# Brain Tumor Sectionalization through Semantic Segmentation Approach

by

Tirthankar Saha Dibbya
20101409
Md. Sayem Khan
20301248
Tasfia Tarannum
20101237
Rahmat Ullah Mahin
20101013

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**
October 2024

# 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:**

<table>
<tr><td>Tasfia Tarannum<br>20101237</td><td>Tirthankar Saha Dibbya<br>20101409</td></tr>
<tr><td>Rahmat Ullah Mahin<br>20101013</td><td>MD. Sayem Khan<br>20301248</td></tr>
</table>

# Approval

The thesis/project titled "Brain Tumor Sectionalization through Semantic Segmentation Approach" submitted by

1. Tirthankar Saha Dibbya (20101409)

2. Sayem Khan (20301248)

3. Tasfia Tarannum (20101237)

4. Rahmat Ullah Mahin (20101013)

Of Summer, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on October, 2024.

**Examining Committee:**

Supervisor:
(Member)

_____

Dr. Md. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Member)

_____

Rafeed Rahman

Senior Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

_____

Dr. Md. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

_____
Sadia Hamid Kazi, Ph.D.

Chairperson
Department of Computer Science and Engineering
Brac University

# Abstract

Accurate brain tumor detection and segmentation from magnetic resonance imaging (MRI) scans are vital for effective diagnosis, treatment planning, and patient monitoring. However, manual segmentation is time-consuming and subject to variability. As a result, it is a necessity for the development of automated solutions. Traditional convolutional neural network (CNN) approaches, such as U-Net and its variants, often face limitations in handling the high-resolution, complex patterns of MRI data. Traditional convolutional models struggle to generalize across diverse tumor characteristics and often fail to capture long-range dependencies, which are crucial for accurate segmentation and their computational inefficiency limits real-time applications. In the case of traditional transformer models, it often relies on fixed positional encodings to understand the spatial relationships between parts of an image.[25] There is also a need for adaptable models that do not rely on fixed positional encoding, enabling them to accurately segment tumors irrespective of their location within the brain. In this research we aim to address these challenges by implementing SegFormer, a Transformer-based architecture, and EfficientNetB4, a convolutional model, to enhance segmentation and classification accuracy for brain tumors.Here, we strategically adapted and integrated pretrained models, specifically SegFormer and EfficientNetB4 in order to create a robust framework for brain tumor segmentation and classification. Like other existing studies we did not focus on a single model. Rather, our approach combines SegFormer's capability for precise pixel-level segmentation with EfficientNetB4's efficient classification to improve overall diagnostic accuracy. In order to handle the specific challenges of high-resolution MRI data, we carefully tuned SegFormer, maintaining fine details and adapting to varying tumor characteristics, while using EfficientNetB4 to accurately distinguish between tumor and non-tumor images. Additionally, our method focuses on computational efficiency and real-time applicability by optimizing the models to ensure fast processing speeds, which is crucial for clinical application. In our experiment we demonstrated that SegFormer achieves superior segmentation performance, with a Dice score of 0.7961 and a Mean Intersection over Union (IoU) of 0.7382 which significantly outperforms other models like LinkNet, U-Net, and U-Net++, which recorded Dice scores of 0.3445, 0.2985, and 0.1575 respectively. Similarly, EfficientNetB4 achieved exceptional classification accuracy, with precision, recall, and F1-scores of 99% for both tumor and non-tumor classes, highlighting its reliability in distinguishing between the two. These results tell us that SegFormer's efficient hierarchical encoder and pixel-level precision, combined with EfficientNetB4's robust classification capabilities, offer a powerful and comprehensive solution for brain tumor segmentation and Lightweight MLP Decoder is computationally efficient for real time application.


**Keywords:** SegFormer, semantic segmentation, tumor detection, convolutional, EfficientNetB4

# Acknowledgement

# 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 Network

DC    Dice Coefficient

FCN   Fully Convolutional Network

HGG   High-Grade Glioma.

IoU   Intersection over Union

LGG   Lower Grade Glioma

MRI   Magnetic Resonance Imaging

# Chapter 1

# Introduction

Brain tumor which is commonly referred to as an intracranial tumor is an unusual growth of tissues where brain cells multiply and grow unrestrained without the observation of mechanisms that control the normal cells[11]. Among them, Gliomas are the most prominent type of adult brain tumor, accounting for 78 percent of malignant brain tumors. According to clinical research, individuals with less aggressive tumors might expect to live for several years after diagnosis, while those with more hostile variants typically have a maximum life expectancy of two years. For both types of tumor, MRI is mostly used to provide us with detailed images of the brain, which is useful both before and after treatment in order to learn about the properties of the disease. Hence, precise segmentation of MRI images would be helpful for treatment planning, tumor growth rate, and survival prediction of patients[19]. The task of segmenting MRI images across multiple modalities has proven to be a difficult challenge, mainly due to the indistinct boundaries of tumors and the partial volume effect that occurs in these images. Multiple differences between the tumors such as tumor size,shape, location create challenges in the effective segmentation of medical image analysis[23]. Deep Learning techniques, specifically Convolutional Neural Networks (CNN), have shown remarkable success in segmenting various tumor components from diverse sequences of MR images in recent times due to their capacity to effectively grasp hierarchical features and spatial details from images.Medical images like MRI images are volumetric, with organs being represented as 3-dimensional entities.For this reason, we have implemented semantic segmentation, which commonly falls under the domain of deep learning and classifies each image pixel into an instance where the instance represents a class[20].In this research, we have used EfifficientNetB4 for the classification of the brain tumor and based on the classification performed semantic segmentation using the SegFormer model on the dataset to segment the brain tumor from the MRI images which help in the identification of the tumor and further treatment planning.

## 1.1 Problem Statement

A brain tumor is a typical accumulation of tissue in which cells grow and proliferate rapidly. These tumors can either be categorized as primary or metastatic. MRI imaging is most commonly used for detecting tumor location, size and morphology. Brain tumors can be enhanced using different modalities of MRI imaging. The initial stage of establishing a diagnosis involves the precise identification and delineation of the tumor and its constituent parts on various imaging modalities known as tumor segmentation, is critical in understanding the nature and extent of the disease, as well as for creating a suitable treatment strategy. Manual segmentation can be inefficient and time-consuming. That is why automatic segmentation is necessary. In recent years, many approaches have been made for automatic segmentation and survival predictions[19]. Different research has different types of approaches for optimization and efficiency. But most of the research has some limitations due to the unavailability of datasets or even lack of adequate technologies. For instance, in the paper [19], proposed a 3-D fully convolutional neural network (CNNs) to segment gliomas for segmentation of gliomas and its components from MRI images and survival prediction through Random Forest Regressor. The paper only deals with 2 out of 3 problems given by the challenge organizer, so the paper doesn't fulfill all requirements. The performance of the model used in this paper is highly reliable on the high-quality trained data like the Brats 2020 challenge dataset which is highly used for brain tumor segmentation tasks, but it may not be able to fully represent the real world's clinical settings. Used 3D CNN architecture has high computational complexity, so it is expensive to train and require a high number of resources.

Moreover, in paper [18] a 3D U-Net based deep learning algorithm which is trained for the brain tumor segmentation task in Brats 2019 competition with the help of brain wise normalization with two patching strategies and a fully connected neural network(FCNN) for solving the task of survival prediction. The survival prediction task only considers patients who have undergone resection surgery, which means that the prediction is limited to a specific subset of patients. During training, the current training platform only allows for the feeding of one input image at a time. This limitation may be restraining the model's capabilities to deal with the multiscale features of the medical images so it can not properly integrate multiscale features.

Again, in the paper [1] by combining multilevel processing and Bayesian model classification, they developed a revolutionary method for precisely segmenting brain tumors. The voxel-based nature of these two sets of techniques imposes a high degree of localization, which leads to a limitation in considering local or global contexts. While they have shown some success on a sizable data set segmenting low-grade gliomas and meningiomas (which are reasonably homogeneous), their success is constrained in the more pertinent GBM (heterogeneous) segmentation situations.

Furthermore, In the paper [12] they used a new multi-atlas segmentation (MAS) framework for MR tumor brain images. However, for brain pictures including diseases, such as tumors, the majority of MAS approaches often produce unsatisfactory

outcomes. This is because diseases might make it difficult to register standard brain atlases to the desired picture. Furthermore, their method has the issue of requiring tumor sites to seem fairly distinct from normal brain areas in MR tumor brain imaging. Otherwise, no tumor zones could be found, and COLOR is reduced to a traditional low-rank method.

Again, another research [27] mentioned that they have a lack of training datasets, which leads to the challenging issue of CNN. They have only used MRI and Brats databases for brain tumor detection. The anatomical locations of the patch must still be classified in patch classifications for deep learning approaches to detect brain cancers. However their dataset does not contain such classifications. This restriction may be removed by using a deep network to analyze the complete image and a suitable loss function to carry out the necessary learning. But this caused a new challenge for those models. MRI pictures have a gigapixel-high resolution. Due to limitations like finite memory and computer resources, processing such high-dimensional pictures in the network is challenging.

In paper [11], researchers proposed a combination of concatenated and connected random forest and multiscale patch-driven active contour models for automated segmentation of brain tumors from MRI images. However the proposed method in the paper has some drawbacks too. A large amount of labeled training data is required to employ the suggested ccRFs-mpAC model. In this work, the segmentation outcomes from algorithms and clinical specialists were combined to annotate the training data. As a result, the ground truth may tend to be systematically distorted by the algorithm result. Furthermore, training the modal-specific feature learning kernels independently utilizing several imaging modalities, and then indiscriminately aggregating the feature maps may not be the optimal choice. This article [6] presented a technique to segment brain tumors using CNN (convolutional neural networks) in 2D MRI images. They used a subset of the Brats dataset. According to them, they only utilized 100 MRI images out of 300 MRI images. In their paper, they acknowledged that using 2D MRI images is one of their limitations.

In[26]a new model was developed, but there was no deep work on pre-processing. The bias field distortion, which causes the intensity of the same tissues to change across the image, is known to affect MRI images.Also, it deals with 2D MRI images but in the real world, most MRI images are 3D which is a limitation of the proposed model in the paper. However, when the intensity problem appears in the study, the introduced model will not perform well, as there were no precautions taken to handle such problems. As a result, the model prediction will work well during well-organized images only.

Here, in paper [9] the same problem of intensity will be seen as there are no deep pre-processing techniques applied. As a result, when the intensity problem arises, there will be chaos. Also, the dataset used is also not that much and there was no data augmentation technique applied. As a result, feature extraction will not be efficient enough to give a good output. Additionally, there are two types of gliomas, LGG and HGG. As there are no proper data augmentation methods, their prediction will not work well in case of multi-modal images.

