

Location aware tumor segmentation on MRI images

by

Kevin Jadiya
202111010

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of

MASTER OF TECHNOLOGY
in
INFORMATION AND COMMUNICATION TECHNOLOGY
to

DHIRUBHAI AMBANI INSTITUTE OF INFORMATION AND COMMUNICATION TECHNOLOGY

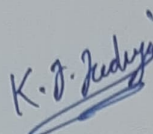


May, 2023

Declaration

I hereby declare that

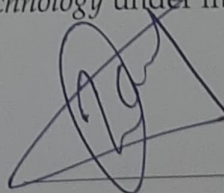
- i) the thesis comprises of my original work towards the degree of Master of Technology in Information and Communication Technology at Dhirubhai Ambani Institute of Information and Communication Technology and has not been submitted elsewhere for a degree,
- ii) due acknowledgment has been made in the text to all the reference material used.



Kevin Jadiya

Certificate

This is to certify that the thesis work entitled **Location aware tumor segmentation on MRI images** has been carried out by **Kevin Jadiya** for the degree of Master of Technology in Information and Communication Technology at *Dhirubhai Ambani Institute of Information and Communication Technology* under my/our supervision.



Prof. Bakul Gohel
Thesis Supervisor

Acknowledgments

I want to express my sincere gratitude and heartfelt appreciation to my esteemed Professor Bakul Gohel for his invaluable support and guidance throughout the journey of completing this thesis. His unwavering dedication, immense knowledge, and profound expertise have been instrumental in shaping this research work. I am deeply indebted to him for his continuous encouragement and belief in my abilities. His guidance has been the driving force behind my academic growth and success. His insightful feedback, constructive criticism, and meticulous attention to detail have played a pivotal role in refining my research methodology and enhancing the overall quality of this thesis. I am genuinely grateful for his time and effort in mentoring me. Moreover, His availability and willingness to address my queries, regardless of how trivial or complex, have fostered an environment of open dialogue and intellectual growth. I would also like to express my deep appreciation for the university and its resources, which have provided me with a conducive environment for learning and research. To my parents, thank you for always standing by me. Your faith in my capabilities, even during moments of doubt, has driven my determination. Your sacrifices and unwavering support have been the foundation for my academic success. I would like to extend my appreciation to my friends who lent a listening ear, offered words of encouragement, and cheered me on from the sidelines.

Contents

Abstract	v
List of Acronyms	vi
List of Tables	vii
List of Figures	viii
1 Introduction	1
1.1 Motivation	3
1.2 Problem Statement	3
1.3 Outline of the Thesis	4
2 Literature Surevey	6
3 Dataset Description	8
4 Proposed Method	11
4.1 Image Registration	11
4.2 Preprocessing the dataset	12
4.2.1 Folder structure of dataset	14
4.2.2 Normalization of dataset	15
4.3 Execution sequence of our proposed method	15
4.4 Model Configuration	19
5 Experiments and Results	22
5.1 Performance Metric	22
5.2 Dice scores of all the models	24
5.3 Comparision with other related work	26
6 Conclusion and Future works	27

Abstract

In our research, we introduce an innovative approach to the segmentation of brain tumors, utilizing a convolutional neural network (CNN) architecture that incorporates localization awareness. This approach represents a significant advancement in tumor segmentation, as it effectively addresses two critical challenges encountered in this field: limited resources and the requirement for precise localization. To overcome these challenges, our methodology leverages 2D slices during training and integrates registration operations for MRI images during application. The proposed method is evaluated extensively on the BRATS-2018 dataset and its augmented dataset version, encompassing distinct variations of CNN-based models. Furthermore, it exhibits computational efficiency during inference, enabling the segmentation of the entire brain in a matter of seconds. The outcomes of our research position our deep learning model as a promising tool with immense potential for both research purposes and clinical applications, offering good segmentation outcomes.

List of Principal Symbols and Acronyms

CNN Convolution Neural Network

DIPY Diffusion Imaging In Python

FLAIR Fluid Attenuation Inversion Recovery

HGG High grade glioma

LGG Low grade glioma

MRI Magnetic resonance imaging

List of Tables

5.1	Dice score on our proposed models when trained on batch size=4 .	24
5.2	Dice score on our proposed models when trained on batch size=10	25
5.3	Dice score on our proposed models when trained on batch size=10 with augmented dataset	25
5.4	Avg dice score comparison of our method and other related models on brats 2018 dataset	26

List of Figures

3.1	All images from dataset	10
4.1	Image registration example	12
4.2	Preprocessing flow	13
4.3	3d images	16
4.4	Slice dataset	16
4.5	Model 1(with control path) flow	17
4.6	Model 2(without control path) flow	17
4.7	Model 1 architecture	20
4.8	Model 2 architecture	21
5.1	Comparison of actual mask and predicted mask	23

CHAPTER 1

Introduction

Image segmentation is the process where we divide an image into meaningful and distinct regions or segments. It aims to partition an image into different regions based on their visual characteristics, such as color, texture, intensity, or object boundaries. The goal of image segmentation is to extract relevant and semantically meaningful information from an image, allowing for a more detailed analysis and understanding of its contents.

The process of image segmentation typically involves the following steps:

Initially, the image may undergo **preprocessing** steps to enhance its quality and remove any noise or artifacts that could affect the segmentation process. Common preprocessing techniques include denoising, contrast enhancement, and image normalization.

Depending on the characteristics of the image and the desired segmentation outcome, different segmentation techniques can be employed. These techniques can be categorized into two types: threshold-based methods and advanced methods.

Threshold-based methods : These methods use a threshold value to divide the image into foreground and background regions. Pixels with intensity values above the threshold are classified as foreground, while those below the threshold are considered background [8].

Advanced methods: These methods utilize more complex algorithms and techniques to segment the image. They may involve edge detection[13], region growing[13] or machine learning-based approaches like watershed transformation[14], active contours (snakes)[14], or convolutional neural networks (CNNs)[10].

In some cases, additional features may be extracted from the segmented regions to provide more detailed information about the objects or regions of interest. These features can include shape, texture, color, or spatial information.

After segmentation, post-processing steps can be applied to refine the results and improve the accuracy. This can involve techniques like noise removal, filling small gaps or holes, morphological operations (such as dilation or erosion), or connected component analysis.

Finally, the segmented regions can be evaluated and validated based on specific criteria or ground truth data. This step helps assess the quality and accuracy of the segmentation results.

The process of image segmentation plays a crucial role in various fields, including computer vision, medical imaging, object recognition, autonomous driving, and many more. It enables applications such as object detection, image understanding, image-based measurements, and subsequent analysis of the segmented regions.

Now lets move from generic image segmentation to our main focus which is brain tumor segmentation.

Gliomas represent the most prevalent primary neoplasms affecting the brain in the adult population. These tumors originate from the glial cells of the brain and are typically classified into distinct grades: High-Grade Gliomas (HGG) exhibit rapid growth and possess greater malignancy, whereas Low-Grade Gliomas (LGG) are characterized by slower growth rates and a more favorable prognosis for patients. Magnetic Resonance Imaging (MRI) plays a crucial role in the assessment of gliomas, enabling the evaluation of disease progression, treatment planning, and overall disease management. Various MRI sequences can be employed to visualize brain tumors, including T1-weighted, T2-weighted, contrast-enhanced T1-weighted images with contrast enhancement (T1ce), and FLAIR images. T2 and FLAIR images primarily highlight the entirety of the tumor region, encompassing infiltrative edema, whereas T1 and T1ce images provide enhanced contrast specifically for the tumor core region, excluding infiltrative edema. Consequently, by integrating these distinct imaging sequences, which offer complementary information, it becomes feasible to analyze different subregions within brain tumors.

