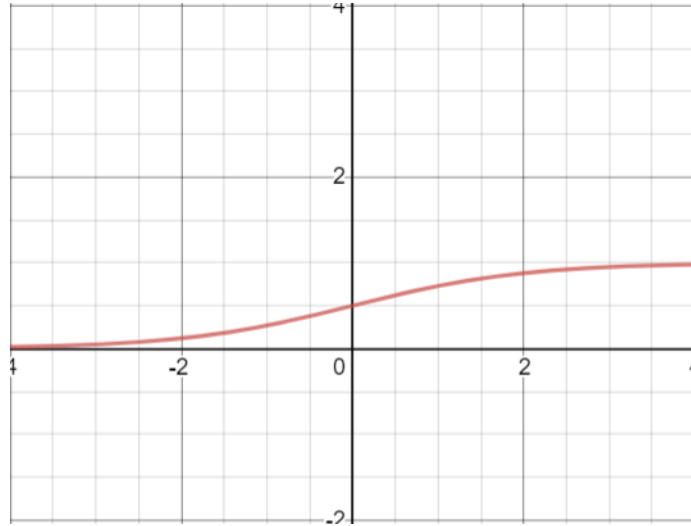


Figure 3.9: Sigmoid function graph.

2. ReLU (Rectified Linear Unit) Activation Function

The ReLU is probably the most used activation function in the world. It is used in almost all convolution neural networks or deep learning. The ReLU is half rectified from the bottom. When the input is less than zero, then the output will always remain zero but the output will be equal to the input when the input is zero or greater than zero [34]. The range of the ReLU function is 0 to infinity. The function for ReLU is given in equation (3.2)

$$f(x) = \max(0, x) \quad (3.2)$$

This function is linear for the values that are greater than zero and it suits well when training a neural network with the help of back propagation. Also, it is a non-linear function as the output will be zero when there are negative values as input. Graphical representation of a ReLU activation function is shown in the Figure 3.10.

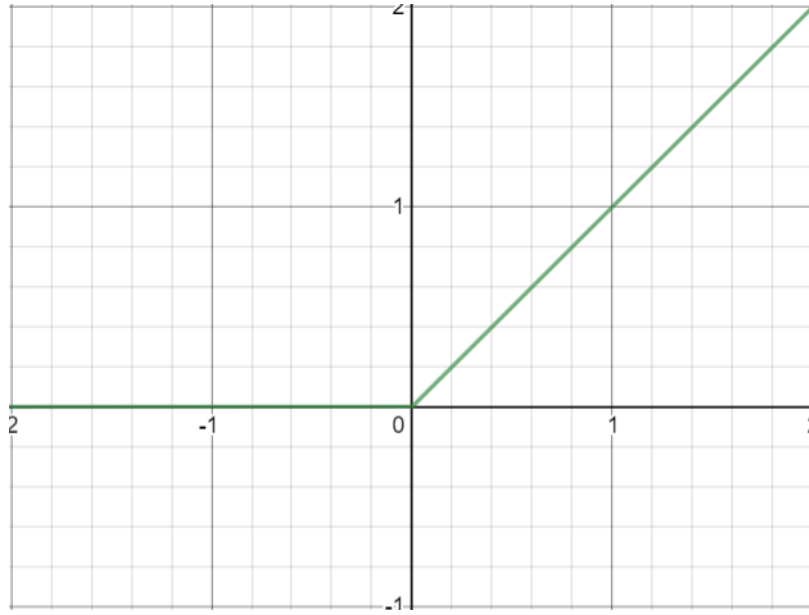


Figure 3.10: ReLU activation function graph.

3.7 Convolutional Neural Network

Convolutional Neural network is a deep learning algorithm which brings a new dynamic to the traditional machine learning approaches with its millions of learnable parameters. It is specially designed to recognize visual patterns directly from pixel images as shown in the Figure 3.11. There are various architectures of CNN such as UNet, LeNet, AlexNet, GoogLeNet and so on. ConvNet basically takes input vector X and produces output Y by performing function F on it as shown in equation (3.3). And W , weight represents the strength of interconnection between neurons of two adjacent layers[40].

$$F(X, W) = Y \tag{3.3}$$

Deep learning CNN models are basically used to train and test data. Each input image will pass it through a series of convolution layers with filters, pooling, fully connected layers and classify an object with probabilistic binary values 0 and 1. CNN

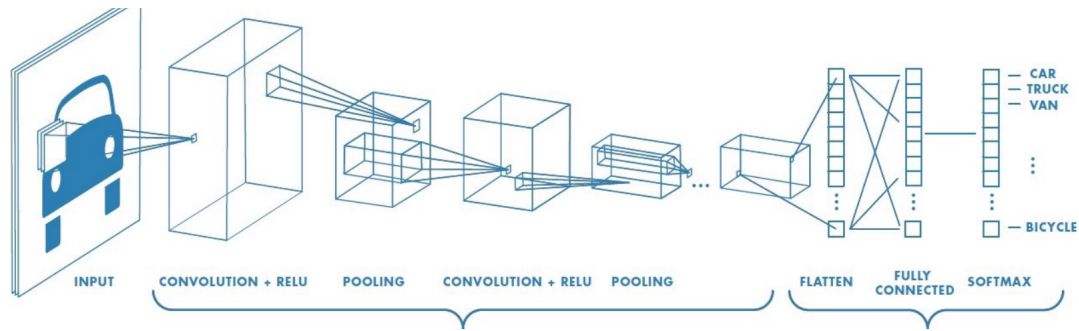


Figure 3.11: Convolutional Neural Network.

A general model of CNN consists of three components or layers: convolutional layer, pooling layer and fully connected layers. The first two layers mainly works for feature extraction and the last layer brings out the final output by mapping the extracted features.

3.7.1 Convolution layer

Convolution layer is the first layer and a key component of CNN. Its purpose is to detect the presence of features in the image received as input by convolutional filtering. Thus, this layer receives multiple images as input and calculates the convolution for each input and get individual feature maps. Contrasting traditional methods, features aren't predefined according to some particular form, rather it's learned by the network during training phase. Convolution layer typically forward the propagation of a training dataset and learnable parameters[46].

3.7.2 Pooling Layer

This type of layer is generally placed between two layers of convolution to reduce the in-plane dimensionality of the feature maps by providing a typical down sampling process and preserving the important characteristics of the image at the same time. Padding is often used to preserve the size of the feature maps or to keep important features from being clipped. Padding refers to the number of pixels added to an image before being processed by kernel. Due to the relevant filters in the pooling layer CNN can capture the spatial and temporal dependencies in an image. The amount of movement between applications of the filter to the input image is referred as the Stride. It has effects on how the filter is applied to the image and the size of the resulting feature map. Depending on the selected pixel values, there are three types of pooling: Maximum pooling, Average pooling, Minimum pooling[17].

3.7.3 Fully Connected Layer

After feeding on the output of the previous layer, this layer applies a Linear combination and then followed by an activation function, such as ReLU, to produce the output[46]. This layer determines the relationship between the position of features in the images and a class. To make CNN efficient and ready for sudden feature occurrence an optimizable feature extractor is applied at each image position.

