Tumor Segmentation from Brain MR images using U-Net & ResNet : A Comparative Study

Thesis submitted in partial fulfillment of requirements
For the degree of
Master of Technology in Computer Technology
of
Computer Science and Engineering Department
of
Jadavpur University

by

Swapnadip Moni
Regn. No.- 154172 of 2020 - 2021
Exam Roll No. - M6TCT23009

under the supervision of

Prof. (Dr.) Jamuna Kanta Sing

Department of Computer Science and Engineering
Jadavpur University
Kolkata, West Bengal, India
2023

Certificate from the Supervisor

This is to certify that the work embodied in this thesis entitled **“Tumor Segmentation from Brain MR images using U-Net & ResNet : A Comparative Study”** has been satisfactorily completed by **Swapnadip Moni** (Registration Number 154172 of $2020 - 2021$; Class Roll No. 002010504006; Examination Roll No. $M6TCT23009$). It is a bonafide piece of work carried out under my supervision and guidance at Jadavpur University, Kolkata for partial fulfilment of the requirements for the awarding of the **Master of Technology in Computer Technology** degree of the Department of Computer Science and Engineering, Faculty of Engineering and Technology, Jadavpur University, during the academic year $2020 - 23$.

Prof. Jamuna Kanta Sing,
Professor,
Department of Computer Science and Engineering,
Jadavpur University.
(Supervisor)

Forwarded By:

Prof. Nandini Mukherjee,
Head,
Department of Computer Science and Engineering,
Jadavpur University.

Prof. Ardhendu Ghoshal,
DEAN,
Faculty of Engineering & Technology,
Jadavpur University.

Department of Computer Science and Engineering
Faculty of Engineering And Technology
Jadavpur University, Kolkata - 700 032

Certificate of Approval

This is to certify that the thesis entitled **"Tumor Segmentation from Brain MR images using U-Net & ResNet : A Comparative Study"** is a bona-fide record of work carried out by **Swapnadip Moni** (Registration Number 154172 of 2020 – 2021; Class Roll No. 002010504006; Examination Roll No. $M6TCT23009$) in partial fulfilment of the requirements for the award of the degree of **Master of Technology in Computer Technology** in the **Department of Computer Science and Engineering, Jadavpur University**, during the period of August 2022 to August 2023. It is understood that by this approval, the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but approve the thesis only for the purpose of which it has been submitted.

Examiners:

(Signature of The Examiner)

(Signature of The Supervisor)

Department of Computer Science and Engineering
Faculty of Engineering And Technology
Jadavpur University, Kolkata - 700 032

Declaration of Originality and Compliance of Academic Ethics

I hereby declare that the thesis entitled **"Tumor Segmentation from Brain MR images using U-Net & ResNet : A Comparative Study"** contains literature survey and original research work by the undersigned candidate, as a part of his degree of **Master of Technology in Computer Technology** in the **Department of Computer Science and Engineering, Jadavpur University**. All information have been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name: Swapnadip Moni

Examination Roll No.: M6TCT23009

Registration No.: 154172 of 2020 – 2021

Thesis Title: Tumor Segmentation from Brain MR images using U-Net & ResNet : A Comparative Study

---

Signature of the Candidate:

ACKNOWLEDGEMENT

I am pleased to express my gratitude and regards towards my Thesis Guide **Dr. Jamuna Kanta Sing, Professor**, Department of Computer Science and Engineering, Jadavpur University, without whose valuable guidance, inspiration and attention towards me, pursuing my project would have been impossible.

I am grateful towards all the members of Alpha Lab, Department of Computer Science and Engineering, Jadavpur University, for cooperating with me in all possible ways and providing me with a good working environment throughout the duration of the work. I express my regards towards parents for bearing with me and for being a source of constant motivation during the entire term of the work.

I would also like to thank my friends Debam Saha, Umesh Pal, Anupam Barat, Sayan Porey, Ranadeep Bhattacharya, Niladri Sekhar Sarkar, Arpita Dafadar, Bipul Dhali, Amartya Roy, Neelotpal Chakrabarty, Sourajit Maity, Saba Nadim for constantly supporting and guiding me through the research work.