The automated segmentation of brain tumors holds great importance in the field of medical image processing, as timely detection of such tumors greatly enhances the prospects of successful therapy and patient survival. Nonetheless, the manual segmentation of brain tumors from vast quantities of MRI data acquired during routine clinical procedures proves to be a labor-intensive and time-consuming endeavor. Hence, there is a pressing need for automated segmentation techniques to expedite the process and improve its efficiency.

1.1 Motivation

For automatic brain tumor segmentation using machine learning models, training these models on MRI data can be a time-consuming process for several reasons. Firstly, the inherent three-dimensional nature and high resolution of MRI data contribute to a large number of features per sample. This substantial feature count significantly extends training times, often requiring several hours or even days to complete a single run. Additionally, the abundance of features can surpass the available memory capacity, leading to the need for data retrieval from disk during training. This further prolongs the training process and adds to the computational load.

Considering these challenges, there is an urgent need to address the issue of reducing computation time and optimizing resource utilization for ML models trained on MRI data. Finding solutions to improve the efficiency of training processes becomes crucial. By minimizing training time, researchers and practitioners can save valuable computational resources and accelerate the overall workflow. This optimization not only benefits the researchers by enabling faster experimentation and analysis but also contributes to enhancing the practical applicability of ML models in clinical settings.

1.2 Problem Statement

Our research aims to address the challenges associated with the segmentation task by introducing a patch-based approach that takes into account the specific location of the image being segmented. Our primary focus is on improving the

accuracy of tumor segmentation by correcting the brain's orientation and location prior to analysis. To achieve this, we will utilize Image Registration techniques, employing a carefully selected template image as a reference point.

By incorporating a localization-aware CNN-based neural network, we seek to leverage the spatial information present in the images and enhance the precision of tumor segmentation. This approach allows us to better capture the distinct features and boundaries of tumor regions within the brain.

Furthermore, we aim to enhance the segmentation process by incorporating control images, which are normal brain images. By including these control images in our model, we enable it to learn the specific characteristics that differentiate tumor regions within normal brain images. This integration establishes a correspondence between normal and abnormal images, contributing to the refinement and optimization of the segmentation process.

Through our comprehensive methodology, we strive to advance the field of medical image analysis and promote improved accuracy in tumor segmentation. By addressing the complexity associated with the segmentation task and incorporating location-specific information, we aim to enhance the performance of segmentation models and ultimately provide more accurate and reliable results in medical imaging.

1.3 Outline of the Thesis

The organization of the thesis is as follows:

Chapter 1 introduces the very basic image segmentation and specifically what is brain tumor segmentation and why it is necessary. We also discuss motivation of our research and formed problem statement.

Chapter 2 presents detailed literature survey of previous methods proposed to solve brain tumor segmentation problem.

In **Chapter 3** we have discussed about dataset we are using for our problem.

Chapter 4 is the one where we discuss various preprocessing techniques we have applied, how our dataset is stored, execution sequence of our proposed method and model configuration.

Chapter 5 discusses the performance metric we've used and the results on our test dataset.

Chapter 6 gives the conclusion of our research work and future scopes.

CHAPTER 2

Literature Surevey

The upcoming section provides pertinent background information regarding prior research efforts on brain tumor segmentation utilizing the BRATS dataset.

According to this study conducted in [7] during the initial stages of the BraTS challenge, the most successful models relied on Random Forest (RF) techniques (BraTS 2012-2014). However, the advent of deep learning revolutionized this trend, and the majority of submissions for BraTS 2015-2018 were based on deep learning models. Recent advancements in deep learning, such as non-linear activation, batch normalization, and dropout, have significantly improved the accuracy of these models and consequently increased their popularity. Since 2015, deep learning-based models, particularly Convolutional Neural Network (CNN) models, have demonstrated the highest performance in the task of segmentation.

CNN models can be developed with either 2D or 3D inputs. In the context of the BraTS dataset, 3D CNNs are anticipated to achieve higher accuracy because they consider valuable information along the z-axis. However, the amount of available GPU memory for training can impose limitations. Consequently, many participants in the BraTS challenge opted to train their models using 2D slices extracted from MRI scans. To address the loss of z-axis information, some participants trained multiple CNNs using slices from different views like coronal, sagittal and axial and subsequently combined the predictions of these CNNs to obtain the final segmentation result. This approach has proven effective in enhancing the accuracy of Dice scores for such models.[17],[4],[3].

The recent advancements of CNN-based models in various image segmentation tasks have motivated researchers to make modifications to further enhance their performance. As a result, networks such as ResNet and DenseNet have been proposed. Several participants in the BraTS17 and BraTS18 challenges leveraged

these architectures to develop their models and reported improved accuracy in terms of Dice scores.[9],[11]

The paper [5] describes the development of a brain tumor segmentation approach utilizing a combination of 3D U-Nets. The method involves training and ensembling six networks, each with varying numbers of encoding/decoding blocks, input patch sizes, and different weights assigned to the loss function. The final prediction probabilities are obtained by averaging the outputs of these ensemble networks.[5]

Automatic glioma segmentation methods in the traditional approach face difficulties in converting prior knowledge into probabilistic maps or selecting highly representative features for classifiers. Nonetheless, convolutional neural networks (CNN) offer an advantage by autonomously acquiring complex features that effectively represent both healthy brain tissues and tumor tissues directly from multi-modal MRI images. The future enhancements and modifications in CNN architectures, along with the incorporation of complementary information from other imaging techniques like Positron Emission Tomography (PET), Magnetic Resonance Spectroscopy (MRS), and Diffusion Tensor Imaging (DTI), hold the potential to enhance existing methods. Ultimately, these advancements could lead to the development of automatic glioma segmentation methods that are clinically acceptable, enabling improved diagnosis.[15]

Based on [12], it was evident that affine transformations remain the most commonly utilized technique in practice. This preference stems from their ease of implementation and their ability to generate anatomically accurate brain tumor examples. However, it is worth noting that there are alternative augmentation methods available that incorporate a combination of different approaches, including elastic transformations.

CHAPTER 3

Dataset Description

This chapter presents details regarding the utilization of the Brats 2018 dataset in our study.

For our study, we utilized the BraTS 2018 dataset [1], which offers a valuable collection of MRI scans obtained from 285 patients with brain tumors. Within this dataset, there were 210 high-grade glioblastomas (HGG) and 75 low-grade gliomas (LGG) available for training purposes. The provided annotations in the dataset focused on three tumor sub-regions, namely the enhancing tumor, peritumoral edema, and the necrotic and non-enhancing tumor core. To simplify the analysis, we combined these annotations into three nested sub-regions: the whole tumor (WT), tumor core (TC), and enhancing tumor (ET). This consolidation allowed for a more comprehensive understanding of the tumor characteristics and their spatial relationship within the brain.