So, after diving through various research, we came to an understanding that all the research papers tend to approach the most effective and efficient way possible for optimized brain tumor segmentation. So, the question that this research tries to answer is:

**How efficient is the Semantic Segmentation as an optimal way for brain tumor segmentation from MRI image"**

This research will look forward to using advanced MRI images for segmentation to predict more accurately brain tumors. So, for using MRI images and precise segmentation this research chose the EfiicientNetB4 model for the classification of brain tumors identifying tumor and non-tumor cells, and chose the SegFormer model to perform segmentation on the tumor cells. By giving each pixel in an image a class or category, semantic segmentation makes it easier to understand the complex information contained inside.

## 1.2    Research Objective

This research aims to present to an accurate semantic segmentation of brain tumor to detect the tumor based on the segmented tumor.Usually brain tumors are abnormal growth of tissues within the brain.Brain tumor segmentation is necessary for diagnosing and treatment planning properly.The main goals of conducting the research are:
1. Collect annotated datasets of brain tumors and perform pre-processing on the dataset

2. Perform classification of tumor and non-tumor cells using the EfficientNetB4 model and train the model with high precision and recall, ensuring reliable detection across diverse cases.

3. Perform segmentation on the tumorous cells found from the classification using SegFormer model and train the model by utilizing SegFormer's efficient hierarchical encoder and self-attention mechanisms to accurately segment brain tumors, achieving superior performance compared to existing CNN-based models.

4. Adapt and optimize the combined model for faster processing, making it suitable for real-time clinical applications.

5.Compare the performance of the SegFormer model with the other three models(UNET,UNET++, LinkNet) based on the dice coefficient and IOU score.

## 1.3    Thesis Organization

This section provides an overview of the structure of our paper. In Chapter 1, we covered the introduction and main objective of our study. In Chapter 2, we cov-

ered brain cancers, brain tumor segmentation, semantic segmentation, and medical picture segmentation. We have also discussed several relevant studies on the segmentation of brain tumors. In Chapter 3, we discussed about the datasets we used for our models and the necessary preprocessing needed to use them. In Chapter 4, we defined our models' structures as well as the procedures for implementation and training. Chapter 5 covers the segmentation results that are examined and a performance comparison of the models is provided.Chapter 6 contains a conclusion and a summary of our paper.

# Chapter 2

# Literature Review

A tumor in the brain is an atypical multiplication of cells that can be categorized as either benign or malignant. The symptoms may vary depending on the location and size of the tumor, but all types of tumors can encompass headaches, seizures, modifications in speech or vision, feeling nauseous, vomiting, and impaired balance or coordination. Treatment options for brain tumors typically involve some combination of surgery, radiation therapy, and chemotherapy, customized based on factors such as the type, size, and location of the tumor, as well as the general health of the patient. As patients' tumor location is an important factor, MRIs are mostly used for brain tumor segmentation. Segmentation is necessary for tumor diagnosis, growth, and treatment planning. Also, clinical experts do manual segmentation which is time-consuming and tends to be not so precise, which is why automated segmentation is necessary. Automatic segmentation of brain tumors based on magnetic resonance images (MRIs) can be a challenging task due to the varied intensity of the tumors.

Over time, many methods have been proposed to address this problem. One such method, as described in [10], is the multi-fractal Brownian model, which uses a stochastic process to describe the spatial distribution of complex structures with multiple scales of variation. The authors propose a novel algorithm to extract the multi-fractal features from the model, which enables better differentiation between tumor and non-tumor tissue. They use data from 14 pediatric patients and show that their method outperforms avant-garde techniques in terms of accuracy and speed. That method aims to accurately map the tumor in the mri image using the healthy reference images. It makes the use of atlas based registration which allows it to segment MRI images carrying tumors.

Another approach, as reported in [2], involves adapting a healthy brain atlas to MR images of tumor patients using a combination of registration techniques and tumor growth simulations. This method aids in accurately mapping tumor growth in the patient's image by comparing it to a healthy reference image. By utilizing atlas-based registration, this method enables implicit segmentation of brain images containing tumors. This is significant not only for accurate segmentation but also for rapid delineation of subcortical structures, which can be helpful in radiotherapy planning.

In this paper [11] ,a new method has been developed for automatic segmentation

glimos from MRI images. That technique utilizing contour model and random forest to generate accurate segmentation also utilizing feature representation technique which is capable to capture local and contextual information from the Mri images.According to the finding, that proposed technique is capable of better segmentation accuracy and computational efficiency comparing with the other techniques available right now. [19] Paper introduces a CNN architecture which is in 3D that utilizes thick and dense connectivity pattern and connection with residual for segmentation of tumor from Mri images. They used hard mining during the training for improving their dice similarity coefficient.

Likewise, in paper [18], for brain tumor segmentation, a 3DUnet based Unet algorithm is introduced that has been trained in a competition name BraTS 2019 competition. To overcome the limitations like needing powerful GPUs they proposed patching strategies and brain-wise normalization.Meanwhile,[1] this paper suggest a new approach for accurately segmenting brain tumor by combining Bayesian classification and multilevel processing. This study solely concentrate on the GBM tumor ,which is the primary tumor that can be found in the nerve system.
Here this paper [12]provides a review of image segmentation methods available for MRI detection of brain tumors and also talks about the therapy planning. This paper provides some machine learning algorithms such as thresholding, region-growing ,CNN and Unet. It also showcased the evaluation metrics to the effectiveness of those techniques.

Additionally, the paper[22] introduced a new technique for segmenting brain tumors using fusion loss ,dynamic decision fusion and deep neural network architecture. This is known as a multi-view dynamic fusion framework. This technique or model was analyzed in two dataset. Overall ,these paper showcased currently ongoing research on brain tumor segmentation and highlighted the challenges in the ongoing experiment.

In this paper [21], those models applied from 2012 and 2018 has been analyzed and showed a significant progression. This paper says that automatic brain tumor segmentation is going to be a better choice for humans by doing more research and advancement in this area. In [4] this paper , based on LIPC ,a brain tumor segmentation model has been proposed. This model utilizes SVM for classifying tumors. The LIPC model achieved high accuracy for the whole ,enhance and tumor core. By advancing this technique it has a potential to segment brain tumors more accurately.

The Tumor-Cut method was presented in the work [5]. It segments brain tumors on contrast-enhanced MRI images for radiosurgery systems using a cellular automata (CA) model. Twenty brain tumor patients underwent testing of the algorithm, which showed excellent accuracy for the total tumor, increased tumor, and core tumor. The discovery has the capacity to precisely identify and distinguish between healthy brain tissues and holds promise for advancement.

Paper[4] provides a new technique to segmenting brain tumors from MRI data that is accurate and efficient. This method consists of two steps: first, utilizing linear programming to estimate the tumor's border; second, employing an optimum mass

transport strategy to refine the boundary prediction. The study emphasizes its potential for therapeutic uses, including arranging radiation therapy and tracking tumor progression.

In the publication [[3], a modified version of the U-Net model was suggested for brain tumor segmentation. In order to enhance the model's performance, the authors corrected class imbalances and applied data augmentation approaches. They therefore produced cutting-edge outcomes on the BRATS 2015 dataset. The study also included a survival prediction algorithm that made use of radio mics. The model demonstrated encouraging results for clinical decision-making by combining a random forest regressor with multilayer perceptrons.

Furthermore, a CNN-based method for segmenting brain tumors was suggested in publication[22]. Using data augmentation approaches, the scientists were able to improve the model's capacity to generalize and achieve high accuracy when testing on MRI images that had tumor locations labeled. The study demonstrates how well CNNs segment brain tumors and advances the development of automated techniques for precise and effective tumor analysis.

Several articles have suggested using machine learning approaches to automate the segmentation of brain tumors. The ZNet model, which uses skip connections, encoder-decoder designs, and data augmentation methodologies to increase segmentation accuracy for brain tumors in 2D MRI images, is one of these approaches that is described in the publication [9]. A die loss function was used in this modified version of the U-Net architecture to address a class imbalance in the dataset. Good performance was shown by the findings, which showed high mean dice similarity coefficient, pixel accuracy, and F1 score values. The approach's potential application to 3D brain volumes was highlighted by the scientists, indicating its use in many medical imaging settings.

Going on to [7], the authors used multi-modal medical imaging data to develop a 3D network architecture for brain tumor and tissue segmentation. To increase segmentation accuracy, our model used volumetric feature alignment and cross-modality feature interaction. The authors obtained higher segmentation accuracy on a brain tumor dataset by mixing characteristics from several modalities and maintaining consistency in the segmentation findings.

Overall, the goal of both studies is to increase the accuracy of brain tumor segmentation by utilizing various methods and architectural designs. While [26]focuses on 2D MRI images and employs a modified U-Net design,utilizes multi-modal medical imaging data and a 3D network architecture. To improve performance, both studies stress the need to combine complementing data and make sure the segmentation findings are consistent.

Several pertinent papers might be noted concerning our particular focus on brain tumor segmentation. A deep learning-based technique for brain tumor segmentation using structural magnetic resonance imaging data was presented by Magadza and Viriri (2021b)[8]. Ding et al. (2020b) [5] created a deep multimodal fusion network

that combines many MRI modalities for the purpose of brain tumor segmentation. The Selective Deeply Supervised Multi-Scale Attention Network was proposed by Rehman et al. (2023)[28]. (SDSAN) method, which segments the tumor area effectively and precisely by utilizing deeply supervised attention multi-scale features.

In addition, the paper [14] tells about semantic segmentation which is very crucial in image processing and computer vision domain. The paper reviewed the progress of semantic segmentation in image segmentation based on deep convolutional neural networks, covering the following eight aspects: a fully convolutional network. These include up-sample ways, FCN joint with CRF methods, dilated convolution approaches, progresses in the backbone network, pyramid methods, Multi-level feature, and multi-stage method, supervised, weakly-supervised and unsupervised method.

In paper [13], they have proposed a semantic segmentation system named ICnet which includes acceleration of network speed without hampering its speed. It helps save operations in multiple resolution and fusion units using an optimal balance of speed and accuracy. It is beneficial for tasks that require fast scene and object segmentation enhancing the practicality of semantic segmentation.

Moreover, in paper [29] conducted a comprehensive study comparing six u-net architectures including two and three layers of variants for the semantic segmentation of teeth in x-ray images. The paper focuses on an in-depth analysis of U-Net models which helps us in choosing segmentation approaches for the ongoing evolution of U-Net models. Among the six U-Net architectures, three-layer variants of R2 U-Net and Dense U-Net give superior performance in terms of the dice coefficient.

Coming to the paper [30], proposes SegFormer, which is a simple yet powerful semantic segmentation method that involves a positional-encoding-free, hierarchical transformer encoder along with a lightweight MLP decoder to avoid the complex designs in previous methods and result in both high efficiency and performance showing strong zero-shot robustness. Its only limitation is that it is not clear if it can work well on a chip-edge device with 100k memory.

