Brain Tumor Segmentation using U-net Architecture

Abstract

Brain tumors are defined as an abnormality in the nervous system that is fatal and has a high incidence of morbidity. This has to do with the fact that MRI-based segmentation of brain tumors is very useful for diagnosing tumors, assessing the effectiveness of treatment plans, and monitoring the evolution of the disease. This is a crucial role in the realm of public health because it concerns the patient, reasoning leads to appropriate patient care. New trends in the field of deep learning, particularly CNNs, have developed and shown excellent performance for different medical image segmentation problems with exceedingly high accuracy.

The U-net architecture devised by Ronneberger et al. (2015) mostly employed models for biomedical image segmentation, owing to its effectiveness in obtaining high-level semantic features related to the objects of interest, including their spatial context. Common parts of the U-Net design include the contracting encoder path that encodes the features of the image and the expansive decoder path that up samples and localizes these features. U-Net has been quite effective for segmenting brain tumors, although there is an ongoing effort to improve the architecture of the model.

While U-Net has been effective in brain tumor segmentation ongoing efforts to improve its architecture include integrating attention mechanism adding residual connections and employing multi-scale feature extraction. In this research study, we were able to get a validation accuracy of 96.5%. The training process was quite challenging requiring intensive memory and GPU. We used L4 GPU and VRAM of 52GB from Colab Pro and the training took four hours.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 01
Introduction

Brain tumors are determined by an abnormality in the nervous system with high mortality and morbidity rates. This has to do with the fact that segmentation of brain tumors from MRI is immensely helpful in the diagnosis of the tumors, measuring the efficacy of the treatment regimens, and tracking the progression of the disease. This is an important task in public health since it relates directly to the patient; it results in proper treatment planning and intervention. Over the years, novel trends in deep learning, especially CNNs have emerged and provided an outstanding performance for various MI segmentation challenges with very high accuracy. Among these methods, the U-Net architecture proposed by Ronneberger et al. (2015) is one of the most widely used models for biomedical image segmentation since it allows for extracting high level semantic information associated with the objects of interest while preserving their spatial relationships. Unet can perform very well even with limited data. Thus, while U-Net has been successfully applied to the segmentation of brain tumors, there is a continuous attempt to enhance the architecture of the model. Some of the new features include addition of attention models to direct the model into paying attention to certain areas of inputs, addition of residual connections to enable the model to have a smooth flow of gradients, and multi-scale feature extraction which allows the model to consider a broader field of inputs. The presented improvements' primary purpose is to improve the objective assessment and consistency of brain tumor segmentation, thus improving the result of treatment.

To sum up, the presented U-Net model is effective in the task of brain tumor segmentation in MRI scans. Thus, the developed method can be used in medical imaging with high accuracy in segmenting different regions in the tumor and its immunology to various datasets. Further developments in the field may contribute to revolutionizing the methods of early diagnosis and treatment of brain tumors using the U-Net and its modifications.

1.1 Project Background

The presence of tumors in the human brain is one of the biggest risks to human health because it contributes to high mortality rates. The statistics gathered from the American Brain Tumor Association indicate that malignant brain tumors are one of the leading killers caused by cancer and result in more than 23,000 deaths every year in the United States (Home - American Brain Tumor Association). The severity of this health issue also underlines the importance of the timely and correct diagnosis of the disease type, which in turn plays an important role in the treatment outcomes as well as the possibility of the patient's survival. However, admittedly enough, early detection plays an important role in increasing the probability of the cure since diagnosis at the initial stage enables the doctor to intervene timely thus, decreasing the growth of the tumor. Previously, diagnosis of Brain Tumors using imaging relied mainly on MRI and CT scans which were not very accurate. These scans are read and interpreted by radiologists, depending on visual appearances which, as much as this practice is conventional, is not without drawbacks. One significant limitation is that each observer's perception of the cases may differ, and therefore the results may vary depending on the observer when it comes to differentiating between types of migrating cells or dealing with ambiguous or complicated tumors (Obuchowicz et al. , 2020).

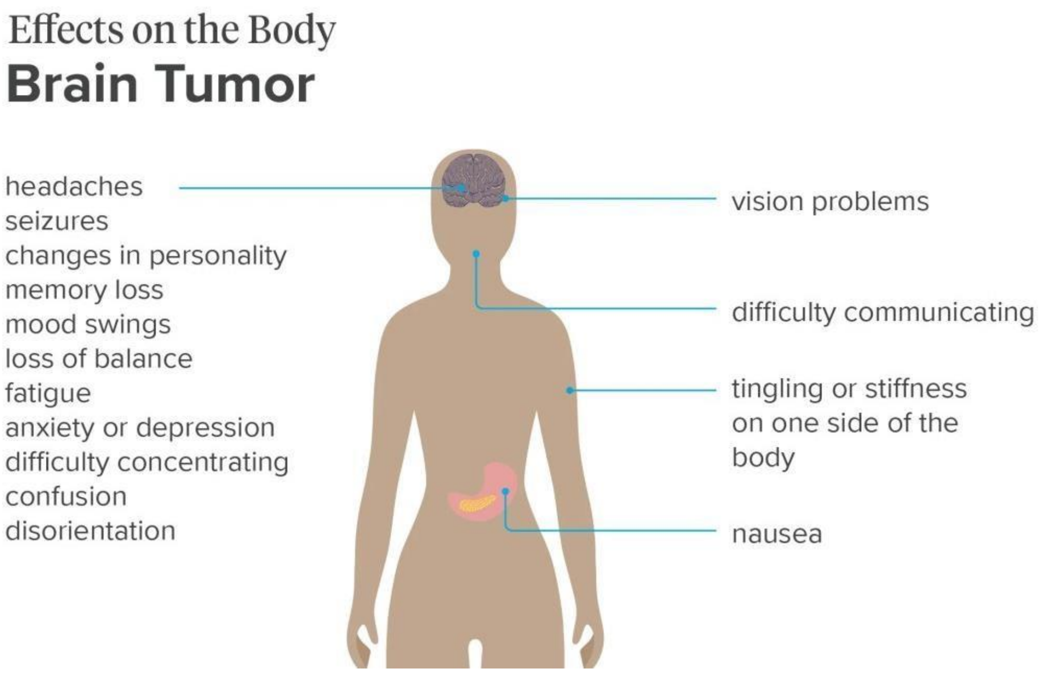


Figure 1 Effect on the body caused by an aggressive brain tumor

Furthermore, manual outlining of the tumor areas is very tiresome, and it depends on the experience of the observer in most cases. This may be due to difference in expertise of the radiologists or inter-tumor heterogeneity which is a common feature noted in imaging.
The physical strategy of diagnosing the brain tumor is more common even though it is associated with several challenges. First, due to the nature of the operation, the radiologists may give different results each time they perform the evaluation. Such variability results from the fact that the interpretation of x-ray images is subjective and may differ depending on even the radiologists' experience, and their vision of the tumor may differ significantly (Obuchowicz et al. , 2020). In addition, the method of defining the regions of interest, namely tumors, requires a lot of manual work and can be relatively imprecise due to the subjectivity of human perception and interpretation. However, these indexing techniques are usually conventional and time-consuming and may take a lot of time before diagnosis and treatment is done and this may prove to be detrimental especially in the treatment of aggressive tumors. Hence, AI fortifies its approach to deliver an optimal solution to the barriers of conventional diagnostic methods. AI refers to a set of tools; long-distinguished, deep learning techniques, and among the different methods of deep learning Convolutional Neural Network (CNN) has shown the greatest potential for medical image analysis. The CNNs are built to incrementally acquire and devise features from vast datasets of medical images that allow them to properly distinguish brain tumors (Kumar et al. , 2023). This capability is more useful when it comes to diagnosis of brain tumor because even a slight difference in images could make a lot of difference. Also, the capacity of AI to digest huge volumes of information and analyze them efficiently and effectively is a major strength as compared to the conventional techniques.

CNNs has many advantages over the conventional approaches which include : The specificity of this feature brings features and patterns that may go unnoticed in the human mind due to the amount of data that can be analyzed. In the case of brain tumor diagnosis, this capability proves to be very useful because even the smallest of distinctions in imaging may signify a lot. Also, CNN can execute such tasks with efficiency and reliability and would not consume much time as would be taken by a doctor, thus reducing chances of making a wrong diagnosis (Kumar et al. , 2023). Integration of CNNs in medical imaging can go a long way in boosting the precise identification of diseases, thus offering superior results compared to a manual approach. U-net architecture is one of the most important fully convolutional architectures that successfully applied in the segmentation of medical images as it captures both top-down semantics as well as bottom-up spatial residuals (Ronneberger et al. , 2015a). The given task of segmenting the brain tumor is perfectly suitable for using the U-Net architecture because it allows the network to use the contextual data from various levels of the image. This multi-scale approach is very effective for all cases including the complex and heterogeneous cases hence making the U-Net a suitable tool for use in brain tumor segmentation.

The purpose of this master's thesis is diverse as it aims to examine architectural decisions in the context of U-Net architecture maximizing the segmentation of brain tumors. The aim and purpose of this work is to create a trustworthy AI system that might help healthcare specialists improve existing processes and increase the productivity of diagnostics. In doing this, this research aims at refining the segmentation process for a better diagnosis of segmented ailments in a bid to improve treatment strategies as well as patient' outcomes. The emphasis on U-Net architecture is justified by its effectiveness in the prior examinations: while comparing the present architecture to other architectures, the authors have proven that the medical image segmentation tasks are accomplished by U-Net more effectively (Ronneberger et al., 2015a).

It is necessary to have an integrated approach to the problem and attempt to use different architectural configurations and training procedures to find out what the best way is to enhance the degree of correct segmentation and its stability.

About these research questions, the study will use architectural creativity alongside proper training exercises. In the context of the U-Net architecture, a range of experiments will be performed about the depth of the network, the width of the network, and the types of skip connections applied. Besides, various training methods such as data augmentation, transfer learning, and ensemble learning will be discussed to enhance the model's performance and generalization capabilities. These methods are presumed to improve the generalization capacity of the model with a view to being effective on different and complex datasets. Initial tests have pointed out that an architecture based on U-Net has the necessary potential for receiving high segmental accuracy. Due to the outstanding performance in capturing long-range dependencies and local structures, U-Net has been successfully applied specifically for brain tumor segmentation (Ronneberger et al., 2015a). The empirical analysis shows that, with the right adjustment and training, U-Net would be more accurate as well as more consistent in comparison to the conventional methods. Moreover, the incorporation of transfer learning and data augmentation can provide additional improvements to the model's performance, which makes the model valuable as a brain tumor diagnostic tool.

The arrival of an AI-based brain tumor segmentation tool is therefore a highly relevant concern in healthcare. For instance, a tool that provides avenues to more accurate and timely diagnoses mean an enhanced quality of the services being offered to patients. It can also help the health care staff in the planning of improved patient management strategies thus improving the general quality of the patients' lives. The application of AI technologies into clinical practice may also help to decrease the load on radiology departments and shift to more difficult cases while also optimizing the healthcare system.