3.8 Autoencoder

Autoencoder is an unsupervised artificial neural network which first learns how to efficiently compress and encode the given input data and then learns how to reconstruct the data back from the reduced encoded representation to a representation that is as similar to the original input as possible[33]. Autoencoder learns how to ignore the noise in the data and by doing so, it reduces the data dimensions [16]. It basically consists of three parts which are encoder, code and the decoder. The encoder part compresses or encodes the input image, produces the code part and then the decoder part reconstructs the input image by applying the code in the hidden layers as shown in Figure 3.12

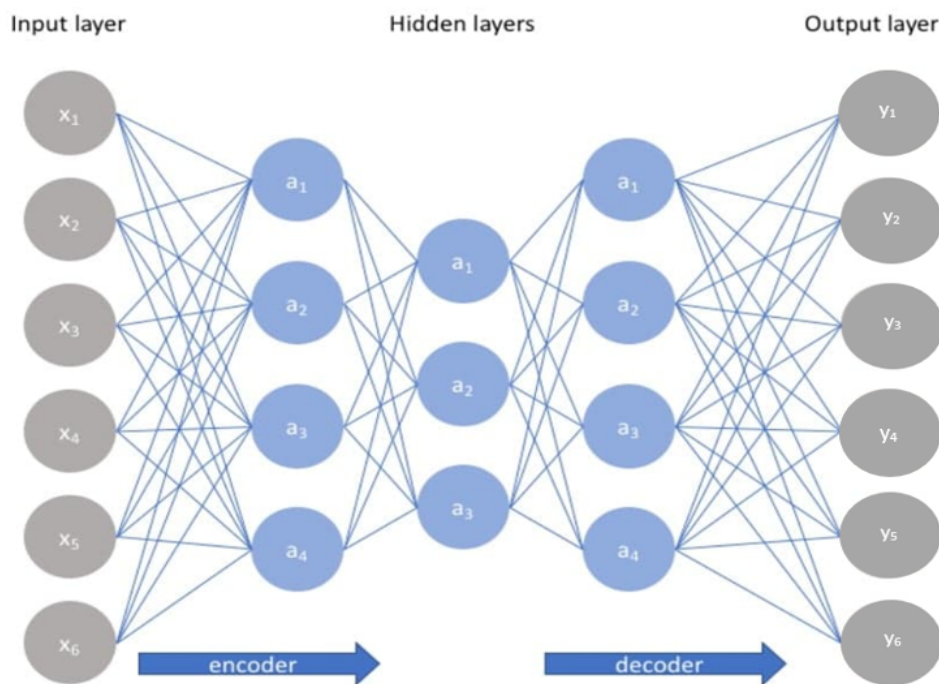


Figure 3.12: Basic Autoencoder Model.

The input and output layers have the same size and from the starting input layers the number of nodes in each layer get reduced[23]. As visualized above, x_1, x_2, \dots, x_6 are the inputs of an autoencoder and these inputs are fully connected with the hidden layers. The hidden layers which are a_1, a_2 and a_3 in this example, need to be restricted in sufficient number of nodes otherwise the model will end up memorizing all the input data. This layer contains the compressed information of the input data and thus the lowest possible dimensions and a smaller number of nodes from the input data. y_1, y_2, \dots, y_6 are the output layers in this example and this layer mainly represents the data which are reconstructed from the encoded form and these are as close as to the original input as possible.

Autoencoder has some important key features such as it is data-specific, lossy and unsupervised etc. Autoencoders are only able to compress or encode the input data only if those are type of data they have been trained on. It cannot compress an input data if that is not the similar kind of data which they learned earlier. For example, if the autoencoder is trained on for handwritten digits then it will not be

able to compress landscape images. Furthermore, an autoencoder does not provide the exact same output as the input data. It will be as close as the input but not the exact same and so it is not wise to apply it in a lossless compression method.

We will use autoencoder method as we need to give the brain images of the subjects as input images and expect to get the segmented tumor images as output. So, autoencoder architecture is the best choice for this purpose.

3.8.1 U-Net

The main idea of the U-Net architecture in image segmentation is that the feature map of the input image is needed to be converted into a vector and then reconstruct a segmented image from this vector[47]. The architecture uses the same feature maps that are used for contraction to expand a vector to a segmented image. This would preserve the structural integrity of the image which would reduce distortion enormously. This architecture is developed upon the Fully Convolutional Network and modified in a way that it creates better segmentation in medical imaging which means the human body parts. The U-Net is a symmetric shaped model and because of its symmetry, the network has a large number of feature maps in the upsampling path which allows to transfer information. The basic concepts and how it works is shown in the Figure 3.13.

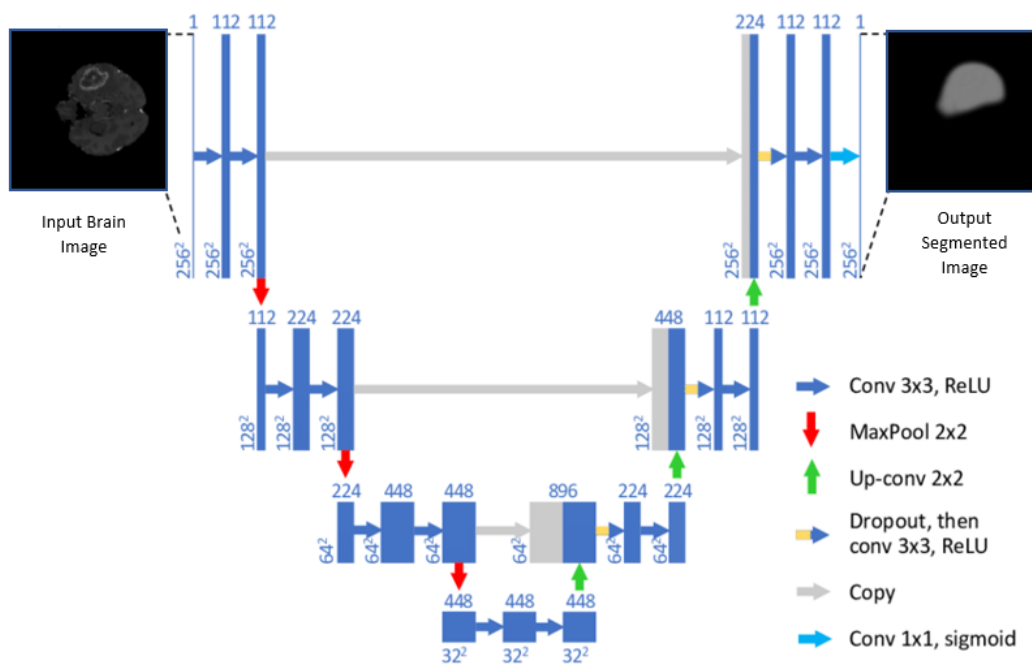


Figure 3.13: U-Net Architecture.

U-Net architecture consists of three sections which are:

- (1) The contraction or downsampling path
- (2) Bottleneck
- (3) The expansion or upsampling path

The contraction section has some contraction blocks. Each block takes an input applies two 3*3 convolution layers followed by a 2*2 max pooling[47]. The number of kernels or feature maps after each block doubles so that architecture can learn the

complex structures effectively. The bottom-most layer is a connection between the contraction layer and the expansion layer. It uses two 3×3 CNN layers and then a 2×2 up convolution layer. Similar to the contraction layer, the expansion layer also consists of several expansion blocks and each block passes the input to two 3×3 CNN layers and then a 2×2 upsampling layer. Also, the number of feature maps used by convolutional layer gets half after each block ends to maintain symmetry. However, every time the input is also get appended by feature maps of the corresponding contraction layer. This action would ensure that the features that are learned while contracting the image will be used to reconstruct it. In the contraction and the expansion layer, the number of blocks is same.

The U-Net architecture assembles the location information from the downsampling path along with the contextual information in the upsampling path to obtain a good general information which combines localization and context and thus it can create a good segmentation map. However, for this advantage of this architecture, we will use U-Net for getting good accuracy in tumor segmentation.

Chapter 4

Proposed Model

In this chapter, we are going to describe about the steps of our proposed model as shown in the Figure 4.1.