The paper [15] addresses the issue of ConvNet scaling and emphasizes maintaining network width, depth, and resolution to enhance both accuracy and efficiency. For this issue, the author introduces a compound scaling method that improves performance also maintaining efficiency. Using this approach they demonstrated that a mobile-size EfficientNetB4 model can perform better than other state-of-the-art models achieving higher accuracy with fewer parameters.

# Chapter 3

# Dataset

Two datasets were used in the research: one dataset was used for classification and another dataset was used for segmentation which contains binary masks and after classification we have segmented the tumor MRI images found from the classification.

## 3.1 Dataset for Classicification

### 3.1.1 Data Collection

The dataset used in this paper was provided by [17], a combination of three datasets: figshare, SARTAJ, and Br35H. The dataset contains a total of 7023 MRI images that are categorized into four differentiable classes that are Glioma, Meningioma,No tumor, and Pituitary.The datasets are collected through MRI scans which are widely used to detect any abnormal functions of brain.

The dataset contains a total of 7023 images are classified into 4 classes those are glioma,meningioma,no tumor,pituitary. Here,glioma tumors are tumors which arises from gilal cells in the brain or spine,meningioma are tumors that arises on membranes of brain covering the brain and spinal cord inside the skull of the brain, no tumor are healthy brain images containing no tumors sourced from Br35H dataset.The MRI images across the dataset varies that's why it is required to resize the images first.As SARTAJ dataset has some mislabeled glioma images those were replaced by figshare dataset images to have more accuracy and maintain data integrity.

Each image is labeled into one of four classes those are no tumor,glioma,meningioma, and pituitary.So,for our research the tumors will be classified into two types which are tumor and non tumor type.There will be no segmentation masks and bounding boxes for non tumors images those are only for images containing tumors.

### 3.1.2 Data Pre-Processing

As a part of pre-processing, we have divided the dataset into train and test datasets and converted the image data into NumPy arrays of type float32. Furthermore, reshaped the NumPy array into a type of 4D array.We have performed batch normalization that stabilizes the training process by normalizing the inputs.We have

modified the model architecture by adding layers like GlobalMaxPooling2D,dense, dropout(0.5), output layer(dense).

## 3.2    Dataset for Segmentation

### 3.2.1    Data Collection

The Brain Tumor Segmentation Dataset consists of MRI scans specifically acquired to identify and categorize brain malignancies, with gliomas receiving special attention. The data collection approach is based on standard medical imaging techniques that employ MRI to provide incredibly precise images of brain tissue.The dataset comes from the Figshare Brain Tumor Dataset provided by[24], which was put together to aid with medical research on brain tumor identification.After being released to the public through the Figshare platform, this dataset has been widely used for both research and teaching.The dataset in kaggle contains 3064 MRI images and their corresponding binary segmentation masks.MRI images were used to capture 2D images of brain which is ideal for for its capacity to differentiate soft tissues within the brain.All the patients identity has been anonymized to protect their privacy.The MRI scans were collected using standard MRI sequences to differentiate between normal brain tissue and brain tumors.The data's were collected during routine scans or diagnostic evaluations usually in a clinical setting.

The dataset contains 3064 MRI images of brain slices, providing a cross-sectional view of the brain showing tumor and non-tumor cells. Each MRI image is given with a binary mask. The MRI images and their corresponding masks are provided in PNG format. Each image and its masks are the same in dimensions. The dataset is designed to support tasks like tumor segmentation and detection which is useful for training deep learning models aimed at automation of identifying brain tumors in medical images. The dataset emphasizes binary tumor detection clarifying the areas on the brain affected by the cancer without further classifying the tumor type.

Throughout the annotation process, segmentation masks are generated that clearly define the boundaries of brain tumors within each MRI image. The binary masks that are provided with each MRI image were created through manual or semi-automated segmentation techniques. The binary masks were denoted by 0 and 1 where 0 means normal brain tissue and 1 means the presence of a tumor. These masks highlighted the tumor region without further dividing the tumor into subcategories like necrotic or enhancing core making it a straightforward binary segmentation mask. Brain tumors may vary significantly in size,shape, and intensity of MRI scans making the annotation process a difficult task. Tumors can have asymmetrical forms, and their contrast with surrounding tissue can vary depending on the MRI sequence used. This variability needs to be carefully annotated.

### 3.2.2 Data Pre-Processing

For pre-processing, we have split our dataset into train and test giving eighty percent for training and twenty percent for testing. Labels (ground truth masks) are converted into binary format dividing by 255. This standardizes the labels so that tumor regions are marked as 1 and non-tumor regions as 0. We normalized the train and test images using feature extractors as a part of image pre-processing. Mean and Standard deviation of the pixel values across the dataset are used for normalization. All images are resized to a consistent dimension (512x512) in order to train the model. We converted the processed images into pixel tensors, in order to be compatible with the SegFormer model.
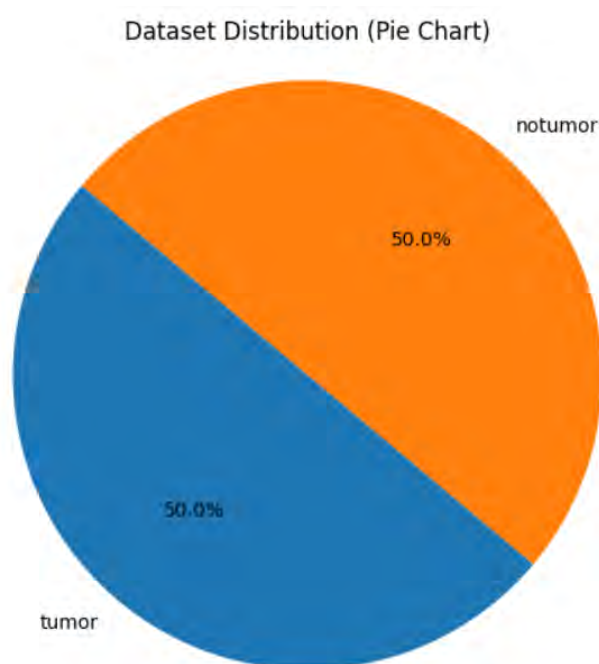


Figure 3.1: Dataset Distribution (Pie Chart)

# Chapter 4

# Methodology

The purpose of this thesis is to design a completely accurate brain tumor segmentation using semantic segmentation. We intend to detect the location of brain tumors more accurately by our supervised semantic segmentation model to help in further treatment planning, and tumor growth prediction.

We want to divide our model into five parts as follows:

1. Take the Brain Tumor MRI dataset for classification which containing tumors and no tumors both and take another dataset the Bain Tumor Segmentation dataset which contains binary masks . Perform necessary pre-processing on the dataset to make it applicable to the model.

2. Split the dataset into train and test sets and load the EfficientNet-B4 model.

3. Train the EfficientNet-B4 model to label tumors and non-tumor and predict tumors or no-tumor.

4. Next, filter the tumor images and segment them using the SegFormer model and generate a segmentation mask.

5. Evaluate the model's performance by comparing its prediction with the original masks.

Firstly, we have chosen the Brain Tumor MRI dataset and Brain Tumor Segmentation dataset for our research to perform semantic segmentation on. Then we imported the necessary modules to run the EfficientNet-B4 model and perform classification.The datasets are not directly applicable to run the model on it so preprocessing on the dataset is necessary.For preprocessing on the dataset first we run the preprocessing function that comes from the library.Also,we have perform data resize and normalization for preprocessing to make it applicable for running it into the EfficientNet-B4 model.Then we divided the dataset into train and test data converting them into numpy array.Input images of(64,64,3) are given into the model and custom layers are added to the EfficientNet-B4 model like GlobalMaxPooling2D for compact representation,dense layer(256 neurons,ReLU),Batch Normalization,Dropout(0.5), output layer(dense) for classification.Then we com-

piled the EfficientNet-B4 model and evaluated the model based on model accuracy,loss,confusion matrix,ROC and PR curves.

For next step, after detecting tumor and no-tumor cells using the trained Efficient-NetB4 model on the segmentation dataset for segmenting the images containing tumor cells we have used SegFormer model.As input, we have given MRI images and tumor masks. For,pre-processing the input images are normalized using feature extractors.Load the SegFormer model and compile the model on the train images.We have trained the SegFormer model using Dice Loss and BCE Loss and used a learning rate scheduler to reduce learning rate at a specified epoch and predict the segmentation mask.We have used AdamW as an optimizer and for evaluation metrics of the SegFormer model we have used Dice Loss, weighted cross entropy, dice score, IoU.
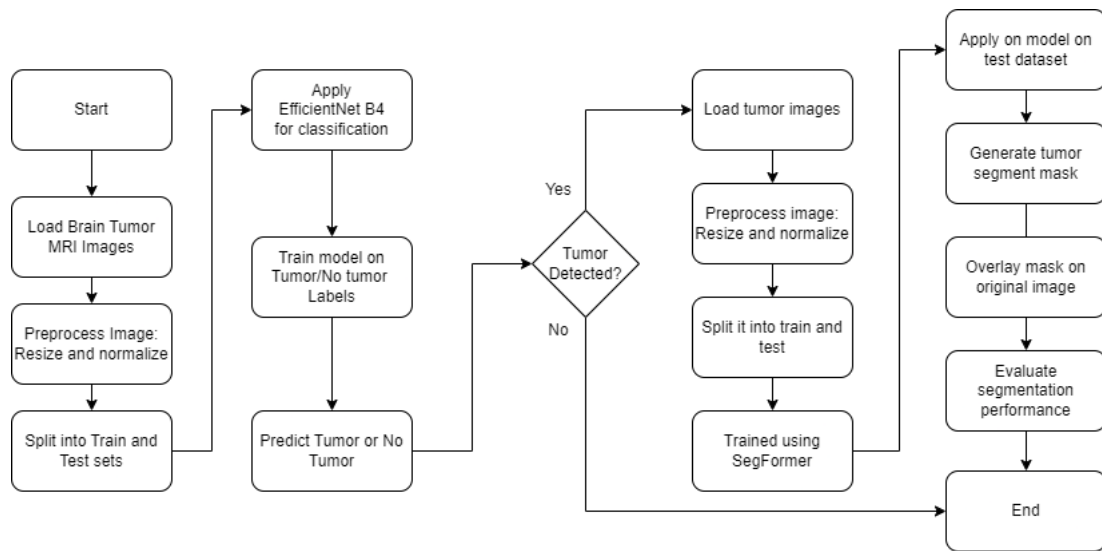


Figure 4.1: Top level overview of the proposed semantic segmentation model

## 4.1 Model Specifications

### 4.1.1 EfficientNetB4

A very effective convolutional neural network architecture called EfficientNetB4 was created for image categorization applications.Contrary to conventional architectures like VGG-19, EfficientNet makes use of an innovative method known as compound scaling to make an efficient use of network depth,width, and resolution.EfficientNet is a new mobile-size baseline developed to better demonstrate the effectiveness of the scaling method.EfficientNet has versions from B0-B7.In our research,we have used EfficientNetB4 architecture which is designed to support various input image sizes and make it compatible with diverse datasets and applications.It's architecture is defined by the optimized use of parameters and computational resources which is achieved by using compound scaling.