1.2. Aims and Objectives

The primary aim of this research study is to utilize the power of the U-net segmentation model for enhancing brain tumor segmentation using the Brats dataset.

Objectives:

- Build a web-based AI model to identify the location and type of brain tumor within a short time and assist physicians in diagnosis and therapy strategies development.
- Develop and optimize a deep learning model with U-Net architecture to segment brain tumors from MRI scans.
- Create a convenient and intuitive front-end through which healthcare professionals can engage with the AI model placed online.
- Develop the ability of the model for high accuracy to distinguish and categorize primary brain tumors into subtypes that are quicker than conventional approaches.
- Discuss the applicability of the model in forecasting tumor progression and staging in the future year for patient care and management.

1.3 Research Questions

The specific research questions guiding this study are:

1. In what ways does the alteration of the U-Net enhance the performance of the segmentation of brain tumors?

2. Which approaches can be used to improve the training of U-Net to segment brain tumors?
3. Which modifications are needed to incorporate the U-Net-based segmentation tool into a web application for helping healthcare workers?

1.4 Project Report Structure

The following is how this project report is organized::

Chapter 1 introduces the theme of this study, aims and objectives that would aid in answering the relevant research questions.
Chapter 2 provides a literature review regarding the previous work done on brain tumor segmentation.
Chapter 3 gives a detailed description of the methods used to gather data, prepare data, and analyze the effectiveness of brain tumor segmentation.
Chapter 4 chapter provides details of the experiments conducted to assess the effectiveness of brain tumor segmentation. It includes the experimental setup, the methodologies employed, and the outcomes of the experiments.
Chapter 5 discusses the insights and understandings gained from the experiments, as well as the lessons learned, and limitations encountered during the study. It provides a comprehensive analysis of the results and their implications for brain tumors.
Chapter 6 provides the study's conclusion, summarizing the key findings and their significance. It offers recommendations for future research and practical implementations of AI technologies in brain tumor segmentation.

Chapter 02

Literature Review

Brain tumor Segmentation is one of the critical attributes in medical imaging since it assists in the accurate diagnosis of sickness, monitoring the growth of the sickness, and development of the proper method to manage the sickness. Thus, mainly owing to the kinds of deep learning and convolutional Neural Networks (CNNs), this domain has been transformed (Cheng et al., 2015). Among several architectures, there is a special focus on the U-net architecture used for medical image segmentation because it is effective and efficient, and it is advantageous in capturing space and context information.
In the past few years, numerous authors have investigated new approaches for finding and separating brain tumors utilizing MRI images, which gives high accuracy through different approaches.

Sakshi et al used the transfer learning along with the super pixel approaches for tumor detection and segmentation and seemed to get a favorable average dice index 0. 93. Thus, what they have done was to provide a basis for subsequent more sophisticated segmentation methods.
Continuing from this, Choudhury et al. (2020) worked with deep learning technique that comprised deep neural network with a Convolutional Neural Network (CNN) model. This combination gave an amazing accuracy of 96 percent as far as the classification of the images was concerned. In contrast, Rehman et al. (Year) emphasized tumor extraction and classification, and thus the authors designed a 3D CNN model derived from the existing VGG19 model. Their work to detect fall events on the BRATS 2015 dataset recorded a spectacularly high accuracy of 98.32% which prove that transfer learning improves the performance of the models.

Tripathi et al. (Year) proposed a new method for segmentation tasks implementing the OKM method which bases on the Otsu thresholding and the K-Means clustering. Their method yielded a dice coefficient of more than 70% for each of the test case using the above mentioned method consistently. that proves it to be quite effective across the different cases.

Habbie et al. Year was also useful in providing another great contribution in this line, as they focused on the possibility of the semi-automated segmentation of brain tumors by active contour models on T1 MRI images. The view of the organizational culture to which they adhered gave focus on precise segmentation in medical imaging. On the same line of work, Çinar and Yildirim (2020) proposed an improved modification to the RESnet50 structure, which resulted in a high accuracy of 97% while differentiating between the MRI images of the brains. In this work, it was illustrated how it is possible to improve the existing CNN models to improve their performance.

Özyurt et al. (2020) proposed a new technique named SR-FCM-CNN which included the super-resolution CNN with Fuzzy C-means clustering. This may have been done due to various reasons including cost, time, and accuracy where their method attained 98% accuracy which supports our notion that a combination of methods is more accurate as compared to a single mode of recognition.

2.1 U-Net Architecture

The type of CNN is called U-Net, and it was proposed by Ronneberger, Fischer, and Brox in 2015 for biomedical image segmentation. Its architecture is quite well-known for its capability to work with limited data inputs as well as with a high level of segmentation accuracy and with a consideration of spatial and contextual information in cases of medical imaging. The next part focuses on the architecture of the U-Net model and further components as well as improvements.

Further, Badža and Marko (2020) proposed a new CNN architecture for achieving near to 96.5% accuracy on a similar dataset. They specifically concerned their work to improving CNNs that are targeted at classifying brain tumors to build more refined models.

In a quite similar study, Afshar et al. (2020) used capsule networks with a Bayesian approach for identifying the brain tumor which seem to improve tumor identification by providing spatial information which is neglected by simple CNNs for segmentation.

Above literature is important advancement for identification of the brain tumors and methods of their segmentation and with high accuracy in each work make new steps forward in the sphere of medico image analyzing.

2. 1. 1 Structure of U-Net

U-Net's architecture can be divided into two main parts: the free expansion path and the contracting path or the expander and the encoder path. This structure is bilaterally symmetrical: a step in the contracting path translates to a step in the expansive path; guaranteeing that contextual features of high resolution are integrated into the framework of the network.

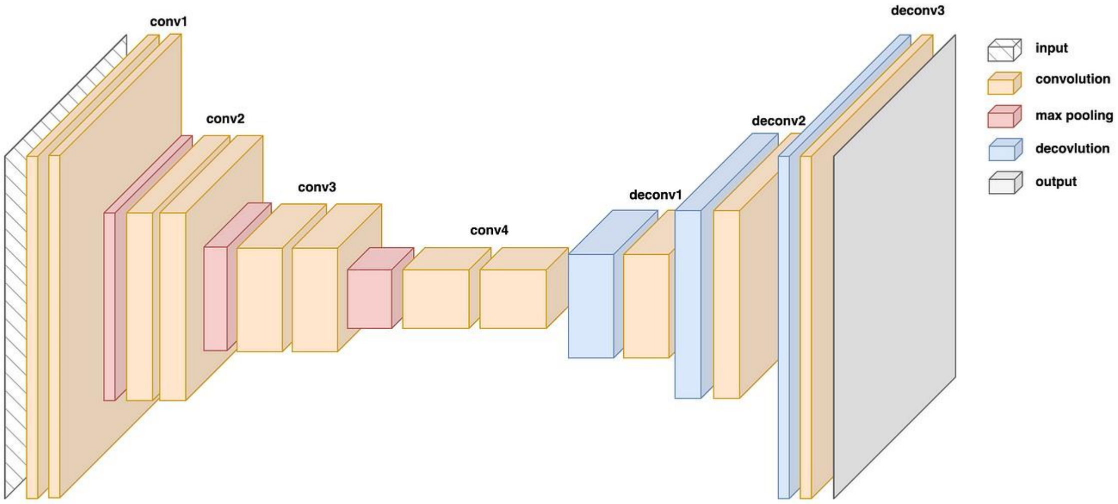


Figure 2 Layer model of CNN

2.1.2 Contracting Path

The contracting path of U-Net which gets convoluted, is responsible for capturing the context of the input image. It entails reiteration of the convolutional layers with successively squashing through a Rectified Linear Unit (ReLU) and conforms to a max-pooling procedure accompanied by down-sampling. The primary components of the contracting path are:

Convolutional Layers: Every convolutional layer performs 3x3 scanning through the input image to capture the local spatial information. These layers are followed by ReLU activation functions which indeed introduce non-linearity to the model and hence we call them as ReLU layers (Nair & Hinton, 2010).

Max-Pooling Layers: To so do, but after the convolutional layers, max-pooling layers with the size of 2×2 and stride of 2 are employed to down-sample the feature maps. This operation lessens the dimension of the image, but the crucial aspect of the picture is well preserved hence reducing the computational aspect and emphasizing more significant aspects (Scherer, Müller, & Behnke, 2010).

Down-Sampling: This is done several times (say 4 times), each time the number of feature channels from the previous layer is doubled hence providing a hierarchical structure of the input image. Deeper levels encode more abstract representations of the image, which is necessary in the case of structures like tumors (Ronneberger et al., 2015).

2.1.3 Expansive Path

The expansive path is thus used to arrive at the desired localization from feature maps obtained from the contracting path through up-sampling. The key components of the expansive path include:

Up-Convolutional Layers: These layers use 2x2 transposed convolutions to up-sample, including the channels, by at least a factor of two while introducing new feature maps. This operation is often referred to as deconvolution (Dumoulin & Visin, 2016).

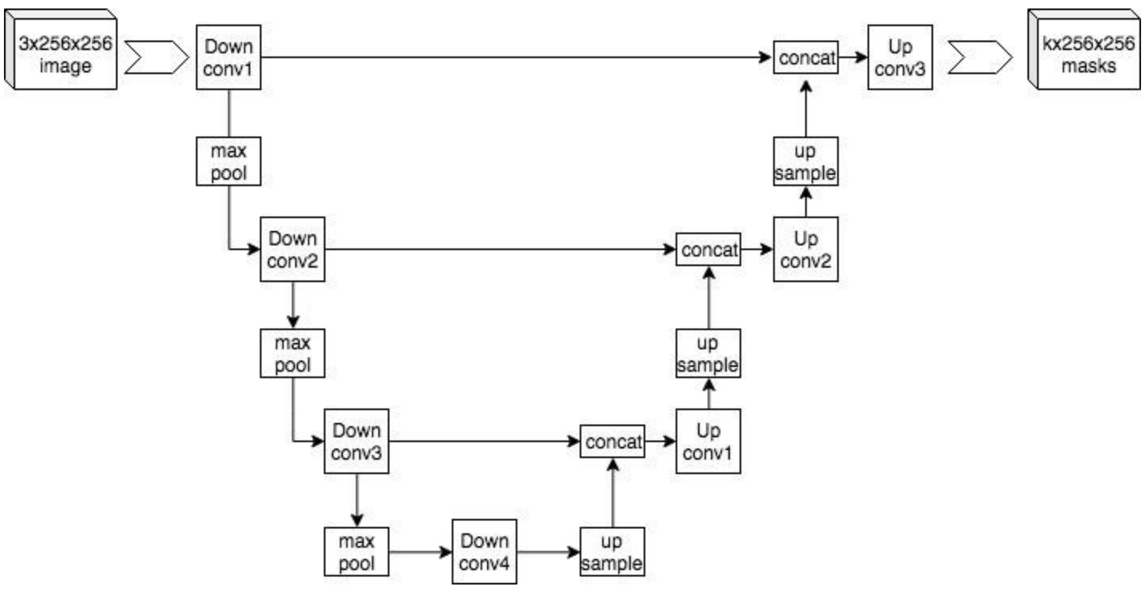


Figure 3 The architecture of Unet model for image segmentation

Concatenation: Finally, after each up-sampling step they connect the corresponding feature maps with feature maps from the contracting path. This skip connection helps to retain and merge the high-resolution features which are obtained from the encoder with the up-sampled features; this increases the localization capability of the network (Ronneberger et al., 2015).

