Brain Tumor Segmentation from MRI images using Convolutional Neural Networks

by

Mushfiqur Rahman Khan
19241006

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

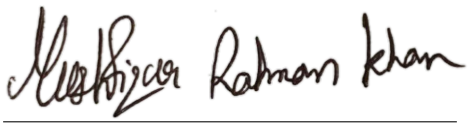
Department of Computer Science and Engineering
Brac University
Spring 2022

Declaration

It is hereby declared that

1. The thesis submitted is my 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:



Mushfiqur Rahman Khan

19241006

Approval

The thesis titled "Brain Tumor Segmentation from MRI images using Convolutional Neural Networks" submitted by

1. Mushfiqur Rahman Khan (19241006)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on May 29, 2022.

Examining Committee:

Supervisor:
(Member)

Ms. Arnisha Khondaker

Lecturer
Department of Computer Science and Engineering
BRAC University

Program Coordinator:
(Member)

Dr. Golam Rabiul Alam, PhD

Associate Professor
Department of Computer Science and Engineering
BRAC University

Head of Department:
(Chair)

Dr. Sadia Hamid Kazi, PhD

Chairperson and Associate Professor
Department of Computer Science and Engineering
BRAC University

Abstract

Brain tumors result from the accumulation of abnormal cells inside the brain. The process of brain tumor segmentation plays a vital role in detecting early-stage brain tumors. There are several challenges in the tumor segmentation process due to the variations in size, location, and intensity of brain tissues. Traditional brain tumor segmentation methods are very time-consuming as segmentation is carried out manually. Automated segmentation methods are necessary for rapid diagnosis and treatment. In our paper, we highlight the complications surrounding a brain tumor and use an encoder-decoder-based approach of CNN algorithms to train segmentation models that will help identify and localize brain tumors with the utmost accuracy. We have used four CNN architectures to train our models, namely U-Net, ResNet50, ResNext50, and EfficientNetB7. For our encoder-decoder models, we used ResNet50, ResNext50, and EfficientNetB7 as the encoder blocks of U-Net in three different models. Hence, we performed three experiments using these four architectures and compared their performance. We obtained the best results using the U-Net + EfficientNetB7 model.

Keywords: Bain Tumor, Convolutional Neural Network, U-Net, ResNet50, ResNext50, EfficientNetB7, Segmentation, MRI image

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis have been completed without any major interruption.
Secondly, to our supervisor Ms. Arnisha Khondaker ma'am for her kind support and advice in our work. She helped us whenever we needed help.
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

List of Figures

List of Tables

Nomenclature

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

CNN Convolutional Neural Networks

DC Dice Coefficient

IoU Intersection over Union

LGG Lower Grade Glioma

MRI Magnetic Resonance Imaging

$ReLU$ Rectified Linear Unit

$TCGA$ The Cancer Genome Atlas

$TCIA$ The Cancer Image Archive

Chapter 1

Introduction

Brain tumors are formed when abnormal cells inside the brain reproduce and multiply in an uncontrollable way. There are various subtypes of brain tumors, but they are mainly classified into two categories pertaining to primary brain tumors: benign and malignant. The benign(noncancerous) tumors are formed slowly, usually isolating themselves from the neighboring normal brain tissue, and are less aggressive. In contrast, malignant(cancerous) tumors are harder to identify and distinguish from the normal brain tissues. Therefore, diagnosing and treating the malignant tumor is more difficult without violating the neighboring brain tissues.
Diagnosis of a brain tumor starts with a doctor asking about the symptoms, family medical history, overall health, etc., followed by a routine physical test. If the doctor suspects presence of a brain tumor, he may ask to do a medical imaging examination. There are several techniques the doctor may choose. Medical imaging techniques such as Magnetic Resonance Imaging(MRI) or the Computerized Tomography Scan (CT scan) helps display detailed images of the tissues inside the brain. Other methods like the Magnetic Resonance Angiogram(MRA) scan the blood vessels inside the brain to identify abnormalities.
MRI is one of the most frequently used techniques in medical imaging. It provides high-resolution, detailed images of the internal structure of the body that aids in diagnosing a wide range of problems. This technique uses magnetic field and computer-generated radio to create comprehensive images of tissues and organs. MRI provides clear and detailed information about the dimension, shape, and location of brain tumors [1]. Therefore, most researchers in the medical field use MRI for brain tumor diagnosis.
Image segmentation is a critical aspect in the medical image processing sector. It is the procedure of breaking down digital images into segments or instances through classical AI-based techniques. Therefore, segmenting brain tumors from MRI images is essential for an accurate representation of the measurement of the tumor, which can assist in better diagnosis and treatment [2].

1.1 Research Problem

According to world health organization (WHO), malignant brain tumor accounts for 1284 new cases and 1144 deaths in Bangladesh alone in the last five years [3]. In the global context, the number of cases and deaths account for 308,102 and 251,329, respectively [4]. In 2007, WHO announced a grading scheme for classifying brain

tumors on a scale of I to IV. Grade I tumors are slow-growing and least malignant. The severeness of the abnormalities increases with each grade. Grade IV tumors are most life-threatening and can have bizarre appearances under a microscope [5].
In recent times, there have been significant advancements in the study of brain tumor segmentation. Traditional machine learning methods such as Random Forest [6] and Support Vector Machine [7] were previously applied to segment brain tumors with a sound success rate. But recently, state-of-the-art deep learning methods with Convolutional Neural Network(CNN) have become widely favorable among researchers. One of these is the U-Net segmentation architecture. It was developed for biomedical image segmentation at the University of Freiburg, Germany, by Ronneberger et al. Since then, researchers have created different variations of U-Net by combining them with other architectures to produce better results. CNNs are deep learning algorithms that can handle large amounts of data. Many researchers prefer working with CNNs as they can provide superior performance in medical imaging compared to traditional machine learning algorithms.
Amid all the breakthroughs, accurate and reproducible segmentation of brain images remains a big challenge. For the most part, brain tumor segmentation methods remain manual, time-consuming, and susceptible to human error. Hospitals and clinics generate a considerable number of brain images every day. However, it is a cumbersome process to manually annotate and segment these images at a reasonable rate. The regions of the brain have vague boundaries between normal and abnormal brain tissues. Although researchers have previously used previous information of location and shape to segment other organs, such as the liver, the location, dimension, and shape of brain tumors vary massively across multiple patients [8]. Due to this, it is challenging to segment the size and location of brain tumors using prior knowledge.

1.2 Research Objective

Our motivation behind writing this paper is to propose and implement a state-of-the-art CNN model that will help systematically diagnose patients with brain tumors correctly with the least amount of error. Our goal is to perform semantic segmentation on MRI images using annotated dataset of brain tumors. We believe proper detection of brain tumors at an early stage will assist clinicians and doctors take necessary steps before the tumor becomes life-threatening. The main objectives of this paper are emphasized below:
1. Collect annotated dataset of brain tumor MRI images and perform pre-processing.
2. Perform training using three models: U-Net + ResNet50 as encoder, U-Net + ResNext50 as encoder), and U-Net + EfficientNetB7 as encoder).
3. Analyze the models' performance using Intersection over Union (or Jaccard Index) and Dice Coefficient and propose the best model.
We organized our paper as follows:
1. In chapter 1, we start with an introduction and the main goals of our paper.
2. In chapter 2, we provided a discussion of background study related to brain tumors, CNN architecture, and image segmentation. We also highlight some related works based on past segmentation research on brain tumors.
3. In chapter 3, we described the datasets that we used in our models and the necessary pre-processing steps that we performed before implementing them.

4. In chapter 4, we defined the structures of our models and defined the implementation steps and training process.
5. In chapter 5, we analyzed the segmentation results and compared the models' performance.
6. In chapter 6, we provided a conclusion and brief summary of our paper.

Chapter 2

Literature Review

2.1 Background Study

The most sophisticated organ in the human body is the brain. It controls memory, thought, touch, emotion, temperature, hunger, breathing, and all the essential functions required to regulate the body. The brain weighs about three pounds and is made up of 60% fat. The remaining 40% is salt, carbohydrate, water, and protein. The brain lies within the shell of the skull bone and is protected by a layer of thick protective fluid. The brain is the control center of the whole nervous system. It controls almost every functions in the human body. The brain makes up the central nervous system with the spinal cord. Information transfers from the brain to the rest of the body through electrical and chemical signals. The three main components of the brain are the cerebrum, brainstem, and cerebellum. Among these, the cerebrum is the largest. It is responsible for processing language, speech and visual information. The brainstem connects the cerebrum to the spinal cord. And lastly, the cerebellum (also known as the "little brain") helps maintain equilibrium and balance.

2.1.1 Cerebrum

The cerebrum is the biggest part of the brain [9]. It is situated in the uppermost part of the brain, and it is divided into the right and left hemispheres. The right hemisphere coordinates the left part, and the right hemisphere coordinates the right part of the body. They are linked together by a bundle of neural fibers known as the "corpus callosum." It is responsible for passing information between the two hemispheres. Each hemisphere is separated into four regions: 1. frontal lobe, 2. parietal lobe, 3. temporal lobe, and 4. occipital lobe. Each of them has specific functions. A pictorial representation of the lobes are visualized in figure 2.1.
The largest among all the lobes is the frontal. It is one of the most common regions of injury in the brain. It is situated in front of the cerebral hemisphere within the anterior cranial fossa [10]. The frontal lobe is responsible for expressing language, coordinating movement, and functions related to cognitive skills such as planning and organizing. Furthermore, the frontal lobe is in command of our behavior and emotional response and defines our personality. Any complication to the frontal lobe may result in a range of problems like movement difficulty, problem-solving difficulty, and changes in mood, behavior, and personality.

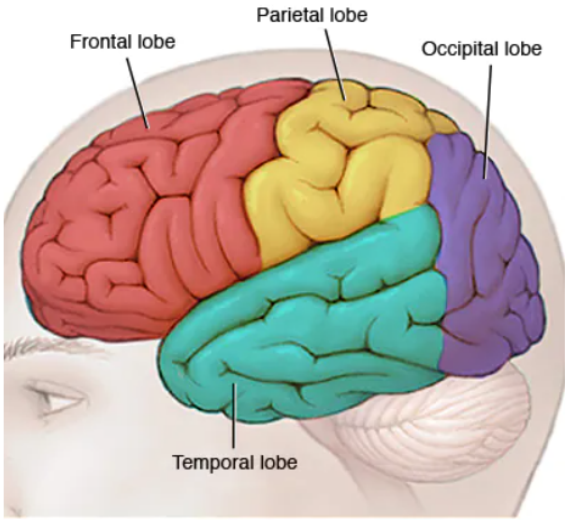


Figure 2.1: Four lobes of the brain

The parietal lobe [11] is situated in front of the occipital lobe, behind the frontal lobe. It is the brain's primary sensory region that is dedicated to carrying out a range of processing. The parietal lobe is essential for perceiving sensations like touch, smell, and sound. Sometimes the parietal lobe works in conjunction with the occipital lobe for processing other tasks like reasoning and spatial navigation. Damage to any region of the parietal lobe might cause a host of problems. For instance, a person may struggle to write, express language, and perceive objects if they suffer any damage in the left parietal lobe. On the other hand, injury to the right parietal lobe may result in a person neglecting self-well-being and any deficit in one side of the body.
The smallest lobe is the occipetal lobe [12]. It rests behind the temporal and parietal lobes. It derives its name as so because it sits below the skull's occipital bone. The occipital lobe's primary function is interpreting visual information and decoding them. But recent research suggests that the occipital lobe can perform other tasks as it receives messages from other regions of the brain. Additionally, the occipital lobe can perform tasks such as determining color, distance, depth, and size. Any damage to the occipital lobe can result in drastic visual complications, most severely blindness. Other complications include epilepsy, difficulty in movement, and recognizing colors and objects.
The temporal lobe sits below the frontal and parietal lobes and in front of the occipital lobe [13]. It is positioned behind the temples of the skull; from which it derives its name. The temporal lobe deals with auditory stimulation, emotion, and memory. With the help of the primary auditory cortex, the temporal lobe interprets sound data and helps understand and recognize language. It also helps remember and recall verbal information. Injury to the temporal lobe can result in impaired long-term memory, difficulty understanding verbal words and languages, and "temporal lobe epilepsy," which is a common cause of seizures.