Swapnadip Moni
MTCT Final Year
Exam Roll No. - M6TCT23009
Regn. No. - 154172 of 2020 – 2021
Department of Computer Science and Engineering,
Jadavpur University.

Abstract

Brain tumors are abnormal growth of cells that need to be detected early for treatment. Magnetic Resonance Imaging (MRI) is a routinely utilized procedure to take brain tumor images. Manual segmentation of tumors is a crucial task and laborious. There is a need for an automated system for segmentation and classification for tumor surgery and medical treatments. This work suggests an efficient brain tumor segmentation and classification based on deep learning techniques. In this thesis work, we proposed a comparative analysis between two brain tumor MRI image segmentation methods U-Net and ResNet. We trained both U-Net and ResNet models using BraTs data and compared the outputs to find out which model works best and gives better accuracy, less loss. In conclusion, this thesis aims to explore the potential of U-Net and ResNet for brain tumor segmentation in medical imaging data. By investigating the effectiveness of U-Net and ResNet architecture, training and evaluating the models, and comparing it with each other, this research seeks to contribute to the field of neuroimaging and facilitate accurate and efficient brain tumor diagnosis and treatment planning. The Ultimate goal is to provide doctors with a reliable tool for improved patient care and management.

Keyword: Brain Tumor, Segmentation, ResNet50, U-Net.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Brain tumor is one of the most challenging and difficult diseases for people in the world. Early diagnosis and accurate diagnosis play an important role in determining appropriate treatment strategies and improving patient outcomes. Non-invasive evaluation of brain tumors has become easier and more efficient with advances in medical imaging technology, particularly magnetic resonance imaging (MRI). However, manual segmentation of brain tumors from MRI scans is still a laborious and time-consuming task for radiologists and neurologists. The complexity and heterogeneity of brain tumors, the presence of anatomical structures, and variability in the imaging protocol poses serious challenges for successful tumor identification. Therefore, there is an increasing need for reliable and efficient automated methods for brain segmentation. Brain segmentation plays an important role in clinical image analysis and is essential for accurate diagnosis, treatment planning and monitoring of brain tumor patients. Segmentation is the process of separating and separating tumors from normal brain tissue in medical imaging such as magnetic resonance imaging (MRI) scans. Dealing with segmentation is a laborious and time-consuming task that requires the expertise of radiologists. Therefore, it is important to develop effective and efficient segmentation Methods. In recent years, deep learning techniques have emerged as powerful tools for medical image analysis, offering the potential for automated and accurate tumor segmentation. Accurate and efficient segmentation of brain tumors from clinical data is an important task in neuroimaging research and medicine. Segmentation allows precise localization of the tumor, aiding in diagnosis, treatment planning and monitoring of brain cancer patients. In recent years, deep learning models have emerged as powerful tools for analyzing medical images, providing the ability and precision to study tumors. Among these models, ResNet (Residual Neural Network) has been found to be effective in many computer tasks. This article aims to explore the effectiveness of ResNet in classifying brain tumors and explore its effectiveness in this particular application.

1.1 Image Segmentation

Image segmentation is a method of separating a digital image into groups called image segments; this reduces the complexity of the image and allows each part to be processed or analyzed. Technically, segmentation is the task of tagging pixels to identify objects, people, or other important features in an image. Image segmentation is an essential part of computer vision techniques and algorithms. It is used in many applications, including medical image analysis, computer vision for autonomous vehicles, face recognition and detection, video surveillance, and satellite imagery analysis. Image segmentation is a widely used technique in digital image processing and analysis. Image segmentation may involve separating the foreground from the background or grouping areas of pixels by similarity in color or shape. For example, one application of image segmentation in medical imaging is to identify and label pixels in images or 3D volumetric voxels that represent tumors in a patient's brain or other organs.