The BraTS dataset includes multimodal scans that are available in NIfTI file format (.nii.gz). Each scan consists of four modalities, namely T1, T1c, T2, and FLAIR. Each modality in MRI has its own unique characteristics and provides valuable information about different aspects of the human body. By combining information from multiple modalities, radiologists and researchers can gain a comprehensive understanding of the structure and function of organs and tissues, aiding in the diagnosis and treatment of various medical conditions. Here are the information of modalities we have used in our dataset:

1. T1-weighted (T1): T1-weighted images provide good anatomical details with high contrast between different tissues. They are often used to visualize the structure and morphology of organs and tissues. T1 images are typically acquired in the axial or sagittal planes.

2. T1-weighted post-contrast (T1CE): T1-weighted post-contrast images are obtained after the administration of a contrast agent, such as Gadolinium. Contrast agents highlight certain structures, such as blood vessels or tumors, which appear brighter in the image. T1CE images help in enhancing the visibility of abnormal tissues, such as tumors and lesions.

3. T2-weighted (T2): T2-weighted images are sensitive to differences in water content and provide good contrast between different soft tissues. T2 images are particularly useful for detecting edema, inflammation, and fluid-filled structures. They are commonly acquired in the axial, coronal, or sagittal planes.

4. Fluid-attenuated inversion recovery (FLAIR): FLAIR images are a variant of T2-weighted images that suppress the signal from cerebrospinal fluid (CSF). By nullifying the signal from CSF, FLAIR images enhance the visibility of abnormalities, such as lesions and brain tissue changes. FLAIR images are useful for detecting white matter lesions and evaluating conditions like multiple sclerosis.

In our study, we unified the annotations with three distinct classes, assigned with labels 1, 2, and 4, into a single tumor class by labeling it as 1. The background region was designated as label 0. This merging of labels facilitated the binary segmentation process, focusing on distinguishing the tumor region from the background. This simplification is depicted in figure 3.1 part (f), illustrating the combined tumor class and background assignment.

By using this dataset and considering the simplified binary segmentation task, our model aims to accurately identify and segment the tumor regions within the brain. The availability of multiple modalities and the comprehensive annotations provided in the BraTS 2018 dataset enable us to train our model effectively and facilitate the development of an accurate and robust tumor segmentation approach.

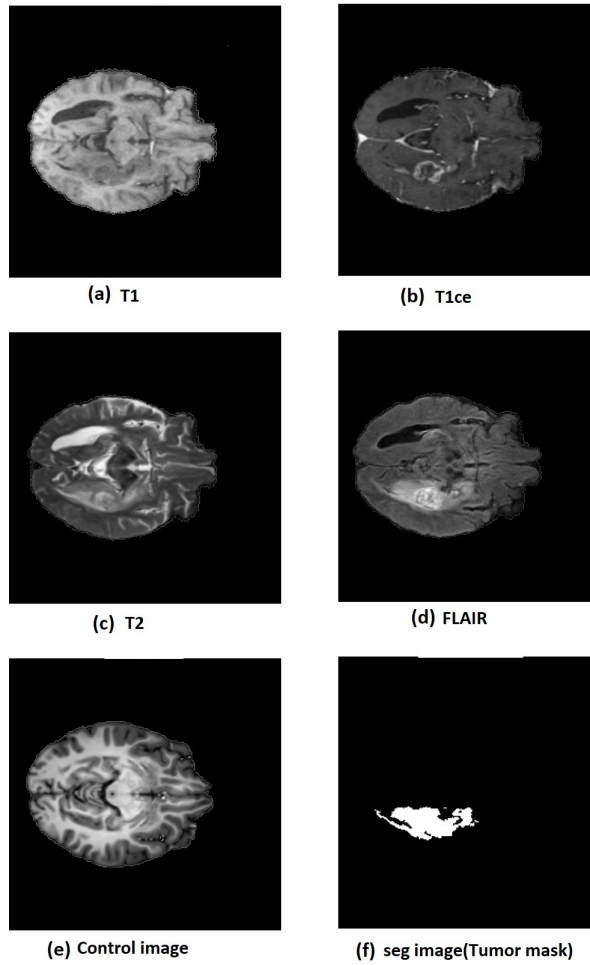


Figure 3.1: a) T1 image (b) T1ce (c) T2 (d) Flair (e) Control Image (f) Seg Image(Tumor mask) - Black: Background and Healthy tissue, White: Tumor tissue

CHAPTER 4

Proposed Method

Our research introduces a novel approach we call location-aware tumor segmentation. Unlike conventional techniques that rely solely on pixel-to-pixel mapping, our approach recognizes the importance of spatial information within MRI data, as the orientations of normal and abnormal brain images may differ significantly. In order to achieve precise and reliable segmentation results, a more sophisticated strategy is required.

A simple pixel-to-pixel mapping approach fails to consider the fact that a voxel at a particular location in one image may not correspond to the same brain region in another image. To address this challenge, we have devised a solution that incorporates the spatial context of the MRI data. Our technique involves performing image registration, specifically affine registration, to correct for variations in orientation and location across the images.

By incorporating affine registration into our methodology, we overcome the limitations of a simplistic pixel-to-pixel mapping and achieve a more robust and accurate tumor segmentation outcome.

4.1 Image Registration

Image registration in medical imaging refers to the process of aligning and matching two or more medical images of the same patient or region of interest. It involves finding a spatial transformation that brings the images into correspondence, allowing for a comprehensive analysis and comparison of the information contained in each image.

Affine registration is a technique used in medical image processing and image

registration to align and match two or more medical images. It applies an affine transformation to spatially transform one image to match the other image. The affine transformation consists of translation, rotation, scaling, and shearing operations.

In affine registration, the goal is to find the optimal set of transformation parameters that minimize the differences between corresponding features or structures in the images being registered. By applying the affine transformation, the images can be aligned in a way that compensates for differences in position, orientation, and scale.

Example of image registration operation on a 2d slice of a patient's MRI image is shown in Figure 4.1 from the brats 2018 dataset

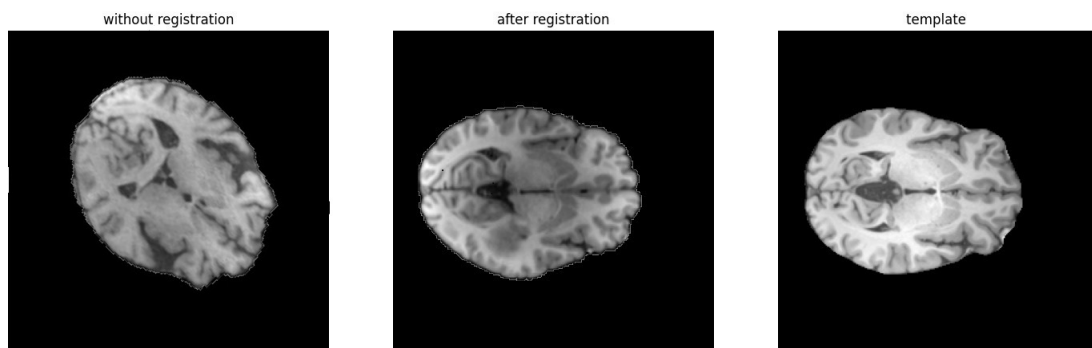


Figure 4.1: Example of how query image will look like without registration and after registration with given template image

4.2 Preprocessing the dataset

In our approach, illustrated in figure 4.2, we employ image registration technique to ensure consistent orientation and location across all MRI images in our dataset. This involves obtaining a transformation matrix for each patient's MRI picture based on the orientation of a fixed template image. By aligning the images to a common reference frame, we establish a unified coordinate system for the dataset.