EfficientNet has seven versions ranging from B0 to B7, each designed with different levels of accuracy, number of parameters, and computational complexity. The primary distinction between these versions lies in the trade-off between accuracy and resource requirements, such as the number of floating point operations per second (FLOPs) and the number of parameters.

EfficientNetB0: This baseline, while using only 5.3 million parameters and having low FLOPs at 0.39B, achieves a Top-1 accuracy of 77.1% and a Top-5 accuracy of 93.3%. Since it is so light, it is well-suited for most applications, though in terms of accuracy, it falls behind the higher variants. EfficiencyNet B1-B3: From B1 to B3, we can observe that accuracy and resource consumption increase systematically. EfficiencyNetB3 reaches 81.6% Top-1 accuracy with 12M parameters and 1.8B FLOPs, outperforming the smaller models like B0and B1 in terms of accuracy, yet requiring more computational power to operate.

EfficientNetB4: EfficientNetB4 has been chosen for the study is that it strikes a perfect balance between accuracy and computational efficiency. EfficientNetB4, with 19M parameters and 4.2B FLOPs, achieves 83.9% Top-1 accuracy and 96.8% Top5 accuracy. Compared to the lower versions, B4 significantly improves the classification accuracy without the excessive computational load seen in larger models. This makes it particularly well-suited for tasks like brain tumor segmentation where both precision and efficiency are critical.

EfficientNetB5-B7: In the case of larger variants, B5B7, it even furthers the accuracy at considerable computational cost due to an unsettling growth in parameters and FLOPs. For instance, EfficientNetB5, while attaining a Top-1 accuracy of 83.6% comparable to its predecessor B4, this is at a big computational budget of 30M parameters and 9.9B FLOPs-more than twice as costly. While B6 and B7 push the accuracy even higher to 84.3 % and 84.4%, respectively, their resource requirements make them less efficient for practical use: 66M parameters and 19B to 37B FLOPs. [16].

This scaling method improves accuracy compared to traditional existing ConvNets which uses single dimension scaling methods.EfficientNetB4 architecture incorporates depthwise separable convolutions, squeeze and excitation blocks, and efficient channel attention techniques which allow the network to identify intricate patterns in the data that are incoming while decreasing the computational complexity. Usually, for downsampling 2X2 pooling windows including a stride of 2 are implemented.EfficientNetB4's convolutional layer usually implements 3X3 kernels with

a stride of 1 making sure of covering the input image. Moreover,to reduce computational complexity while balancing with expressive power depthwise separable convolutions are implemented. Pooling procedures are usually executed as maxpooling is implemented in its architecture to downsample feature maps which results in enhancing feature extraction and spatial hierarchy learning. The activation mechanism across the network is served by Rectified Linear Unites(ReLU) which introduces nonlinearity and augments the model's capacity to comprehend intricate linkages among the data. The last classification layer of EfficientNetB4 creates probabilities for a variable quantity of output classes which is contingent over the specific classification task.EfficientNetB4 demonstrates enhanced performance and efficiency by outperforming standard designs in many classification tasks.Its scalable architecture and efficient use of computational resources makes it a desired option for practical applications where precision and efficiency are important.

In this research, we have added custom layers. Here, GlobalMaxPooling2D to reduce the spatial dimensions, resulting in a more compact representation. After that, Fully connected layer with 256 neurons (Dense) are added with ReLU activation function to increase the model's ability to capture complex patterns of tumor images. Then, batch normalisation is used to ensure faster and more stable training. Dropout with a rate of 0.5 is added to prevent overfitting. After that, Fully connected layer with Softmax activation function are used to provide outputs in probability distribution format.
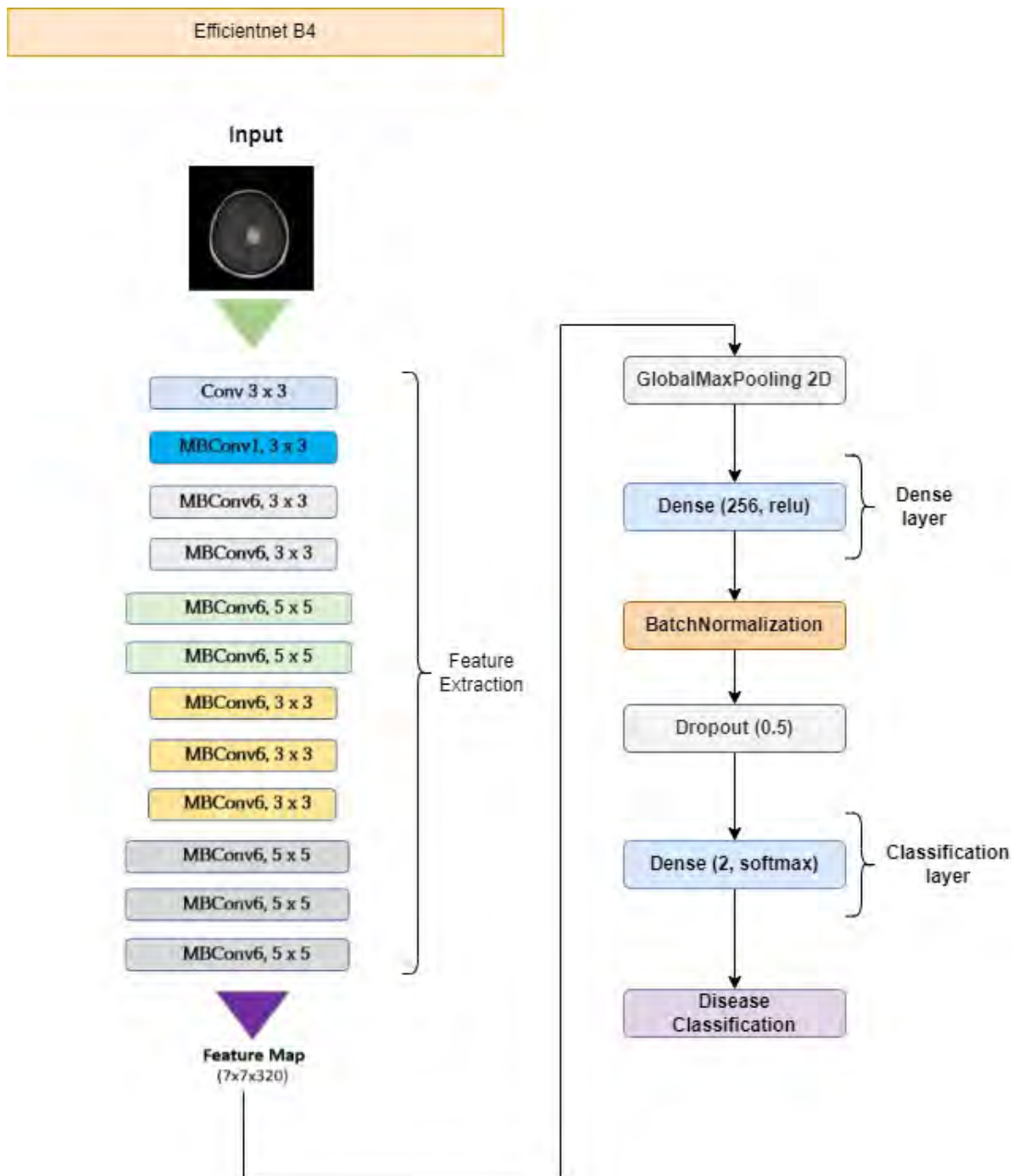
Figure 4.2: EfficientNetB4 Architecture

## 4.1.2  SegFormer

SegFormer is an efficient,robust and powerful framework without hand-crafted and these tend to be computationally demanding modules composed of two major modules, which are a hierarchical transformer encoder for generating high-resolution coarse features and low-resolution fine features with a lightweight All-MLP decoder to fuse these multi-level features to produce the final segmentation mask. In Seg-Former, given an input image, size $H \times W \times 3$ first divide into patch sized $4 \times 4$, here smaller patches help to dense the prediction mask, that is used as an input to the hierarchical. It obtains the multi-level features of the original image resolution through the transformer encoder. These are then fed into the All-MLP decoder to predict the segmentation mask.
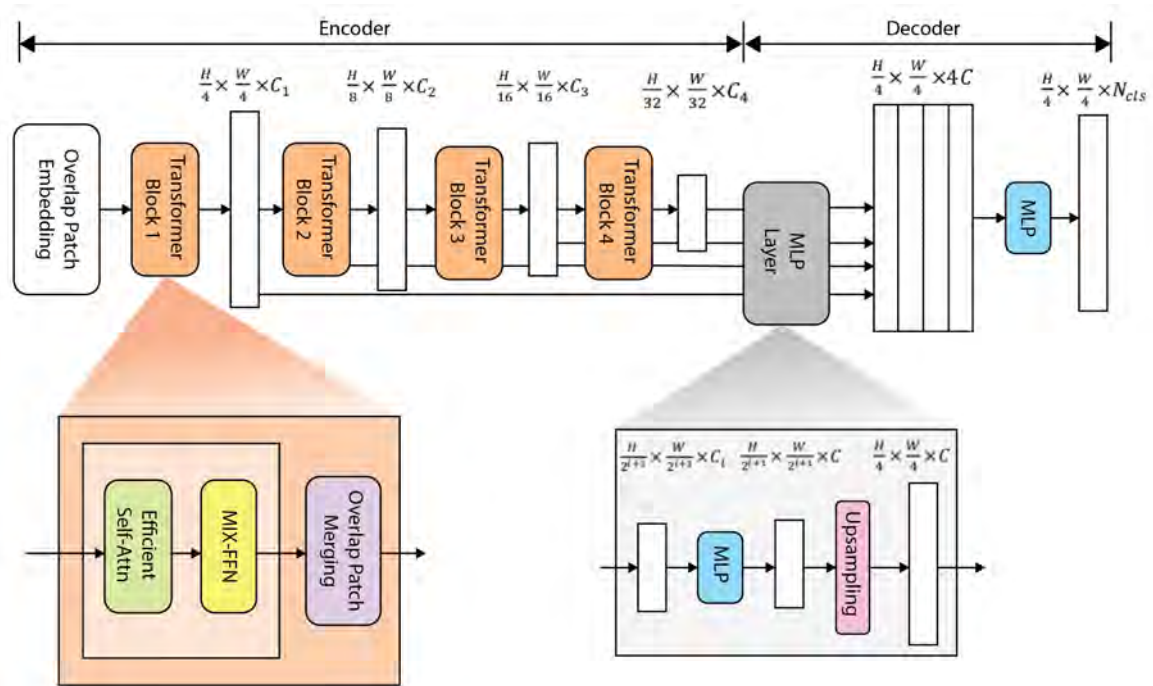