1.2 Brain Tumor Image Segmentation

In brain MRI analysis, segmentation images are often used to measure and visualize brain anatomy, identify brain changes, identify pathological areas, and for surgical planning and Image guided interventions. In the last few years, many segmentation methods with varying levels of accuracy and complexity have been developed. To get an MRI image of a patient they are inserted into a tunnel with a magnetic field inside. This causes all protons in the body to 'align' themselves so their quantum spin is the same. A pulse of the oscillating magnetic field is then used to disrupt this alignment. When the protons return to equilibrium they send out an electromagnetic wave. Based on fat content, chemical composition, and importantly type of stimulation (i.e. sequences) used to disrupt the protons, different images will be obtained. Four common sequences that are obtained are T1, T1 with contrast (T1C), T2, and FLAIR.

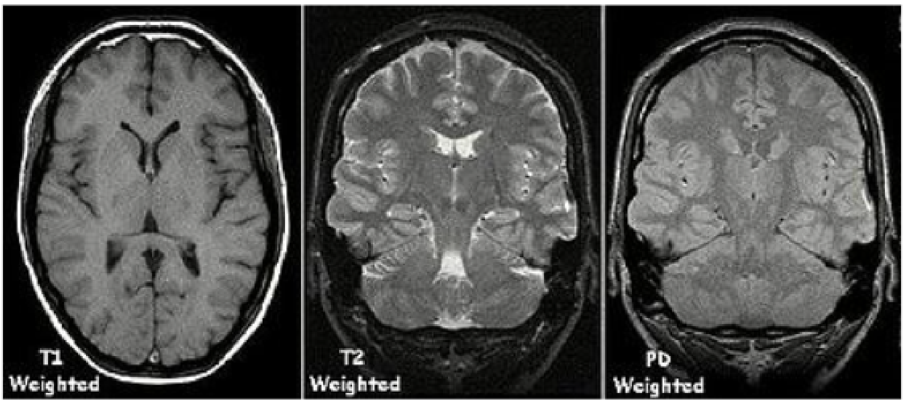


Figure 1.1: Brain MRI Scans

1.3 Brain MRI Image Segmentation Techniques

Brain segmentation is the process of identifying brain tumors and automatically labeling them according to tumor type. Mastering tumor segmentation from brain MRI is time-consuming and error-prone. Rapid and accurate Brain MRI image segmentation techniques are needed. Rapid and accurate brain segmentation techniques are needed. Convolutional neural networks (CNNs) have recently shown promising results in computer vision for image segmentation and classification tasks. U-Net.[1], SegNet.[2], ResNet.[3], VGG19.[4], P-Net.[5] are the most popular CNNs for image Segmentation. Brain MRI image segmentation is a vital task in medical image analysis, as it plays a crucial role in diagnosing and treating various neurological disorders. Deep learning models have proven to be highly effective for this task, and several different architectures have been explored. Here are some commonly used deep learning models for brain MRI image segmentation:

1. **U-Net:** U-Net is a popular architecture widely used for medical image segmentation. It consists of an encoder-decoder structure with skip connections. The encoder captures the context and the decoder produces the segmentation map. U-Net has been successfully applied to various brain segmentation tasks, such as whole brain segmentation, tumor segmentation, and lesion segmentation.

2. **DeepLab:** DeepLab is a deep convolutional neural network (CNN) architecture that utilizes atrous convolutions to capture multi-scale contextual information. It employs a fully convolutional network with a final upsampling layer to generate the segmentation map. DeepLab has been used for brain tumor segmentation and other brain structure segmentations.
3. **3D U-Net:** While U-Net operates on 2D images, 3D U-Net extends the architecture to handle 3D volumetric data, such as brain MRI volumes. It incorporates 3D convolutions and pooling operations to capture spatial information along all three dimensions. 3D U-Net has shown promising results for brain tumor segmentation and brain structure segmentation in volumetric MRI scans.
4. **V-Net:** V-Net is another 3D CNN architecture specifically designed for medical image segmentation. It utilizes a 3D encoder-decoder structure with skip connections, similar to U-Net. V-Net incorporates residual learning and dense connections to improve the learning capability and capture fine details in the segmentation task.
5. **Attention U-Net:** Attention mechanisms have been introduced to enhance the performance of segmentation models. Attention U-Net utilizes self-attention modules to selectively focus on relevant image regions during the encoding and decoding stages. This allows the model to assign different weights to different image features and improve segmentation accuracy.
6. **Fully Convolutional Networks (FCN):** FCN is an early deep learning architecture for semantic segmentation. It replaces the fully connected layers in traditional CNNs with convolutional layers to allow input images of arbitrary sizes. FCN has been adapted and applied to brain MRI image segmentation tasks.These are just a few examples of deep learning models commonly used for brain MRI image segmentation. Researchers continue to develop and explore new architectures and variations to improve the accuracy and efficiency of segmentation tasks in medical imaging.