To implement this image registration process, we utilize the powerful capabilities of the dipy library[2]. Dipy is a popular Python library that provides various tools and algorithms for diffusion MRI analysis and image registration. It offers

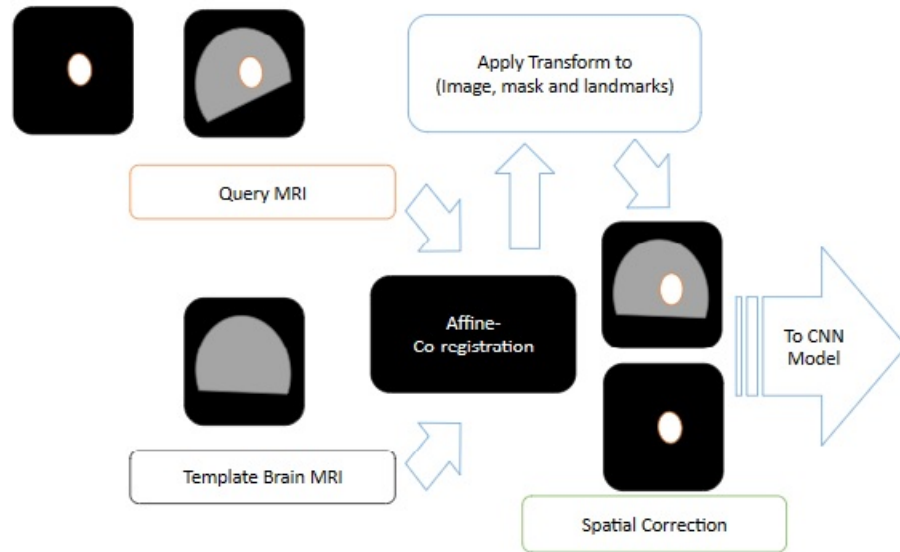


Figure 4.2: Preprocessing flow: based on template image orientation, We'll apply affine registration on query image, and based on calculated transformation, we'll get the transformed image.

a comprehensive set of functions for image transformation, registration, and manipulation, making it an ideal choice for our task.

Firstly, we obtain a fixed template image that serves as the reference image for orientation and location. The transformation matrix is then calculated by aligning each patient's MRI image to this fixed template. This matrix captures the necessary translation, rotation, and scaling parameters required to transform the patient's image to match the template's orientation and location.

Next, we apply this transformation matrix to both the MRI image and its associated mask. By performing the transformation on both the image and the mask simultaneously, we ensure that the spatial relationship between the image and the corresponding segmentation mask is preserved. This step is crucial for accurate analysis and subsequent segmentation tasks.

As a result of this image registration procedure, we create a modified image dataset in which all the images share the same orientation and location. This standardized dataset eliminates any variations or inconsistencies in image positioning, enabling fair and reliable comparisons across patients. This alignment facilitates subsequent analysis and segmentation tasks, ensuring consistent and accurate re-

sults.

To ensure compatibility with our image registration process, we applied zero padding to the original images from the Brats 2018 dataset. The original images had dimensions of 240x240x150, but we extended them using zero padding to achieve a uniform image size of 256x256x256. This adjustment was necessary as our image registration technique relies on a static image size of 256x256x256.

4.2.1 Folder structure of dataset

In order to optimize runtime performance, it may not be efficient to load the entire 3D image into memory. To address this issue, we implemented a strategy where we organized the images into separate folders based on their content. This approach allowed us to enhance the efficiency of image retrieval and processing.

More specifically, we created image-wise folders, where each folder contained a set of 10 carefully selected image slices. These slices were strategically chosen to include a slice from the tumor's centroid, as well as the adjacent slices immediately above and below it. The selection of these specific slices aimed to capture the essential tumor region while minimizing the memory footprint required for processing.

By adopting this approach, we were able to load and process a smaller subset of the 3D image at a time, reducing the memory usage and enhancing runtime efficiency. It made training process much faster.

Folders structure of the dataset is made as follow:

img_0,img_1,...,img_n : MRI image / features slices are stored in each img folder

mask_0,mask_1,...,mask_n : MRI image mask/ labels slices are stored in each mask folder

fixed_0,fixed_1,...,fixed_n : MRI image control image but one particular fix image common for all corresponding input image

var_0,var_1,...,var_n : MRI image control image but variable image for corresponding input image

Here n represent number of input image of different patients available in our dataset. In our case n= 285.

The naming of all the sliced data ensures their correspondence with each other. The folders are numbered at the end of their names, allowing us to access corresponding images across multiple folders.

4.2.2 Normalization of dataset

Due to the inclusion of four modalities (T1, T1-CE, T2, and FLAIR) in our dataset, we encounter a challenge related to varying contrast levels. This contrast discrepancy can lead to the loss of gradient information during the training phase. To address this issue, we employ a standardization technique to preprocess the images. By performing image-level standardization, we normalize the image data by subtracting the mean value and dividing it by the standard deviation. Formulation is as follows:

$$\hat{X} = \frac{X - \mu}{\sigma} \quad (4.1)$$

In the given context, μ represents the mean of the image, σ represents the standard deviation, X represents the image matrix, and \hat{X} denotes the normalized image matrix.

4.3 Execution sequence of our proposed method

To begin with, we will sequentially load 3D MRI images of all four modalities for each patient. These images, depicted as gray-colored cubic volumes in Figure 4.3, serve as our input data. Additionally, there will be a common mask for each variant of MRI images, represented by blue-colored cubic volumes in Figure 4.3, which will act as our label. For Model Architecture 1, we will also load a control image volume, denoted as a light green-colored cubic volume in figure 4.3, which represents a normal image.

After loading the image, we will first do preprocessing as mentioned in section 4.2 and will save the slice data in corresponding folders.

Once the preprocessing step is complete, we proceed to obtain the training dataset, which consists of image slices sized (256,256,4). This dimension signifies that we have 2D image slices with dimensions (256,256) from all four modalities. Subsequently, we combine the individual slices to create image batches with a size of (n,256,256,4), where 'n' represents the batch size. This is illustrated in Figure 4.4.