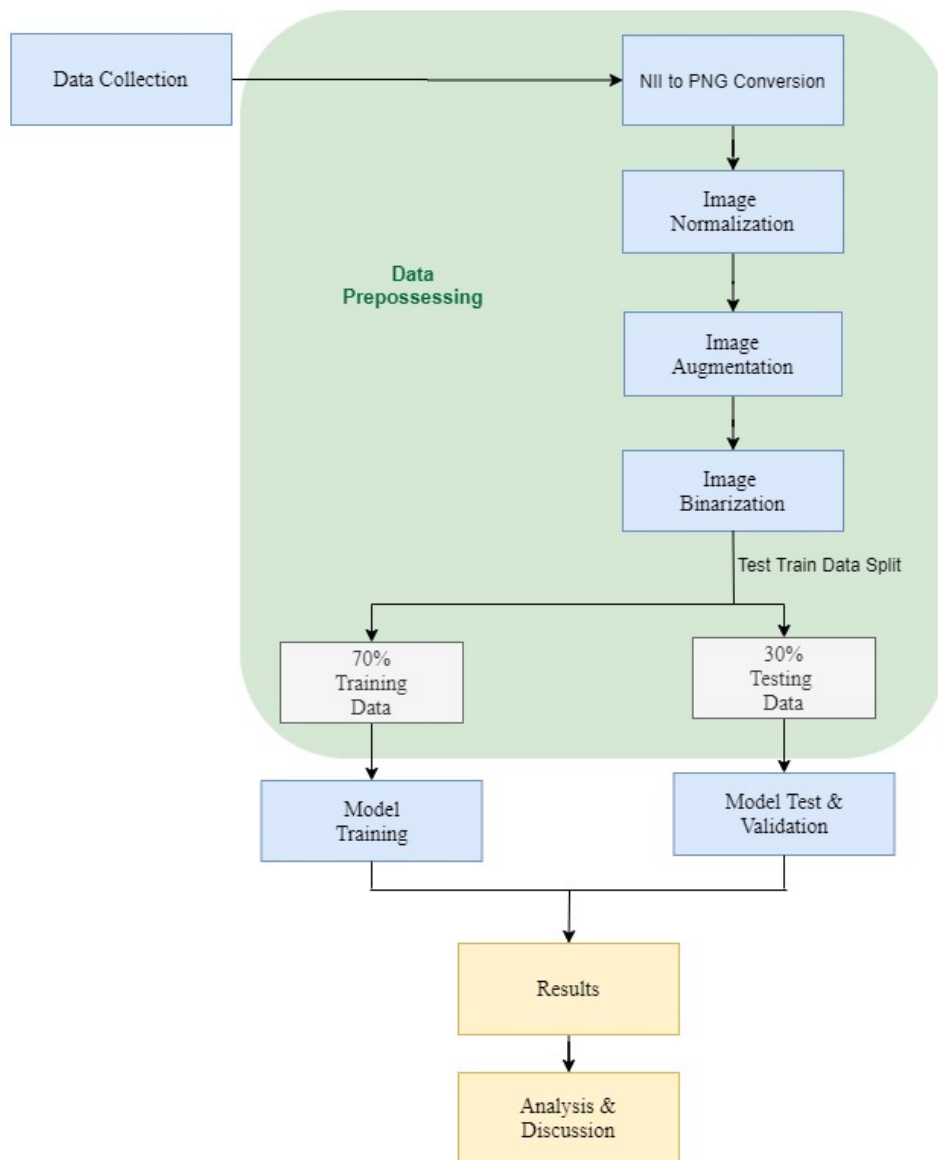


Figure 4.1: Workflow of Proposed Model.

Our segmentation approach was performed in several steps to get higher accuracy rate in segmenting the tumor region from the MRI brain images. To begin with, we collected the MRI dataset of human brain images that contains tumor to work on our research. We categorized and selected training data as well as validation data of some subjects. After that, we did the pre-processing for our dataset which includes image normalization method to convert all the input images into a range of pixels values that are more familiar and normal to work further on, image augmentation method to artificially create training images through different ways of processing and image binarization method to scale the segmented image to scale into same type of pixels. In the next step, we split the total dataset into two part where the larger part is for training purpose and the other part to test. Several processes and algorithms were applied to train and test the data. Finally, after following each step sequentially, we were able to segment the tumor region from the input brain images with a very good accuracy.

4.1 Data Collection

The brain tumor related 3D MRI dataset in which we are working on, are provided by The Medical Image Computing and Computer Assisted Intervention Society (MICCAI Society) [14], [31], [38], [32], [36]. All of the brain tumor segmentation multimodal scans are available as NIfTI files (.nii.gz). There are two types of data provided in which one type is the training data and the other one is the validation data. The training data is also separated by the two types of glioma tumor's files and between those one is High-grade glioma and the other one is Low-grade glioma files. There are 210 subjects for the High-grade gliomas and 75 subjects of the Low-grade gliomas. Also, In the data validation file there are images of 67 subjects. In each of the subject's file there are four types of images and those are 1. native (T1) and 2. post-contrast T1-weighted (T1Gd), 3. T2-weighted (T2), and 4. T2 Fluid Attenuated Inversion Recovery (FLAIR) volumes.

All of the images were acquired with different clinical protocols and various scanners from multiple (19) institutions. The provided data are distributed after their pre-processing, i.e. co-registered to the same anatomical template, interpolated to the same resolution ($1mm^3$) and skull-stripped. All the imaging datasets have been segmented manually, by one to four raters, following the same annotation protocol, and their annotations were approved by experienced neuro-radiologists. Annotations comprise the GD-enhancing tumor (ET - label 4), the peritumoral edema (ED - label 4), and the necrotic and non-enhancing tumor core (NCR/NET - label 1). The dataset contains the 3D axial, sagittal and coronal MRI images for each subject as given in Figure 4.2.

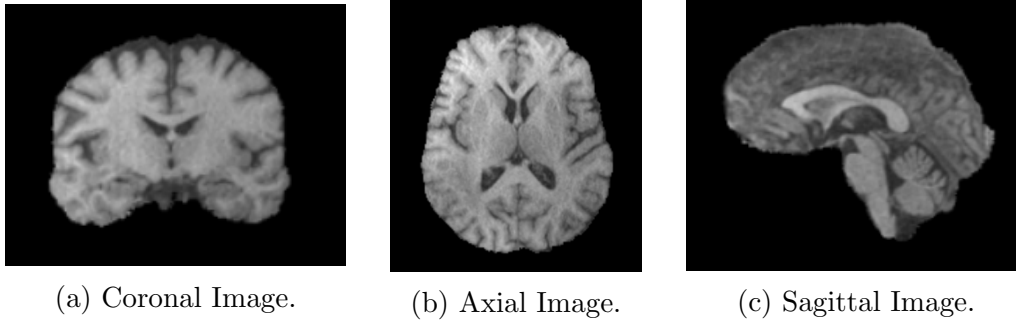


Figure 4.2: Coronal, Axial, Sagittal 3D MRI images of a brain volume.

The size of the input image is $240 * 240 * 155$ where the length and width are 240 and there are 155 slices of a single image. The training dataset contains the ground truth segmentation labels so that we can train the input data with that. An additional 3D MRI dataset without the ground truth tumor segmentation labels was also provided for validation and testing. We have used the T1ce MRI modalities of each subject's brain image slices as inputs and the segmented images of the tumor of those subjects as outputs in our model to train the model about the key feature of the tumor. By learning the key features from the training, the model will be able to identify tumors when it will be applied into any new brain images that contains tumor.

4.2 Data Pre-processing

Data Preprocessing is that step in any Machine Learning process in which the data is transformed, or encoded, to get it to such a state that the computer can now easily interpret it. In other words, the algorithm can now easily interpret the features of the data. As machines do not understand the images or texts and so our dataset of images needs to be pre-processed. As our dataset is a collection of 3D MRI images so, we have to transform the data and select the feature of the data as well to use in our model smoothly. We have gone through some pre-processing steps for the 3D MRI dataset and those methods are described below.