Deep learning models have proven to be highly effective for this task, and several different architectures have been explored.

Chapter 2

Literature Survey

Ramya D et al.[6], 2023 proposed an improved U-Net architecture for use in the BraTS20 and BraTS21 challenge's brain tumor segmentation problem. The accuracy was improved by modifying the loss function. Comprehensive ablation research to investigate Deep Supervision loss, Cross-Entropy, Decoder Attention, and Residual Connections to determine the best model architecture and learning schedule is performed. Multiple convolutional channels have been experimented with, and post-processing techniques to find the ideal spot for the U-Net encoder's depth have also been undertaken. The proposed technique outperforms every U-Net variant and produces superior outcomes while incurring a minimal loss.

Wenbin Wu et al.[7], 2023 proposed a devised novel Inner Cascaded U-Net and Inner Cascaded U2-Net as improvements to plain cascaded U-Net for medical image segmentation. The proposed Inner Cascaded U-Net adds inner nested connections between two U-Nets to share more contextual information. To further boost segmentation performance, they proposed an Inner Cascaded U2-Net, which applies residual U-block to capture more global contextual information from different scales.

Jian Wu et al.[8], 2022 proposed that this work introduces and validates a deep-learning-based fitting method, which can rapidly provide accurate and robust estimation of cytological features of brain tumor based on the IMPULSED (imaging microstructural parameters using limited spectrally edited diffusion) model fitting with diffusion-weighted MRI data.

Qing Xu et al.[9], 2023 invented a new model DCSAU-Net that is a deeper and more compact split-attention U-Net for medical image segmentation. Highlights of this model is Low-level features are collected using optimized depth-wise separable convolution.A lightweight multi-scale split attention block is used for deep feature extraction.Notable performance improvements on complex images are achieved with a compact model.

Bumshik Lee et al.[10], 2023 proposed a patch-wise U-Net architecture for the automatic segmentation of brain structures in structural MRI. In the proposed brain segmentation method, the non-overlapping patch-wise U-Net is used to overcome the drawbacks of conventional U-Net with more retention of local information. In their proposed method, the slices from an MRI scan are divided into non-overlapping patches that are fed into the U-Net model along with their corresponding patches of ground truth so as to train the network.

Akash Saxena et al.[11], 2023 did a comprehensive study on numerical simulation and analysis of deep learning-based semantic segmentation for complex background images using an improved Resnet 50 network. The objective of this research was to enhance the accuracy and computational efficiency of semantic segmentation models, particularly in challenging scenarios with complex backgrounds. The proposed methodology incorporated several modifications to the Resnet 50 architecture, including skip connections, residual attention modules, and spatial pyramid pooling, to improve its ability to capture fine-grained details, focus on salient regions, and incorporate multi-scale contextual information.

Aggarwal et al.[1], 2023 did research work that provides an efficient method for brain Tumor segmentation based on the Improved Residual Network (ResNet). Existing ResNet can be improved by maintaining the details of all the available connection links or by improving projection shortcuts. These details are fed to later phases, due to which improved ResNet achieves higher precision and can speed up the learning process.

Menze et al.[12], 2015 In this paper reported the set-up and results of the Multimodal Brain Tumor Image Segmentation Benchmark (BRATS) organized in conjunction with the MICCAI 2012 and 2013 conferences. Twenty state-of-the-art tumor segmentation algorithms were applied to a set of 65 multi-contrast MR scans of low- and high-grade glioma patients—manually annotated by up to four raters—and to 65 comparable scans generated using tumor image simulation software. Quantitative evaluations revealed considerable disagreement between the human raters in segmenting various tumor sub-regions (Dice scores in the range 74%–85%), illustrating the difficulty of this task. They found that different algorithms worked best for different sub-regions (reaching performance comparable to human inter-rater variability), but that no single algorithm ranked in the top for all sub-regions simultaneously. Fusing several good algorithms using a hierarchical majority vote yielded segmentations that consistently ranked above all individual algorithms, indicating remaining opportunities for further methodological improvements.