Figure 4.3: SegFormer Architecture

**Hierarchical Transformer Encoder**

Unlike Vision Transformer models that only produce a single-resolution feature on the other hand,SegFormer creates multi-level features.In SegFormer it uses both high-resolution and low-resolution features.Moreover,it divides the image into smaller patches to dense the prediction mask and process them to generate features at multiple levels.That is how SegFormer represents hierarchical features.[25]
The images that splited into smaller patches for better feature representation are overlapping in nature which means some part of the patches are overlap with each other to emit losing important details of about the images content and the method that is used for join the patches assist in maintaining local and global information of the image that helps in achieving better segmentation results.To the end,the patch

merging uses parameters like K,S and P where K defines patch size,S defines the distance between the patches,P defines the extra spaces that were added during processing of the image.[25]

We know that self-attention in transformers is a process that allows SegFormer model to focus on various parts of the image to understand it better but computing this for large image can be slow. As a result,to make it faster and make use of fewer resources before performing self-attendtion SegFormer decrease the amount of information in a sequence.SegFormer don't use Positional encoding rather it learns spatials relationships during training to learn where things are in the image.It uses a combination of 3X3 convolutional with transformer layers to learn about where things are located in the image keeping the model efficient ensuring it doesn't require many parameters and still performs well.[25] The self attention is computed as:

$$Attention(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_{head}}}\right) V \tag{1}$$

Here Q,K,V are the input sequence that have the same dimensions NXC where N=HXW which is the length of the sequence.To reduce the computational complexity for large images a reduction ratio of R used to decrease the length of the sequence as follows:

$$\hat{K} = \text{Reshape}\left(\frac{N}{R}, C \cdot R\right)(K) \tag{2}$$

$$K = \text{Linear}(C \cdot R, C)(\hat{K}) \tag{3}$$

Where K is the sequence that is to be reduced and reshaped with the shape of $\frac{N}{R} \times (C \cdot R).C_{\text{in}}$ dimensional tensor as input and $C_{\text{out}}$ dimensional tensor as output. In SegFormer,Mix-FFN is used which uses 3 X 3 Conv in feed-forward network that considers the impact of zero padding to leak location information.Mix-FFN can be put together as:

$$x_{\text{out}} = \text{MLP}(\text{GELU}(\text{Conv}_{3\times3}(\text{MLP}(x_{\text{in}})))) + x_{\text{in}} \tag{4}$$

Where $x_{\text{in}}$ defines the feature from the self-attention module. Here, Mix-FFN mixes $3 \times 3$ convolution and a MLP into every FFN.

**Lightweight All-MLP Decoder**

SegFormer avoids the need for complex, computationally expensive components that are usually found in other methods by using a lightweight decoder that only depends on MLP layers and this simple yet effective design is made possible due to the hierarchical Transformer encoder that gives a larger effective receptive field in comparison with traditional CNN encoders.It encompasses four steps.First,to unify the channel dimensions of multi-level features$F_{\text{i}}$ from the MiT encoder passed through a MLP layer.Second, the features are up-sampled up to one-fourth of the original resolution and mergedThird,another MLP layer fuses the merged features F. Lastly,a final MLP layer predicts the segmentation mask.The decoder is formulated as:

$$\hat{F}_i = \text{Linear}(C_i, C)(F_i), \forall i \tag{5}$$

$$\hat{F}_i = \text{Upsample}\left(\frac{W}{4} \times \frac{H}{4}\right)(\hat{F}_i), \forall i \tag{6}$$

$$F = \text{Linear}(4C, C)(\text{Concat}(\hat{F}_i)), \quad \forall i \tag{7}$$

$$M = \text{Linear}(C, N_{\text{cls}})(F), \tag{8}$$

Where M defines the predicted segmentation mask and Linear refers to the linear layer with separate input and output vector dimensions respectively.

From the SegFormer model using AdamW optimizer with a learning rate of 0.000006 we computed the loss,dice coefficient,mean iou score.The Mean IoU score came as 0.7382.

# Chapter 5

# Results Analysis

We conducted classification and segmentation tasks for our brain tumor dataset. For the classification task we have utilized EfficientNetB4 and for the segmentation task we employed Unet ,Unet++, Linknet and Transformer based SegFormer model for our dataset. All the models were set up and validated within the PyTorch and Keras environment assisted by GPU support

## 5.1 Classification Module

First on our data we had tumor and non tumor brain mri images. The dataset had 8000 MRI images, 4000 with tumor and 4000 without tumor. We split the data in 80:20 . So, 3200 tumor and 3200 no tumor for the training part and 800 tumor and 800 no tumor images for the validation part.
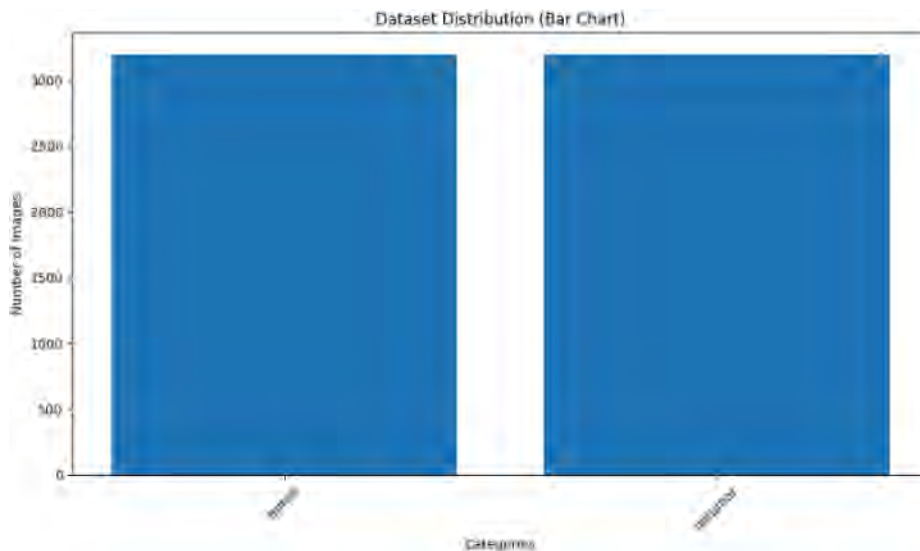


Figure 5.1: Training data distribution chart
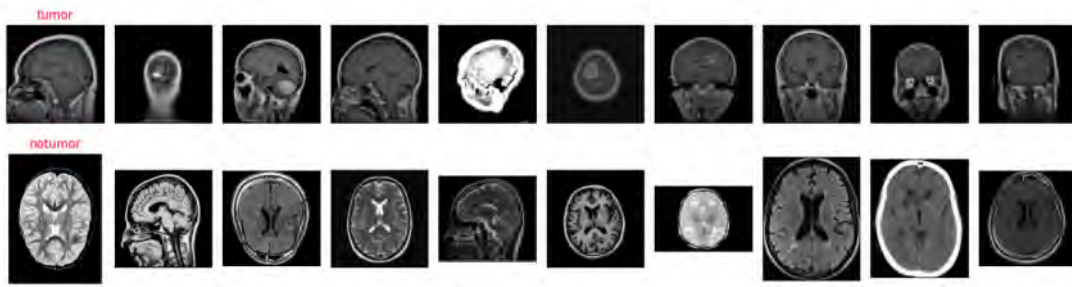
Figure 5.2: Tumor images (up) and Non-tumor image(down)

The shape of each image is (img size, img size, 3) where img size is height and weight and 3 represents the number of channels. We used the EfficientNetB4 model for our classification task. Using keras, we've used the EfficientNetB4 to train with our dataset. We used pre-trained ImageNet weights for our respective model. With 30 epochs and a batch size of 64 we trained our model with our dataset
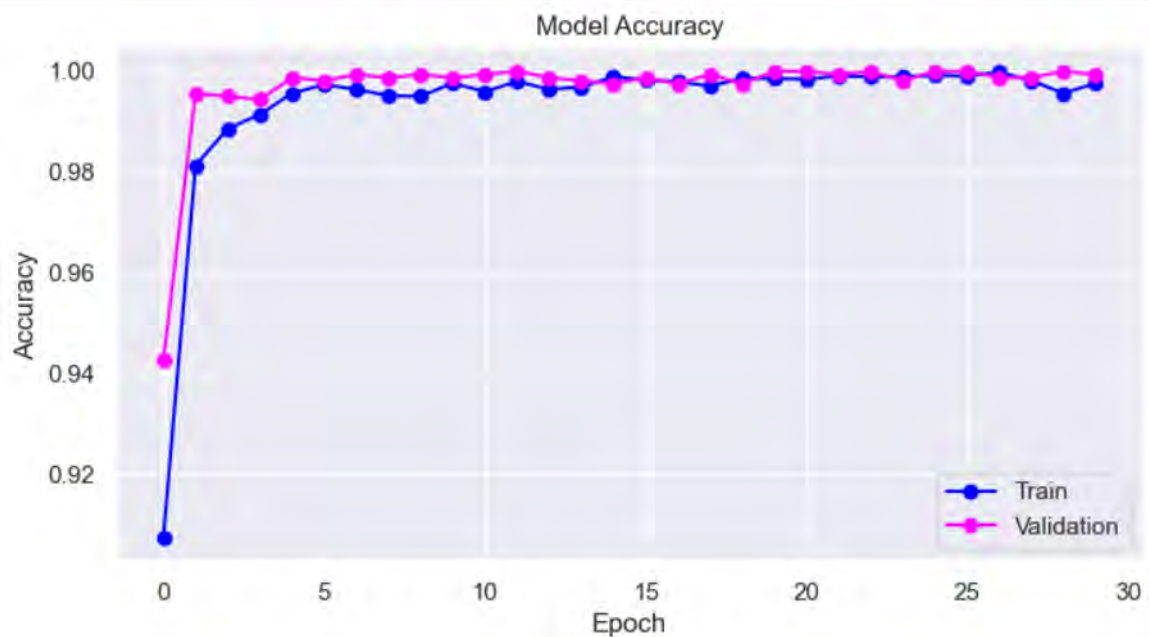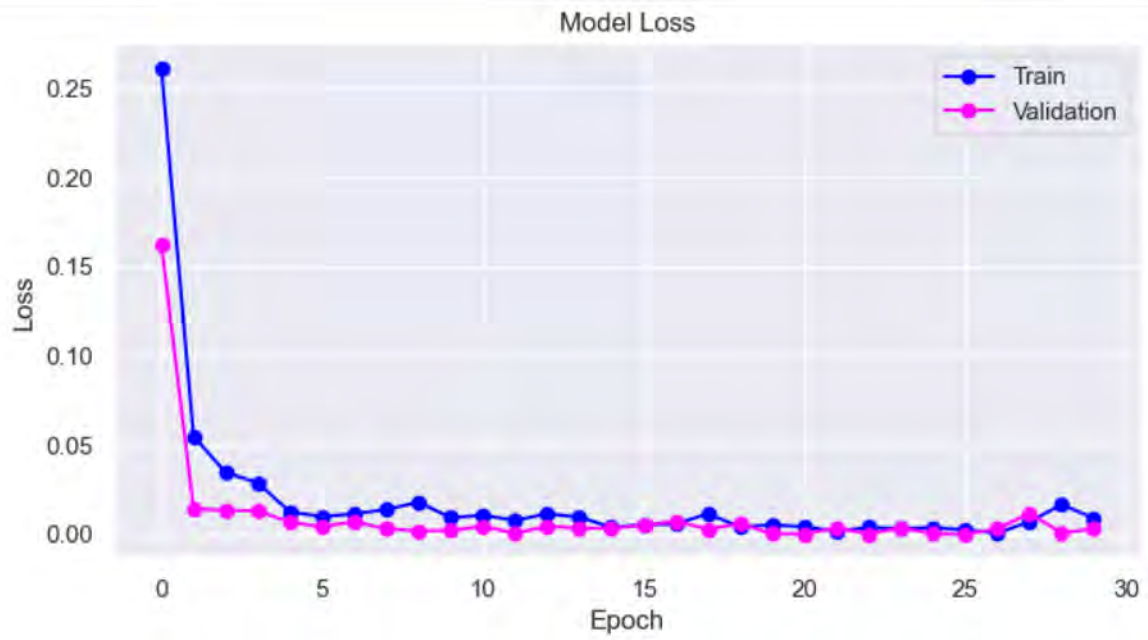