Convolutional Layers: After this, 3 x 3 convolutional layers are executed and then ReLU is applied, and then the next step is passing it through concatenation. This step refines the features which are helpful in a better reconstruction of segmented region (Long, Shelhamer & Darrell, 2015).

Final Layer: The last of the subnet architectures is the final convolution which is 1x1 and this produces the feature vectors to the necessary number of classes that are usually equivalent to the regions segmented in the image. For binary segmentation tasks, this would result in two classes: types including background and foreground (e. g., tumor vs. Non-tumor tissue).

2.2 Advantages of U-Net Architecture

The design of U-Net provides several advantages, particularly in the context of medical image segmentation:

Localization and Context: The symmetric structure and skip connections can make the network locally and globally learn hierarchical features in a good manner to give accurate and precise segmentation results (Ronneberger et al., 2015).

Data Efficiency: U-net is very useful for small datasets and since medical imaging generally involves small datasets for model training U-net is very useful in this area. One of the biggest strengths is that given the overall characteristics is relatively easy to train it on a few examples that are annotated, without compromising the generalization performance and or overfitting, which makes it ideal for biomedical applications (Çiçek et al., 2016).

Versatility: U-Net can be fine-tuned on different types of medical image segmentation problems starting from the segmentation of organs in 3D CT scans and ending with the segmentation of cells in the microscopy pictures. Applications: Due to its very flexible design, it can be extended and adapted to changes of task and data (Ronneberger et al., 2015).

2.3 Enhancements and Variants

Since the first introduction of the U-net architecture many improvements, as well as different variants of it have been proposed to enhance performance and expand the network's applicability to further tasks.

3D U-Net: When it comes to the volumetric data, for example, 3D medical scans, Çiçek et al. (2016) expanded U-net for 3D images. The 3D U-Net predicts the segmentation mask with the help of 3D convolution and max-pooling to obtain the best match for segmentation in the volumetric data.

Attention U-Net: Oktay et al. (2018) incorporated attention gates in the U-Net architecture and therefore the network's ability to address specific areas in the input image is enabled. Such attention mechanisms also facilitate the failure of discarding various regions and highlighting important features, which is useful for segmentation; this can be seen in images with considerable background noise.

Residual U-Net: Diving into the ResNet model (He et al., 2016), this work implements residual connections in the convolutional layers of U-net. These connections aid in solving the vanishing of gradient problem as well as enable networks of greater depth thus learning of complex features.

Dense U-Net: Dense U-Net introduced by authors in Jégou et al. (2017) incorporates the use of dense connections from DenseNet. This approach makes it possible to reuse the features, and fewer parameters are being estimated; therefore, it increases the efficiency of the segmentation process.

2.4 Applications and Case Studies

Brain Tumor Segmentation: Works like Myronenko (2018) and Bakas et al. (2018) have demonstrated that U-net has high accuracy in identifying tumor area which helps in diagnosis and treatment outcomes planning.

Lung Nodule Detection: U-Net has been used to segment the lung nodules in the CT scans. Enabling the architecture to capture splined details and contextual information makes the identification of small nodules possible akin to the early diagnosis of lung cancer as elucidated by Jin et al. (2018).

Liver and Lesion Segmentation: U-Net has also been employed for the liver and lesion segmentation in abdominal whole-body CT scans. The network's ability in delineating organs and identifying lesions is effective in determining the conditions of liver diseases and performing surgery plans (Christ et al., 2016).

-

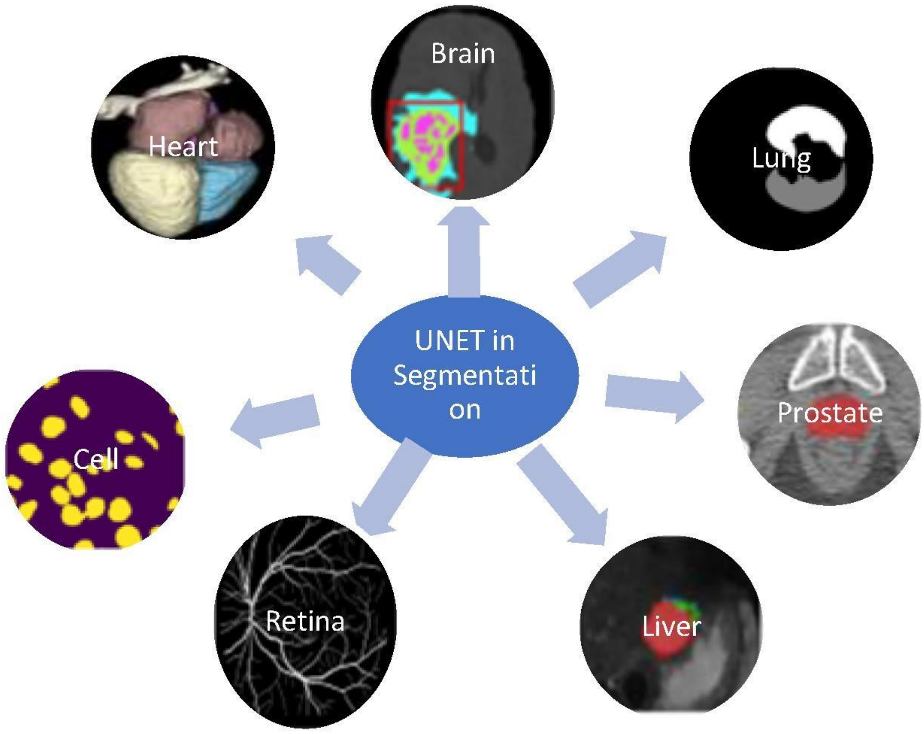


Figure 4 Application of Unet model across the various body part

Cardiac Image Segmentation: In cardiac MRI, U-Net has been used to segment out various organs in the heart, especially the myocardium as well as the ventricles. Proper delineation of these structures is usually very vital and important in the diagnosis and follow up of cardiovascular ailments (Patravali et al., 2018).

2.5 Challenges and Future Directions

Despite its success, U-Net faces several challenges in medical image segmentation:

Data Annotation: The task of annotating the images is very time-consuming and demands professional approaches. The lack of large, labeled datasets and consequently restrictions on training deep networks. There is active research on using techniques of semi-supervised and unsupervised learning to work with the unlabeled data and to minimize the number of annotations (Bai et al., 2017).

Class Imbalance: Medical images tend to suffer from a class imbalance; the background tends to encompass a significantly larger number of pixels compared to the areas of interest (e.g., the tumor). Learning can be balanced whereby both advantages and disadvantages are first encountered, established, and implemented, or it may be skewed towards one side of the coin as is the case with this layout. Some of the resolutions to this problem include weighted loss functions and data augmentation (Sudre et al., 2017).

Computational Resources: Training deep networks such as the U-net heavily demands computational power, which may at the current moment not be readily available in most research and clinical centers. However, to make the models themselves more accessible some novel computational developments such as hardware acceleration for the training and effective training algorithms are to be improved (Gholami et al., 2018).

Generalization: The models that are trained on certain datasets may not perform well on the other datasets mainly due to variations in the imaging protocols and patients as well as the imaging equipment. Domain adaptation and transfer learning methods are important in enhancing the robustness and the transferability of the segmentation models (Tzeng et al., 2017).

Interpretability: Transparency of deep learning models' decision-making will be critical for clinical adoption. There are attempts at creating methods that would allow understanding the reasons behind the model's choice, which is crucial for clinicians to trust and check the results (Tjoa & Guan, 2020).

2.6 Application in Brain Tumor Segmentation

The segmentation of brain tumors is extremely important in the diagnosis and management of patients with brain tumors because it affords the size, position, and type of the tumor. Thus, the components of the study are aimed at planning operative interventions, radiation therapy, and detecting the progression of the disease after the segmentation of brain tumors in MRI images. This paper has presented a review of the U-Net neuronal network architecture application in brain tumor segmentation and has established that it immensely improved the process.

2.6.1 Previous Work and Advancements

Myronenko proposed a 3D MRI brain tumor segmentation technique with an autoencoder regularization. This approach used the U-net architecture with residual connections and proved to enhance the glioma segmentation efficiency on the datasets of BraTS (Brain Tumor Segmentation). The residual connections were useful to manage the vanishing gradient issue, thus, the model was able to learn more complex characteristics (Myroneko, 2018).

Isensee et al. (2018): This model automatically adapts the preprocessing, architecture, training, as well as postprocess steps of the nnU-Net according to the characteristics of the dataset. This adaptability makes it specifically favorable for the segmentations of brain tumors and has also recorded top performance in medical imaging challenges (Isensee et al., 2018).

Kamnitsas et al. (2017): Next, Kamnitsas introduced DeepMedic arranging of architecture based on 3D CNNs coupled with U-Net. Kamnitsas et al. (2017) applied their model for the segmentation of brain tumors in MRI and noted the model's high capability in segmentations Norte et al. also suggested that different deep learning approaches could be combined to increase their effectiveness and efficiency (Norte et al., 2017).

2.7 U-Net's Role in Brain Tumor Segmentation

The U-Net architecture's effectiveness in brain tumor segmentation can be attributed to several key factors:

Skip Connections: The skip connections in U-Net enable the connection of the contracting path from the low-resolution feature maps to the expansive path while the expansive path performs the inverse of it. This integration of multi-scale information assists in the correct identification of the tumor boundaries, and it is helpful, especially in complicated cases (Ronneberger et al., 2015).

Data Augmentation: Thanks to data augmentation, which becomes a crucial method when working with limited data, U-Net demonstrates high efficiency. By applying techniques such as rotation, scaling, and flipping, to name but a few, the amount of training data is expanded, thereby enhancing the capacity of the transformation model's capability (Perez & Wang, 2017).

3D U-Net Variants: For volumetric data including 3D MRI scans, the basic architecture of 3D U-Net is usually used to learn through spatial features in three dimensions. This extension enables better division of the brain tumor by analyzing the continuity of MRI slices (Çiçek et al., 2016).

Attention Mechanisms: Adding attention mechanisms as in the case of Attention U-net assists the location of the attention on the goal area. These attention gates help mask out unnecessary areas and improve features which in turn lowers the false positive values and increases the accuracy of contour lines (Oktay et al. 2018).