B.K.Tripathy et al.[13], 2022 analyzed various semantic segmentation network architectures such as U-Net, ResNet, and PSPNet used for the purpose of tumor segmentation from 3-D MRIs. We compare and analyze various parameters (such as accuracy, loss, etc.) of each of these network architectures.

Parasa Rishi Kumar et al.[14], 2023 worked on Multi-class Brain Tumor Classification and Segmentation using Hybrid Deep Learning Network Model Based on convolutional neural network (CNN), a Hybrid Deep Learning Network (HDLN) model is proposed in this research for classifying multiple types of brain tumors including glioma, meningioma, and pituitary tumors. The Mask RCNN is used for brain tumor classification. They used a squeeze-and-excitation residual network (SE-ResNet) for brain tumor segmentation, which is a residual network (ResNet) with a squeeze-and-excitation block. A publicly available research dataset is used for testing the proposed model for experiment analysis and it obtained an overall accuracy of 98.53%, 98.64% sensitivity and 98.91% specificity.

All these methodologies, proposed by researchers including the U Net and ResNet50 model for Brain MRI image segmentation were studied and then we worked on both models. We trained both models on BraTs Dataset. And then performance of both the models were analyzed and compared.

Chapter 3

Proposed Methodology

In our suggested approach firstly we tried the U-Net model for brain MRI image segmentation. Then we used ResNet50 as a backbone and on top of that we used the U-Net model. We replaced the encoder part of the U-Net architecture with a pre-trained ResNet50 model. Firstly loaded the ResNet50 model with pre-trained weights without the top classification layer. Connected the output of the ResNet50 model to the decoder part of the U-Net. Free-zed the weights of the ResNet50 layers to prevent them from being updated during training. Then we trained the model and validated its performance. Comparing quantitative metrics such as Accuracy, Precision, Loss, Dice-coefficient helps analyze the strength and weakness of each approach.

3.1 Proposed Methodology of using U-Net and then combining U-Net and ResNet50 architectures

3.1.1 Image Data Descriptions

All BraTS multimodal scans are available as NIfTI files (.nii.gz) -¿ commonly used medical imaging format to store brain imaging data obtained using MRI and describe different MRI settings:

- **T1:** T1-weighted, native image, sagittal or axial 2D acquisitions, with 1–6 mm slice thickness.
- **T1c:** T1-weighted, contrast-enhanced (Gadolinium) image, with 3D acquisition and 1mm isotropic voxel size for most patients.
- **T2:**T2-weighted image, axial 2D acquisition, with 2–6 mm slice thickness.
- **FLAIR:** T2-weighted FLAIR image, axial, coronal, or sagittal 2D acquisitions, 2–6 mm slice thickness.

Data were acquired with different clinical protocols and various scanners from multiple (n=19) institutions. All the imaging datasets have been segmented manually, by one to four raters, following the same annotation protocol, and their annotations were approved by experienced neuro-radiologists. Annotations comprise the GD-enhancing tumor (ET — label 4), the peritumoral edema (ED — label 2), and the necrotic and non-enhancing tumor core (NCR/NET — label 1), as described both in the BraTS 2012-2013 TMI paper and in the latest BraTS summarizing paper. The provided data are distributed after their pre-processing, i.e., co-registered to the same anatomical template, interpolated to the same resolution (1 mm$\hat{3}$) and skull-stripped.

3.1.2 U-Net Architecture

U-Net, evolved from the traditional convolutional neural network, was first designed and applied in 2015 to process biomedical images. As a general convolutional neural network focuses its task on image classification, where input is an image and output is one label, but in biomedical cases, it requires us not only to distinguish whether there is a disease, but also to localize the area of abnormality.

In Figure 3.1, U-Net is an architecture for semantic segmentation. It consists of a contracting path and an expansive path. The contracting path follows the typical architecture of a convolutional network. It consists of the repeated application of two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for down sampling. At each down sampling step we double the number of feature channels. Every step in the expansive path consists of an up-sampling of the feature map followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In total the network has 23 convolutional layers.