Figure 5.3: Model accuracy graph of EfficientNet-B4

Figure 5.4: Model Loss of EfficientNet-B4

Our Model achieved an accuracy score of 99.937 percent and while training the model loss started high but over time it decreased for both train and validation dataset, which indicates that model was learning effectively and showing better result for unseen data.
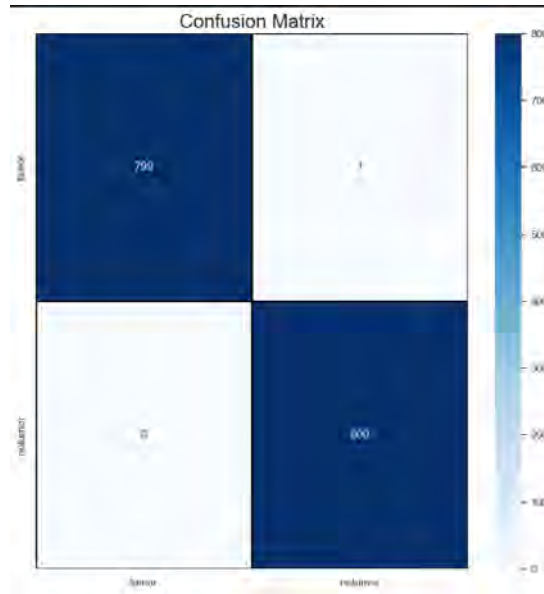
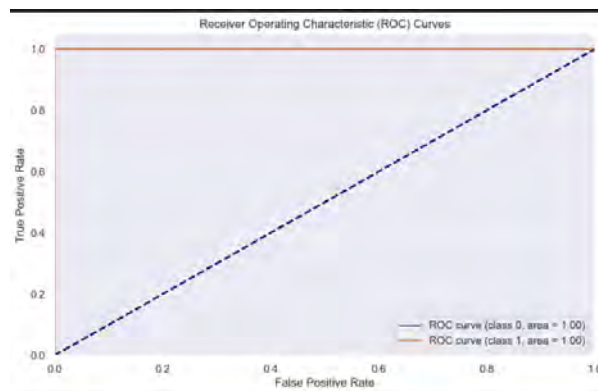Figure 5.5: Confusion Matrix of EfficientNet-B4
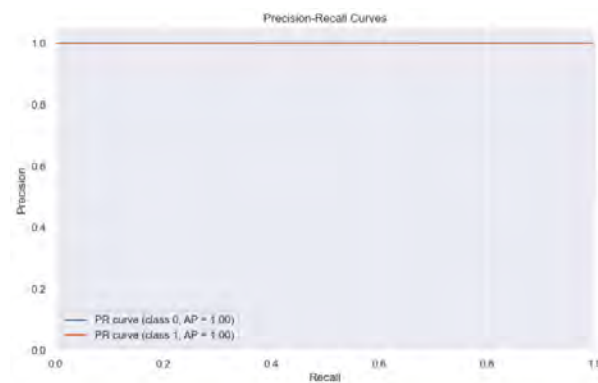


Figure 5.6: ROC Curve of EfficientNet-B4



Figure 5.7: Precision Recall Curve of EfficientNet-B4

For classification tasks ,the Confusion matrix is a useful tool to evaluate the performance of a classification model. From our confusion matrix we can see that the model was able to correctly dentify the 799 'tumor' cases (True positive) out of 800. The model missed 1 'tumor' case ,predicting it as 'no tumor' (False Negative) out of 800. The model did not incorrectly identify any 'no tumor' case as tumor (False Positive). Lastly , The model correctly identified 800 'no tumor' cases (True Negative).

Table 5.1: Precision, Recall, F1-score of EfficientNetB4 classification model

|  | Precision | Recall | F1-score |
|---|---|---|---|
| **Tumor** | 1.000 | 0.999 | 0.999 |
| **No Tumor** | 0.999 | 1.000 | 0.999 |

Our precision score for the Tumor class was excellent , with a score of 100% . That means,every image that was predicted as a tumor was actually a tumor. On the other hand the score for No tumor class came 99.9% , which means only one image predicted as no tumor was a tumor.Our recall score for the Tumor part came 99.99% as only one Tumor it couldn't recognize. And for the NoTumor class it was able to recognize every ,so the score came 100%.Our F1 score for both Tumor and NoTumor class came 99% . F1 score is the harmonic mean of precision and recall ,providing a single metric that balances both. An F1 score of 99% for both Tumor and NoTumor classes indicates an excellent performance by the model, balancing both precision and recall effectively.Indeed, our EfficientNetB4 model shows exceptional performance in the classification task,making it a reliable tool for aiding in medical diagnoses. This level of accuracy can potentially improve patient outcomes by ensuring timely and accurate detection of brain tumors.

## 5.2    Segmentation Module

For the Segmentation task ,we used multiple models. We have used UNet, UNet++,LinkNet and transformer based Segformer. For the UNet , UNet++ and LinkNet we have used pre-trained ResNet50 as a backbone and for the Segformer we used pre-trained Mit-B0 which was trained on a large dataset.

For our LinkNet model we splitted our dataset in 80% for training and 20%testing .Additionally, 15% of our training data was again splitted for validation purposes. The training and testing dataset were resized to a fixed dimension of 224*224. This resizing is done to maintain uniformity across the data before feeding it to the model for training and evaluation.
We have used ResNet50 as backbone using its pre-trained weights to train our model. ResNet50 is a pretrained model which has been trained on a large dataset. The model was compiled with the optimizer Adam with a learning rate set to 0.001. Our model was trained using a batch size of 32 over 70 epochs ,with 15%of the training data was reserved for the validation task. Unfortunately, our LinkNet model Performance didn't meet our expectations. The Dice score we achieved was 0.3445 while the mean IoU score came 0.2344 ,indicating room for improvement in segmentation
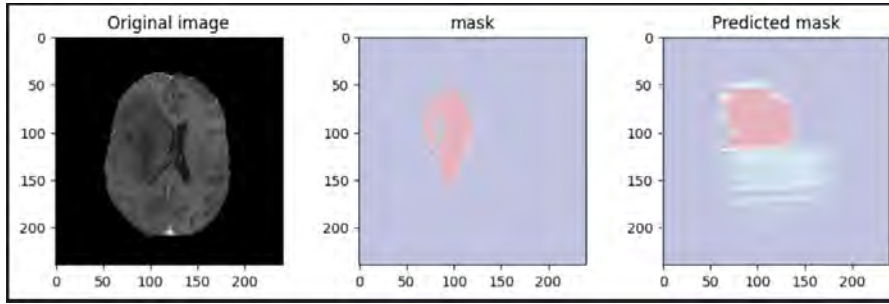
performance.



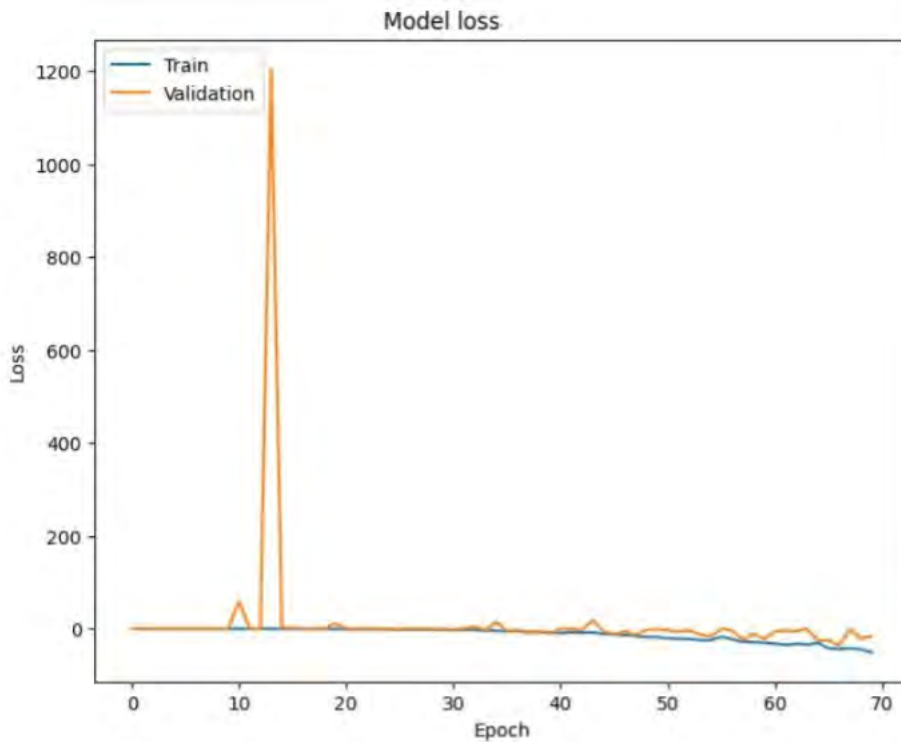Figure 5.8: Predicted output of LinkNet



Figure 5.9: Model loss graph of LinkNet

For our UNet model we splitted our dataset in 80%for training and 20% testing . Additionally,15%of our training data was again splitted for validation purpose. The training and testing dataset were resized to a fixed dimension of 224*224. This resizing is done to maintain uniformity across the data before feeding it to the model for training and evaluation.