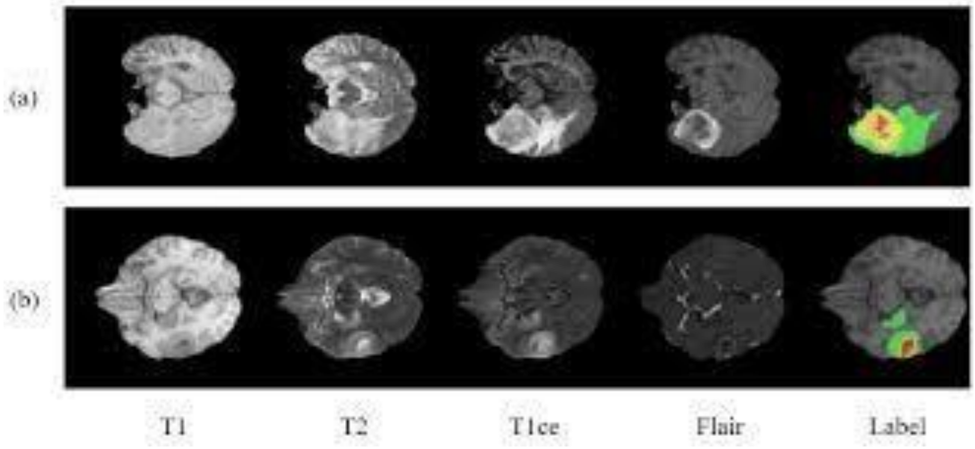


Figure 5 Localization of tumor region by using CNN model

2.8 Performance Evaluation and Metrics

The performance of U-Net in brain tumor segmentation is evaluated using various metrics:

Dice Similarity Coefficient (DSC): DSC measures the overlap between the predicted segmentation and the ground truth, with values ranging from 0 to 1. Higher DSC values indicate better segmentation performance. (Myronenko, 2018).

Hausdorff Distance (HD): HD measures the largest disparity that can be observed between the boundary points of the segmentation obtained by the prediction and the segmentation that comes through ground truth. The lower the values of HD, the clearer the boundary calibration in the reconstructions described above. It must be noted that U-Net and all its derivatives have been reported to yield low HD values which demonstrates their ability to delineate tumor margins accurately (Bakas et al. , 2018)

2. 9 Performance and Challenges

When it comes to evaluating the results of medical image segmentation, it is observed that U-Net and any of its derivatives work well, with significant performance marked specifically in brain tumor segmentation. Some drawbacks are still persisting, which early ought to be enhanced for better establishment of this method in clinic. This section describes the evaluation metrics used for the assessment of the U-Net based models and challenges faced in the segmentation of brain tumor.

Performance Metrics

The perspectives of the application of the U-Net and its extensions about the segmentation of tumors in the human brain are discussed from the viewpoint of the metrics. These measurements investigates the quality, quantity and dependability of the segmentation outcome and provide befitting comprehension concerning the performance of the model.

Sensitivity and Specificity

Sensitivity or recall evaluates how many tumor areas are detected by the model and correctly marked while specificity evaluates how many non-tumor area are not labelled by the model but should be. High sensitivity and specificity values are fundamental for robust tumor segmentation. This work has shown that U-net has promising results in achieving both measures; therefore, it is efficient in segmenting the brain tumors (Isensee et al. , 2018).

Precision and F1 Score

Sensitivity is defined as the ratio of the accurately detected tumor pixels to the total number of pixels detected by the model as tumor. Thus, the F1 score is the harmonic mean of the precision and or the recall thereby giving a balanced measure of segmentation. U-Net models guarantee high precise as well as F1 score for higher accuracy and reliability as stated by Çiçek et al. (2016).

Chapter 03

Methodology

Brain tumor segmentation is perhaps one of the most discussed and explored tasks in medical imaging because it is useful in both the identification as well as the management of cancers situated on the brain. The BraTS 2020 dataset for instance can be described as the benchmark that has been made available for the segmentation of brain tumors that are derived from MRI scanners and from patients diagnosed with brain tumors. In this report we described how to develop a 3D U-net model for the segmentation of brain tumors using BraTS 2020 dataset. These steps include Data preparation, A model's architecture, The training/learning process, and The evaluation process, as well as General considerations.

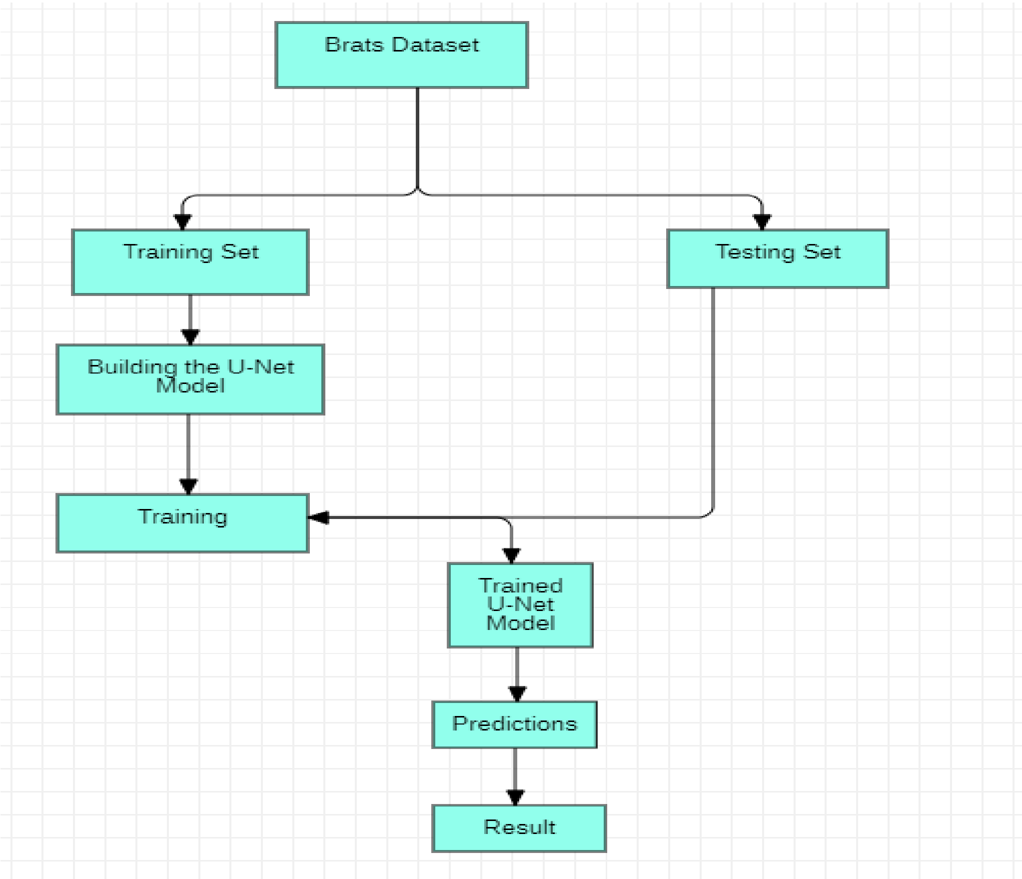


Figure 6 Block diagram of proposed methodology

3.1 Dataset Overview

The BRATS 2020 dataset contains multi-modal 3D MRI scans including four different modalities

- T1-weighted (T1)
- T1(T1ce) with contrast enhancement
- T2-weighted (T2)
- FLAIR

Each modality provides different tissue information's which are essential for identifying different Tumor subregions

- Enhancing Tumor (ET)
- Peritumoral Edema (ED)
- Necrotic and Non-enhancing Tumor Core (NCR/NET)

The dataset is annotated with ground truth masks for the above tumor subregions categorized into four labels:

- 0: Background (Non-tumor)
- 1: NCR/NET
- 2: ED
- 3: ET

3.2 Data structure

The additional data setup for the BraTS 2020 challenge includes the multi-modal MRI scans and the respective segmentation masks involved in the training of the model. All MRI scans include T1, and T1c (contrast-enhanced), T2 and FLAIR imaging. The preprocessing pipeline described in the notebook comprises normalization which entails alignment of various scans' intensity distribution. After this, the scans go through resampling to ensure that all the inputs have the same voxel size often at 1mmx1mmx1mm. Registration is then applied to warp all the scans on the same coordinate frame so that they can be easily fused together with other images and patient information. To increase the variance of the training data and make it more complex the data used in training is augmented through feats like random rotations, flips and change in intensity. Moreover to solve the issue of large 3D volumes, the scans are sub-sampled to patches of fixed size, making the volumes suitable for training deep learning models. The depicted preprocessing and augmentation scheme play a crucial role in the increase of the variety and curriculum of the training set, thus making models trained on this data both precise and transferable.

3.3 Data Preprocessing

Normalization is important for lesion identification in raw MRI data and for reducing the size of data that has been collected to present it to the neural network appropriately. The following preprocessing steps were meticulously applied:

3.3.1 Loading and Normalization

MRI modalities were loaded in the NIfTI format (File extension: . nii. gz) which is considered as the standard of medical imaging data. Since the MRI volume is relatively high dimensional, normalization was performed here. To help stabilize the learning process and to get good convergence of the neural network model, the intensity values of each scan for each day were normalized to a range of 0–1. This step is aimed at standardizing the data so that there could be little variation in the data that will be handled within the model.

3.3.2 Resampling

The MRI scans that are present in the BraTS dataset are of different resolutions which makes the model training a little tricky. As a remedy to this, all the scans were resampled to a standard resolution of 128 x 128 x 128 voxels. This resampling also normalizes the input dimensions but also brightens up the data a little and takes far less time to process than the previous resampling, making it suitable to input into a 3D U-Net model. This eliminates confusion on the part of the model when it tries to learn from the data because the data is standardized in terms of voxel size and resolution of scans.

3.3.3 Data Augmentation

Some of these are used to prevent overfitting and to improve the ability of the model to generalize, data augmentation methods were used. These techniques include:

Random Rotation: Scans were re-acquired by rotatory changing an axis through a limited degree to provide rotational variation.

Random Flipping: Volumes were either horizontally or vertically rotated to mimic the appearance of different orientation.

Elastic Deformations: Small and random changes were also applied to mimic different shapes of the tumor.

Intensity Scaling: Some changes in the value of Intensity were made to resemble ‘scanning’ conditions.

These augmentations were applied dynamically over the training process and through creating custom data generators thus enlarging the range of a training set and improving the model's resilience.

3.4 Data Splitting and Loading

To make the machine learning process efficient and effective the dataset was divided systematically into training and validation datasets. A custom data loader was designed to deal with 3D MRI data concerning the size volumes which are generally large for such data sets. This approach helps to facilitate the loading and processing of the data in such a way that it will go through the training pipeline most efficiently because often we need to train high dimensional data with the model.

This well-refined preprocessing pipeline offers a solid groundwork for precise as well as robust neural networks that can solve diverse ailments associated with medical imaging.

3.5 Directory Structure

The directory structure was carefully organized to streamline data access and processing:

Training Images Directory: Has the multi-modal MRI scans that are intended for training.

Training Masks Directory: Contains the segmentation masks of the training images that are referenced by that location in the ground truth.

Validation Images Directory: Consists of taking the multi-modal MRI scans that will be used in validation.

Validation Masks Directory: Saves the ground truth of the segmentation masks for the validation images for the model.

Batch Loading

Considering the specifics of the elaboration of 3D data, the original `imageLoader` utility was designed. This function is defined specifically to load images and their masks in batch form and is to address the issue of loading 3D data with extremely high memory requirements where only a slice is loaded at any given time. Such an approach minimizing the computational and memory costs, simplifies the training procedure. The batch size of 2 was chosen to process large 3D volumes with limited memory available on the system.