2.1.2 Cerebellum

The cerebellum (also known as the "little brain") is an essential component of the brain that regulates voluntary movement [14]. It also helps maintain balance, posture, equilibrium, and other motor functions. This small section of the brain takes up just about 10% of the entire volume of the brain but carries about 50%-80% of the brain's total neurons [15]. The cerebellum is situated below the occipital and temporal lobes. A tough layer of dura matter known as the "tentorium cerebella" separates the cerebellum from the lobes. The cerebellum consists of two lateral hemispheres: the cerebellar cortex and the cerebellar nuclei. The cerebellar cortex is formed by tightly folded grey matter, and it contains most of the brain's neurons. Similarly, embedded in white matter are the cerebellar nuclei. There are four cerebellar nuclei in total: emboliform, globose, dentate, and fastigial nuclei.
The cerebellum carries out several crucial functions in the brain. First and foremost, it helps maintain balance [16]. It helps maintain necessary adjustments in posture and equilibrium by regulating muscle tension. The cerebellum works in conjunction with other parts of the brain and sends signals back and forth to convey messages. To initiate the movement of a muscle, the cerebellum sends signals to the spinal cord about the impulse of the intended movement. In return, the cerebellum receives information from other muscles about any movement occurring at some given time. The cerebellum does not directly influence motor commands. Instead, it modifies the command to make movements more accurate and adaptive. If the cerebellum detects any discrepancies in movement, it will send information directly to the anterior horn and cerebral cortex of the spinal cord to fix the discrepancy immediately so that movement can continue to carry out accurately.
As the cerebellum is such a vital part of carrying out motions of the muscles, cerebral damage can lead to severe movement disorder. Much of the known information about cerebellar functions stems from examining subjects with cerebellar damage. Ataxia is one of the most common forms of cerebral disorder. This condition makes victims lose their muscle control and coordination. Maintaining speech also becomes difficult. Ataxia can be caused by excessive damage to the nerve cells, blood clot, or tumors in the cerebellum. Research also shows that it can be caused by toxins in the brain or by a stroke.

2.1.3 Brainstem

The brainstem is the area of the brain that sits between the cerebral hemispheres and the spinal cord that carries out several vital roles for the brain [17]. It resides in the lower part of the brain, and it is made up of three components: the midbrain, pons, and medulla oblongata. Each component carries out unique tasks that help regulate several essential functions like breathing, heartbeat, blood pressure, and a few others. They also accommodate a collection of cranial nerve nuclei and are pathways for several essential neural routes. These functions are possible due to the brainstem's unique anatomy.
The brainstem is a small but vital structure that acts as a bridge between the brain and other parts of the body. Therefore, several neurological conditions may arise if any damage occurs within the brainstem. Brainstem stroke is a common condition that patients suffer. It stops receiving blood supply from the several arteries it is attached to, namely the basilar artery, vertebral artery, and pontine

artery. Several movement disorders are also associated with brainstem injury, most commonly Parkinson's disease.

2.1.4 Brain tumor

Brain tumors have been a subject of research for many years as it they are one of the most complicated diseases to treat. Brain tumors are masses of abnormal cells(visualized in figure 2.2) that form inside the brain. It can develop in any part of the brain when left unchecked by the mechanism that controls normal cell growth. WHO have classified over 130 different kinds of brain tumors, but they are mainly termed as two groups, primary and metastatic [18], for their unique originated locations.

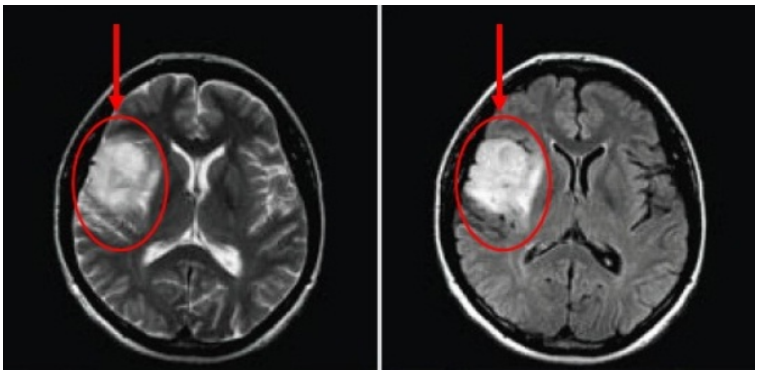


Figure 2.2: MRI image of brain tumor

Primary brain tumors occur in the brain and spinal cords and their immediate surrounding regions. They tend to stay within these regions without spreading to other organs of the body. Primary brain tumors are considered very rare, but they are becoming more prevalent every day as human life expectancy increases and doctors are coming up with new ways to diagnose them. Children and adults alike can suffer from primary brain tumors. But the way they develop in children and adults can differ in some ways. For instance, the tumor may appear in different sections of the brain. Additionally, it may originate from different brain cells. Treatment and outcome may also vary between children and adult patients.
The main factors that lead to primary brain tumors are still unknown. But there are some proven risk factors that researchers believe lead to a primary brain tumor. One such risk factor is prolonged exposure to radiation and radiation therapy, which is believed to increase the probability of brain tumor development. Another risk factor is gene inheritance like Von Hippel-Lindau syndrome, Li-Fraumeni syndrome, neurofibromatosis, and other conditions. Some non-proven risk factors include smoking, head injury, hormone therapy, cell phone usage, and alcohol consumption.
Primary brain tumors can be further categorized into two categories: benign and malignant [18]. The world health organization(WHO) had introduced a grading system in 2008 to describe a tumor's abnormality level under a microscope. They are described as follows:
Low Grade:
WHO Grade I: These are the least malignant kinds of tumors that patients can get rid of through surgery. As these tumor cells are slow-growing, patients can expect long-term survival before the tumor reaches any critical stage.

WHO Grade II: These are also relatively slow-growing, but they may spread to surrounding tissues. Nevertheless, the chances of the tumor growing back are higher than grade I tumors.
High Grade:
WHO Grade III: These are somewhat malignant compared to grade I and grade II tumors. They spread quickly and grow at a relatively faster rate. There is a very high chance that the tumor returns with a higher grade.
WHO Grade IV: These are the most malignant and aggressive variants. They do not look like normal cells under the microscope. They spread and multiply rapidly. These tumors almost always re-occur after surgery.
In contrast to primary brain tumors, metastatic brain tumors (also known as secondary brain tumors) form when cancer from other parts of the body spreads to the brain through the bloodstream. Tumors that develop in this manner are called "brain metastases." They are more common than primary brain tumors, and they tend to be more prevalent in adults than children. Brain metastases may occur from any cancer, but most commonly from primary cancer in the lung, breast, colon, kidney, and melanoma.
Meningiomas are a commonly occurring primary brain tumor of the benign variant. They originate from the meninges, and they account for 10 to 15 percent of all brain tumors. These tumors are rarely malignant, but sometimes they are difficult to detect as many patients never show any symptoms. Occasionally, there may be seizures or vision and hearing impairment symptoms. Similarly, gliomas are the most commonly occurring primary brain tumor of the malignant variant [19]. They account for 30 percent of all brain tumors and about 80 percent of all malignant tumors. They form from the glial cells inside the brain and spinal cord. These tumors are primarily prevalent in adults. But, some low-grade gliomas might form in children as well. According to the American Cancer Society, there are three types of gliomas: astrocytomass, oligodendrogliomas, and ependymomas.

2.1.5 Image Segmentation

Image segmentation is one of the most critical applications that help partition images into multiple segments in image processing and computer vision. Image segmentation aims to modify or simplify an image so that it is easier to analyze and process. It is an extension of image classification and is typically used to localize objects in images by classifying labels to every pixel that shares specific characteristics. Brain MRI segmentation is used to visualize and measure the brain's anatomical structures, analyze brain changes, outline pathological areas, image-guided interventions, and surgical planning.
Image segmentation can be broadly categorized into two types: semantic segmentation and instance segmentation. Semantic segmentation [20] is the process of classifying pixels in an image belonging to a particular label. It has several real-world applications such as self-driving vehicles, medical image processing, aerial image processing, scene understanding, and many others. Semantic segmentation only cares about categorizing the classes of each pixel and does not take into account splitting instances of the same object. On the other hand, instance segmentation [21] separates each instance of an object class and attaches a unique label to it. It can also provide the locations of each detected object. This segmentation method

can be applied in images with several distinct objects of the same type and needs accurate localization.

2.1.6 Convolutional Neural Networks

Neural networks are a subset of machine learning used for solving artificial intelligence(AI) related problems [22]. Every neural network has of an input layer, one or many hidden layers, and an output layer. The layers are connected by a series of artificial neurons (or nodes), with each connection having a corresponding weight and threshold. When an output of any particular neuron is higher than a set threshold value, that neuron gets activated, and data is sent over to the next layer.
There are several types of neural networks for solving different tasks. But, the most commonly used neural network used for computer vision applications is convolutional neural networks(CNN). Before CNN, extracting features for object detection and classification was time-consuming and manual. However, convolutional neural networks provide a more efficient approach to image classification, detection, and segmentation tasks. The basic building blocks of a CNN architecture is visualized in figure 2.3 below.

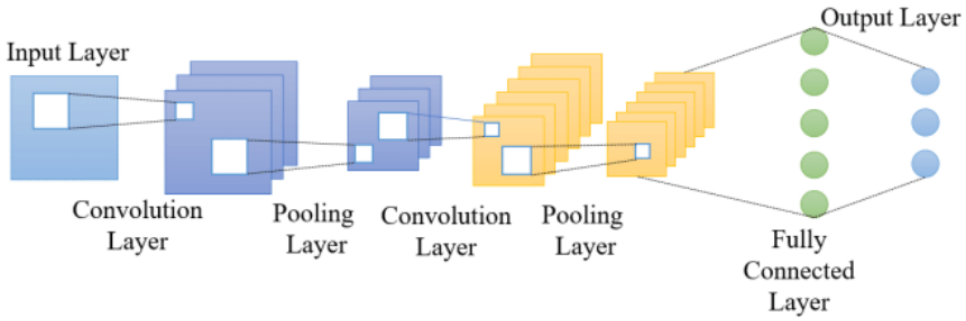


Figure 2.3: Basic architecture of a CNN

Researcher Yann LeCun first introduced CNN in 1989 [23]. The earliest version of CNN was named LeNet (after LeCun), a hand-written recognition program with a 7-layer convolutional network. It quickly found mass relevance in banking and postal services to recognize numbers and alphabets in checks and letters. But, the ability to recognize higher resolution images required a massive amount of data and more layers of the network. Therefore, it was only constrained to processing lower resolution images. Fast-forward to the present, there are now a large number of CNN architectures with applications in several fields. Some of these architectures are U-Net, VGG, Inception, ResNet, ResNext, DenseNet, MobileNet and many others. CNN architectures are superior to other neural networks in their performance with images. They contain mainly three types of layers: 1. Convolutional layer, 2) Fully connected(FC) and 3) Pooling layer.
The convolutional layer is the first layer and the primary building block in any CNN network. The goal of the convolutional layer is to detect features in the images received as inputs. It receives several images as input, and it contains special filters that are applied to the inputs. Over time, repeated application of the same filters in the inputs results in activation of feature maps. This indicates where each feature

is in each image [24]. Next, the pooling layer works on the feature maps received from the convolutional layer. Its aim is to lower the number of parameters and the number of calculations done by the network. Additionally, it helps reduce the size of the images while conserving their important attributes. This makes the network more efficient and avoids overfitting. Lastly, all the outputs from the convolutional and pooling layers are pushed inside the fully connected layer. This layer is always the final layer of a CNN. Then, through the application of linear combination and activation function, it produces the desired output [25].