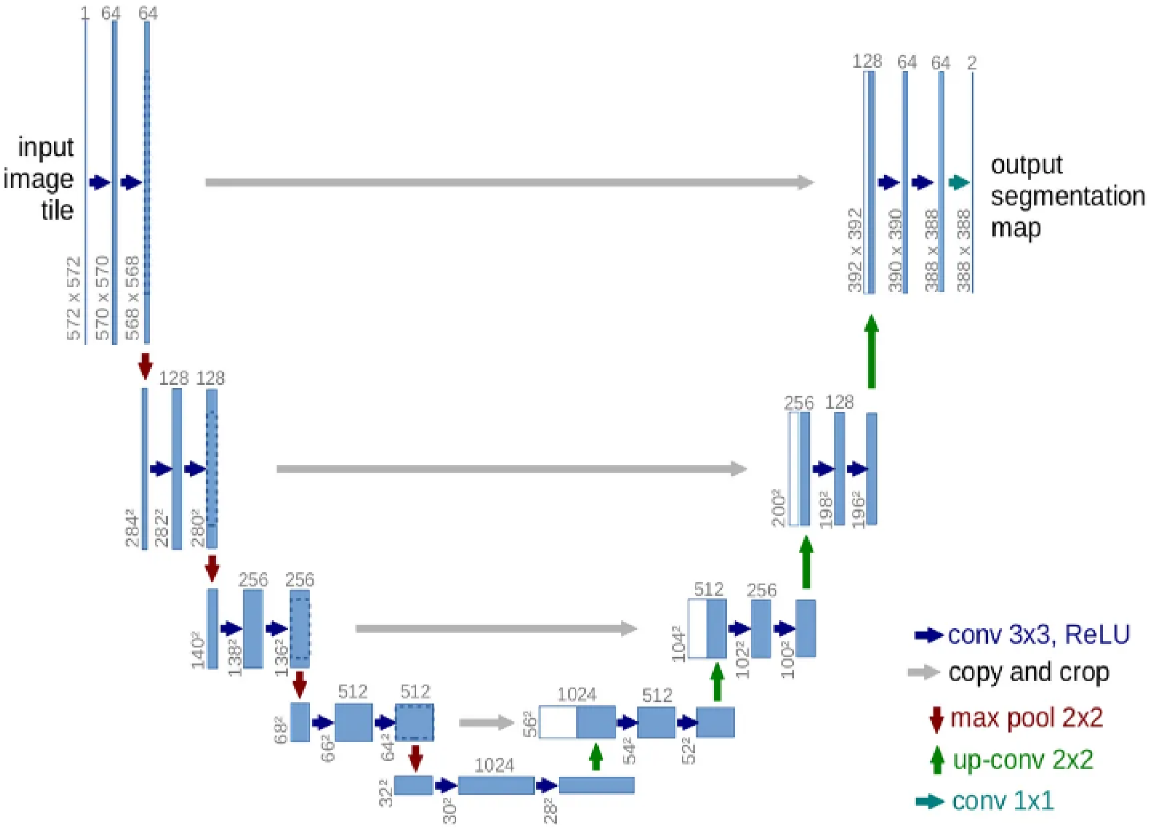


Figure 3.1: U-Net Architecture

Overview of U-Net: First sight, it has a "U" shape. The architecture is symmetric and consists of two major parts — the left part is called contracting path, which is constituted by the general convolutional process. The right part is an expansive path, which is constituted by transposed 2d convolutional layers.

Contracting Path: In Figure 3.2 contracting path follows the formula:

conv_layer1 – > conv_layer2 – > max_pooling – > dropout(optional)

First part of the U-Net model is:

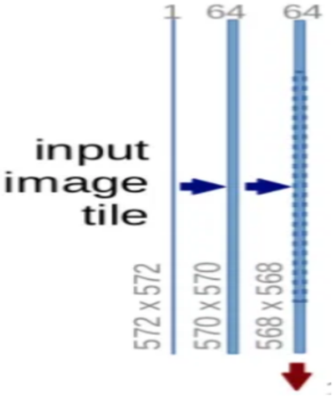


Figure 3.2: Connecting Path

Each process constitutes two convolutional layers, and the number of channel changes from $1 -- > 64$, as the convolution process will increase the depth of the image. The red arrow pointing down is the max pooling process which halves down size of image(the size reduced from $572x572 -- > 568x568$ is due to padding issues, but the implementation here uses padding= “same”).