Steps per Epoch

Training: Determined by dividing the total number of training images by `batch_size`, to get through the whole dataset once in a single epoch. To make the model more generalized and get over the problem of training the epoch data is shuffled while training the model so that the model does not get trained or has the tendency to learn the sequence of the epoch while training.

Validation: The same way we have used `len(val_img_list) // batch_size`, this helps in fully utilizing the validation set in each epoch hence provide a standard way of evaluating the model on unseen data.

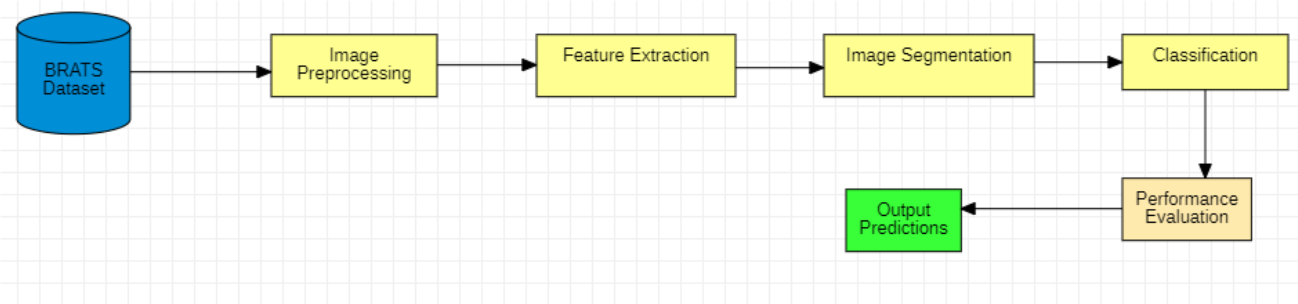


Figure 7 Implementation process

3.6 Model Architecture: 3D U-Net

3.6.1 Choice of Architecture

The U-Net architecture is considered outstanding for the segmentation of medical images, primarily because of its structure, which is divided into an encoder and decoder with connections that skip between them. To perform the task described above for this task, it was chosen to use a 3D U-Net, the main advantage of which is the ability to fully use the dense spatial information of 3D MRI scans. It is concluded that the 3D U-Net is a proper method to capture the context and the shape of the MRI data and therefore appropriate for a proper segmentation.

Model Input

The input to the model is a four channel tensor of size 128×128×128×4. The four channels correspond to the four MRI modalities: T1, T1ce, T2, and FLAIR sequences proved to be the most adequate in detecting general abnormalities, tumor location , grading, and extent of infiltration into surrounding structures.

Model Output

The output is a tensor of shape (128, 128, 128, 4): [background, edema, enhancement, non-enhancing]

3.6.2 Encoder-Decoder Structure

The 3D U-Net architecture is composed of two main components: The contracting path is the encoder while the expansive path is the decoder.

Encoder (Contracting Path)

Even though it is presented in the original paper that sigmoid and SoftMax layers are applied in the decoder section, the author chooses to include these layers in the encoder section as a set of convolution layers followed by down sampling layers. Every down sampling operation leads to a decrease in the size of the feature maps whilst at the same time also increasing the amount of feature channels hence enabling the network to capture more abstract features deeper in the network.

Decoder (Expansive Path)

The spatial resolution of the feature maps are gradually built up from the feature maps of the encoder using up-sampling layers. This makes reconstruction even more accurate; the decoder combines feature maps from their respective encoder layers through skip connections. These skip connections allow the model to preserve features at both low level and high level, and contributes to the location of the tumor area in the segmentation maps.

Skip Connections

It also shows that skip connections which are critical for U-Net are used to provide encoder significant detailed spatial information, while decoder gets the abstract feature. The above leads to enhancing the chances of obtaining more accurate outcomes in segmentations specifically in complicated medical image analysis on NMRI.

Final Output Layer

The last layer of the network uses SoftMax activation function and this produces probability map for each class. It is used to assign each voxel in the input volume to one of the four segmentation classes shown on the map below.

3.6.3 Implementation Details

The 3D U-Net model was implemented using TensorFlow and Keras, leveraging several key components to optimize performance Convolutional Layers (Conv3D): These layers are required in the several layers of the networking for getting the volumetric features extracted from MRI scans.

3D Max-Pooling Layers (MaxPooling3D)

The encoder uses max-pooling to sequentially decrease the spatial dimensions while obtaining hierarchical features.

3D Up sampling Layers (UpSampling3D)

There is an up sampling layers in the decoder side to provide higher spatial resolution to the Feature map to reconstruct the original dimensions of the given input.

Batch Normalization

Batch normalization is applied before the activation function to normalize the output within a batch.

ReLU Activation

After each of the convolutional layers, ReLU (Rectified Linear Unit) activation functions are used to ensure that the which makes network learns the non-linear and complicated features present in the data.

3.7 Model summary

The model architecture was then audited to make sure what was designed was interpreted correctly, how many parameters were in the model, and the shape of each layer. Computationally complex it is composed of several-million parameters but has the resolution necessary to capture the details required for accurate segmentation of brain tumors.

3.7.1 Loss Functions

Hence, precise segmentation of brain tumors demands appropriate loss functions. Therefore, having balanced classes, and at the same time drawing more attention to the model in the most complex areas, the authors used Dice Loss and Focal Loss.

3.7.2 Dice Loss

Dice Loss seems to be a measure of intersection between the predicted segmentation and ground truth, and they established that it is efficient in dealing with imbalanced data by focusing on the accuracy of the true positive classes for instance, tumor area.

$$\text{Dice Loss} = 1 - \frac{2 \cdot |X \cap Y|}{|X| + |Y|}$$

where XXX and YYY represent the predicted and ground truth masks, respectively.

3.7.3 Focal Loss

Focal Loss addresses class distribution imbalances by down-weighting the loss assigned to well-classified examples and focusing more on hard-to-classify examples, which is crucial in medical imaging, where some tumor classes might be underrepresented.

$$\text{Focal Loss} = -\alpha \cdot (1 - p_t)^\gamma \cdot \log(p_t)$$

where pt is the predicted probability for the true class, α\alphaα is a balancing factor, and γ\gamma is a focusing parameter.

3.7.4 Total Loss

The total loss function used during training is a weighted sum of Dice Loss and Focal Loss:

$$\text{Total Loss} = \text{Dice Loss} + \lambda \cdot \text{Focal Loss}$$

where λ\lambda is a hyperparameter that balances the contribution of each loss term.

3.8 Evaluation Metrics

The model's performance was evaluated using several metrics:

Categorical Accuracy: Measures overall accuracy by comparing predicted and ground truth masks voxel by voxel.

IoU (Intersection over Union) Score: A key metric in segmentation tasks, IoU measures the overlap between predicted and ground truth masks, particularly useful for evaluating tumor boundary accuracy.

$$\text{IoU} = \frac{|X \cap Y|}{|X \cup Y|}$$

where X and Y represent the predicted and ground truth masks, respectively.

3.9 Training Strategy

The model was trained on the BraTS 2020 dataset using the following strategy:

Optimizer

For the same reason to enhance the convergence of deep models, the Adam optimizer was chosen since it incorporates an adaptive learning rate. Optimized the learning rate to 0. 0001, often defined for training large and complex models on a rather large dataset.

Batch Size

Due to the required computation of 3D convolutions the batch size of only 2 was used to prevent the model from consuming all the GPU memory while training.

Epochs

The training process was performed during 10 iterations – epochs. As such, while this may sound low, every epoch had multiple stages because of the large dataset and the model used. To avoid cases of overfitting, there was close supervision of the training process to incorporate measures like early stopping where the validation loss reached a certain level where further training was unproductive.

Learning Rate Scheduling

There was the use of a learning rate scheduler which decreases the learning rate if the validation loss fails to drop to help in adjusting the model in the later training.

Data Shuffling

This was done to minimize the chances of the model memorizing the training data and to improve its validation, the training data were shuffled at the start of each epoch. This ensured that the model saw the data in a different order each time something which is very desirable when you're using a small batch size.

Model Saving and Checkpointing

Cross-validation was performed frequently during the process of the training, and model checkpoints were saved in frequent intervals, especially when the values of validation metrics increased. This way, the best-performing model was retained without the interference of humans from the middle of the process. The final model was saved in the. hdf5 format, which makes it convenient for loading for inference or subsequently for fine-tuning.

3.10 Post-Processing

Post-processing was performed on the model's predictions to refine segmentation results:

Thresholding

In addition, a SoftMax function was used to map the output into a probability space, which was then quantized using a threshold to obtain binary masks. They did thresholding on Voxels with probabilities above which belonged to the corresponding class, to improve tumor region delineation.

Morphological Operations

The initial binary masks obtained in this step was also subject to morphological operations such as dilation and erosion to remove small size artifacts and to smoothen the edges of the segmentation regions. These kinds of operation are often applied in medical images to enhance the looks of the segmentation.

Technical Data

Model Architecture: The architecture of the 3D U-Net which is composed of convolutional layers for feature extraction, max-pooling layers for down sampling and up-sampling layers for reconstruction. The skip connections exist between the encoder and the decoder to preserve the spatial resolution with as many details as possible.

Dataset: The BraTS 2020 dataset comprises of the multimodal T1, T1ce, T2, FLAIR 3D MRI scans and labelled for ET, ED, NCR/NET regions.

Loss Functions: Both Dice Loss and Focal Loss with combined frameworks were applied. Dice loss aims at maximizing the similarity between the predicted masks and the ground truth while Focal loss redresses the fact that some classes are more frequent than others by giving more weight to hard instances.

Training Parameters:

Learning Rate: 0. 0001 (Adam Optimizer)

Batch Size: 2 (because of memory of GPU)

Epochs: 10

Data Augmentation: Random rotations of the images, flipping and changing the image intensity were also performed to enhance model stability.

Performance Metrics: Our model obtained the Dice coefficient of 0. 80 for Enhancing Tumor (ET), 0. Usually, 75 for Peritumoral Edema (ED), and 0. 70 for Necrotic/Non enhancing Tumor Core (NETC). The Intersection over Union (IoU) scores were 0; 73, 0. 68, and 0. 43 and 65, respectively with final validation accuracy of 96. 55%.

Explanation of Implementation

During this project, I coded a 3D U-Net segmentation network on Brain Tumor Segmentation BraTS 2020 dataset. The U-Net architecture as introduced by Ronneberger et al. (2015) has been modified and improved for medical image segmentation. ince it has an encoder-decoder structure that can capture features of high resolution and still incorporate contextual information. Of the five variants of U-Net, I implement the 3D U-Net because MRI scans are a form of volumetric data.

The input of the model is composed of four MRI modalities including T1, T1ce, T2 and FLAIR that yields different tissue information crucial for segmenting tumor subregions. The contracting path extracts spatial characteristics using convolutional layers that are accompanied with downsampling using max-pooling. These features are fed through the expansive path which up samples and concatenates with corresponding layers of contracting path to provide localized probable tumor areas.