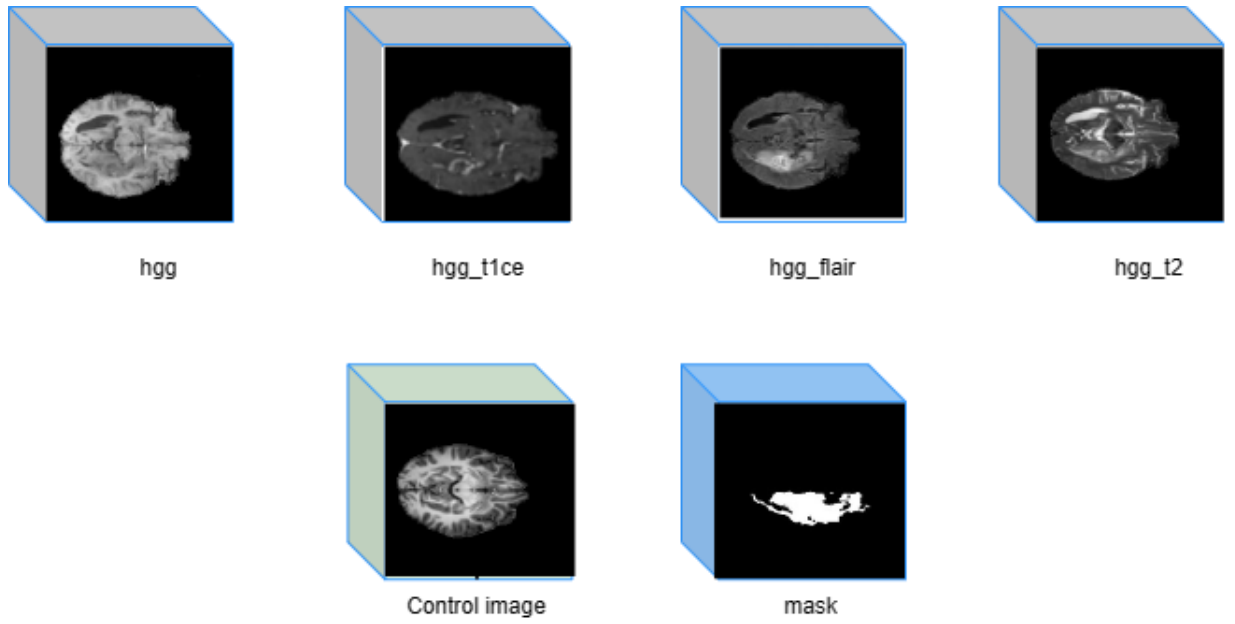


Figure 4.3: 3d input volume of each type of images

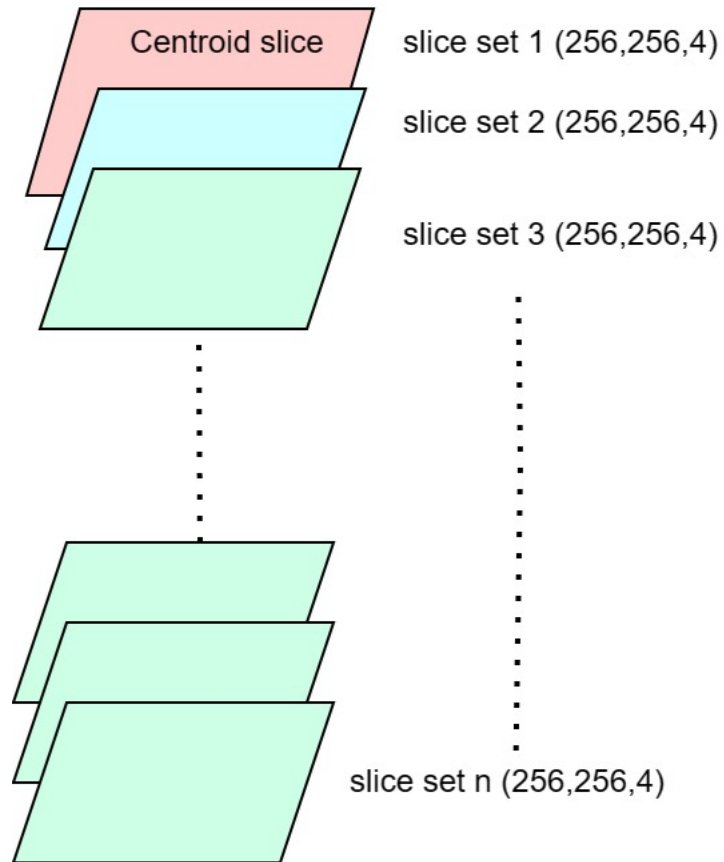


Figure 4.4: 2d slices to make batch of images for training data

Our approach is mainly divided into two models based on its structure and dataset:

1. **Model with control path** :In this case, our convolutional neural network (CNN)-based model will be structured with two separate paths. Each path will receive distinct inputs, namely the query image and the control image. The query image is sourced from the Brats 2018 dataset, while the control image is an MRI image devoid of any abnormalities, as illustrated in the accompanying figure 4.5
2. **Model without control path** : In this scenario, our CNN-based model will adopt a single path architecture, incorporating normal CNN blocks at intermediate stages. The input for the model will consist solely of the query image, as depicted in the accompanying figure 4.6

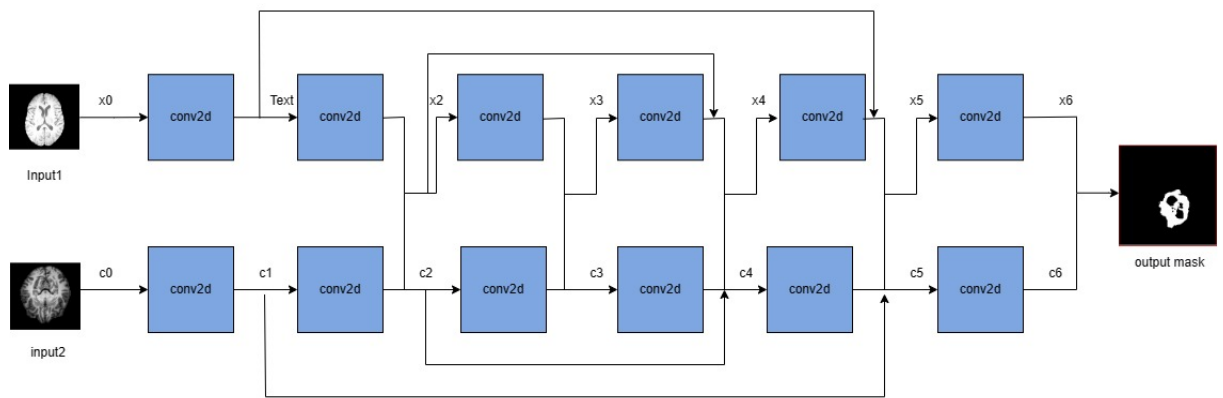


Figure 4.5: Model flow with control path

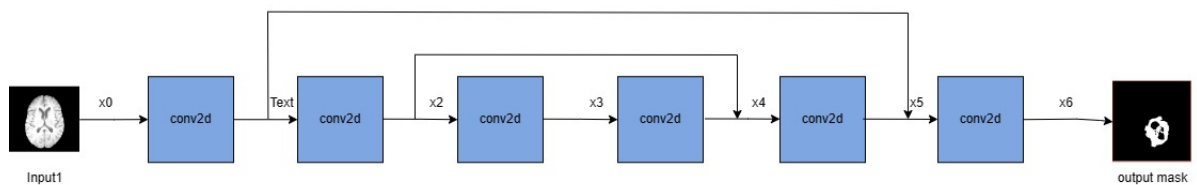


Figure 4.6: Model flow without control path

In our approach, we introduce a modification by incorporating two input images into our model. This addition is motivated by the objective of training our model to effectively differentiate between normal and abnormal images. By providing two input images, we enable the model to gain a comprehensive understanding of the features and patterns present in both types of images. Consequently, this enhanced input configuration enhances the model's ability to accurately discern between normal and abnormal instances, leading to improved classification performance and more reliable diagnostic outcomes.

At present, our CNN-based model utilizes 2D slices extracted from MRI images as input, with a size of $(256, 256, 4)$. The model processes these slices and generates a segmented 2D mask as output, with a size of $(256, 256, 1)$. This configuration allows the model to analyze and extract relevant information from the MRI slices, ultimately producing a mask that highlights and delineates specific regions of interest within the images.