We have used ResNet50 as backbone using its pre-trained weights to train our model. ResNet50 is a pretrained model which has been trained on a large dataset. The model was compiled with the optimizer Adam with a learning rate set to 0.001. Our

model was trained using a batch size of 16 over 60 epochs ,with 15% of the training data was reserved for the validation task.Unfortunately, our UNet model Performance didn't meet our expectations. The Dice score we achieved was 0.2985 while the mean IoU score came 0.1800 ,indicating room for improvement in segmentation performance.
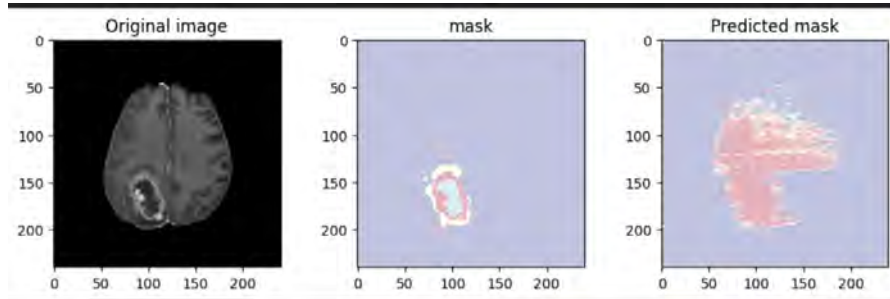


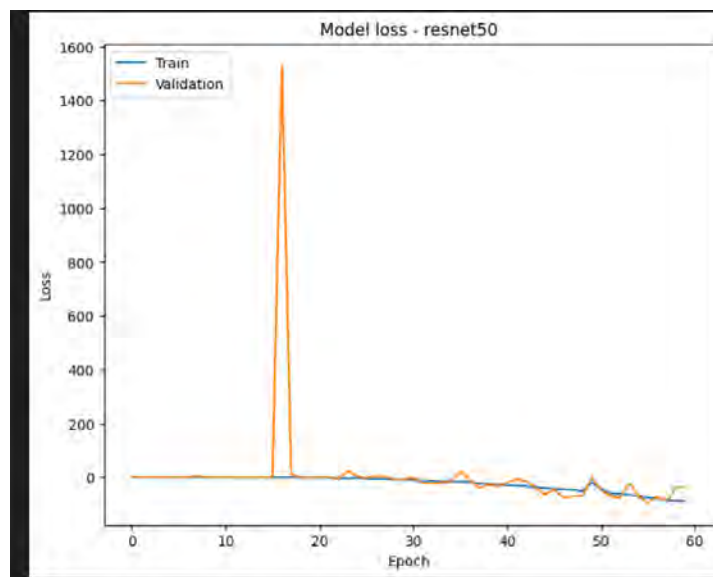Figure 5.10: Predicted output of U-NET



Figure 5.11: Model loss of U-NET

For our UNet++ model we splitted our dataset in 80% for training and 20% testing .Additionally, 15% of our training data was again splitted for validation purposes. The training and testing dataset were resized to a fixed dimension of 224*224. This resizing is done to maintain uniformity across the data before feeding it to the model for training and evaluation. We have used ResNet50 as backbone using its pre-trained weights to train our model. ResNet50 is a pretrained model which has been trained on a large dataset. The model was compiled with the optimizer Adam with a learning rate set to 0.001. Our model was trained using a batch size of 16 over 60 epochs ,with 15% of the training data was reserved for the validation task. Unfortunately, our UNet++ model Performance didn't meet our expectations. The

Dice score we achieved was 0.1575 while the mean IoU score came 0.089 ,indicating room for improvement in segmentation performance.
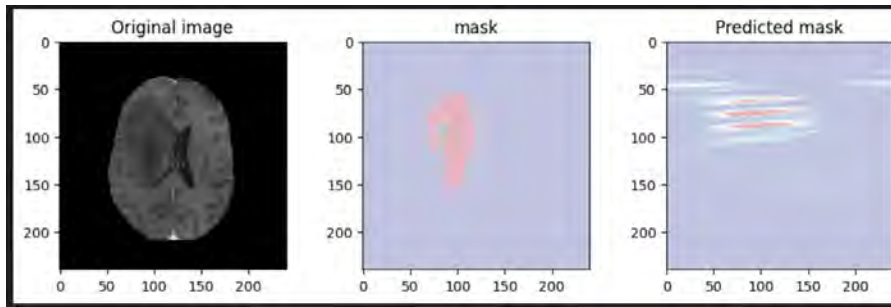


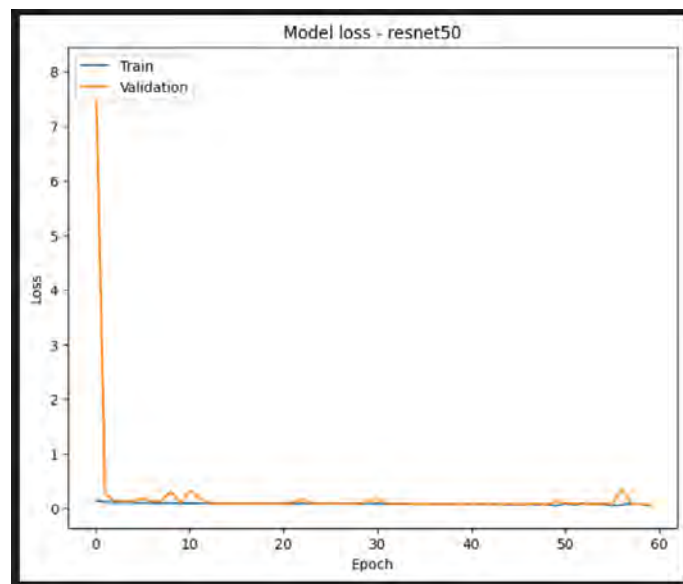Figure 5.12: Predicted output of U-NET++



Figure 5.13: Model loss graph of UNet+

For the SegFormer ,we split the dataset in 80%for training and 20% for testing. Then again we divided the train data 80% for training and 20% for validation. Images in the data were normalized with a mean of 103.84 and a standard deviation of 7.92 .For optimal learning in image segmentation this is crucial. For our SegFormer Custom model we chose MIT-B0 as a backbone, which is a lightweight and efficient transformer suitable for processing the medical images. To optimize our SegFormer model performance for brain tumor sectionalization ,we conducted a careful hyperparameter tuning throughout the training process. We used AdamW as the optimizer , a variant of Adam that incorporates weight decay to prevent over-fitting. Learning rate was set to a very small value of 0.000006. A batch size of 8 was selected for our training and validation. And the model was trained for a total of 150 epochs giving it enough time to learn the patterns required for the accurate brain tumor segmentation.SegFormer model result came excellent, The average Dice score of the SegFormer model came 0.79 and mean iou score came 0.7382
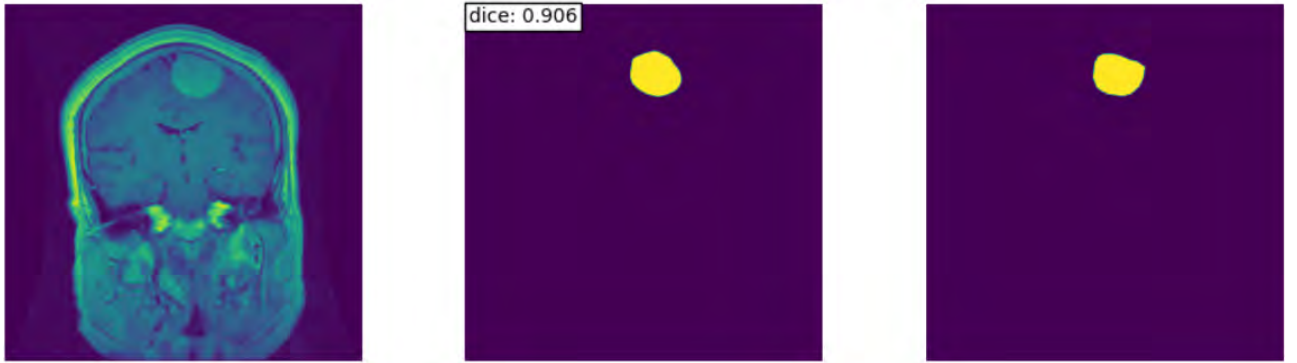
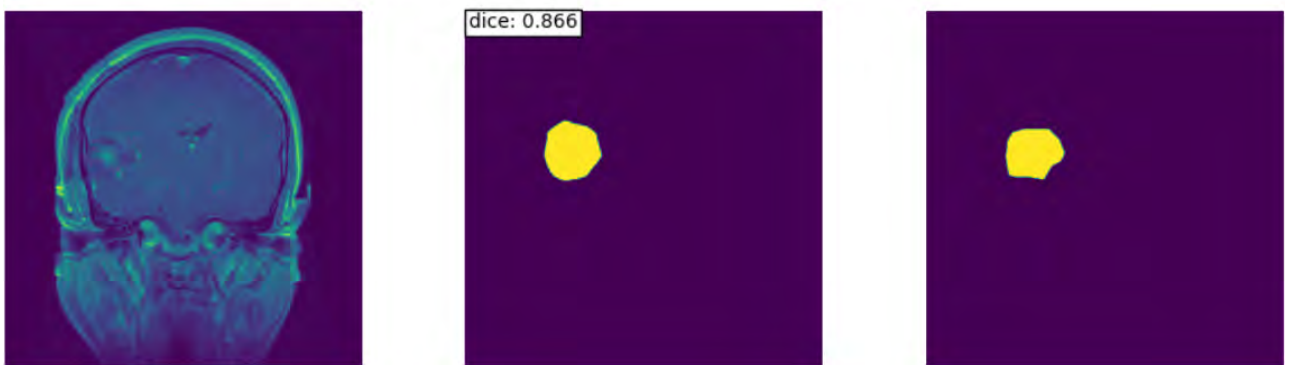Figure 5.14: Predicted Output of SegFormer
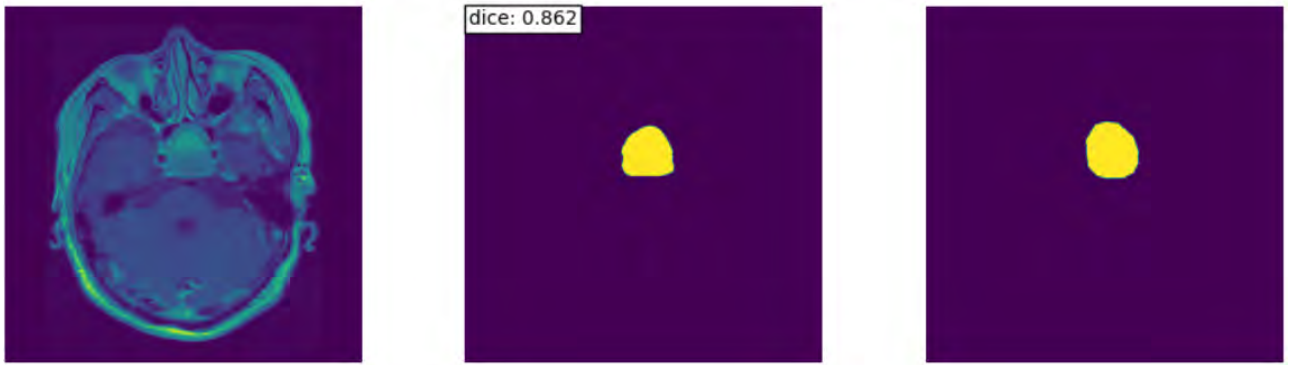