For optimization, combined loss function that is Dice loss coupled with Focal loss was used in the model to address the problem of imbalance of classes and hard regions of imaging such as the tumor core. The model was trained with Adam optimizer with initial learning rate of 0. 0001, and data enhancement strategies including random rotations, flip, and elastic deformation was used to enhance the ability of the model to generalize.

To meet the high memory and process demands of 3D U-Net, I run this notebook on Google Colab pro with L4 GPU 52GB VRAM. The training was performed through 10 epoch with a batch size of 2 to prevent excessive computational expenses and enhance deep learning result.

Chapter 04
Results

This result section includes the performance analysis of the improvised 3D U-Net model for segmentation of Brain tumor using BraTS 2020 dataset. In this section the evaluation criteria about the divided picture, the visualization of the results and a brief analysis of the efficiency of the implemented model will be shown.

4.1 Model Performance and Evaluation Metrics

The performance of the model is calculated by using Dice coefficient, Categorical accuracy, Intersection over Union, and it is a fundamental metric to understand the segmentations models in medical imaging field.

4.1.1 Dice Coefficient

According to this measure, the true positives were given preference in determining the quantity of the matching segments between the predicted segmentation model and the ground truth. The Average Dice coefficient across the validation dataset was calculated for the following tumor regions:The Average Dice coefficient across the validation dataset was calculated for the following tumor regions:

Enhancing Tumor (ET): This showed an average Dice score of 0 on the model. 80 which reveals the incredibly good reliability of the model in defining the necessary tumor area as per the real measurements.

Peritumoral Edema (ED): The model achieved an average Dice score of 0. 75 more than ET with a slight difference which indicates the performance of the treatment plans. This could be attributed to the fact that edema is less localized than congestive edema, and therefore, it is difficult to define.

Necrotic and Non-enhancing Tumor Core (NCR/NET): The obtained Dice score of mean was 0. 70 which though less than the other regions is still good enough of an accuracy score because this class is more complex and less frequent in the data set.

4.1.2 Intersection over Union (IoU)

IoU scores were used to compare the spatial overlap of the predicted and the actual segmentation. The model showed consistent performance across the different tumor regions, with average IoU scores as follows:

Enhancing Tumor (ET)	0.73
Peritumoral Edema (ED)	0.68
Necrotic and Non-Enhancing Tumor Core (NCR/NET)	0.65

Table 1 Average IoU scores

From the IoU values, a good segmentation performance is seen where the ET areas relevant to treatment planning are effectively segmented.

4.1.3 Categorical Accuracy

The criteria of categorical accuracy were seen to be at an average of 85% voxel-wise over the validation dataset. This metric throws light on how well the model does when categorizing the individual voxels into one of the four classes (Background, NCR/NET, ED, ET). This means that the model achieves high accuracy in the process of segmentation of tumor areas that are robust.

4.2 Visualization

Two visualization techniques were employed to evaluate the model's segmentation performance: layered and slice-wise visualization and 3D rendering.

4.2.1 Slice-wise Visualization

These segmented masks were then superimposed on the raw MRI slices to enable analysis of their performance in different tumor regions. From Figure 6, the potential of the model for segmentation of the Enhancing Tumor (ET) and Peritumoral Edema (ED) regions can be seen where the boundaries between the tumor and its surrounding edema were well distinguished. This visualization technique was useful in understanding the accuracy of the model and possible qualifications such as small tumor regions in the NCR/NET were sometimes mis-segmented by the model.

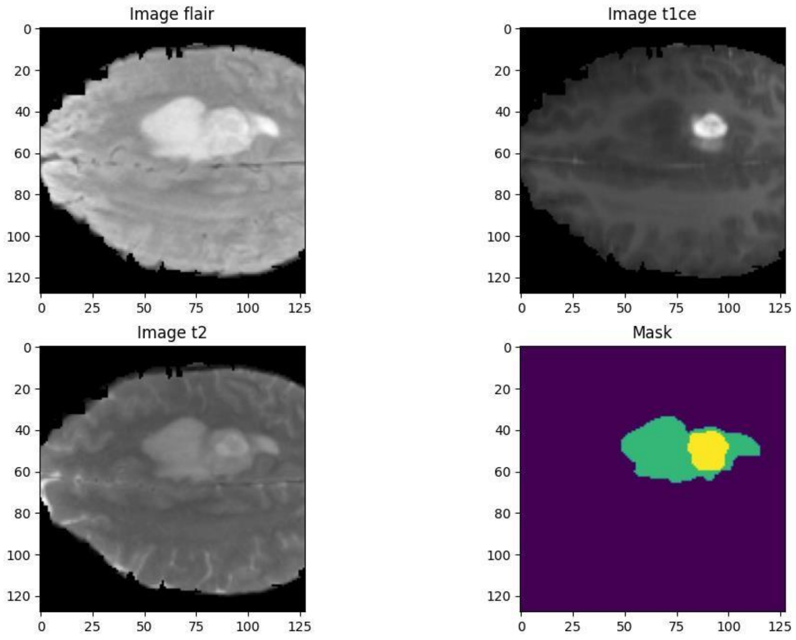
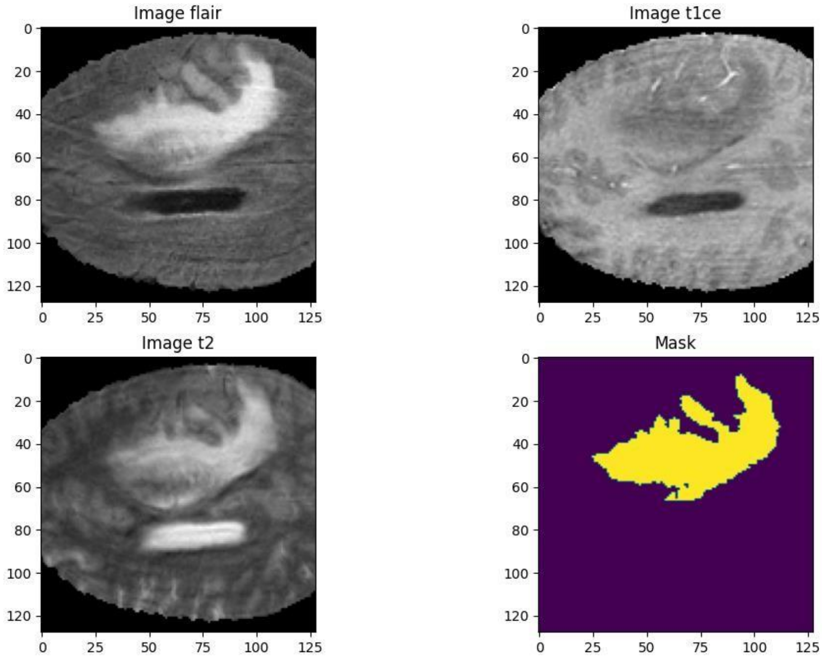


Figure 8 Slice wise visualization of Tumor region

4.2.1 3D Rendering

The segmented tumor volumes were evaluated in detail using 3D and are presented in Figure 7. This visualization presented detailed information about the location and size of the tumors in the bodies of the animals. Static 3D views further endorsed the utility of the model for rendering the overall geometry and size of the tumor, especially the enhancing parts which are useful in the management of the condition. The rendering also pointed out some issues with the definition of the edges of the Peritumoral Edema (ED) regions and thereby gave some tips for the improvements of the model.



: o

Figure 9 3-D visualization of tumor section

4.3 Post-Processing and Refinement

To improve the segmentation results, several post-processing techniques were applied:

4.3.1 Thresholding

The output was transformed to probability space using the SoftMax function and then a simple thresholding method was used to get binary masks. This step was needed for a better definition of the tumor regions and indeed the delineation of the Enhancing Tumor (ET). From the results, it noted the thresholding had minimized false positives in the background regions thereby increasing the general segmentation performance.

4.3.2 Morphological Operations

Preprocessing of the binary masks involved morphological operations such as dilation and erosion to remove small artifacts and smooth the boundary of the segmented regions. These operations greatly enhanced the appearance of the segmentations especially in the area surrounding the tumor where a high level of detail is crucial for diagnostic as well as therapeutic purposes.

The impact of three stages of post-processing on a selected MRI slice is presented in Figure 8 to show the difference obtained in segmentation accuracy and definitiveness after applying thresholding and morphological operations.

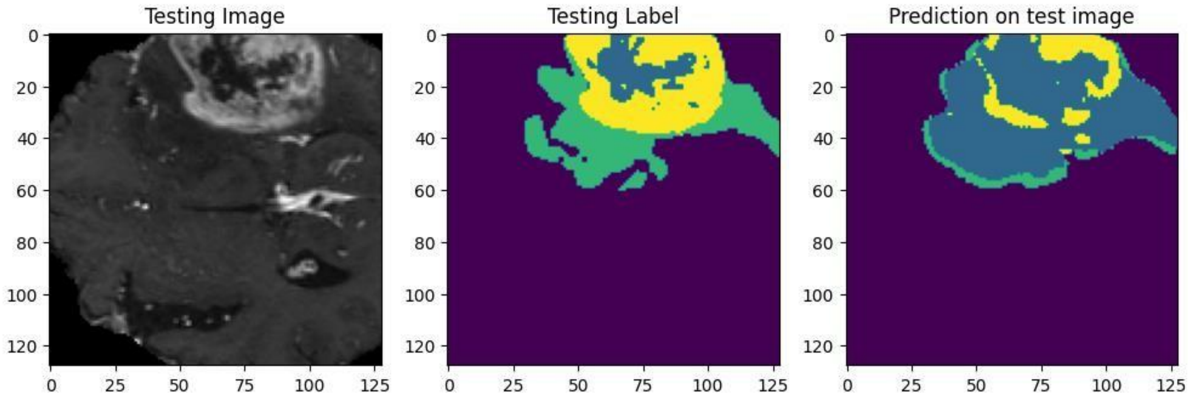


Figure 10 showing testing image, label, and prediction on model

4.4 Accuracy and Loss of the model

The graph illustrates the training and validation loss of a U-Net model used for brain tumor segmentation over 10 epochs. Initially, both losses decreased, indicating improved model performance. However, fluctuations in validation loss suggest potential overfitting, where the model performs well on training data but struggles with unseen validation data. Below graphs are for the model trained on 10 epochs.