4.2.1 NII to PNG conversion

As the 3D image dataset is in NII file extension which is a medical image formatted files that contains too many information encoded in it and these files have voxel sizes which does not have any range like pixels have. So, we need to flatten each file into a PNG file for every single slice to make it prepared for our next steps (in 2D shape) and make it more visual friendly. Image files like PNG, JPEG contains pixel size and with these pixel values we can work on our further methods. Also, some input images were in .nii.gz format and we imported a library called nibabel to extract the NII file from that format first. Then, we applied the same procedure to convert the data into PNG format. As we cannot work with 3D images in our model so, we did this conversion.

4.2.2 Image Normalization

Image normalization is a common process that adjusts the range of pixel intensity values in image processing. Its primary function is to transform an input image into a set of pixel values that are more familiar or normal values to work on. It is known that every image (grayscale or RGB) contains pixels that has a range from 0 to 255. On the contrary, the 3D MRI images have voxels which does not have any particular range. As our model is based on convolutional neural network that works with images and so we needed to scale the MRI images into 0 to 1 by dividing it with the maximum value of the input image. After that, we multiplied that value with 255 to convert that value into normal pixel sized images. We did this so that the value of each voxel of the MRI images come to the range of 0 to 255 that represents the general pixel values. After scaling the images, we were able to plot the images and work on it further. The formula that we used for scaling the images is given in equation (4.1).

$$imageArray = \frac{imageArray}{np.max(imageArray)} * 255 \quad (4.1)$$

Here, `imageArray` is the input image we took as input and `np.max(imageArray)` is the maximum value of that input image. From the division part the value generated into 0 to 1 and then we needed to multiply that value with 255 to make it in the range of general pixel value. We applied this method in the T1-weighted 3D MRI images of the dataset as we worked with that modality.

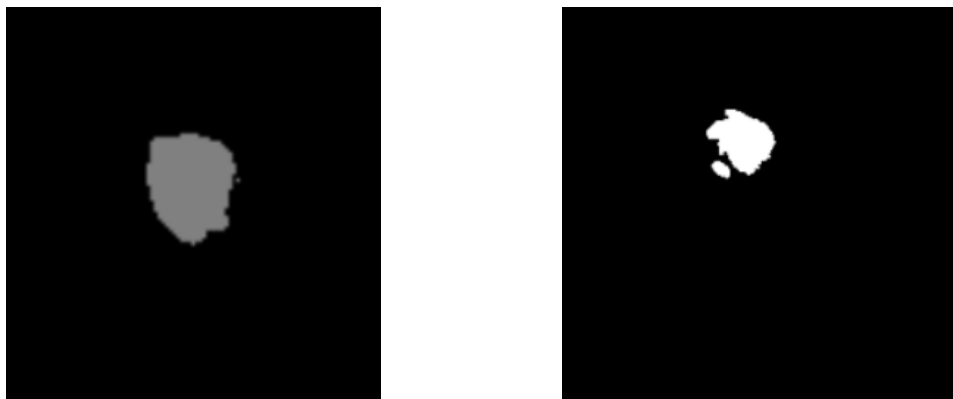
4.2.3 Image Augmentation

Image augmentation is a technique for the artificial expansion of the dataset and it normally works with few training data. It can increase the size of the dataset without acquiring new images. Some of the parameters of this technique are zoom, shear, rotation and so on. These parameters are used to increase the data sample to make it more efficient in deep learning network. We applied data augmentation using pixel-level image transformations. In this technique the geometrical shape of the input image and all the geometrical features remain unchanged but the pixel intensity values are changed either locally or across the whole image. As our dataset was obtained from different institutions and also from different scanners so, so pixel intensities, intensity gradients may be inherently heterogeneous. We used pixel-level operation to perform shifting, scaling of pixel intensity values and also to modify the image brightness. We selected the non-zero voxels of the image and normalized every input image to hold zero mean and unit std. After that, in the input image channels, we attached a random intensity shift and scale to help the neural network to learn valuable deep features of their original scaled images on their own.

4.2.4 Image Binarization

Image binarization is the process of converting a grayscale image into black and white (0 and 1) image with the help of a threshold value. As the grayscale images contain noises and so the autoencoders can not able to recognize the image properly. To prevent this problem, image binarization technique is improved to avoid those background noises in the image and convert it into two colors only. In the image

binarization technique a threshold value is selected and all the pixels that have the higher values than the threshold value are classified as white where the lower values than the threshold value are classified as black. In our model, we applied the image binarization technique in the given segmented tumor input images so that our autoencoder model could able to learn the key features of the tumor more efficiently and apply it on the test data. The threshold value for our dataset was 256 and we classified the image into two color (black and white) based on this value. We fixed the threshold value to 256 and applied a loop for 155 slices of each segmented input image. The values which were higher than 256 were classified as white color and the values which were lower than 256 were classified as black color. After applying image binarization technique the tumor of the segmented image was in white color and the rest of the image was in black color as shown in Figure 4.3.



(a) Segmented image before binarization. (b) Segmented image after binarization.

Figure 4.3: Segmented image before and after applying Image Binarization.

4.2.5 Test-Train Data Split

We had the training dataset of 210 patients which contained brain images and segmented images of high-grade glioma. Also, we had the validation dataset of 66 patients which contained only the brain images. So, we applied our U-Net autoencoder model to the training dataset and after that, we applied the model to the rest 66 validation dataset to check the accuracy of the model. We split the training data and testing data into 70% and 30% respectively to ensure that the model learned efficiently from the large amount of input data and could apply the model in the remaining data and obtain good accuracy.

Training Data

We had 210 patient's T1-weighted 3D MRI brain images with ground truth segmentation labels. The segmented images were used to help the applied U-Net autoencoder model to learn the key features of the tumor. The amount of the training data had been kept 70% so that the model could learn more detailed parts as well to ensure better segmentation. The input size of the training dataset was (256, 256, 1).

Testing Data

We had the validation dataset combined of 66 cases that contained 3D MRI brain images without the ground truth segmentation labels. We applied these cases to check the accuracy of our U-Net autoencoder model. The amount of the validation data was 30% to obtain the desired output. We intended to give the validation data as input and our model would segment the tumor region from the input images using its previous learning experience and give us the output of a segmented tumor image.

4.3 Model Training

For training the model, we used the autoencoder based convolutional neural network architecture and we followed the supervised approach for the data-driven features learning. In this process, the three layers of the autoencoder model which are input layers, hidden layers and the output layers are fully connected and the connection weights are updated through backpropagation algorithms. The input layers will be activated by following an activation function and in this model, we followed the sigmoid and ReLU activation function. We applied the ReLU activation function in the input layers and sigmoid function in the output layers. ReLU function gives an output of zero when the input is less than zero and gives the output as same as the input when the input is greater than zero. Sigmoid function gives an output closer to 0 when the input is very much less than 0 and gives an output closer to 1 when the input is much greater than 0.

To train our autoencoder model, we kept the 3D MRI T1ce image slices as the inputs of the model and the segmented image slices as the outputs of the model. we put the epoch number 20 so that the model would learn more efficiently and the value of the loss function would reduce. We put the batch size to 20 so that the model could learn with specific details and did not miss any major or minor key features from the given input image. Furthermore, We kept the learning rate ($1e - 4$) to adjust the weights of our network with respect to loss gradient and to minimize the loss function after each epoch. The learning rate had been kept so low to ensure that we would not miss any local minima and also to get the higher accuracy. The equation between learning rate and weights of the input is given (4.2).

$$newWeight = existingWeight - learningRate * gradient \quad (4.2)$$

from the equation, we can see that the lower the value of the learning rate was, the smaller changes in the weights occurred and thus the model could learn more efficiently. This can be understood by plotting a graph of epoch number vs. loss and from that we can see that after each epoch the loss got reduced in Figure 5.6.