2.2 Related Works

A computer-aided detection(CAD) method was proposed by B. Devkota et al. that uses mathematical morphological reconstruction(MMR) for brain tumor segmentation and diagnosis at an early stage [26]. They used four different types of brain tumor MRI images: Meningioma, Glioma, Metastatic Adenocarcinoma, and Sarcoma. In their research, they first applied a median filter as pre-processing step for the images. Next, they implemented a Fuzzy C-Means algorithm for segmentation. Then they used statistical features for extracting features which was followed by Principal Component Analysis(PCA) for reduction of features. Finally, they used Support Vector Machine(SVM) with a non-linear kernel for classification.
R. Pitchai et al. developed a deep learning-based method that uses a combination of Fuzzy K-Means clustering and Artificial Neural Networks(ANN) to segment brain tumor MRI images [27]. First, they used a wiener filter for noise reduction during the pre-processing step. Then they extracted the important features from the filtered images using the CSOA algorithm. Next, they used ANN for classifying the images. Finally, they applied the Fuzzy K-Means algorithm for segmenting the tumor region. Few major drawbacks of their method is that their algorithm is susceptible to noise and computationally complex.
HNTK Kaldela et al. introduced a method that implements CNN and Faster Region-based CNN (Faster R-CNN) to classify and segment brain tumors [28]. They performed their test on 218 images using the python and Keras libraries. To increase computational speed, they down-sampled the original images to size 128*128 from 512*512. Next, they analyzed the accuracy for the classification of glioma and meningioma, where they acquired an accuracy of 94%. Then, they applied faster R-CNN on the classified images to locate the tumor region. Finally, they compared the results with the ground truth labels, where they acquired a performance level of 94.6% for tumor region extraction.
A novel approach was proposed by X. Zhao et al. for segmenting brain tumor that integrates Fully Convolutional Neural Networks(FCNN) and Conditional Random Fields(CRF), where they implemented three segmentation models [29]. They used 2D image slices and patches acquired in coronal, axial, and sagittal views to train the three models they implemented. For pre-processing, they implemented N4ITK and intensity normalization on the MRI images. They further improved the performance of the models by post-processing the segmentation results in which they removed the little 3D-connected regions and rectified false labels using thresholding. Their methods were performed on the datasets of BRATS 2013 and 2015. A major drawback of their algorithm is that it requires huge memory demand makes a trade-off between accuracy and resource consumption.

L. Ramya et al. proposed a robust segmentation algorithm to segment 2D brain tumors using morphological operators [30]. First, they used image denoising as a pre-processing step using a partial differential equation. Next, they used a seed region growing segmentation technique to detect the tumor. They also employed skull removing procedure using morphological operators to increase the brain tumor detection accuracy. One drawback of their algorithm is that it is prone to noise.
S. Alqazzaz et al. proposed the fully convolutional neural network called SegNet for automated brain tumor segmentation [31]. They used 285 patient data containing HGG and LGG images from the BRATS 2017 dataset. In their approach, they ran four MRI modalities separately. They used normalization and N4ITK during pre-processing to remove unwanted noise. Their architecture consisted of encoders and decoders, with each having thirteen convolutional layers with 3x3 filters. The decoder extracted high dimensionality features, which were then fed into softmax layers which classified each pixel separately. Their evaluated results achieved 85%, 81%, and 79% accuracy for complete, core, and enhanced tumor regions.

Chapter 3

Dataset

The dataset is downloaded from Kaggle.. The dataset consists of MRI images with FLAIR abnormality segmentation masks. The dataset is publicly available in The Cancer Genome Atlas(TCGA) [32]. In the TCGA lower-grade glioma(LGG) collection, there are 120 patients with image data, having at least a FLAIR sequence and genomic data cluster. The images were collected from The Cancer Image Archive(TCIA) [33], and they are the medical image data counterpart of the TCGA patient data. Patient data and tumor genomic cluster is included in the 'data.csv' file. Ten of the patients did not have genomic cluster data available hence they needed to be excluded. Patient data comprised of total 3929 image slices. We can see sample image with their corresponding tumor masks of 10 patients in figure 3.2. Out of these 2556 images contained positive diagnosis and 1373 images contained negative diagnosis(visualized in the bar chart in figure 3.1). All the images are in 256 x 256 format. The list of the 110 patients is from the following institutions:
1. 16 patients from Thomas Jefferson University, Philadelphia, PA (TCGA-CS).
2. 45 patients from Henry Ford Hospital, Detroit, MI (TCGA-DU).
3. 1 patient from the University of North Carolina, Chapel Hill, NC (TCGA-EZ).
4. 14 patients from Case Western Reserve University, Cleveland, OH (TCGA-FG).
5. 34 patients from Saint Joseph Hospital and Medical Center, Phoenix, AZ (TCGA-HT).

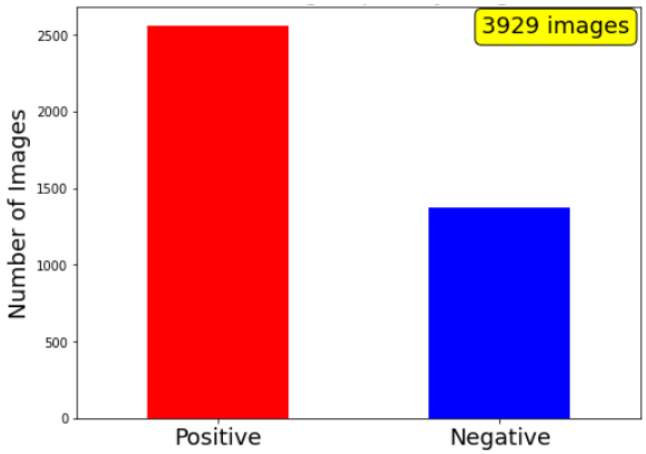


Figure 3.1: Patient data grouped by diagnosis. Positive diagnosis(Red bar) and Negative diagnosis(Blue bar).

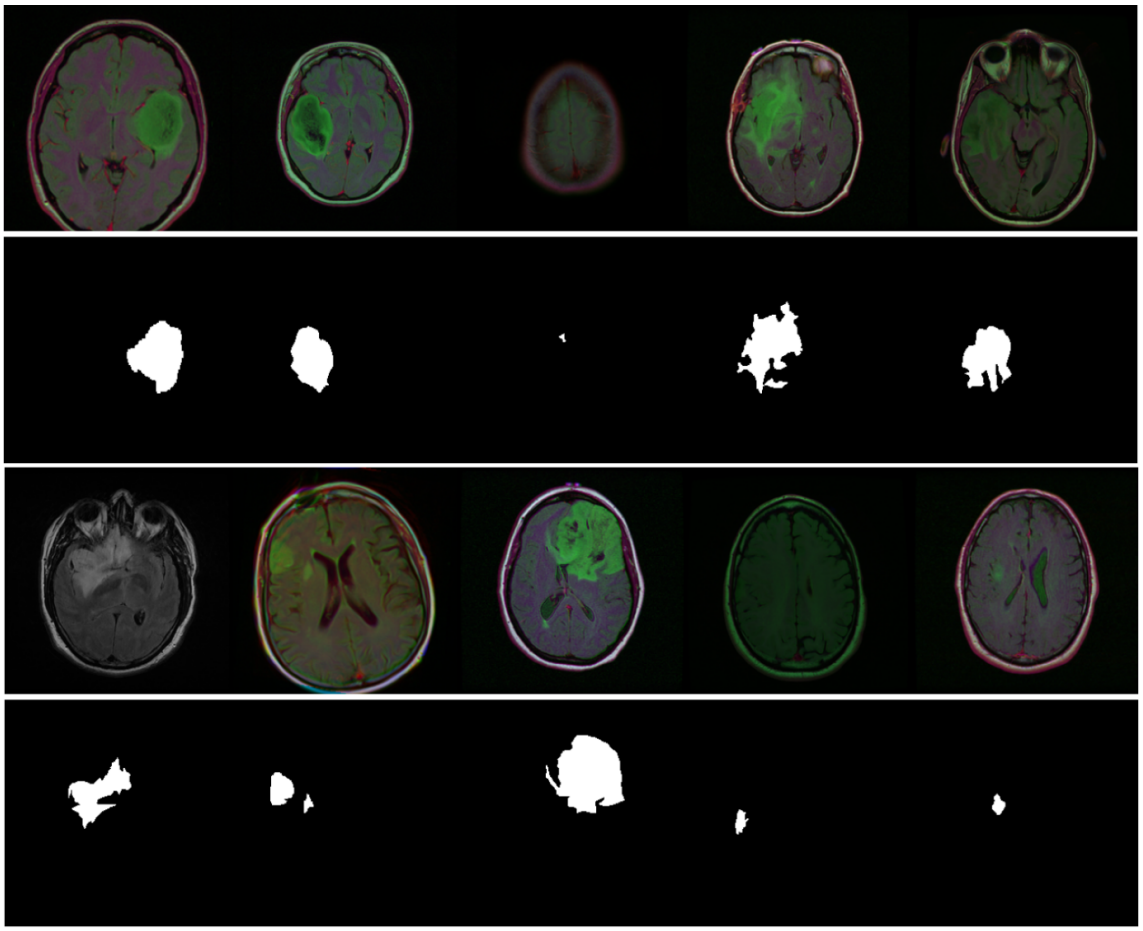


Figure 3.2: MRI images from the dataset with corresponding tumor masks. MRI images(1st and 3rd rows) and corresponding Tumor masks(2nd and 4th rows)

3.1 Data Pre-processing

Data pre-processing is a crucial step before starting our implementation of the models. It transforms the raw data into a well-organized format that is easily readable by the machine. This step is crucial because datasets often contain inconsistent, missing, or noisy data. Application of our models in these types of datasets would fail to identify patterns adequately. As a result, it will lead to false predictions and overall incorrect data statistics. Also increasing the size of the overall dataset allows the algorithm tow work with a more diverse amount of data that helps improve training and testing accuracies.
In our pre-processing steps, the images were first scaled to a common frame of reference. Image scaling refers to the resizing of digital images. Then, we performed some image augmentation techniques. In this process, we artificially increased the size of the dataset without actually acquiring new data. We applied simple transformation functions to generate these artificial sets of data. We applied channel dropout, color jitter, horizontal flip, vertical flip, and random rotations. Examples of three of the augmentations are shown in figure 3.3. Finally we applied data distribution where we distributed the data into three separate sets. 70% of the data was distributed in training set, 15% in the validation set, and 15% in the testing set.