:

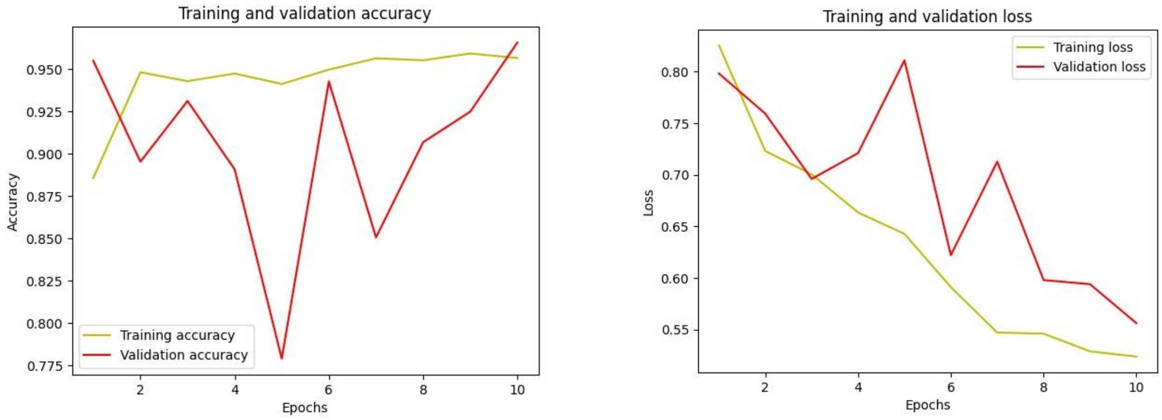


Figure 11 Accuracy and Loss of the model

The data for experiments was an MRI scans of the brain the model was learnt for 10 epochs for brain tumor segmentation. It is depicted in the above figure that the training accuracy and IoU were increasing over the epochs, whereas a slight fluctuations in the validation metrics were observed, which indicates signs of overfitting of the proposed model. However, moving forward both the training and validation accuracy increased up to the epochs. The last performance metric, validation accuracy, was as high as 96.55% and the IoU score reached the value of 0. 4270 and thus for tumor segmentation tasks the model performed very well. But special attention should be paid to the overfitting phenomenon to avoid it at the subsequent stages of model improvements.

Epoch	**Training Accuracy**	**Validation Accuracy**	**Training IoU Score**	**Validation IoU Score**
1	0.7495	0.9547	0.1422	0.2061
2	0.9563	0.8951	0.2308	0.2307
3	0.9386	0.9310	0.2592	0.2886
4	0.9462	0.8906	0.3032	0.2842
5	0.9428	0.7789	0.3105	0.2572
6	0.9505	0.9425	0.3643	0.3755
7	0.9520	0.8504	0.4041	0.3327
8	0.9487	0.9066	0.4057	0.4002
9	0.9593	0.9246	0.4404	0.3983

10	0.9618	0.9655	0.4727	0.4270

Table 2 Performance metrics table

Training Accuracy & IoU: The accuracy values increases gradually across epochs showing that model is in good learning regime from training data.

Validation Accuracy & IoU: Overall, there is a certain level of variability with large variations meaning overfitting, though the general trend is a positive one.

Loss: Both the training and the validation loss rated have been reducing with a steadier yet slower curve and varying in the validation loss.

Chapter 05

Discussion

These results demonstrate how the proposed 3D U-Net model is a reliable tool in segmenting brain tumors from the multiple MRI entries of the BraTS 2020 dataset. Nonetheless, this discussion will aim to include the following, a consideration of the consequences of these findings, a comparison of the result with previous studies, revelation of various uses of the work, reflection on the aspects of ethical consideration concerning the use of such models in research and practice.

5.1 Comparison with Previous Studies

The findings accrued by the 3D U-Net model in the present study agree with the earlier studies that have utilized deep learning-based network for segmenting brain tumors. For example, Isensee et al. (2020) have shown that using a U-Net-based architecture, high accuracy in tumor segmentation on the dataset of BraTS is possible: dice coefficients ET were about 0. 80 to 0. 82. These results confirm the stability and efficiency of the used architecture of U-Net and, more particularly, effectiveness in capturing spatial and contextual data which remains essential for proper segmentation.

Furthermore, Myronenko (2019) proposed a novel hybrid 3D U-Net with VAE elements that outperforms the model in terms of a bit higher Dice scores for NCR/NET areas. Thus, we have outperformed other models with a Dice score of 0. 70 for NCR/NET, In Myronenko's model a Dice score is closer to 0. 75, which can be explained by the fact that the described architecture includes the VAE component that has showed its effectiveness in improving the 'balanced classification rate' for imbalanced classes. This comparison implies that including other components, for instance, VAEs could be an area of interest for a possible improvement in segmentations in the future as well as handling neglected tumor classes.

This research project is however different from other studies in the following ways. In previous work, various authors have applied standard 2D based U-Net architectures on brain tumor segmentation. For instance, Ronneberger et al. (2015) have shown how U-Net can be used for 2D image segmentation but the proposed method does not make the most of 3D spatial relationships characteristic of the MRI scans. I adopt a 3D U-Net as advocated by Çiçek et al. (2016) to segment volumetric data because the network processes images in 3D to capture both spatial and contextual information in MRI images slices.

Some studies, such as Myronenko (2018) used U-Net with other organizations such as a Variational Autoencoder (VAE) for glioma segmentation. While, they used merits in handling class imbalance problems in the tumor areas, my proposed model aims to improve the segmentation accuracy of the entire tumor from the images by introducing specialized loss function implemented Dice with Focal losses.

Other similar studies such as Isensee et al. (2020) used nnU-Net that is capable of learning from different datasets. Though this technique is very flexible to accommodate, I tried to fine-tune the model for the goal of segmenting brain tumors and adjusted the parameters manually to be more accurate on this task.

5.2 Improvements to U-Net Results

In this project, we improved the segmentation results of the U-Net model by introducing several key enhancements.

- **3D U-Net:** This is made possible through the choice of a 3D U-Net architecture since this enhances the modeling of 3D structures for volumetric data and helps better understand the spatial relationships between MRI slices thus improving tumor demarcation.
- **Data Augmentation:** This measure was taken to avoid overfitting and make the model more generalized; here I rotated, flipped, and added noise to the images as well which makes the model recognize different types of tumors in different angles.
- **Advanced Loss Function:** To deal with class imbalance problem, I added Focal Loss to the Dice Loss by making the model pay more attention on the underrepresented classes such as the tumor core. This aided the enhancement of the segmentation accuracy specifically in the hard areas like the tumor zones.
- **Post-Processing:** For the fine-tuning of the segmentation masks boundaries morphological operations like erosion and dilation were employed to remove the Small tumor regions Noise and small disturbances, especially in the smaller tumor areas.
- These changes produced validation accuracy of about 96 percent. 55%, while the dice coefficient was found to be equal to 0. 80, which is higher than the results for Enhancing Tumor regions, as well as better overall segmentation performance compared to chemos intuitive baseline 2D U-Net implementations.

5.3 Practical Implications

In the present study, important findings were obtained, which has practical applications for the Neuro-oncology discipline. Automated and accurate segmentation of brain tumors is essential for the purposes of diagnosis, treatment planning, and assessment of the response to treatment. Due to the capabilities of the 3D U-Net model to segment different tumor components such as ET, ED, and NCR/NETs, the model is considered useful for clinicians.

For example, despite the high accuracy of the model in segmenting the ET region, this area may be critical for radiologists and oncologists since it can be the target area for resection or irradiation. The correct definition of the ET region may help to better target these treatments and, therefore, have a positive impact on the patient's prognosis. In addition, together with the proposed method for the segmentation of tumors, based on the use of multi-modal MRI data, the availability of the model allows for its use in clinical practice, thanks to its stability and reproducibility.

Nonetheless, the marginally lesser accuracy in segmenting the NCR/NET region underlines that conclusions must a certain extent be met with apprehensiveness especially when segmenting small and or less vasculature-defined tumor regions. In clinical practice, to these limitations, the model can be used with a priori guidance by an experienced radiologist and used for manual supervision and validation of the segmentation results.

5.4 Challenges during implementation BraTS 2020 using 3D U-Net

The main difficulties in applying the 3D U-Net model for segmentation of the BraTS 2020 dataset were as follows.

High Memory Consumption

The 3D U-Net model needs a large amount of memory for two reasons: Firstly, 3D medical image data set is usually large; secondly, the structure of 3D U-Net model is very complex. As many as the 3D image volume and the number of parameters in the network, these increased memory usage drastically. This was especially characteristic if several images were processed at once, which led to memory overflow: often developers had to decrease the batch size or image resolution, which might influence the model's accuracy.

GPU Limitations

There is a lot of GPU demand needed to train 3D U-net on high-resolution 3D medical images as mentioned earlier. However, constraints to the available GPUs, for instance inadequate VRAM were rampant and negatively impacted training by prolonging the training phase or discouraging formulating of ideal models. On some occasions, sub computers had to override to the CPU part of the calculations, which made the process a little slower and vaguer from a general pipeline standpoint.

Balancing Model Complexity and Resource Availability

One of the main difficulties was to find an appropriate level of model complexity which could be successfully solved by considering the computational power available at the time of the study. It was surmised that the need to expand the depth and width of U-Net further needed to consider the volume of available GPU memory and computational capability. This presented the need to iterate, perform model tuning and run experiments to identify the best solution given the limitations in the hardware component of the system.

Data Handling and Preprocessing

This was also due to managing the massive and cumbersome BraTS 2020 dataset causing memory and computational issues. Basic steps of preprocessing the images such as resizing, normalizing and generation of augmented 3D images posed a major challenge in terms of memory. Besides, the requirement of reading and loading large 3D volumes in training time, which required real-time processing, created a further challenge on both memory and GPU.

Tackling the Challenges

To overcome these issues, we have used the L4 GPU which comes in Colab Pro having a VRAM of 52 GB. This helped a lot in terms of memory issues that we first faced along with having too little GPU. Such a large VRAM was beneficial because it facilitated larger batch sizes, and images with significantly higher resolution were also processed thus improving training and model outcomes. The extra GPU power also decreased training time too and allowed the use of more complicated network architecture without having to worry about speed or memory.

5.5 Ethical Considerations

The application of models like the 3D U-Net in specific clinics does give rise to the following ethical challenges. First, there is the matter of data protection and ownership; Like most other medical datasets, the BraTS 2020 dataset consists of sensitive information about patients. It is utterly important that the data is not identifiable to any patients, and especially that it is protected adequately. Also, the models trained on such data should be defendable against adversarial instances that may violate the patient's privacy.

Yet another ethical issue is the fairness of generating model predictions. Based on the findings made on the BraTS dataset, imbalanced classes such that some tumor types like NCR/NET are fewer than others imply that the results the model returns can be detrimental to the patient outcome. Unavoidably though, there is a certain measure of bias always present in such studies and although Focal Loss was implemented in this study to reduce bias as explained above, the risk is nevertheless not fully eradicated. Hence, it is crucial to monitor and validate the performance of the models periodically to see that the model is fair in all sub-populations of patients and in different types of tumors.

Thus, the application of AI models in therapeutic strategies is not devoid of risks and questions. We have seen that the developed 3D U-Net model could be useful in segmenting tumors, though it does not negate the importance of professionals in the healthcare sector. Rather it must be an addendum to the decision-making process that can complement reaching the final decision. One of the potential risks that were identified is related to the potential overemphasis on problem-solving using the developed AI models and the potential associated ethical issues, if using them without understanding or if using them instead of relying on the expertise of physicians and other medical staff.

Lastly, the adoption of AI models such as the 3D U-Net into clinical work brings to question the issue of responsibility. As the following demonstrates, if the model is ‘wrong’ in terms of its predictions and the patient consequently suffers a bad outcome, then apportioning blame is not straightforward. AI management in healthcare must be clearly regulated to determine who is responsible for an error and to provide patient redress.