Our primary CNN-based approach can be categorized into three variations, each dependent on the dataset being utilized. Each three variations with different datasets are described below:

1. Model with the Registered dataset:

- (a) **Fixed Control Image on Registered Dataset** : In this model, we maintain a fixed control image throughout the training process. At each training iteration, the same control image is provided consistently. This approach ensures that the model receives a constant reference point for comparison and enables it to learn and adapt based on the consistent presence of the control image.
- (b) **Variable Control Image on the Registered Dataset**: In this approach, during the training process, we introduce a random control image selected from a pool of 15 available control images. This randomness allows for a diverse set of control images to be utilized, enhancing the model's ability to generalize and adapt to various scenarios during training.
- (c) **Without Control Image on the Registered Dataset** : It adopts a different approach. In this case, the model does not incorporate a control

path. Instead, it solely relies on the query image as input, derived from the registered dataset.

2. Model with the unregistered dataset:

- (a) **Fixed Control Image on Unregistered Dataset** : similar to variation a) but with unregistered dataset
- (b) **Variable Control Image on the Unregistered Dataset** : similar to variation b) but with unregistered dataset
- (c) **Without Control Image on the Unregistered Dataset** : No control input, only query image will be provided from the unregistered dataset

4.4 Model Configuration

To train our model, we employed the 5-fold cross-validation technique. For each fold, we conducted testing on all 256 2D slices of the test image. Subsequently, we calculated the Dice score by comparing the actual 3D volume with the combined predicted 3D volume obtained from the predicted 2D slices.

Model configuration information is as follow:

No. of total parameters in model 1: 50,145

No. of total parameters in model 2: 17,009

Optimizer : Adam

Loss function : Dice loss

No of epochs for training: 300

Training set : 10 slices from each 228 3d image volumes

Test set: 57 3d image volumes

Model evaluation technique: 5 fold cross validation

Figure 4.7 shows model with control path and Figure 4.8 shows model architecture without control path with all the parameters in each layers. We can see our model consist of convolution blocks, batch normalization and some skip connections.