4.3.1 Adam Optimizer

Adam is an optimization algorithm that can be used in training data to update the network weights iteratively. In Adam optimizer, a learning rate is fixed for each network weight or parameters and it computes different adaptive learning rates for each parameter by estimating the first and second moments of the gradients. We applied Adam optimizer for our input dataset with initial learning rate of $\eta = 1e - 4$

and we gradually decreased the value of the learning rate after each epoch. The equations that had been used to update the weight values for each parameter are shown below:

Exponentially moving averages of gradient along the parameter:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (4.3)$$

Exponentially moving averages of squared gradient along the parameter:

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (4.4)$$

Weight update:

$$w_t = w_{t-1} - \eta \frac{m_t}{\sqrt{v_t} + \epsilon} \quad (4.5)$$

Here, w_t is the updated model weight, η is the step size, m_t and v_t are moving averages, g_t is gradient value and β_1 , β_2 , ϵ are hyper-parameters of the algorithm.

4.4 Model Test and Validation

For model validation, we took the 10% of images from around 1500 number of slices of tumor images to provide an unbiased evaluation of the model fit on the training dataset while tuning the model hyperparameters. We applied this to describe the evaluation of models while data preparation. This evaluation becomes more biased when the skill on the validation dataset is incorporated into the model configuration. Additionally, we used the 20% of images from the around 1500 number of slices of tumor images to test the model. This helped us to provide an unbiased evaluation of the final proposed model to fit on the training dataset. We used this to describe the evaluation of the final model when comparing it to other final model.

Chapter 5

Results and Analysis

5.1 Result

In this chapter, we will try to illustrate the experimental results of our proposed model and analyze the results further for a proper understanding on it.

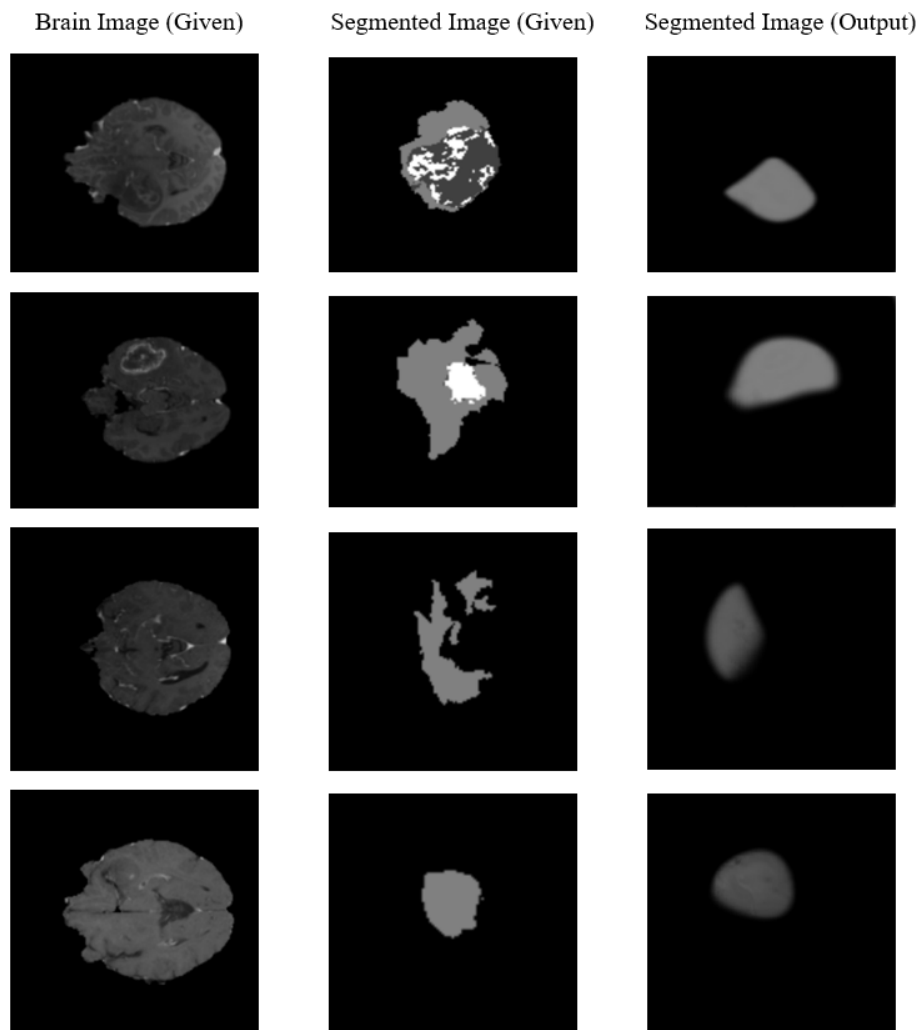


Figure 5.1: Semantic Segmentation example from given brain and segmented tumor images (T1ce).

In this paper, the semantic segmentation of brain tumor from 3D MRI images using U-Net Autoencoder is proposed on the basis of 210 MRI images from the patient. Later on, the model is applied on 66 new 3D MRI brain images to test the segmentation accuracy and performance of the model. Several data pre-processing methods were applied to prepare the data for the model. We implemented our neural network in Tensorflow and used several libraries such as keras, nibabel, numpy etc. to execute our codes.

To begin with, we apply our Autoencoder based U-Net model in those patient’s brain images which had the ground truth segmentation label along with the brain images. We apply it on the random four patient’s dataset to see whether our network can able to segment tumors from the brain images or not. This approach was successful as we could extract the tumor region from the given brain images along with the segmented images with high accuracy as given in Figure 5.1. The output segmented images match with the ground truth segmented tumor images well.

This approach provides high accuracy as well as validation accuracy and the loss value along with the validation loss value are low as given in Table 5.1. We applied 20 epochs and observed that after each epoch the accuracy and the validation accuracy were increasing whereas the loss value and the validation loss value was decreasing. We take these four subjects randomly and the average accuracy of this proposed approach is 97.12 percent which is pretty good.

Table 5.1: Accuracy, Loss, Validation Accuracy and Validation Loss for the four subjects (after 20 epochs).

Subject	Accuracy	Loss	Validation Accuracy	Validation Loss
1	95.83 %	4.34 %	95.66 %	4.31 %
2	95.6 %	4.81 %	98.53 %	1.13 %
3	98.64 %	2.92 %	94.48 %	3.35 %
4	98.4 %	1.12 %	98.41 %	1.16 %

As visualized above, it is understood that the model is ready to be applied on the 66 new 3D MRI brain images to segment the tumor region from the input brain images. We train the model with the 210 given 3D MRI segmented tumor images to learn the key features of the tumor and apply the model on the new brain images. The proposed approach was successful as we could extract the tumor region from only the given 3D MRI brain images with very good accuracy and performance as well. We used the T1ce modality for the input brain images and apply our model which was trained before with the 210 examples of the 3D MRI segmented tumor images. The model was able to identify the tumor region from the given input brain images and obtained that tumor region as output of the model as shown in Figure 5.2.