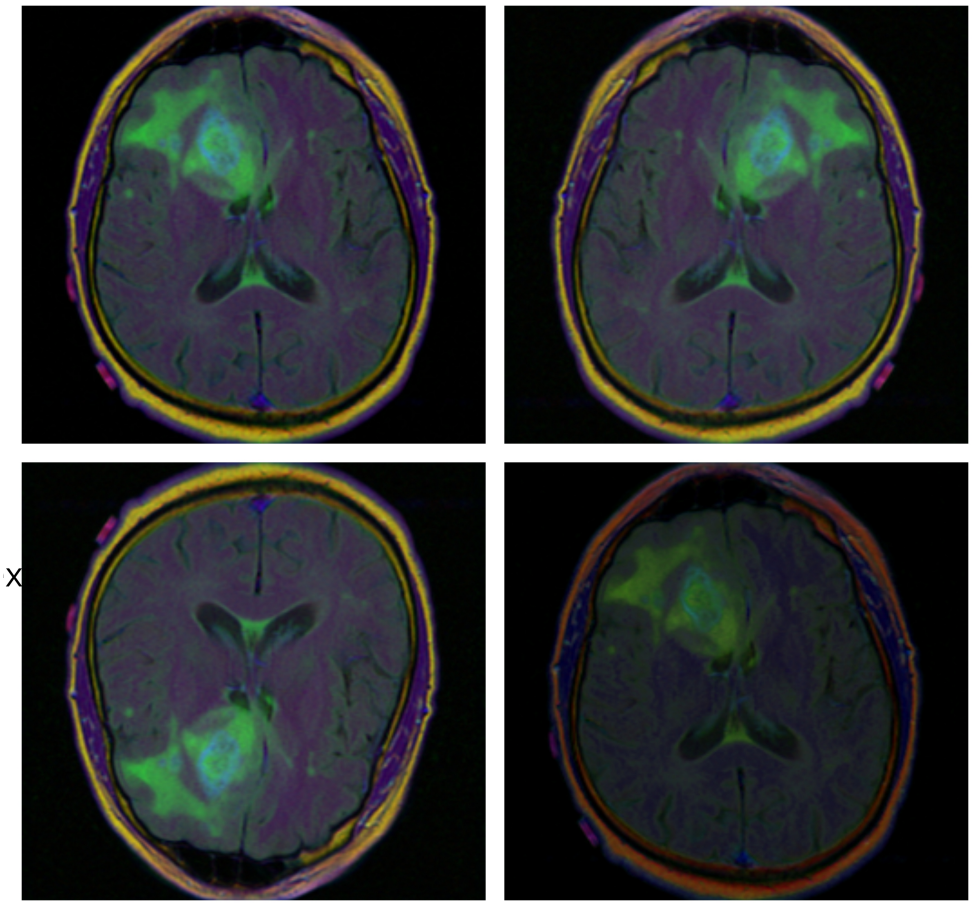


Figure 3.3: Original(Top Left). Horizontal Flip(Top Right). Vertical Flip(Bottom left). Color Jitter(Bottom right).

Chapter 4

Model Selection and Implementation

The following section provides description of the architecture of the models. We have chosen four CNN architectures to implement and test: U-Net, ResNet50, ResNext50, and EfficientNetB7 to implement in our encoder-decoder models.

4.1 U-Net

One the most versatile CNN architecture for image segmentation is the U-Net architecture. The network is built upon a fully convolutional network. It consists of two main paths, contracting (left side) and expansive (right side) [34]. Due to its U-shaped architecture, the contracting(encoder) and expansive(decoder) paths are more or less symmetric. An input image enters through the contracting path and goes through repeated applications of 3x3 convolutions. After each convolution, the rectified linear unit(ReLU) function is activated. Next, a 2x2 max pooling operation of stride 2 is applied to each image to reduce its spatial dimensionality. At each down-sampling step, the number of feature channel is doubled.
On the expansive path, every step consists of up-sampling the feature maps followed by 2x2 transpose convolution, which cuts down the number of feature channels in half. Next, feature maps also concatenate with corresponding cropped feature maps from the contracting path. The cropping is done because some border pixels are lost in each convolution. Then, 3x3 convolutions are applied repeatedly, with each application followed by ReLU operation. A 1x1 convolution is applied in the final layer which helps to map each feature to the preferred number of classes. Figure 4.1 visualizes the detailed architecture of the U-Net.

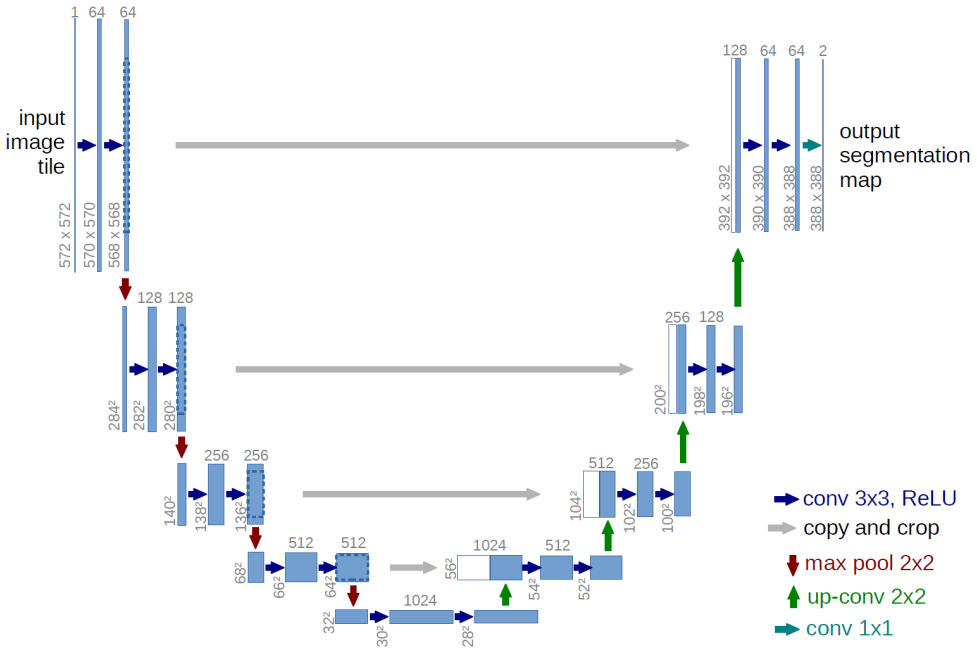


Figure 4.1: Detailed Architecture of U-net

4.2 ResNet50

Residual Network(ResNet) is another commonly used architecture alongside U-net for image detection and segmentation tasks. They are capable of working with many layers, and they are often written with a number that specifies the number of layers in the network (e.g., Resnet18, ResNet50, ResNet101, etc). ResNet was initially introduced to overcome the vanishing gradient problem in deep neural networks. The authors of the original ResNet paper demonstrated that adding more and more layers in a network degrades performance rather than improves it [35]. That is why they introduced the residual blocks that let the model skip over one or more layers (shown below in figure 4.2).

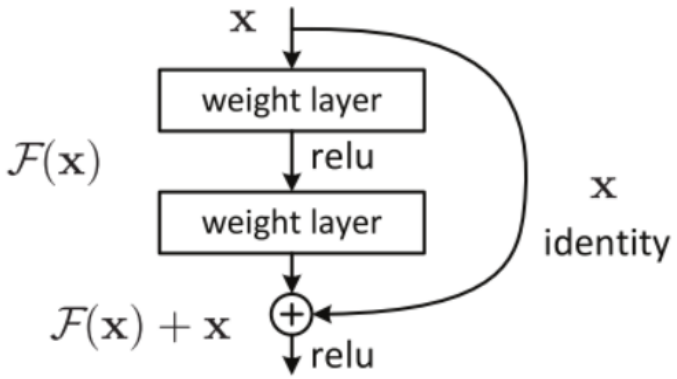


Figure 4.2: Basic Residual learning blocks

For our experiment, we used the ResNet50 architecture as an encoder for one of our models. The architecture of the ResNet model contains four main stages. In the

first layer, there is a 7x7 convolution followed by a 3x3 max pooling layer of stride 2 (1 layer). Next, in every four stages, there are three convolution layers of 1x1, 3x3 and 1x1, with a filter size of 64 for the first two and 256 for the last one (the filter size doubles in every stage). The first stage repeats itself three times (9 layers), the second stage four times (12 layers), the third stage six times (18 layers), and the fourth stage three times (9 times). The final layer ends with average pooling and a fully connected layer with a softmax activation function (1 layer). In total, this gives us 1+9+12+18+9+1 = 50 layers. The architecture of the ResNet50 model is visualized below in figure 4.3.

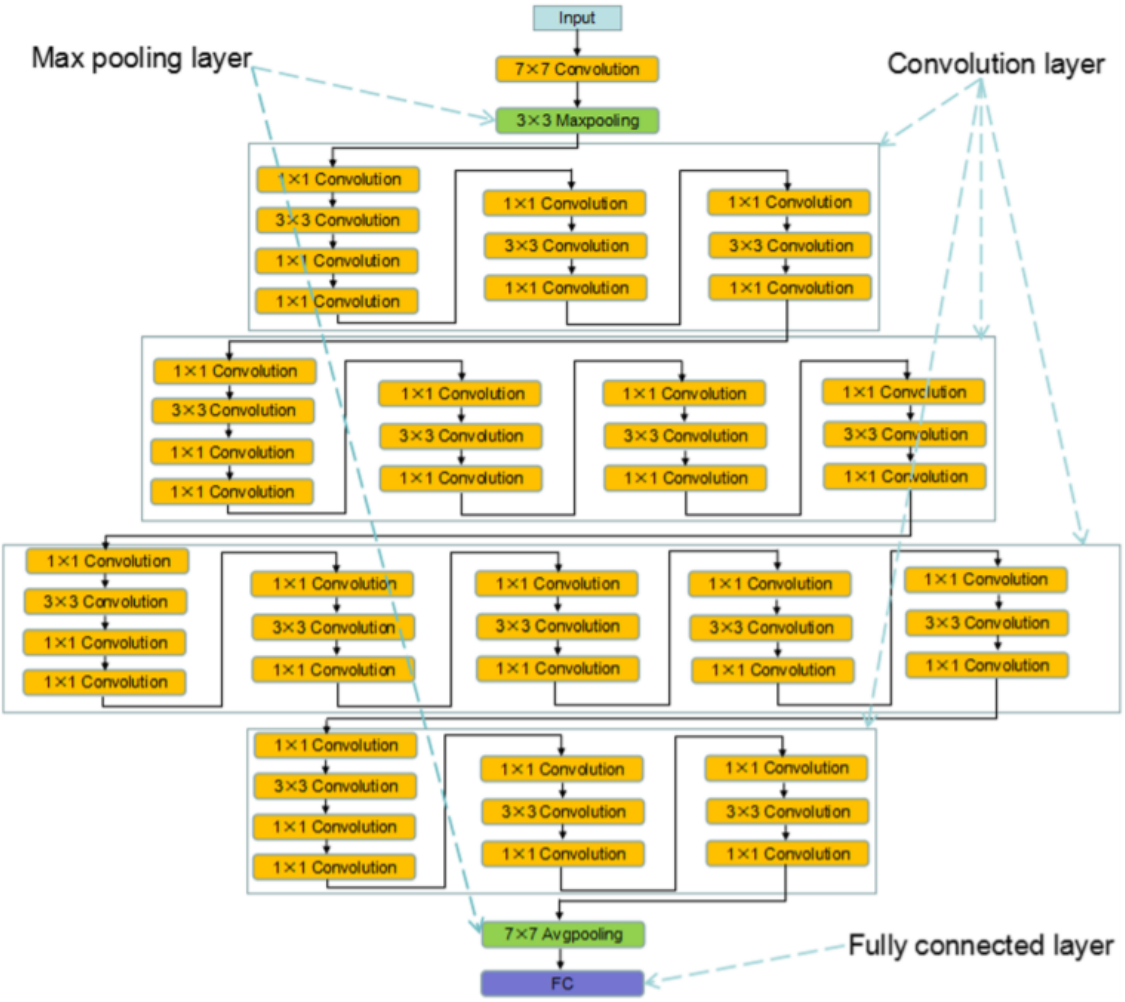


Figure 4.3: Detailed Architecture of ResNet50

4.3 ResNext50

ResNext50 is a homogenous, multi-branch CNN architecture that incorporates fewer number of hyperparameters compared to conventional ResNets [36]. This is accomplished by adding a new dimension, cardinality, alongside the width and depth dimensions. Cardinality is defined as the precise size of the set of transformations. In figure 4.4, we can see the typical structure of a ResNet and a ResNext block. The ResNext block has a cardinality of 32, which means the same transformation