The process is repeated 3 more times:

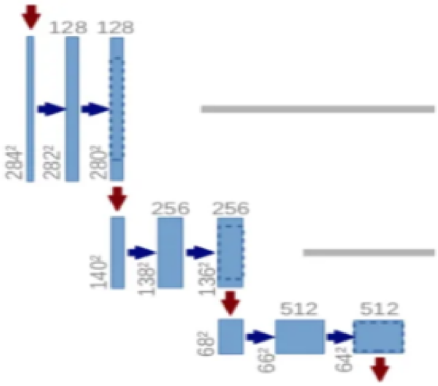


Figure 3.3: Repetition of Convolution Layers

And in Figure 3.4 now we reach the bottom-most part

Figure 3.4: Bottom most part

Still 2 convolutional layers are built, but with no max pooling. The image at this moment has been resized to 28x28x1024. Now let's get to the expansive path.

Expansive Path: In the expansive path, the image is going to be up-sized to its original size. The formula follows:

conv_2d_transpose $->$ concatenate $->$ conv_layer1 $->$ conv_layer2

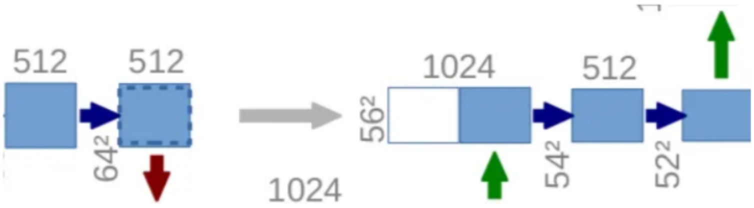


Figure 3.5: Expansive Path

Transposed convolution is an up-sampling technique that expands the size of images. Basically, it does some padding on the original image followed by a convolution operation. After the transposed convolution, the image is up-sized from 28x28x1024 $->$ 56x56x512, and then, this image is concatenated with the corresponding image from the contracting path and together makes an image of size 56x56x1024. The reason here is to combine the information from the previous layers in order to get a more precise prediction. Same as before, this process is repeated 3 more times Now we've reached the uppermost part of the architecture, the last step is to reshape the image to satisfy our prediction requirements.

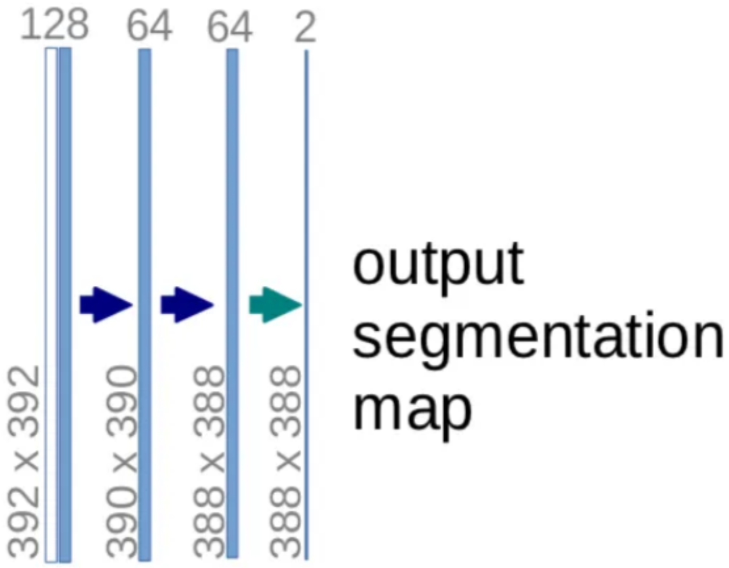


Figure 3.6: Output Segmentation Map

The last layer is a convolution layer with 1 filter of size 1x1(notice that there is no dense layer in the whole network). And the rest left is the same for neural network training.