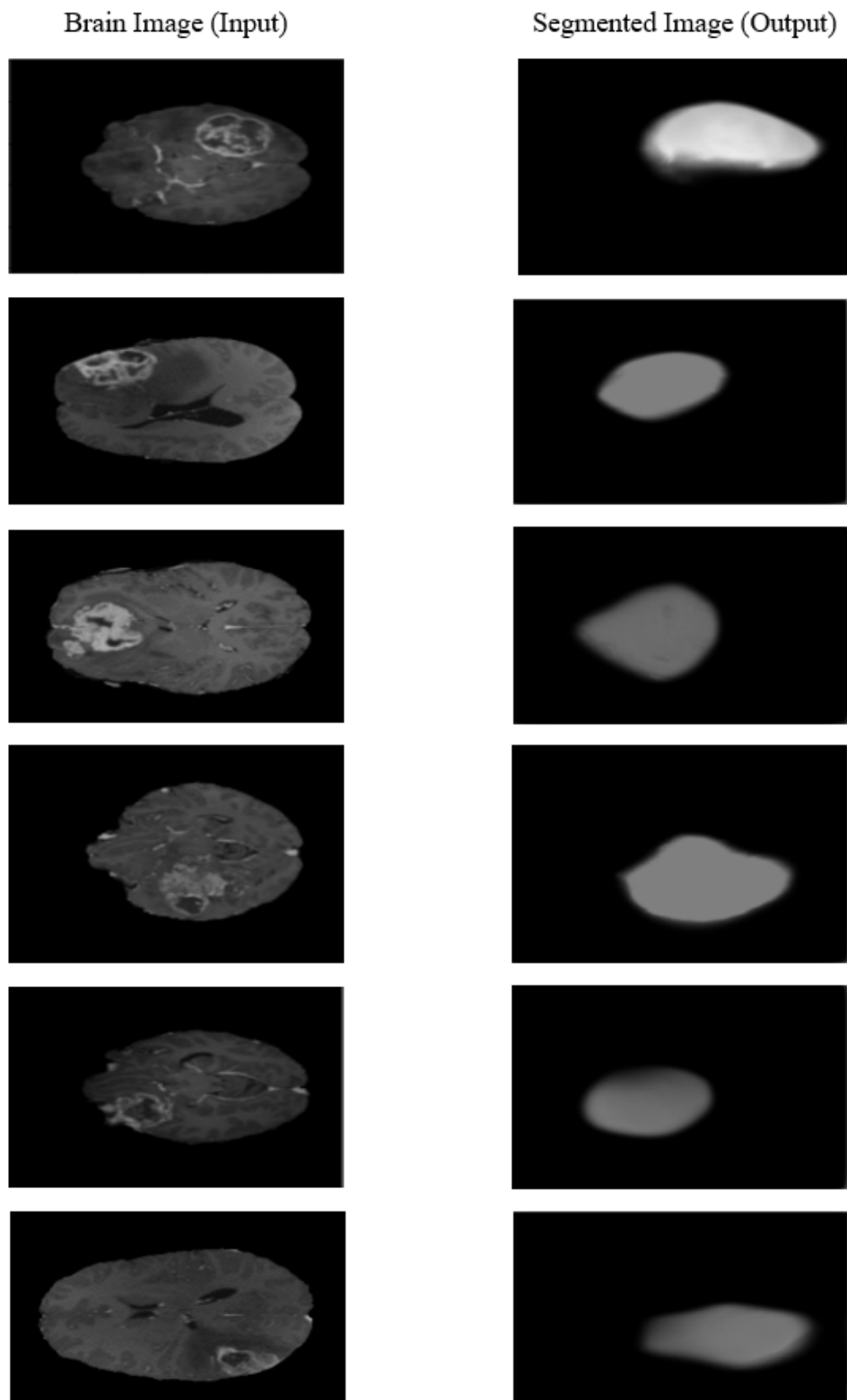


Figure 5.2: Semantic Segmentation of tumor region from 3D MRI (T1ce) input brain images.

This approach provides high accuracy and on the contrary, the loss value is very low. We applied 20 epochs and observed that after each epoch the accuracy whereas the loss value and decreasing. We take the 66 datasets from the new validation dataset which does not provided the ground truth segmentation labels. The average accuracy of this proposed approach is 96.06 percent which is pretty good. The accuracy and the loss value of random six patient’s data is shown in the Table 5.2.

Table 5.2: Accuracy and Loss value of the random six subjects (after 20 epochs).

Subject	Accuracy(%)	Loss(%)
1	97.39	2.49
2	95.92	4.16
3	95.73	3.92
4	94.29	5.2
5	95.62	4.02
6	97.42	2.34

5.2 Analysis

To measure the performance of this proposed model, a bar diagram is plotted to show the accuracy of the random six patient’s data which is pretty good as given in Figure 5.3 and the loss value of the model which is too low as given in Figure 5.4. The measurement is taken after training the model for 20 epochs and after the 20 epochs finished, we calculate the accuracy and the loss value.

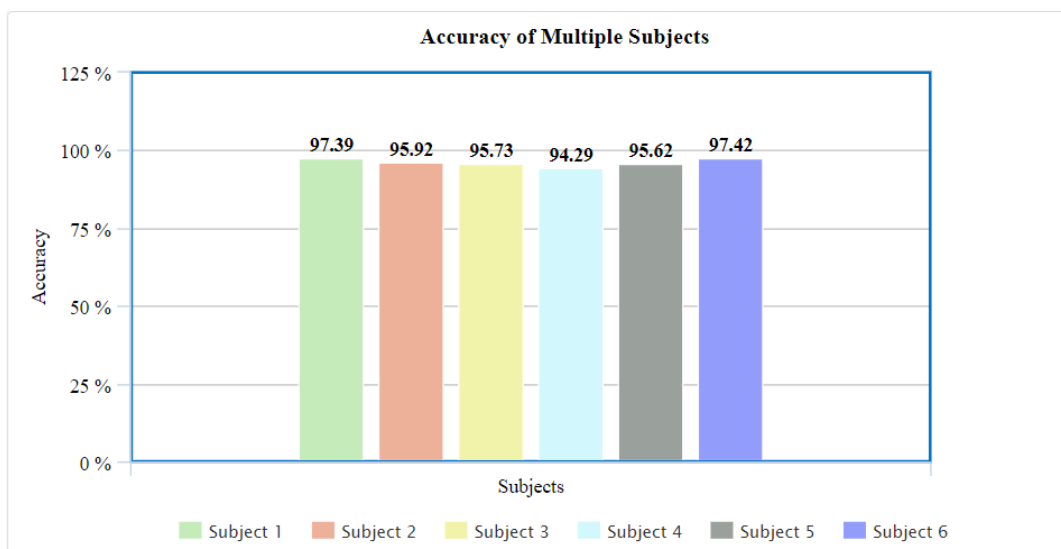


Figure 5.3: Accuracy of the model from random six subject’s data.

As visualized above, we can see that the accuracy of this proposed model is always higher than 94 % and it has obtained highest accuracy of 97.42 % which is very good value compared to other related methods.

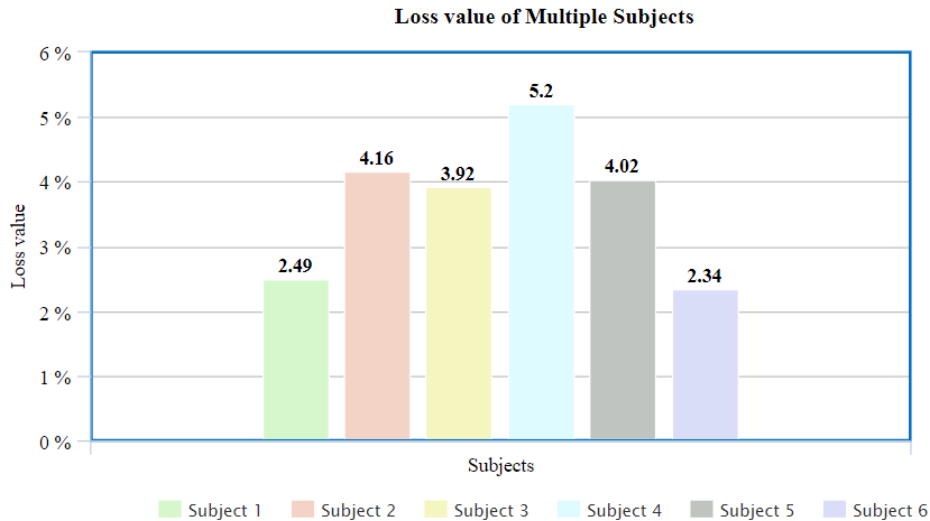


Figure 5.4: Loss Value of the model from random six subject's data.

It is shown in the bar diagram that the loss value is too low for our proposed model which is in between 2.49 % to 5.2 %. Thus, it shows that the model is excellent in performance.