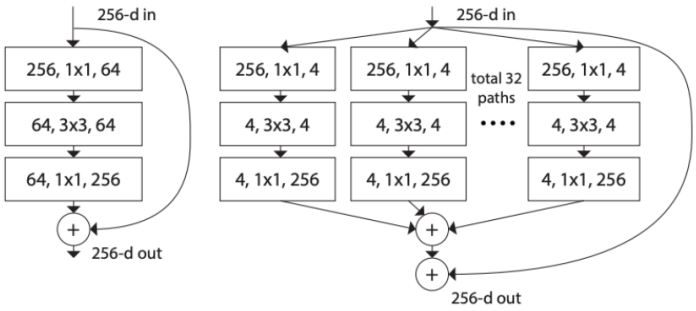


Figure 4.4: ResNet block (left) and ResNext block (right)

is applied 32 times. Finally, the result is aggregated at the end. Each path of a ResNext50 contains a 1x1, 3x3, and 1x1 convolution at the convolution layer. The internal dimension (depth) is 4, which means cardinality repeats four times. Summing up the dimensions of each 3x3 convolution gives a total dimension of 128. The dimensions are then increased to 256, and they are added together along with the skip connections. Furthermore, unlike in conventional ResNets, the neurons in one path are not connected to the neurons in another path. This method was suggested by Xie et al. in their paper "Aggregated Residual Transformations for Deep Neural Networks [36]." They concluded that increasing the cardinality instead of the number of layers of depth and width produces better results. The detailed architecture of ResNext50 is displayed below in figure 4.5.

stage	output	**ResNeXt-50 (32×4d)**
conv1	112×112	7×7, 64, stride 2
conv2	56×56	3×3 max pool, stride 2 [1×1, 128 3×3, 128, C=32 1×1, 256] ×3
conv3	28×28	[1×1, 256 3×3, 256, C=32 1×1, 512] ×4
conv4	14×14	[1×1, 512 3×3, 512, C=32 1×1, 1024] ×6
conv5	7×7	[1×1, 1024 3×3, 1024, C=32 1×1, 2048] ×3
	1×1	global average pool 1000-d fc, softmax
# params.		**25.0**×10^6
FLOPs		**4.2**×10^9

Figure 4.5: Detailed Architecture of ResNext50

4.4 EfficientNetB7

EfficientNet is a relatively new CNN architecture proposed by Tan et al. in 2019. Unlike conventional scaling methods, this model is more efficient, as its name suggests, as it can uniformly scale the network's width, depth, and resolution using a compound coefficient for increased performance [37]. The model was then gradually scaled up to develop a family of EfficientNet by using different compound coefficients. This systematic scaling of the network enhances the overall execution of the model while balancing the dimensions of the width, depth, and resolution with a constant ratio. The compound scaling method uses theta for scaling the width, depth, and image resolution in a proper way.
Currently, there are eight variants of EfficientNet ranging from EfficientNetB0 to EfficientNetB7. The basic building block of the EfficientNet architecture is based on the inverted bottleneck residual blocks of MobileNetV2(MBConv Blocks) along with squeeze and excitation blocks. The different networks of the EffcientNet family have a varying number of these blocks. The architecture of EffcientNetB7 is illustrated in figure 4.6.

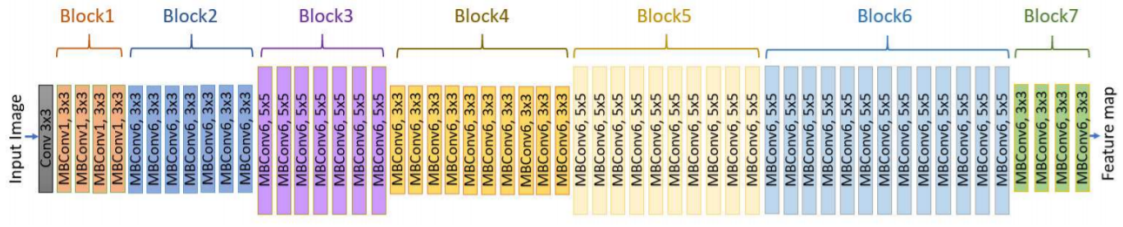


Figure 4.6: Detailed Architecture of EffcientNetB7

4.5 Proposed methodology and training process

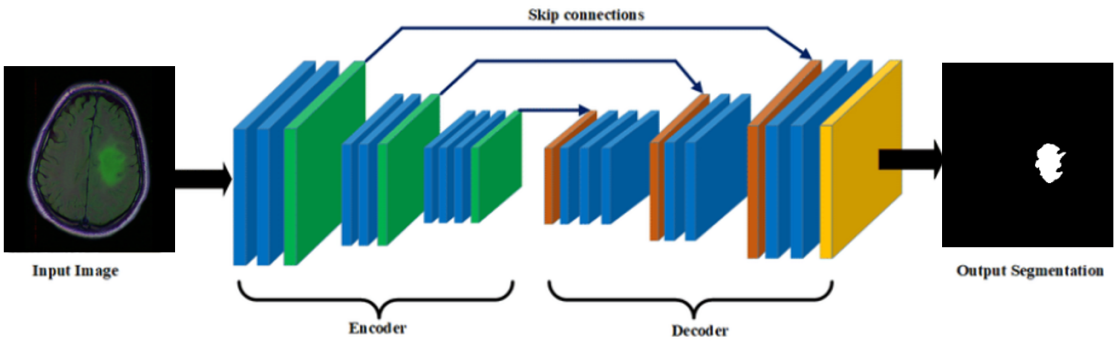


Figure 4.7: Encoder-Decoder Architecture

For our models, we implemented an encoder-decoder architecture. The encoder is the contracting part, and the decoder is the expansive part. Our three encoders (ResNet50, ResNext50 and EfficientNetB7) for our three different models are pre-trained on the ImageNet dataset. Using pre-trained encoders helps the model converge more efficiently. The input images are fed into the pre-trained encoder in each

models, which consist a series of residual blocks as their primary elements. These residual blocks assist the encoder in extracting the essential features from the input images, which are next passed to the decoder. The decoder, U-Net, gradually upscales the features that were encoded back to the original resolution and recovers spatial information. Fig 4.7 demonstrates a classic structure of an encoder-decoder architecture. Data flows from left to right, and each block represents the output from the convolution layer. The arrows represent the skip connections, and different networks might have different activation functions, layer numbers, and skip connections.

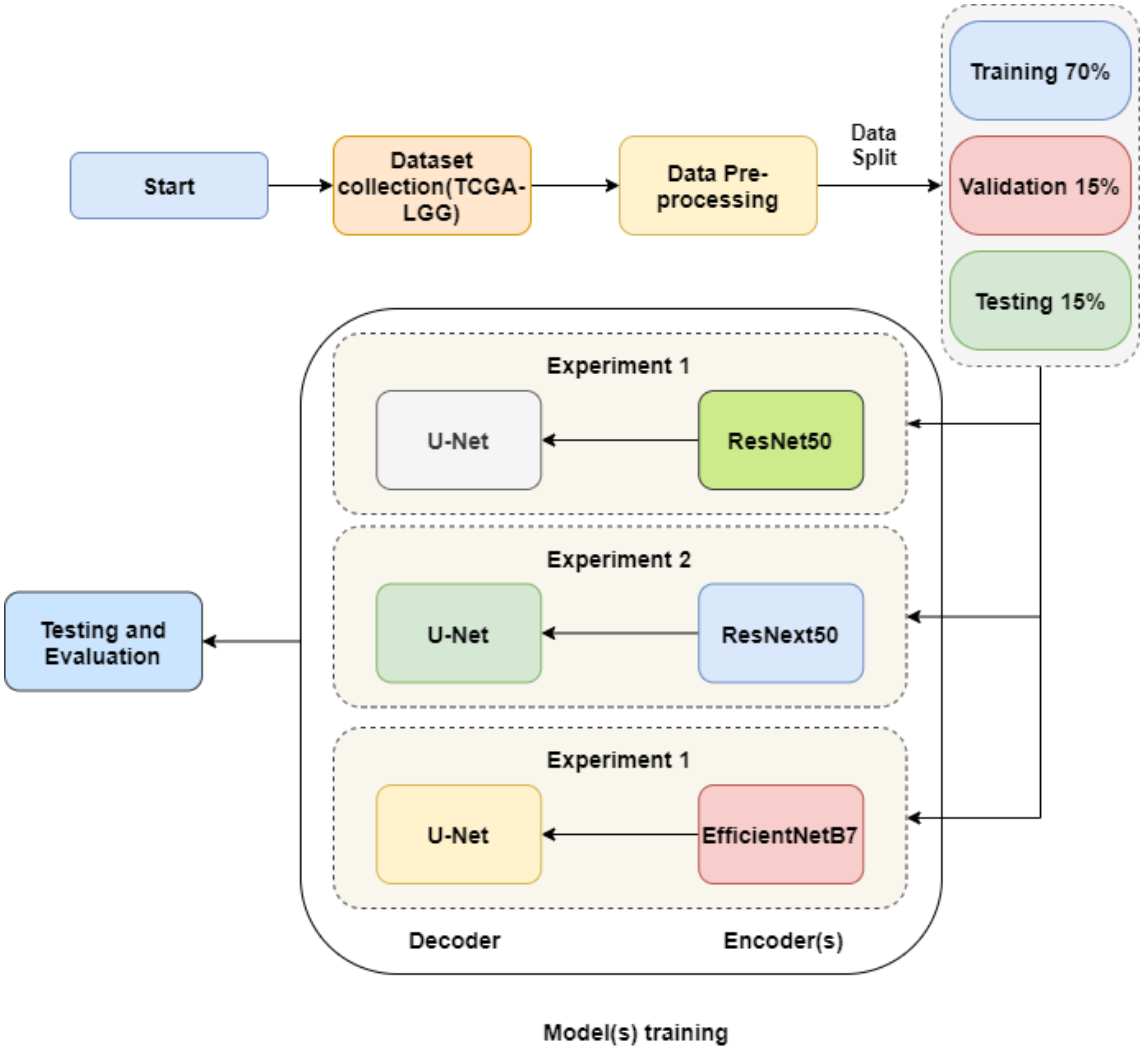


Figure 4.8: Workflow of proposed model(s)

Figure 4.8 describes the workflow of our three models. In our experiments, we trained our three models using ResNet50, ResNext50, and EfficientNetB7 as encoders for the contracting path instead of a conventional set of convolution layers. For the decoder, we used the expansive path of the U-Net. Unlike traditional U-Nets, these three models are asymmetrical. All three segmentation models were developed in the Pytorch environment. Patient data of all 110 patients included a total of 3929 image slices. Out of these images, 70% have been used for training the dataset. The other 30% have been used as 15% for testing and 15% for validation accordingly. For image inputs, we used an input size of 256*256, alongside batch size of 16 and

channel depth of 3. We decided to use 20 epochs in our experiments as we found no sign of overfitting or underfitting while working with this number. The training is carried out using the Adam optimizer (with learning rate as 0.001). Adam is an optimization algorithm that iteratively updates the network weights based on training set data. For loss function, we used binary cross-entropy. We used the sigmoid as our activation function as its output is always between 0 and 1. This function works best for binary classification problems, i.e., classifying tumor and non-tumor pixels.