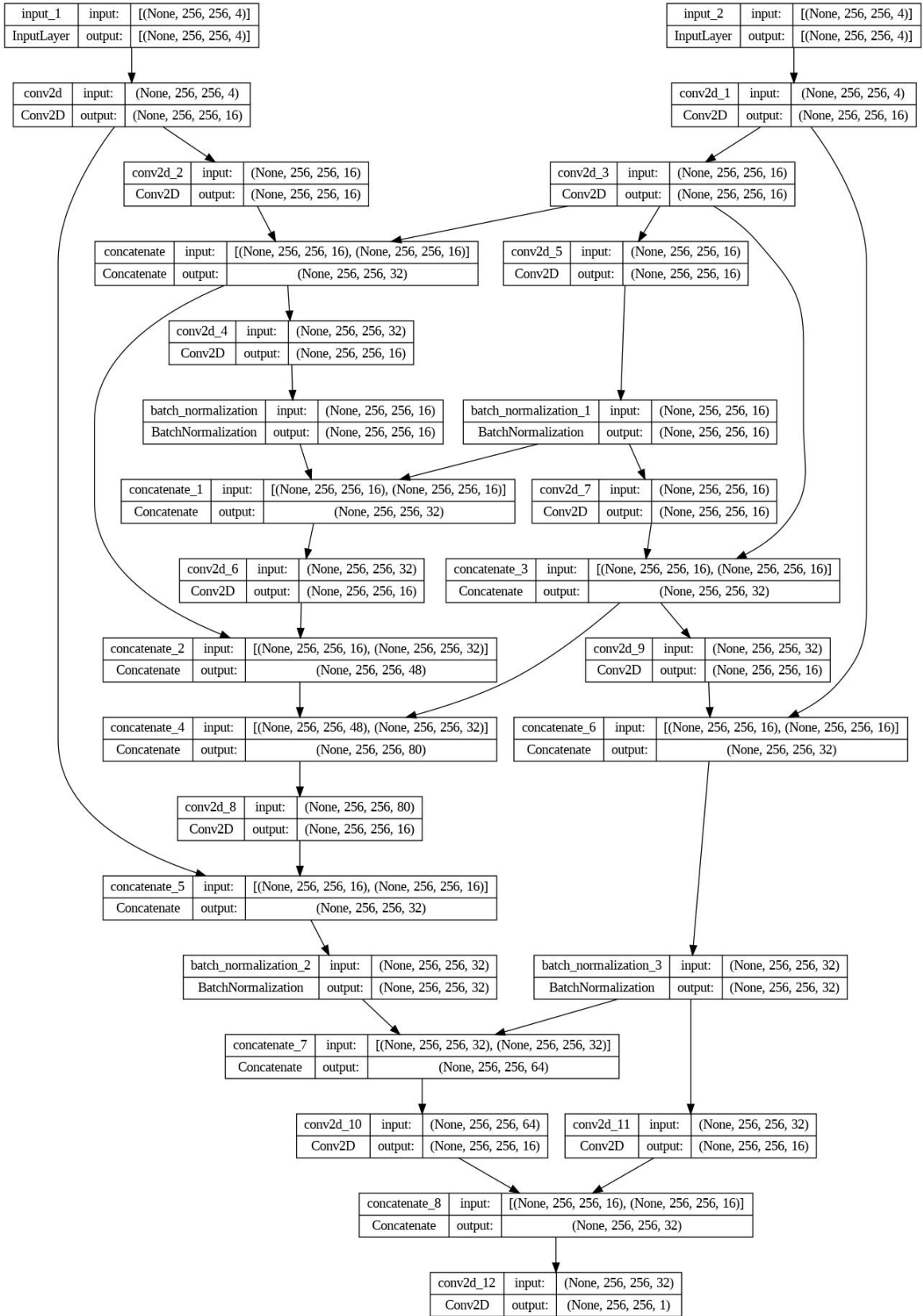


Figure 4.7: Model architecture with control path

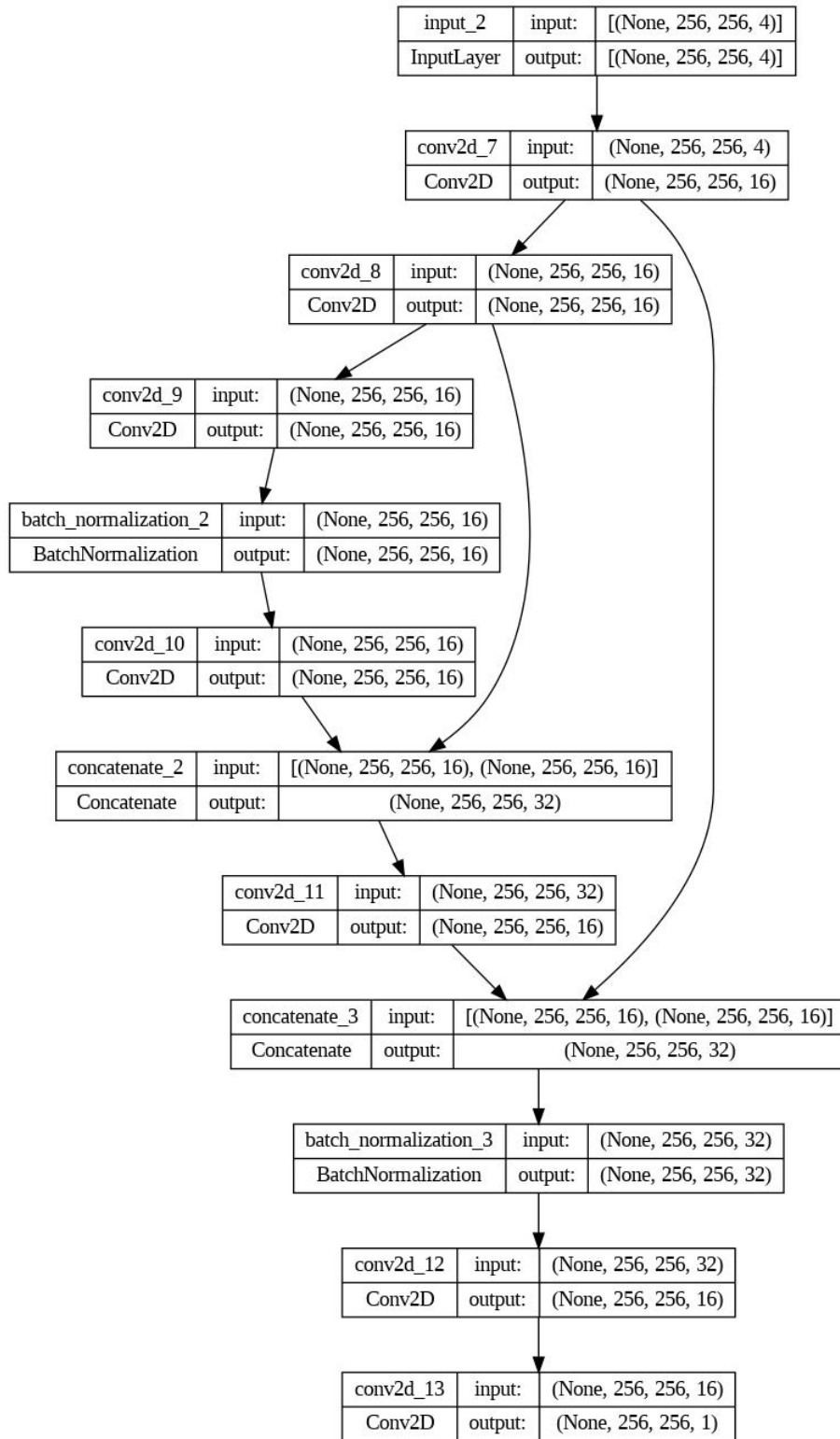


Figure 4.8: Model architecture without control path

CHAPTER 5

Experiments and Results

The figure 5.1 displays a comparison between the actual mask and the predicted mask for a set of test images. The actual mask represents the ground truth segmentation of the target object or region within each image, while the predicted mask represents the segmentation output obtained from a trained model or algorithm.

5.1 Performance Metric

The Dice score, referred to as the Dice coefficient or Dice similarity coefficient, is a widely used metric for assessing how well image segmentation tasks perform. Its purpose is to gauge the resemblance between the predicted segmentation mask and the ground truth mask of an image.

The formula for calculating the Dice score between two image masks is as follows:

$$\text{Dice score} = \frac{2 \times |Y \cap Y \text{ pred}|}{|Y| + |Y \text{ pred}|}$$

where:

- Y represents the pixels (or voxels) in the True mask.
- $Y \text{ pred}$ represents the pixels (or voxels) in the predicted mask.
- $|Y|$ refers to the total number of pixels (or voxels) in mask A.
- $|Y \text{ pred}|$ refers to the total number of pixels (or voxels) in mask B.
- $|Y \cap Y \text{ pred}|$ represents the number of pixels (or voxels) that are correctly classified as both A and B (the intersection of Y and $Y \text{ pred}$).

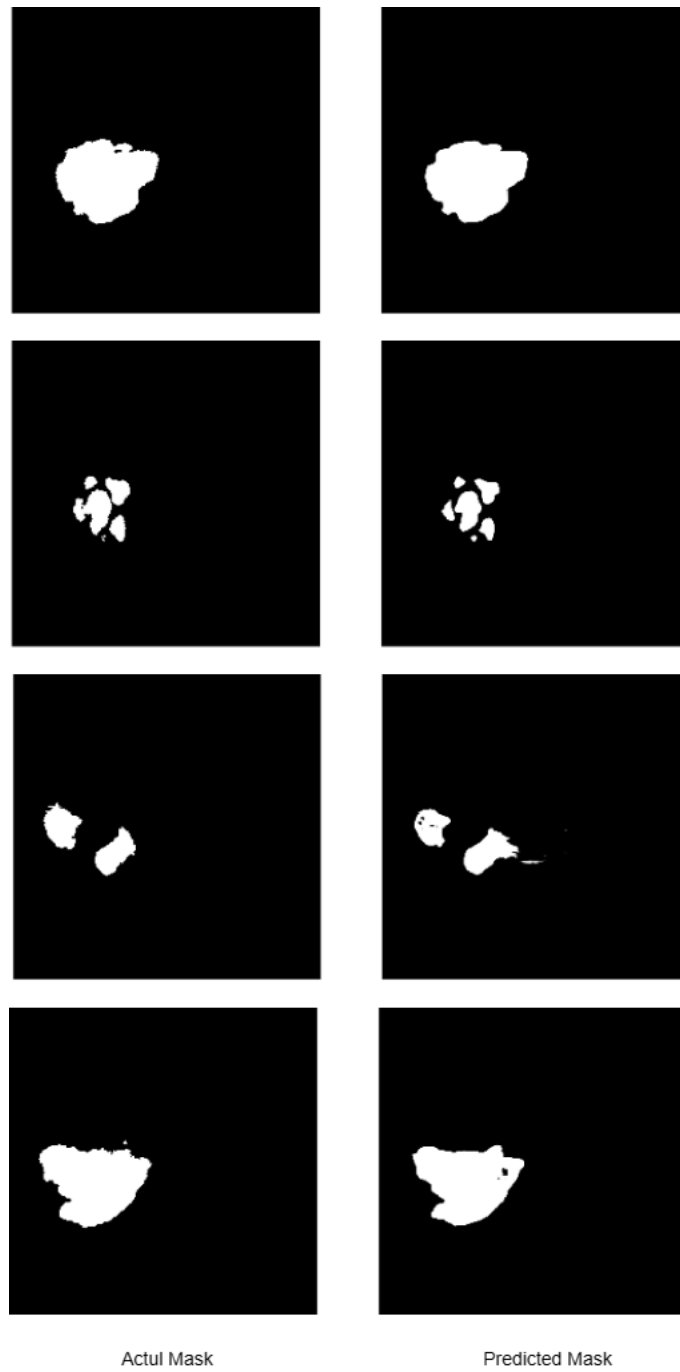


Figure 5.1: Comparison of actual mask and predicted mask

The Dice score is extensively employed in image segmentation tasks due to its ability to provide a reliable assessment of segmentation accuracy, particularly in situations involving imbalanced classes or when the foreground objects represent a small fraction of the image. It takes into account both false positives (pixels inaccurately labeled as foreground) and false negatives (pixels inaccurately labeled as background) during evaluation, making it well-suited for tasks that require precise boundary delineation.

By using the Dice score, researchers and practitioners can compare different segmentation models or techniques, optimize model parameters, and monitor the progress of their segmentation algorithms during development. It helps to quantify the performance of the segmentation task and guide improvements in the algorithm or model design.