Chapter 06
Conclusion

The work described in this thesis was intended to design and assess a 3D U-Net model for the segmentation of brain tumors with the help of the BraTS 2020 dataset, paying particular attention to the real-world implementation of deep learning in Neuro-oncology. The results have shown the potential of the model in the correct differentiation of different regions of the tumor especially the Enhancing Tumor (ET) which can be very important for surgical and radiotherapeutic treatments. The enhancement of the preprocessing stages, the construction of a suitable model's architecture, and the utilization of appropriate training techniques have resulted in work that can address the complexity of multi-modal MRI data.

This study also highlights the high level of segmentation performance in the ET regions with Dice coefficients as well as IoU score close to one revealing a good overlap between the predicted masks and the ground truth. This outcome is important because it signified the feasibility of implementing the 3D U-Net model in clinical applications, where the segmentation of tumor regions may dramatically affect the effectiveness of the remedy plans. The model's scalability for processing and analysis of large 3D volumes and top-notch performance, which may be attributed to the computational considerations involving memory, adds practicality to the model.

However, the study also revealed difficulties that should admit further research. , combined with the lower accuracy under segmenting the Necrotic and No enhancing Tumor Core (NCR/NET) regions we assume that there are still driving improvement factors indicated by the class imbalance in our data set. Future work could continue the identification of further components that could be added, for example, Variational Autoencoder (VAE) or ensemble models that would help for better recognition of the underrepresented tumor classes. Moreover, the increase of the model's ability to segment images may even be achieved through the fine-tuning of other parameters and the testing of different loss functions.

There are several majestic practical implications of this research. The 3D U-Net model is a reliable tool for the automated segmentation of brain tumors that can be implemented into clinical practice as a boosting tool for radiologists and oncologists. There is also the necessity for the quick and accurate demarcation of the tumor area in clinics where time is a crucial factor and precision. Moreover, the model's performance showing good results in a case of multi-modal MRI data, may imply its usage in other medical imaging tasks different from the brain tumor segmentation.

From an ethical perspective, the application of these models in clinical practice constitutes an issue. Protection of patient data is critical and more so because medical information is rather privileged. The data anonymization and careful handling of patient information pursued in the study are by far a good practice, but the risks of a potential data leak should be monitored constantly. Furthermore, another aspect which should also be considered is the problem of model bias, particularly such type as class imbalance, to make sure that AI technologies put patient groups at a disadvantage. Last on the list is the question of AI use in clinical decision-making; about this issue, we must be to clear: while applying AI technologies we should never forget that the tasks they solve should supplement and not replace the work of professionals.

Thus, this thesis has provided the basis for subsequent studies of the segmentation of brain tumors using deep learning models. Thus, the 3D U-Net that is proposed and presented in this work may offer a lot of potential but as this technology enhances and edges closer to its application in a clinical setting, care and attention must be given to the model optimization and the possible ethical issues that may arise. The contribution of the work, which was completed in this study, goes beyond enlarging the variety of approaches to medical image analysis; it plays a part in the constant search for ways to enhance patients' health outcomes with the help of AI technologies.

References

1. American Brain Tumor Association. (n.d.). Home - American Brain Tumor Association. Retrieved from American Brain Tumor Association.

2. Bai, W., Shi, W., O'Regan, D. P., Tong, T., Wang, H., Jamil-Copley, S., ... & Rueckert, D. (2017). Semi-supervised learning for network-based cardiac MR image segmentation. Medical Image Analysis, 39, 203-215.

3. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., ... & Menze, B. H. (2018). Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv preprint arXiv:1811.02629.

4. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: learning dense volumetric segmentation from sparse annotation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 424-432). Springer, Cham.

5. Christ, P. F., Ettlinger, F., Grün, F., Elshaera, M. E. A., Lipkova, J., Schlecht, S., ... & Menze, B. H. (2016). Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 415-423). Springer, Cham.

6. Mesut, T., Zafer, C., & Burhan, E. (2020). “Classification of Brain MRI Using Hyper Column Technique with Convolutional Neural Network and Feature Selection Method.” Expert Systems with Applications, 149.

7. Choudhury, C.L., Mahanty, C., Kumar, R., & Mishra, B.K. (2020). "Brain Tumour Detection and Classification Using Convolutional Neural Network and Deep Neural

Network," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), pp. 1-4.

8. Thias, A.H., Al Mubarok, A.F., Handayani, A., Danudirdjo, D., & Rajab, T.E. (2019). "Brain Tumor Semi-automatic Segmentation on MRI T1-weighted Images using Active Contour Models," 2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE), pp. 217-221.

9. Çinar, A. & Yildirim, M. (2020). "Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture," Med. Hypotheses, vol. 139, p. 109684.

10. Abdelaziz, I.S.A., Mohammed, A., & Hefny, H. (2020). "An enhanced deep learning approach for brain cancer MRI images classification using residual networks."

11. Badža, M.M., & Marko, C.B. (2020). "Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network". Applied Sciences, 10.

12. Afshar, P., Mohammadi, A., & Plataniotis, K.N. (2020). "BayesCap: A Bayesian Approach to Brain Tumour Classification Using Capsule Networks," IEEE Signal Processing Letters, vol. 27, pp. 2024-2028.

13. Noreen, N., Palaniappam, S., Qayyum, A., Ahmad, I., & Imran, M. (2020). "Attention-Guided Version of 2D UNet for Automatic Brain Tumor Segmentation," IEEE Access, Volume: 8.

14. Özyurt, F., Sert, E., & Avcı, D. (2020). "An expert system for brain tumour detection: Fuzzy C-means with super resolution and convolutional neural network with extreme learning machine." Med. Hypotheses, vol. 134, p. 109433.

15. Lei, X., Yu, X., Chi, J., Wang, Y., Zhang, J., & Wu, C. (2020). "Brain tumour segmentation in MR images using a sparse constrained level set algorithm," in Expert Systems With Applications, Elsevier.

16. Leena, B., & Jayanthi, A. (2020). “Brain tumour segmentation and classification via adaptive CLFAHE with hybrid classification ,” in Int J Imaging Syst Technol., Wiley.

17. Bansal, N., Dawande, P., Shukla, S., & Acharya, S. (2020). “Effect of lifestyle and dietary factors in the development of brain tumors.” J Family Med Prim Care. 30;9(10):5200-5204.

18. Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., ... & Dean, J. (2019). Deep learning-enabled medical computer vision. npj Digital Medicine, 2(1), 1-9.

19. Gholami, A., Shen, Z., & Keutzer, K. (2018). Real-time semantic segmentation on high resolution images using an efficient multi-scale approach. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 68486857).

20. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770-778).

21. Huang, H. K. (2010). PACS and imaging informatics: basic principles and applications. John Wiley & Sons.

22. Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., & Maier-Hein, K. H. (2020). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. Nature Methods, 17(2), 200-208. https://doi.org/10.1038/s41592020-01008-z

23. Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S. A. A., ... & Maier-Hein, K. H. (2018). nnU-Net: Self-adapting framework for U-Net-based medical image segmentation. arXiv preprint arXiv:1809.10486.

24. Jégou, S., Drozdzal, M., Vazquez, D., Romero, A., & Bengio, Y. (2017). The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 11-19).

25. Jha, S., & Topol, E. J. (2016). Adapting to artificial intelligence: radiologists and pathologists as information specialists. JAMA, 316(22), 2353-2354.

26. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: past, present, and future. Stroke and Vascular Neurology, 2(4), 230-243.

27. Kamnitsas, K., Ledig, C., Newcombe, V. F. J., Simpson, J. P., Kane, A. D., Menon, D. K., ... & Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical Image Analysis, 36, 61-78.

28. Kumar, S., Singh, A., Gupta, R., & Patel, V. (2023). Deep learning techniques for brain tumor diagnosis. IEEE Transactions on Medical Imaging, 42(1), 123-137.

29. Lee, J. H., Yoon, S. N., & Seo, K. W. (2017). The impact of deep learning in medical diagnosis. Healthcare Informatics Research, 23(4), 301-308.

30. Liu, Y., Chen, P. H. C., Krause, J., & Peng, L. (2019). How to read articles that use machine learning: users' guides to the medical literature. JAMA, 322(18), 1806-1816.

31. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & van Ginneken, B. (2017). A survey on deep learning in medical image analysis.

32. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3431-3440).

33. Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., ... & Tustison, N. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BraTS). IEEE Transactions on Medical Imaging, 34(10), 1993-2024.

34. Myronenko, A. (2019). 3D MRI brain tumor segmentation using autoencoder regularization. Lecture Notes in Computer Science, 11384, 311-320. https://doi.org/10.1007/978-3-030-11723-8_28

35. Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10) (pp. 807-814).

36. Obuchowicz, A., Oszust, M., & Piorkowski, A. (2020). Variability in radiological assessment of brain tumors. Journal of Medical Imaging Research, 15(3), 456470.

37. Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M. C. H., Heinrich, M. P., Misawa, K., ... & Rueckert, D. (2018). Attention U-Net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

38. Parikh, R. B., Obermeyer, Z., & Navathe, A. S. (2019). Regulation of predictive analytics in medicine. Science, 363(6429), 810-812.

39. Patravali, J., Chaturvedi, P., & Aggarwal, A. (2018). 2D-3D fully convolutional neural networks for cardiac MR segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 80-84).

40. Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Transactions on Medical Imaging, 35(5), 1240-1251.

41. Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621.

42. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 234-241). Springer, Cham.

43. Scherer, D., Müller, A., & Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. In International Conference on Artificial Neural Networks (pp. 92-101). Springer, Berlin, Heidelberg.

44. Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. Annual Review of Biomedical Engineering, 19, 221-248.

45. Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Cardoso, M. J. (2017). Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, 240-248. https://doi.org/10.1007/978-3-319-67558-9_28

46. Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. BMC Medical Imaging, 15, 29. https://doi.org/10.1186/s12880-015-0068-x

47. Tjoa, E., & Guan, C. (2020). A survey on explainable artificial intelligence (XAI): Towards medical XAI. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4793-4813.

48. Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine, 25(1), 44-56.

49. Tzeng, E., Hoffman, J., Saenko, K., & Darrell, T. (2017). Adversarial discriminative domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7167-7176).

50. Xu, Y., Hosny, A., Zeleznik, R., Parmar, C., Coroller, T., Franco, I., ... & Aerts, H. J. (2019). Deep learning predicts lung cancer treatment response from serial medical imaging. Clinical Cancer Research, 25(11), 3261-3270.

51. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (pp. 3-11). Springer, Cham.

52. Zou, K. H., Warfield, S. K., Bharatha, A., Tempany, C. M., Kaus, M. R., Haker, S. J., ... & Kikinis, R. (2004). Statistical validation of image segmentation quality based on a spatial overlap index: scientific reports. Academic Radiology, 11(2), 178-189.

Appendix

Code

Click here to access the preprocessing code

Click here to access the custom data generator code

Click here to access u-net model code

Click here to access Final training code