Chapter 5

Performance Comparison and Analysis

We will now evaluate our segmentation results and compare each model's performance. All three models were set up and validated using the Pytorch environment assisted by GPU support. Google colab was used as the IDE. Google colabs's cloud computing environment provide the following hardware configuration: Intel(R) Xeon(R) CPU @ 2.20GHz, NVIDIA Tesla T4 GPU, 12 GB RAM, 37 GB disc space.
We used three evaluation metrics for our result analysis: Dice coefficient, Intersection over Union, and loss function. The Dice coefficient is calculated as twice the area of overlap between the images divided by the total number of pixels in the images. The formula is DC = $(2 \mid A \cap B \mid)/(\mid A \mid + \mid B \mid)$, where A and B are classes representing tumor and non-tumor, respectively. On the other hand, Intersection over Union(IoU) is the area of intersection between the predicted segmented region and the ground truth divided by the union area between them. Lastly, the loss function defines how much the predictions deviate from the actual results.
The first experiment was performed on both the augmented and non-augmented dataset using our U-Net model with ResNet50 encoder with all the optimization and hyperparameters mentioned in the previous section. The training and validation dice curves for the model is shown in the graph in Fig 5.1. It exhibits promising results as the mean Dice score achieves approximately 95.8% and 91.8%, for training and validation respectively. Additionally, the mean IoU score for training and validation reached 93.4% and 88.8%, respectively. As we can see from the graph of the dice curve, the model converges smoothly for the most part while training for 20 epochs, except for the middle region between 3 and 5 epochs when it starts to diverge downwards.

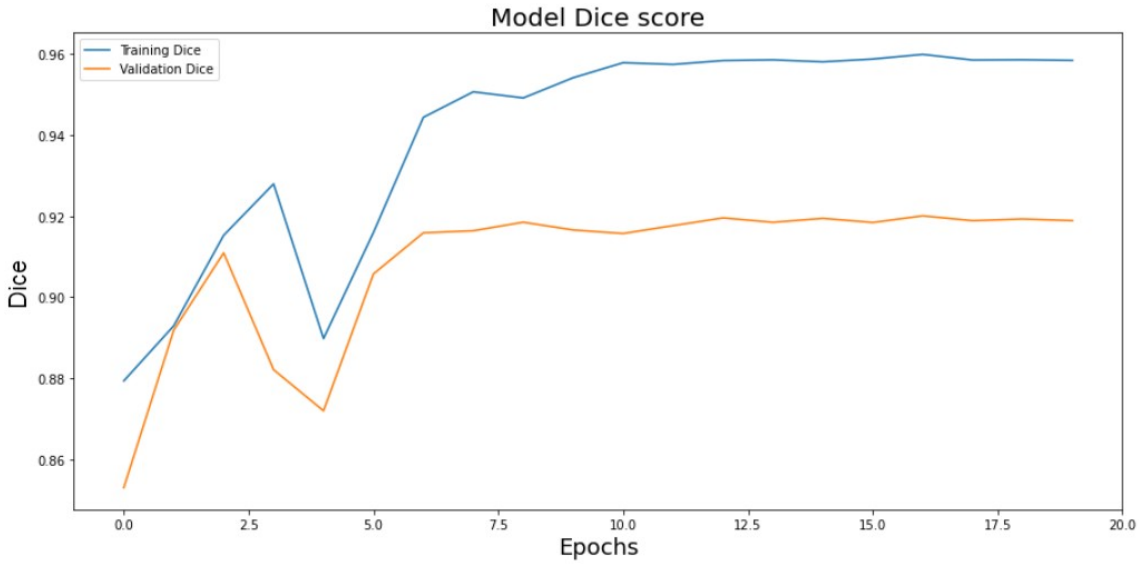


Figure 5.1: Training and Validation Dice curves for U-Net + ResNet50 model. Training (Blue curve) and Validation (Orange curve).

The second experiment was performed on the dataset using our U-Net model with the ResNext50 encoder. As we can see from the graph for Dice in Fig 5.2, the training and validation curve displays the results with respect to the 20 epochs that we used for our test. If we notice the dice curve, we can see that the model training was not as smooth as the previous model as the curve kept going downwards during several epochs. But nevertheless, this model achieved slightly better performance compared to our first model. The mean Dice scores in this model managed to reach a score of 96.4% and 92.2% for training and validation, respectively. Similarly, the IoU result slightly improved from the first model with scores of 94.1% and 89.2% for training and validation, respectively.

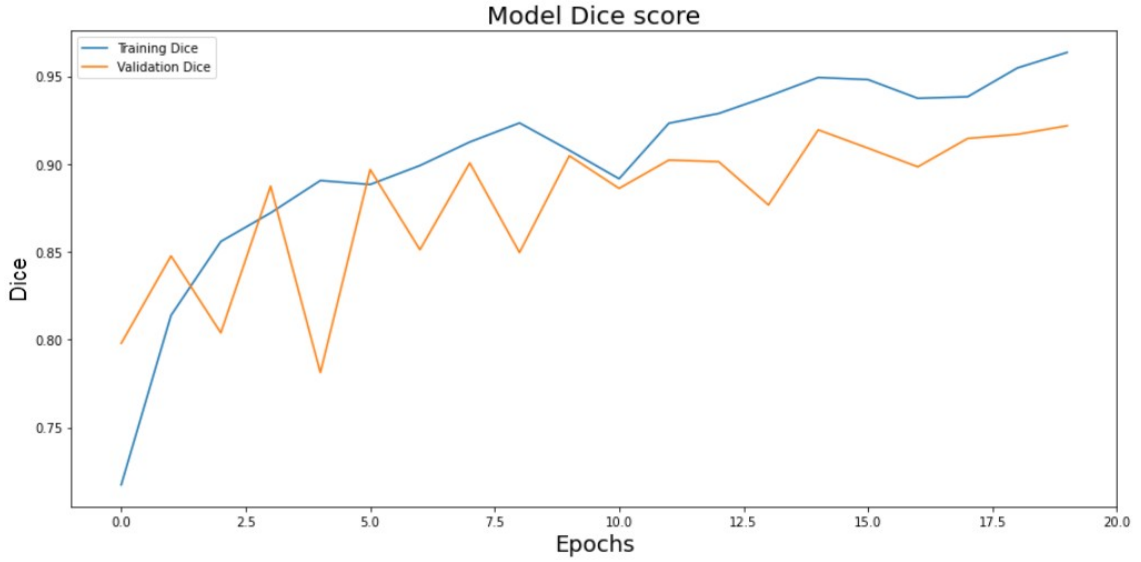


Figure 5.2: Training and Validation Dice curves for U-Net + ResNext50 model. Training (Blue curve) and Validation (Orange curve).

Our final experiment was conducted on the dataset using our U-Net model with the

EfficientNetB7 encoder. This model gave us our best results out of all three models. We have illustrated the dice curves in the graphs in Fig 5.3. The mean Dice scores for training and validation sets reached about 97.5% and 93.3%, respectively. This result is supported by the fact that the curve converges more smoothly compared to the other two models with little to no deviation. Furthermore, the IoU scores for training and validation sets reached about 95.6% and 90.5%, respectively.

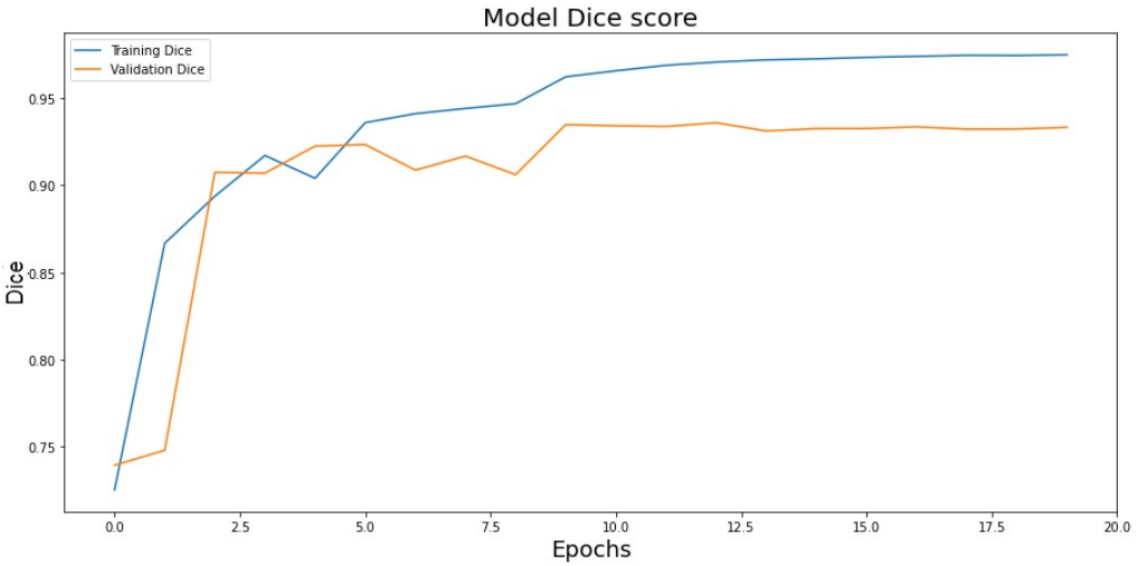


Figure 5.3: Training and Validation Dice curves for U-Net + EfficientNetB7 model. Training (Blue curve) and Validation (Orange curve).

The loss curves for all three models are visualized in Fig 5.4. We can see from the graph that the model with the EfficientNetB7 encoder converges more smoothly than the other two models. The loss reaches as low as 0.25% and 0.66% on the training and validation sets for our U-Net + EfficientNetB7 model, outperforming the other two models by a significant margin. On the other hand, U-Net + ResNet50 reached a loss margin of 0.35% and 0.71% on training and validation sets, respectively. Furthermore, our U-Net + ResNext50 achieved a loss margin of 0.33% and 0.69% on the training and validation sets, respectively. This supports the fact that the learning process of our U-Net + EfficientNetB7 model outperformed the other two models by a significant margin.