Figure 5.15: Predicted Output of SegFormer

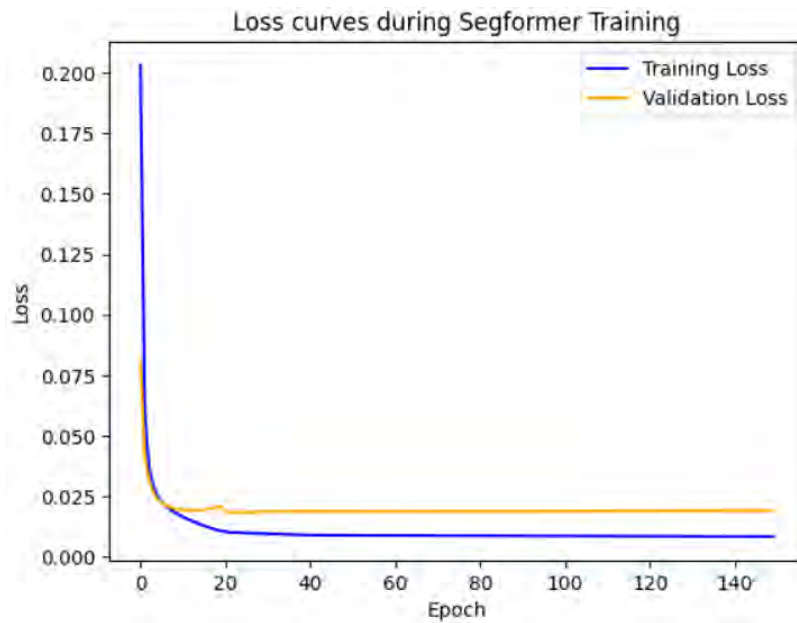Figure 5.16: Predicted Output of SegFormer



Figure 5.17: Loss Curve graph of SegFormer

It is evident on the model loss graph that, over time this model loss was significantly decreasing for both training and validation sets.

Table 5.2: Performance Comparison of Segmentation Models

|  | SegFormer | LinkNet | UNet | UNet++ |
|---|---|---|---|---|
| **Dice Score** | 0.7961 | 0.3445 | 0.2985 | 0.1575 |
| **Mean IoU** | 0.7382 | 0.2344 | 0.1800 | 0.089 |

The Dice score for the SegFormer achieved the highest score among all the other models. The dice score for the SegFormer model is 0.7961. The mean Intersection over Union (IoU) for SegFormer also came significantly higher than the LinkNet,UNet and UNet++, with a score of 0.7382. While LinkNet performed well compared to the other CNN models, it couldn't surpass the SegFormer's performance. From The Above Table ,it is very clear that the transformer-based SegFormer outperformed all CNN based models that includes LinkNet, UNet and UNet++.The Dice Score and mean IoU score for SegFormer are significantly higher than all the other architectures

# Chapter 6

# Conclusion and Future Works

In this paper, we present Brain Tumor Sectionalization through Semantic Segmentation approach. This paper presents an in-depth exploration of brain tumor segmentation using state-of-the-art deep learning models using two different datasets 1)Brain Tumor MRI dataset and 2)Brain Tumor Segmentation dataset. Here we take no tumors and tumors from the Brain Tumor MRI dataset and Brain Tumor Segmentation dataset for segmentating the tumirs.We divided our dataset into train and test images and fed into the EfficieNetB4 model and performed resizing,normalization to make it compatible for the model.After performing classification,we used our trained EfficientNetB4 model to detect tumor and no tumor images from the segmentation dataset and then segment the MRI images containing tumors using SegFormer model.We performed image feature extractors as a part of preprocessing.We evalued the models based on dice coefficient,model accuracy,model loss,IoU score,ROC curves.However,our models have performed the best in terms of dice coefficient and IoU score that's why we have proposed this model for semantic segmentation of brain tumor.We also compared our model's performance with U-Net,U-Net++,LinkNet models trained on the Brats20 dataset.

In addition,various semantic image segmentation techniques have been implemented over the years but it still remains a challenging task to segment overall. One segmentation technique may work on one image but may not work on another image of the same type.That's why it is required for researchers to continue to experiment with different version of same architecture to achieve better accuracy which can help in saving a person's life.There is a lot to improve in the architecture of the models to acquire better results.Further research and experiments need to carry on to improve the performance of the models.We have presented an in-depth review of several brain tumor types using both classification and segmentation.However, we need to experiment using other state-of-the-art models to improve our models performance.We also can large and diverse 3D datasets to achieve better results for our future works.We should optimize our model architecture for improved speed,accuracy and efficiency especially for real time clinical appplications.We should incorporate more advanced techniques like self-supervised learning and transfer learning for further enhaching the model.Semantic segmentation of brain tumor using MRI images has already proven to achieve great success the medical and hopefully in future, it continues to do so. Papers like this are essential for improving the existing segmentation methods including developing more powerful and accurate segmetation methods in future.

# Bibliography

[1] J. J. Corso, E. Sharon, S. Dube, S. El-Saden, U. Sinha, and A. Yuille, "Efficient multilevel brain tumor segmentation with integrated bayesian model classification," *IEEE transactions on medical imaging*, vol. 27, no. 5, pp. 629–640, 2008.

[2] S. Bauer, C. May, D. Dionysiou, G. Stamatakos, P. Buchler, and M. Reyes, "Multiscale modeling for image analysis of brain tumor studies," *IEEE transactions on biomedical engineering*, vol. 59, no. 1, pp. 25–29, 2011.

[3] A. Hamamci, N. Kucuk, K. Karaman, K. Engin, and G. Unal, "Tumor-cut: Segmentation of brain tumors on contrast enhanced mr images for radiosurgery applications," *IEEE transactions on medical imaging*, vol. 31, no. 3, pp. 790–804, 2011.

[4] M. Huang, W. Yang, Y. Wu, J. Jiang, W. Chen, and Q. Feng, "Brain tumor segmentation based on local independent projection-based classification," *IEEE transactions on biomedical engineering*, vol. 61, no. 10, pp. 2633–2645, 2014.

[5] 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, 2014.

[6] 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.

[7] 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.

[8] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.

[9] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3*, Springer, 2018, pp. 287–297.

[10] J. C. C. C. Joseph, "Multi-fractal texture estimation for detection and segmentation of brain tumors," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 17, pp. 1563–1568, 2018.

[11] C. Ma, G. Luo, and K. Wang, "Concatenated and connected random forests with multiscale patch driven active contour model for automated brain tumor segmentation of mr images," *IEEE transactions on medical imaging*, vol. 37, no. 8, pp. 1943–1954, 2018.

[12] Z. Tang, S. Ahmad, P.-T. Yap, and D. Shen, "Multi-atlas segmentation of mr tumor brain images using low-rank based image recovery," *IEEE transactions on medical imaging*, vol. 37, no. 10, pp. 2224–2235, 2018.

[13] H. Zhao, X. Qi, X. Shen, J. Shi, and J. Jia, "Icnet for real-time semantic segmentation on high-resolution images," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 405–420.

[14] X. Liu, Z. Deng, and Y. Yang, "Recent progress in semantic image segmentation," *Artificial Intelligence Review*, vol. 52, pp. 1089–1106, 2019.

[15] M. Tan, "Efficientnet: Rethinking model scaling for convolutional neural networks," *arXiv preprint arXiv:1905.11946*, 2019.

[16] M. Tan, "Efficientnet: Rethinking model scaling for convolutional neural networks," *arXiv preprint arXiv:1905.11946*, 2019.

[17] M. Nickparvar, *Brain tumor mri dataset*, Accessed: 2024-10-22, 2020. [Online]. Available: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset.

[18] F. Wang, R. Jiang, L. Zheng, C. Meng, and B. Biswal, "3d u-net based brain tumor segmentation and survival days prediction," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Revised Selected Papers, Part I*, Springer, 2020, pp. 131–141.

[19] V. K. Anand, S. Grampurohit, P. Aurangabadkar, *et al.*, "Brain tumor segmentation and survival prediction using automatic hard mining in 3d cnn architecture," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6*, Springer, 2021, pp. 310–319.

[20] S. Asgari Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, and G. Hamarneh, "Deep semantic segmentation of natural and medical images: A review," *Artificial Intelligence Review*, vol. 54, pp. 137–178, 2021.

[21] Y. Ding, W. Zheng, J. Geng, *et al.*, "Mvfusfra: A multi-view dynamic fusion framework for multimodal brain tumor segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 4, pp. 1570–1581, 2021.

[22] W.-W. Lin, C. Juang, M.-H. Yueh, *et al.*, "3d brain tumor segmentation using a two-stage optimal mass transport algorithm," *Scientific reports*, vol. 11, no. 1, p. 14 686, 2021.

[23] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using mri multi-modalities brain images," *Scientific Reports*, vol. 11, no. 1, p. 10 930, 2021.

[24] N. Tomar, *Brain tumor segmentation*, Accessed: 2024-10-22, 2021. [Online]. Available: https://www.kaggle.com/datasets/nikhilroxtomar/brain-tumor-segmentation.

[25] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "Segformer: Simple and efficient design for semantic segmentation with transformers," *Advances in neural information processing systems*, vol. 34, pp. 12 077–12 090, 2021.

[26] M. A. Ottom, H. A. Rahman, and I. D. Dinov, "Znet: Deep learning approach for 2d mri brain tumor segmentation," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 10, pp. 1–8, 2022.

[27] T. A. Soomro, L. Zheng, A. J. Afifi, *et al.*, "Image segmentation for mr brain tumor detection using machine learning: A review," *IEEE Reviews in Biomedical Engineering*, 2022.

[28] Y. Zhuang, H. Liu, E. Song, and C.-C. Hung, "A 3d cross-modality feature interaction network with volumetric feature alignment for brain tumor and tissue segmentation," *IEEE Journal of Biomedical and Health Informatics*, 2022.

[29] R. Zannah, M. Bashar, R. B. Mushfiq, A. Chakrabarty, S. Hossain, and Y. J. Jung, "Semantic segmentation on panoramic x-ray images using u-net architectures," *IEEE Access*, 2024.

[30] R. Zannah, M. Bashar, R. B. Mushfiq, A. Chakrabarty, S. Hossain, and Y. J. Jung, "Semantic segmentation on panoramic x-ray images using u-net architectures," *IEEE Access*, 2024.