5.2 Dice scores of all the models

The following Table 5.1 shows the average dice score of all the folds, calculated for each models with batch size = 4, meaning we've taken four slices out of ten slices for training.

Model Type	Average dice score
Fixed control(registered)	0.8111 ± 0.03
Variable control(registered)	0.8153 ± 0.02
Without control(registered)	0.8076 ± 0.02
Fixed control(Unregistered)	0.8098 ± 0.02
Variable control(Unregistered)	0.8188 ± 0.02
Without control(Unregistered)	0.8057 ± 0.02

Table 5.1: Dice score on our proposed models when trained on batch size=4

Table 5.2 shows the average dice score of all the folds,calculated for each models with batch size = 10 meaning we've taken all the 10 slices out for training.

Table 5.3 shows the average dice score of all the folds,calculated for 3 models containing unregistered and augmented dataset in which we have rotated and translated the dataset with batch size = 10 meaning we've taken all the 10 slices out for training.

Model Type	Average dice score
Fixed control(registered)	0.8063 ± 0.04
Variable control(registered)	0.8146 ± 0.03
Without control(registered)	0.8193 ± 0.03
Fixed control(Unregistered)	0.8165 ± 0.02
Variable control(Unregistered)	0.8050 ± 0.03
Without control(Unregistered)	0.8141 ± 0.03

Table 5.2: Dice score on our proposed models when trained on batch size=10

Model Type	Average dice score
Fixed control(Uregistered)	0.8223±0.02
Variable control(Unregistered)	0.7940±0.04
Without control(Unregistered)	0.8126±0.03

Table 5.3: Dice score on our proposed models when trained on batch size=10 with augmented dataset

While our hypothesis has proven to be accurate and reliable in numerous cases, it is important to acknowledge that there are instances where it may not hold true. Embracing these exceptions as valuable learning opportunities, we gain a deeper understanding of the complexities within our research. By recognizing the deviations from our hypothesis, we can refine our approach, explore alternative explanations, and uncover new insights that contribute to the advancement of knowledge in our field.

5.3 Comparison with other related work

Quantitative evaluation with the BraTS 2018 validation of our method and other research are shown in Table 5.4.

Model Type	Average dice score
Zhou et al [18]	0.75
Ensemble 3D Unet[6]	0.84
Cascaded CNN [16]	0.80
Variable control(registered) batch-size=4	0.81
Fixed control(Unregistered)	0.82

Table 5.4: Avg dice score comparison of our method and other related models on brats 2018 dataset

The crux of our research revolves around conducting a thorough comparative analysis of results through the incorporation of a control path. Furthermore, an essential aspect of our investigation involves assessing the effectiveness of the registration process on the obtained results. So we are less focused on comparing with other model's dice score. Though we have mentioned few in above table.

CHAPTER 6

Conclusion and Future works

During our study, we have thoroughly examined the effectiveness of our proposed approach. While the obtained results may not have provided the optimal solution for our specific problem statement and the initial hypothesis we formulated, we are determined to enhance our analysis by incorporating cross-data analysis techniques.

Thus far, our primary focus has revolved around evaluating the advantages of including control images and performing binary class segmentation. We have observed the impact of these factors on the segmentation process, which has helped us gain valuable insights into their contributions to the overall accuracy of our model. However, we recognize the need to expand our analysis to encompass all four class labels. By considering each class label individually, we aim to gain a more comprehensive understanding of the segmentation results and the capabilities of our approach across different tissue types.

Furthermore, we are particularly interested in exploring the potential benefits of training our model directly on 3D data. This approach has the potential to capture the spatial information and inter-slice relationships more effectively, which could potentially lead to improved segmentation accuracy. By investigating the impact of this training strategy, we aim to assess whether it results in an enhanced Dice score, which serves as a crucial measure of the segmentation quality. Through this analysis, we can determine the feasibility and benefits of leveraging 3D data directly in our approach.

In conclusion, while our current study has shed light on the functionality of our proposed approach, there is still room for improvement and further analysis. By incorporating cross-data analysis, exploring all class labels, and investigating the benefits of training on 3D data, we aim to advance our understanding of glioma

segmentation and develop a more accurate and reliable automatic segmentation method for clinical applications.

References

- [1] Brats dataset–2018. <https://www.med.upenn.edu/sbia/brats2018/data.html>.
- [2] Dipy–image registration library. https://dipy.org/documentation/1.0.0./examples_built/affine_registration_3d/.
- [3] P. Ahmad, S. Qamar, L. Shen, and A. Saeed. Context aware 3d unet for brain tumor segmentation, 2020.
- [4] M. Amian and M. Soltaninejad. Multi-resolution 3d cnn for mri brain tumor segmentation and survival prediction, 2019.
- [5] X. Feng, N. Tustison, and C. Meyer. Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features, 2018.
- [6] X. Feng, N. J. Tustison, S. H. Patel, and C. H. Meyer. Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features. *Frontiers in Computational Neuroscience*, 14, 2020.
- [7] M. Ghaffari, A. Sowmya, and R. Oliver. Automated brain tumor segmentation using multimodal brain scans: A survey based on models submitted to the brats 2012–2018 challenges. *IEEE Reviews in Biomedical Engineering*, 13:156–168, 2020.
- [8] S. Jyothi and K. Bhargavi. A survey on threshold based segmentation technique in image processing. 26. K. Bhargavi, S. Jyothi, 3, 11 2014.
- [9] K. Kamnitsas, W. Bai, E. Ferrante, S. McDonagh, M. Sinclair, N. Pawlowski, M. Rajchl, M. Lee, B. Kainz, D. Rueckert, and B. Glocker. Ensembles of multiple models and architectures for robust brain tumour segmentation, 2017.
- [10] B. Kayalibay, G. Jensen, and P. van der Smagt. Cnn-based segmentation of medical imaging data. *arXiv preprint arXiv:1701.03056*, 2017.

- [11] A. Myronenko. 3d mri brain tumor segmentation using autoencoder regularization, 2018.
- [12] J. Nalepa, M. Marcinkiewicz, and M. Kawulok. Data augmentation for brain-tumor segmentation: A review. *Frontiers in Computational Neuroscience*, 13, 12 2019.
- [13] S. Saini and K. Arora. A study analysis on the different image segmentation techniques. *International Journal of Information & Computation Technology*, 4(14):1445–1452, 2014.
- [14] Y. Tan, L. H. Schwartz, and B. Zhao. Segmentation of lung lesions on ct scans using watershed, active contours, and markov random field. *Medical physics*, 40(4):043502, 2013.
- [15] K. Venu, P. Natesan, N. Sasipriyaa, and S. Poorani. Review on brain tumor segmentation methods using convolution neural network for mri images. In *2018 International Conference on Intelligent Computing and Communication for Smart World (I2C2SW)*, pages 291–295, 2018.
- [16] G. Wang, W. Li, S. Ourselin, and T. Vercauteren. Automatic brain tumor segmentation based on cascaded convolutional neural networks with uncertainty estimation. *Frontiers in Computational Neuroscience*, 13, 2019.
- [17] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan. A deep learning model integrating fcnn and crfs for brain tumor segmentation. *Medical Image Analysis*, 43:98–111, 2018.
- [18] C. Zhou, C. Ding, X. Wang, Z. Lu, and D. Tao. One-pass multi-task networks with cross-task guided attention for brain tumor segmentation. *IEEE Transactions on Image Processing*, 29:4516–4529, 2020.