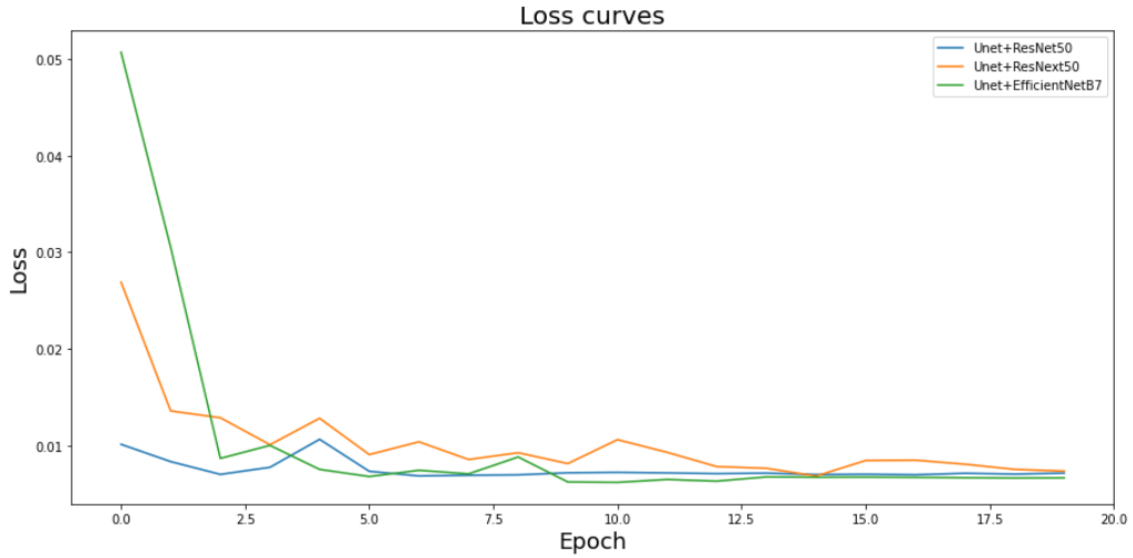


Figure 5.4: Loss curves for all three models. U-Net + ResNet50 (Blue curve, U-Net + ResNext50 (Orange curve), and U-Net + EfficientNetB7 (Green curve).

We created a table to compare the results of our experiments with our augmented dataset against the non-augmented dataset(Table 5.1). Experiment 1 is our model U-Net with ResNet50 encoder, experiment 2 is our model U-Net with ResNext50 encoder, and experiment 3 is our model U-Net with EfficientNetB7 encoder. From the comparison we can observe there is a slight increase in IoU and Dice scores in the augmented models. Although the improvements are not that significant, we can concur that augmenting the dataset output better performance.

	Without Augmentation			With Augmentation		
	U-Net+ ResNet50	U-Net+ ResNext50	U-Net+ EfficientNetB7	U-Net+ ResNet50	U-Net+ ResNext50	U-Net+ EfficientNetB7
Mean Training Iou	92.7%	93.4%	94.8%	93.4%	94.1%	95.6%
Mean Validation IoU	88.6%	89.1%	90.1%	88.8%	89.2%	90.5%
Mean Training Dice	95.2%	95.8%	96.8%	95.8%	96.4%	97.5%
Mean Validation Dice	91.5%	92.0%	93.0%	91.8%	92.2%	93.3%
Mean Training Loss	0.40%	0.38%	0.31%	0.35%	0.33%	0.25%
Mean Validation Loss	0.72%	0.70%	0.68%	0.71%	0.69%	0.66%

Table 5.1: Performance of individual models on the training and validation sets and comparison of the models' performances on the augmented dataset vs the non-augmented dataset.

We performed testing on 589 images which is 15% of the entire dataset. This set of data was not implemented in the training phase and once again we compare the performance of the model on the augmented dataset vs the non-augmented dataset. We have visualized the results in table 5.1 for proper understanding. The test dataset delivered excellent results, with U-Net + EfficientNetB7 providing the best scores from the other two models. In Fig 5.5, we visually compared the predicted tumor mask with the original image mask. The original mask, also known as ground truth, is highlighted in red outline, and the predicted mask is highlighted in green outline.

	Without Augmentation			With Augmentation		
	U-Net+ ResNet50	U-Net+ ResNext50	U-Net+ EfficientNetB7	U-Net+ ResNet50	U-Net+ ResNext50	U-Net+ EfficientNetB7
Mean Iou	88.4%	88.8%	89.7%	89.0%	89.5%	90.5%
Mean Dice	91.3%	91.7%	92.8%	91.6%	92.0%	93.2%
Mean Loss	0.69%	0.64%	0.65%	0.65%	0.62%	0.60%

Table 5.2: Performance of individual models on the test set and comparison of the models' performances on the augmented dataset vs the non-augmented dataset.

We will now compare the performances of our proposed models with other state-of-the-art research findings using the TCGA-LGG dataset. Astaraki et al. [38] used this dataset on a hybrid CNN segmentation model based on autoencoders to detect anomalies in the brain. Their proposed model achieved 90% and 87% dice scores for training and validation sets. Chen et al. [39] used the same dataset on a 3D CNN model to perform dense volumetric brain tumor segmentation. Their proposed method achieved 90.5% dice score on the validation set. Huang et al. [40] used this dataset to perform brain tumor segmentation using a novel V-Net network structure. Their proposed model achieved an average dice score of 79%. Our model outperformed the baseline U-Net model as well. Dong et al. performed segmentation on LGG using the base U-Net model, which managed to achieve a dice score of 86% [41]. Our proposed segmentation models achieved the best performance out of all the methods mentioned above using the TCGA-LGG dataset. Our best performing model using the U-Net + EfficientNetB7 architecture obtained 93.2% dice scores.

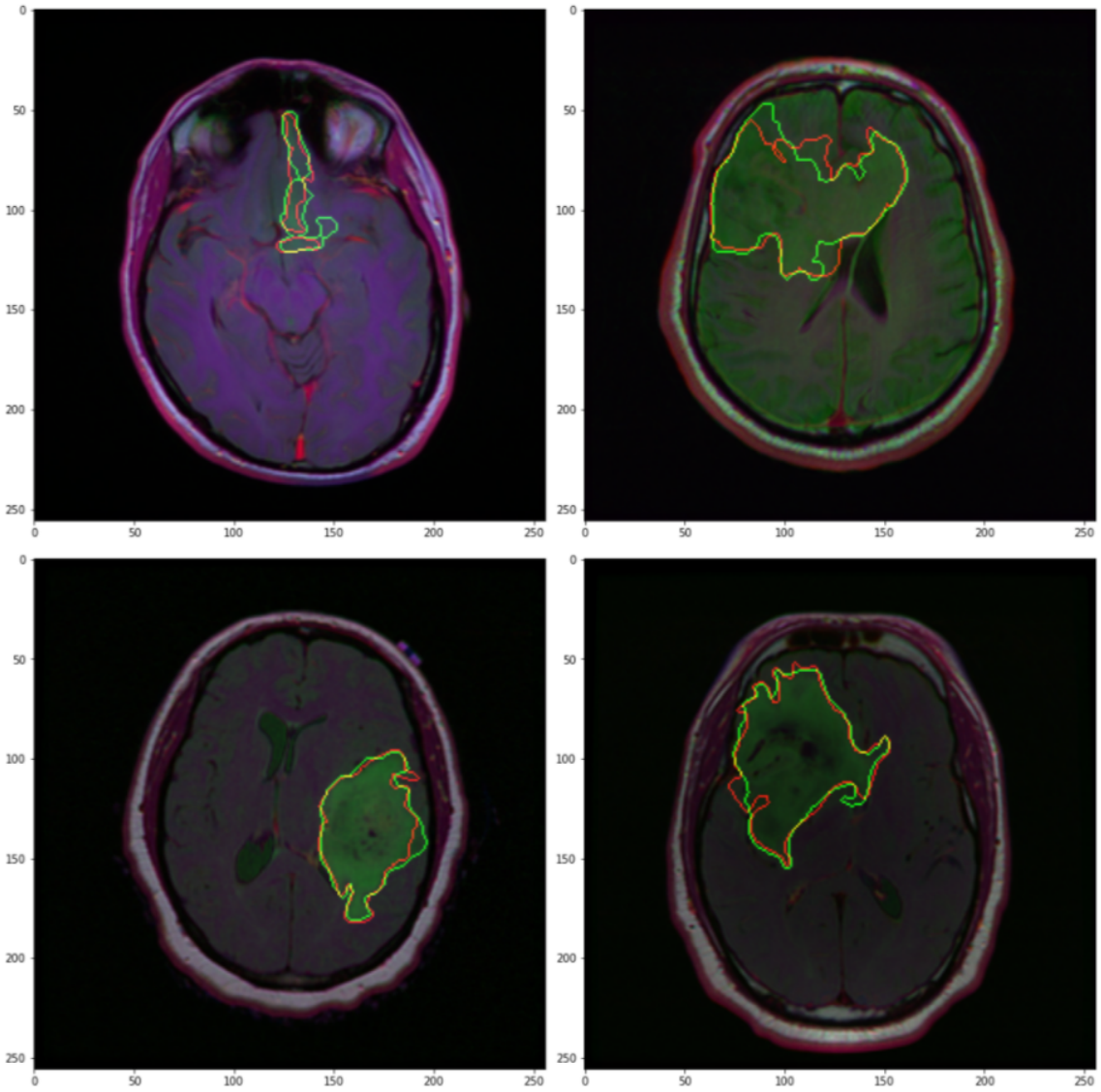


Figure 5.5: Visualization of Ground Truth(Red) and Predicted Tumor Region(Green) from our segmentation results.

Chapter 6

Conclusion and Future Works

In this paper, we experimented with three hybrid models of U-Net using the U-Net as decoder and ResNet50, ResNext50, and EfficientNetB7 as encoders. First, we acquired our images from the cancer imaging archive consisting of 110 patient data. Next, we performed few pre-processing steps, such as image resizing, augmentation, and distribution. We then fed our images into the models, starting with U-Net+ResNet50 first, U-Net+ResNext50 second, and U-Net+EfficientNetB7 last. We evaluated our segmentation results on the augmented and non-augmented dataset using the IoU and Dice coefficient metrics, two widely used metrics for assessing segmentation models. All three models generated promising outputs. However the model U-Net + EfficientNetB7 generated the best performance in the augmented dataset, which is why we propose this model for brain tumor segmentation. We have observed that this model converges significantly well compared to the other two models. We also compared our models' performances with the baseline U-Net model and other segmentation models trained on the TCGA-LGG dataset.
Numerous image segmentation techniques for brain tumors have been implemented throughout the years, but it remains a challenging task overall. One segmentation technique may work on one image but not on a different image of the same type. More research must be done on this topic. That is why researchers must continue to experiment with different versions of the same architecture for obtaining better segmentation accuracy, which can aid in saving a person's life. There is still a lot of room for improvement in the overall design of the models to achieve better results. Further research and study need to be explored to improve the performances of the models. We have presented an in-depth review of several brain tumor types and their complications. We need to experiment using other state-of-the-art CNN architectures as encoders alongside the U-Net decoder to improve our models. We also can use more diverse and large datasets to get better results for our future works. Brain tumor segmentation based on MRI has already proven to achieve great success in the medical field and continues to do so. Such studies like this paper are essential for refining existing segmentation methods as well as developing more powerful segmentation methods in the future.

Bibliography

[1] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of mri-based brain tumor segmentation methods," *Tsinghua Science and Technology*, vol. 19, no. 6, pp. 578–595, 2014. DOI: 10.1109/TST.2014.6961028.

[2] B. H. Menze, A. Jakab, S. Bauer, *et al.*, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015. DOI: 10.1109/TMI.2014.2377694.

[3] World Health Organization, *The Global Cancer Observatory*, Globocan, 2020. [Online]. Available: https://gco.iarc.fr/today/data/factsheets/populations/50-bangladesh-fact-sheets.pdf.

[4] W. H. Organization, *The Global Cancer Observatory*, Globocan, 2020. [Online]. Available: https://gco.iarc.fr/today/data/factsheets/cancers/31-Brain-central-nervous-system-fact-sheet.pdf.