Moreover, the accuracy of this proposed approach was increasing after each epoch. We applied a total number of 20 epochs and the accuracy was getting better and better till 20 epochs as shown in the graph of Figure 5.5. The value of the batch size was 20 which was enough to train the model effectively.

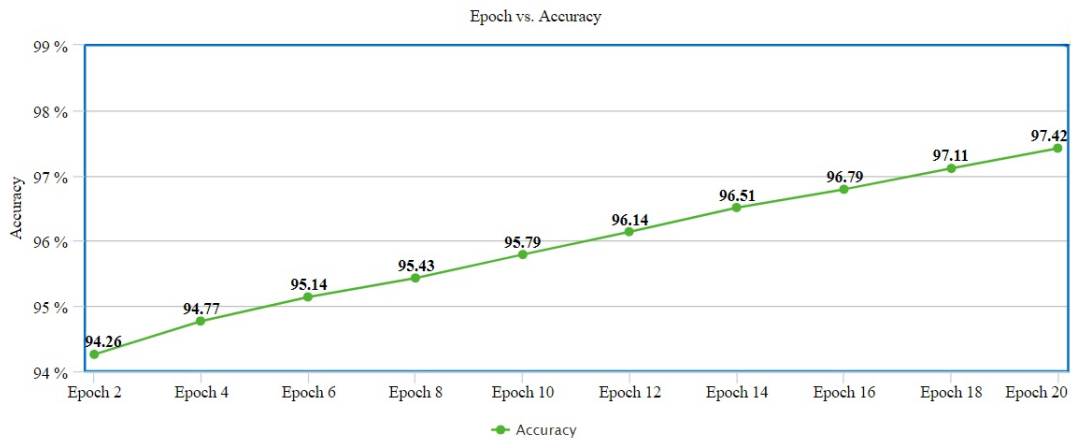


Figure 5.5: Epoch vs. Accuracy graph.

Additionally, the loss value of the proposed algorithm was decreasing after each epoch as shown in graph of Figure 5.6. The epoch number was 20 and the batch size was 20 for the model.

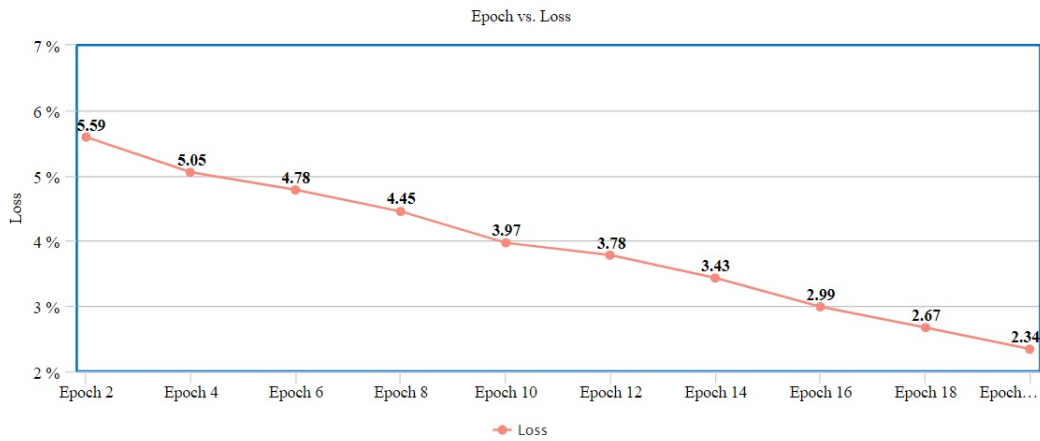


Figure 5.6: Epoch vs. Loss graph.

From the line diagram above, we can observe that the loss value is decreasing after each epoch and at the end of 20 epoch the value is 2.34 %. This value can be further minimized by increasing the number of epochs.

In the next chapter, we will summarize by briefly describing about all the procedures from the beginning and discuss about our approach on the proposed model and the effectiveness of it too. Moreover, we will describe about our future plan for this research as well.

Chapter 6

Conclusion

In this research, we used multimodal 3D MRI brain images to create a semantic segmentation network to segment brain tumors. T1-weighted MRI images were taken into account for the input part. We have applied several pre-processing techniques to prepare the dataset for the further analysis and classification. Image conversion, image normalization, data augmentation, image binarization were some of the pre-processing techniques that we have been applied in the dataset. Moreover, we applied neural network based autoencoder technique to learn our model about the features of tumor image. U-Net autoencoder model was selected for this purpose. After the model learned about the segmented tumor image, we applied the model for the new 3D MRI brain images to check whether the model could segment the tumor region or not. 70 percent of the dataset were used to train the proposed model and the rest 30 percent of data were applied to test the data. To train the model, high portion of dataset have been used so that the model could learn about the key features of brain tumor efficiently. Adam optimization method was applied to the training data to update the network weights iteratively. We have tried various data post-processing techniques to have better segmentation predictions but it did not help much. In this process, it helped for some images but some other image segmentation accuracy got worse than before.

Additionally, to improve the performance of our result, we increased the network depth further but it was not beneficial to do that. Also, we increased the network width which means we added the number of features to improve the result and we were successful in that procedure. The result was consistently improved and the accuracy got better after applying this method. Firstly, we have applied the proposed model into the training dataset which have the ground truth segmentation labels along with the brain images to see whether the model can detect tumor from that approach. The approach have been successful as the model could extract the tumor from the given MRI brain images and the accuracy is 97.12 % (after 20 epochs). Finally, we have applied the model to the new dataset which does not have the ground truth segmentation labels. The proposed model have been able to extract the tumor region from those 3D MRI brain images as well with the accuracy of 96.06 % which is pretty high as compared to the other related approaches.

6.1 Future Work

The future work for our research will include the extension of this proposed methodology to detect different types of tumor in all parts of the human body. In this project, we were only able to visualize the segmented tumor from a given brain image but for the future work the actual size of the tumor can be measured to assist in further tumor analysis and surgical planning. In addition to that, for faster implementation and better accuracy in future work, pre-trained open source CNN model can be used.

Bibliography

- [1] J. F. Schenck, “The role of magnetic susceptibility in magnetic resonance imaging: Mri magnetic compatibility of the first and second kinds”, *Medical physics*, vol. 23, no. 6, pp. 815–850, 1996.
- [2] V. S. Khoo, D. P. Dearnaley, D. J. Finnigan, A. Padhani, S. F. Tanner, and M. O. Leach, “Magnetic resonance imaging (mri): Considerations and applications in radiotherapy treatment planning”, *Radiotherapy and Oncology*, vol. 42, no. 1, pp. 1–15, 1997.
- [3] C. Leighton, B. Fisher, G. Bauman, S. Depiero, L. Stitt, D. Macdonald, and G. Cairncross, “Supratentorial low-grade glioma in adults: An analysis of prognostic factors and timing of radiation”, *Journal of Clinical Oncology*, vol. 15, no. 4, pp. 1294–1301, 1997.
- [4] H. S. Phillips, S. Kharbanda, R. Chen, W. F. Forrest, R. H. Soriano, T. D. Wu, A. Misra, J. M. Nigro, H. Colman, L. Soroceanu, *et al.*, “Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis”, *Cancer cell*, vol. 9, no. 3, pp. 157–173, 2006.
- [5] D. Lu and Q. Weng, “A survey of image classification methods and techniques for improving classification performance”, *International journal of Remote sensing*, vol. 28, no. 5, pp. 823–870, 2007.
- [6] G. Beattie, “What we know about how the human brain works”, in *How public service advertising works*, World Advertising Research Center, 2008.
- [7] R. B. Buxton, *Introduction to functional magnetic resonance imaging: principles and techniques*. Cambridge university press, 2009.
- [8] C. R. Cloninger, “Evolution of human brain functions: The functional structure of human consciousness”, *Australian and New Zealand Journal of Psychiatry*, vol. 43, no. 11, pp. 994–1006, 2009.
- [9] M. Elmaoğlu and A. Çelik, “Mri pulse sequences”, in *MRI Handbook*, Springer, 2011, pp. 47–68.
- [10] P. Natarajan, N. Krishnan, N. S. Kenkre, S. Nancy, and B. P. Singh, “Tumor detection using threshold operation in mri brain images”, in *2012 IEEE International Conference on Computational Intelligence and Computing Research*, IEEE, 2012, pp. 1–4.
- [11] N. Gordillo, E. Montseny, and P. Sobrevilla, “State of the art survey on mri brain tumor segmentation”, *Magnetic resonance imaging*, vol. 31, no. 8, pp. 1426–1438, 2013.