[5] D. N. Louis, H. Ohgaki, O. D. Wiestler, *et al.*, "The 2007 WHO classification of tumours of the central nervous system," *Acta Neuropathologica*, vol. 114, no. 2, pp. 97–109, Jul. 2007. DOI: 10.1007/s00401-007-0243-4. [Online]. Available: https://doi.org/10.1007/s00401-007-0243-4.

[6] A. Pinto, S. Pereira, H. Correia, J. Oliveira, D. M. L. D. Rasteiro, and C. A. Silva, "Brain tumour segmentation based on extremely randomized forest with high-level features," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, Aug. 2015. DOI: 10.1109/embc.2015.7319032. [Online]. Available: https://doi.org/10.1109/embc.2015.7319032.

[7] S. Bauer, L.-P. Nolte, and M. Reyes, "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization," in *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2011, pp. 354–361. DOI: 10.1007/978-3-642-23626-6_44. [Online]. Available: https://doi.org/10.1007/978-3-642-23626-6_44.

[8] G. Wang, S. Zhang, H. Xie, D. N. Metaxas, and L. Gu, "A homotopy-based sparse representation for fast and accurate shape prior modeling in liver surgical planning," *Medical Image Analysis*, vol. 19, no. 1, pp. 176–186, Jan. 2015. DOI: 10.1016/j.media.2014.10.003. [Online]. Available: https://doi.org/10.1016/j.media.2014.10.003.

[9] D. Gupta, "Chapter 1 - neuroanatomy," in *Essentials of Neuroanesthesia*, H. Prabhakar, Ed., Academic Press, 2017, pp. 3–40, ISBN: 978-0-12-805299-0. DOI: https://doi.org/10.1016/B978-0-12-805299-0.00001-4. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128052990000014.

[10] R. Suryadevara, H. Fadel, S. K. Michelhaugh, S. Mittal, and P. Parajuli, "Tumors of the central nervous system," in *Nanotechnology-Based Targeted Drug Delivery Systems for Brain Tumors*, Elsevier, 2018, pp. 1–26. DOI: 10.1016/b978-0-12-812218-1.00001-4. [Online]. Available: https://doi.org/10.1016/b978-0-12-812218-1.00001-4.

[11] J. Petras, "Connections of the parietal lobe," *Journal of Psychiatric Research*, vol. 8, no. 3-4, pp. 189–201, Aug. 1971. DOI: 10.1016/0022-3956(71)90018-5. [Online]. Available: https://doi.org/10.1016/0022-3956(71)90018-5.

[12] M. Mendez, "Chapter 4 - essential behavioral neuroanatomy," in *The Mental Status Examination Handbook*, M. Mendez, Ed., Philadelphia: Elsevier, 2022, pp. 24–38, ISBN: 978-0-323-69489-6. DOI: https://doi.org/10.1016/B978-0-323-69489-6.00004-8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780323694896000048.

[13] T. Hackman and E. H. Toh, "Chapter 102 - petrosal approaches," in *Operative Otolaryngology: Head and Neck Surgery (Second Edition)*, E. N. Myers, R. L. Carrau, D. E. Eibling, *et al.*, Eds., Second Edition, Philadelphia: W.B. Saunders, 2008, pp. 1023–1043, ISBN: 978-1-4160-2445-3. DOI: https://doi.org/10.1016/B978-1-4160-2445-3.50106-0. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9781416024453501060.

[14] R. Singh, "Cerebellum: Its anatomy, functions and diseases," in *Neurodegenerative Diseases - Molecular Mechanisms and Current Therapeutic Approaches*, IntechOpen, Jan. 2021. DOI: 10.5772/intechopen.93064. [Online]. Available: https://doi.org/10.5772/intechopen.93064.

[15] J. Wilkinson, "Chapter 7 - cerebellum," in *Neuroanatomy for Medical Students (Second Edition)*, J. Wilkinson, Ed., Second Edition, Butterworth-Heinemann, 1992, pp. 150–163, ISBN: 978-0-7506-1447-4. DOI: https://doi.org/10.1016/B978-0-7506-1447-4.50011-6. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780750614474500116.

[16] M. V. Johnston, "Development, structure, and function of the brain and neuromuscular systems," in *Pediatric Critical Care*, Elsevier, 2006, pp. 767–779. DOI: 10.1016/b978-032301808-1.50051-1. [Online]. Available: https://doi.org/10.1016/b978-032301808-1.50051-1.

[17] D. Doherty and A. Poretti, "Disorders of cerebellar and brainstem development," in *Swaiman's Pediatric Neurology*, Elsevier, 2017, pp. 199–207. DOI: 10.1016/b978-0-323-37101-8.00027-8. [Online]. Available: https://doi.org/10.1016/b978-0-323-37101-8.00027-8.

[18] J. L. Villano and E. E. Vokes, "Primary and metastatic brain tumors," in *Oncologic Therapies*, Springer Berlin Heidelberg, 2003, pp. 569–585. DOI: 10.1007/978-3-642-55780-4_42. [Online]. Available: https://doi.org/10.1007/978-3-642-55780-4_42.

[19] V. P. Collins, "Gliomas," in *Brenner's Encyclopedia of Genetics*, Elsevier, 2013, pp. 334–336. DOI: 10.1016/b978-0-12-374984-0.00652-5. [Online]. Available: https://doi.org/10.1016/b978-0-12-374984-0.00652-5.

[20] D. Guo, Y. Pei, K. Zheng, H. Yu, Y. Lu, and S. Wang, "Degraded image semantic segmentation with dense-gram networks," *IEEE Transactions on Image Processing*, vol. 29, pp. 782–795, 2020. DOI: 10.1109/tip.2019.2936111. [Online]. Available: https://doi.org/10.1109/tip.2019.2936111.

[21] J. Yi, P. Wu, M. Jiang, Q. Huang, D. J. Hoeppner, and D. N. Metaxas, "Attentive neural cell instance segmentation," *Medical Image Analysis*, vol. 55, pp. 228–240, Jul. 2019. DOI: 10.1016/j.media.2019.05.004. [Online]. Available: https://doi.org/10.1016/j.media.2019.05.004.

[22] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities.," *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, Apr. 1982. DOI: 10.1073/pnas.79.8.2554. [Online]. Available: https://doi.org/10.1073/pnas.79.8.2554.

[23] Y. LeCun, B. Boser, J. S. Denker, *et al.*, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, Dec. 1989. DOI: 10.1162/neco.1989.1.4.541. [Online]. Available: https://doi.org/10.1162/neco.1989.1.4.541.

[24] A. G'eron and a. O. M. C. Safari, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition*. O'Reilly Media, Incorporated, 2019. [Online]. Available: https://books.google.com.bd/books?id=O2VJzQEACAAJ.

[25] H. Wu, Q. Wang, L. Zhou, and J. Meng, "VHF radio signal modulation classification based on convolution neural networks," *MATEC Web of Conferences*, vol. 246, X. Lei, S. Cai, Y. Yu, and H. Li, Eds., p. 03 032, 2018. DOI: 10.1051/matecconf/201824603032. [Online]. Available: https://doi.org/10.1051/matecconf/201824603032.

[26] B. Devkota, A. Alsadoon, P. Prasad, A. Singh, and A. Elchouemi, "Image segmentation for early stage brain tumor detection using mathematical morphological reconstruction," *Procedia Computer Science*, vol. 125, pp. 115–123, 2018. DOI: 10.1016/j.procs.2017.12.017. [Online]. Available: https://doi.org/10.1016/j.procs.2017.12.017.

[27] R. Pitchai, P. Supraja, A. H. Victoria, and M. Madhavi, "Brain tumor segmentation using deep learning and fuzzy k-means clustering for magnetic resonance images," *Neural Processing Letters*, vol. 53, no. 4, pp. 2519–2532, Aug. 2020. DOI: 10.1007/s11063-020-10326-4. [Online]. Available: https://doi.org/10.1007/s11063-020-10326-4.

[28] H. N. T. K. Kaldera, S. R. Gunasekara, and M. B. Dissanayake, "Brain tumor classification and segmentation using faster r-CNN," in *2019 Advances in Science and Engineering Technology International Conferences (ASET)*, IEEE, Mar. 2019. DOI: 10.1109/icaset.2019.8714263. [Online]. Available: https://doi.org/10.1109/icaset.2019.8714263.

[29] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation," *Medical Image Analysis*, vol. 43, pp. 98–111, Jan. 2018. DOI: 10.1016/j.media.2017.10.002. [Online]. Available: https://doi.org/10.1016/j.media.2017.10.002.

[30] L. Ramya and N. Sasirekha, "A robust segmentation algorithm using morphological operators for detection of tumor in mri," in *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2015, pp. 1–4. DOI: 10.1109/ICIIECS.2015.7192927.

[31] S. Alqazzaz, X. Sun, X. Yang, and L. Nokes, "Automated brain tumor segmentation on multi-modal MR image using SegNet," *Computational Visual Media*, vol. 5, no. 2, pp. 209–219, Apr. 2019. DOI: 10.1007/s41095-019-0139-y. [Online]. Available: https://doi.org/10.1007/s41095-019-0139-y.

[32] N. Pedano, A. E. Flanders, L. Scarpace, *et al.*, *Radiology data from the cancer genome atlas low grade glioma [tcga-lgg] collection*, 2016. DOI: 10.7937/K9/TCIA.2016.L4LTD3TK. [Online]. Available: https://wiki.cancerimagingarchive.net/x/BANR.

[33] K. Clark, B. Vendt, K. Smith, *et al.*, "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1045–1057, Jul. 2013. DOI: 10.1007/s10278-013-9622-7. [Online]. Available: https://doi.org/10.1007/s10278-013-9622-7.

[34] V. Zyuzin, P. Sergey, A. Mukhtarov, *et al.*, "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 (USBEREIT)*, IEEE, May 2018. DOI: 10.1109/usbereit.2018.8384554. [Online]. Available: https://doi.org/10.1109/usbereit.2018.8384554.

[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[36] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, http://arxiv.org/pdf/1611.05431, 2017, pp. 5987–5995. DOI: 10.1109/cvpr.2017.634. [Online]. Available: https://app.dimensions.ai/details/publication/pub.1095850372.

[37] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *ArXiv*, vol. abs/1905.11946, 2019.

[38] M. Astaraki, C. Wang, G. Carrizo, I. Toma-Dasu, and Ö. Smedby, "Multimodal brain tumor segmentation with normal appearance autoencoder," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Springer International Publishing, 2020, pp. 316–323. DOI: 10.1007/978-3-030-46643-5_31. [Online]. Available: https://doi.org/10.1007/978-3-030-46643-5_31.

[39] C. Chen, X. Liu, M. Ding, J. Zheng, and J. Li, "3d dilated multi-fiber network for real-time brain tumor segmentation in MRI," in *Lecture Notes in Computer Science*, Springer International Publishing, 2019, pp. 184–192. DOI: 10.1007/978-3-030-32248-9_21. [Online]. Available: https://doi.org/10.1007/978-3-030-32248-9_21.

[40] H. Huang, W. Zhang, Y. Fang, J. Hong, S. Su, and X. Lai, "Overall survival prediction for gliomas using a novel compound approach," *Frontiers in Oncology*, vol. 11, Aug. 2021. DOI: 10.3389/fonc.2021.724191. [Online]. Available: https://doi.org/10.3389/fonc.2021.724191.

[41] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using u-net based fully convolutional networks," *ArXiv*, vol. abs/1705.03820, 2017.