- [12] C. Benson and V. Lajish, “Morphology based enhancement and skull stripping of mri brain images”, in *2014 International Conference on intelligent computing applications*, IEEE, 2014, pp. 254–257.
- [13] M. Madheswaran and A. S. Dhas, “An adroit naive bayesian based sequence mining approach for prediction of mri brain tumor image”, in *Fifth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, IEEE, 2014, pp. 1–7.
- [14] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, *et al.*, “The multimodal brain tumor image segmentation benchmark (brats)”, *IEEE transactions on medical imaging*, vol. 34, no. 10, pp. 1993–2024, 2014.
- [15] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition”, *arXiv preprint arXiv:1409.1556*, 2014.
- [16] W. Wang, Y. Huang, Y. Wang, and L. Wang, “Generalized autoencoder: A neural network framework for dimensionality reduction”, in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2014, pp. 490–497.
- [17] D. Yu, H. Wang, P. Chen, and Z. Wei, “Mixed pooling for convolutional neural networks”, in *International conference on rough sets and knowledge technology*, Springer, 2014, pp. 364–375.
- [18] S. Budday, P. Steinmann III, and E. Kuhl, “Physical biology of human brain development”, *Frontiers in cellular neuroscience*, vol. 9, p. 257, 2015.
- [19] D. R. Kaeli, P. Mistry, D. Schaa, and D. P. Zhang, *Heterogeneous computing with OpenCL 2.0*. Morgan Kaufmann, 2015.
- [20] A. Kendall, V. Badrinarayanan, and R. Cipolla, “Bayesian segnet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding”, *arXiv preprint arXiv:1511.02680*, 2015.
- [21] J. Schmidhuber, “Deep learning in neural networks: An overview”, *Neural networks*, vol. 61, pp. 85–117, 2015.
- [22] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
- [23] A. Gensler, J. Henze, B. Sick, and N. Raabe, “Deep learning for solar power forecasting—an approach using autoencoder and lstm neural networks”, in *2016 IEEE international conference on systems, man, and cybernetics (SMC)*, IEEE, 2016, pp. 002 858–002 865.
- [24] A. E. K. Isselmou, S. Zhang, and G. Xu, “A novel approach for brain tumor detection using mri images”, *Journal of Biomedical Science and Engineering*, vol. 9, no. 10, pp. 44–52, 2016.
- [25] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation”, in *2016 Fourth International Conference on 3D Vision (3DV)*, IEEE, 2016, pp. 565–571.

- [26] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, “Brain tumor segmentation using convolutional neural networks in mri images”, *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.
- [27] V. V. Priya, “Shobarani. “an efficient segmentation approach for brain tumor detection in mri””, *Indian Journal of Science and Technology*, vol. 9, no. 19, pp. 1–6, 2016.
- [28] B. Romera-Paredes and P. H. S. Torr, “Recurrent instance segmentation”, in *European conference on computer vision*, Springer, 2016, pp. 312–329.
- [29] S. Shanmuganathan, “Artificial neural network modelling: An introduction”, in *Artificial neural network modelling*, Springer, 2016, pp. 1–14.
- [30] J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, “A distinctive approach in brain tumor detection and classification using mri”, *Pattern Recognition Letters*, 2017.
- [31] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, “Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features”, *Scientific data*, vol. 4, p. 170117, 2017.
- [32] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, J. Freymann, K. Farahani, and C. Davatzikos, “Segmentation labels and radiomic features for the pre-operative scans of the tcga-lygg collection”, *The Cancer Imaging Archive*, vol. 286, 2017.
- [33] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, “Deep features learning for medical image analysis with convolutional autoencoder neural network”, *IEEE Transactions on Big Data*, 2017.
- [34] T. Kessler, G. Dorian, and J. H. Mack, “Application of a rectified linear unit (relu) based artificial neural network to cetane number predictions”, in *ASME 2017 Internal Combustion Engine Division Fall Technical Conference*, American Society of Mechanical Engineers Digital Collection, 2017.
- [35] W. Mengqiao, Y. Jie, C. Yilei, and W. Hao, “The multimodal brain tumor image segmentation based on convolutional neural networks”, in *2017 2nd IEEE International Conference on Computational Intelligence and Applications (IC-CIA)*, IEEE, 2017, pp. 336–339.
- [36] G. Shukla, S. Bakas, S. Rathore, H. Akbari, A. Sotiras, and C. Davatzikos, “Radiomic features from multi-institutional glioblastoma mri offer additive prognostic value to clinical and genomic markers: Focus on tcga-gbm collection”, *International Journal of Radiation Oncology• Biology• Physics*, vol. 99, no. 2, E107–E108, 2017.
- [37] A. Wanto, A. P. Windarto, D. Hartama, and I. Parlina, “Use of binary sigmoid function and linear identity in artificial neural networks for forecasting population density”, *International Journal Of Information System & Technology*, vol. 1, no. 1, pp. 43–54, 2017.

- [38] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. T. Shinohara, C. Berger, S. M. Ha, M. Rozycki, *et al.*, “Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge”, *arXiv preprint arXiv:1811.02629*, 2018.
- [39] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation”, in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818.
- [40] S. Indolia, A. K. Goswami, S. Mishra, and P. Asopa, “Conceptual understanding of convolutional neural network-a deep learning approach”, *Procedia computer science*, vol. 132, pp. 679–688, 2018.
- [41] M. R. Islam and N. Rishad, “Effects of filter on the classification of brain mri image using convolutional neural network”, in *2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT)*, IEEE, 2018, pp. 489–494.
- [42] Y. Kene, U. Khot, and I. Rizvi, “A survey of image classification and techniques for improving classification performance”, *Available at SSRN 3349696*, 2018.
- [43] J. R. McFaline-Figueroa and E. Q. Lee, “Brain tumors”, *The American journal of medicine*, vol. 131, no. 8, pp. 874–882, 2018.
- [44] A. Myronenko, “3d mri brain tumor segmentation using autoencoder regularization”, in *International MICCAI Brainlesion Workshop*, Springer, 2018, pp. 311–320.
- [45] R. Vinoth and C. Venkatesh, “Segmentation and detection of tumor in mri images using cnn and svm classification”, in *2018 Conference on Emerging Devices and Smart Systems (ICEDSS)*, IEEE, 2018, pp. 21–25.
- [46] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: An overview and application in radiology”, *Insights into imaging*, vol. 9, no. 4, pp. 611–629, 2018.
- [47] V. Zyuzin, P. Sergey, A. Mukhtarov, T. Chumarnaya, O. Solovyova, A. Bobkova, and V. Myasnikov, “Identification of the left ventricle endocardial border on two-dimensional ultrasound images using the convolutional neural network unet”, in *2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT)*, IEEE, 2018, pp. 76–78.
- [48] M. M. Thaha, K. P. M. Kumar, B. Murugan, S. Dhanasekeran, P. Vijayakarthish, and A. S. Selvi, “Brain tumor segmentation using convolutional neural networks in mri images”, *Journal of medical systems*, vol. 43, no. 9, p. 294, 2019.
- [49] K. Meena, S. Pavitra, N. Nishanthi, and M. Nivetha, “A novel method using convolutional neural network for segmenting brain tumor in mri images”,