3.1.3 ResNet50 Architecture

In the Figure 3.7 ResNet50 architecture is considered to be among the most popular Convolutional Neural Network architectures around. Introduced by Microsoft Research in 2015, Residual Networks (ResNet in short) broke several records when it was first introduced.

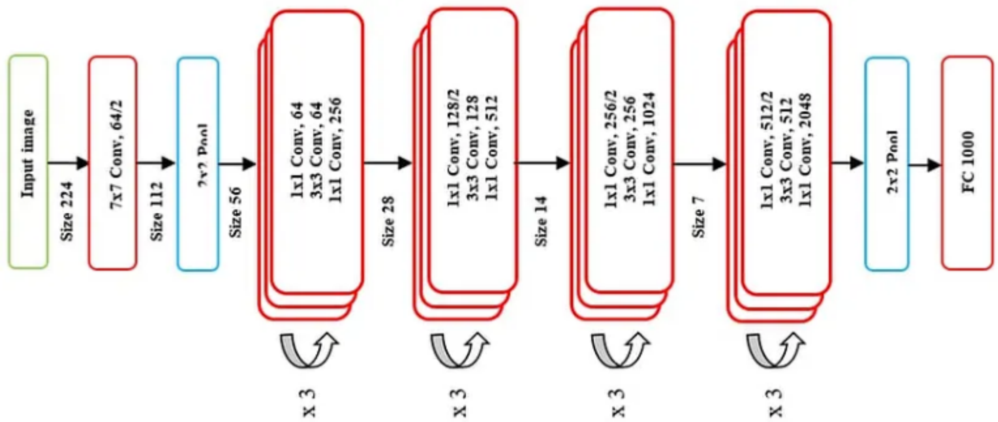


Figure 3.7: ResNet50 architecture

The ResNet-50 architecture can be broken down into 6 parts

1. Input Pre-processing
2. Cfg[0] blocks
3. Cfg[1] blocks
4. Cfg[2] blocks
5. Cfg[3] blocks
6. Fully-connected layer

Different versions of the ResNet architecture use a varying number of Cfg blocks at different levels.

Cfg[0] Block: This block contains 1 Conv Layer and 2 Identity Layers. For helping numerical stability, we specify a kernel constraint which makes sure that all weights are normalized at constant intervals. Between 2 subsequent layers, we also include a BatchNormalization layer.

Cfg[1] Block: This block contains 1 Conv Layer and 2 Identity Layers. This is similar to the Cfg[0] blocks, with the difference mainly being in the number of Out_channels in the Conv and Identity layers being more.

Cfg[2] Block: This block contains 1 Conv layer and 5 Identity layers. This is one of the more important blocks for ResNet as most versions of the model differ in this block-space.

Cfg[3] Block: This block contains 1 Conv Layer and 2 Identity Layers. This is the last set of Convolutional Layer blocks present in the network.

Classifier Block: This block contains an Average Pooling Layer, a Dropout Layer and a Flatten layer. At this block, the feature map is finally flattened and pushed into a Fully Connected Layer which is then used for producing predictions. A Softmax activation is applied to generate logits/probabilities.

Build ResNet Model: Now we take all the blocks and join them together to create the final ResNet Model. In our entire process, we have used the Keras Functional API, which is a best-practice for Tensorflow.

3.1.4 Using Resnet50 as a backbone and combining U-Net

Unet is an extension of an encoder-decoder fully convolutional network. The convolutional layers perform feature extraction by employing the filters in these layers to learn low and high dimensional features as they iteratively get trained. The intuition behind Unet is to encode the image passing it through a CNN as it gets down-sampled and then decode it back or up-sample it to obtain the segmentation mask. The features to be detected in the mask depends on the learned weight filters, up-sampling & down-sampling blocks (which can also be made learn able) and the concatenations and skip connections. The backbone is the architectural element which defines how these layers are arranged in the encoder network and they determine how the decoder network should be built.

The backbone used are often Vanilla CNNs such as VGG, ResNet, Inception, EfficientNet etc which performs encoding and down-sampling by itself. These networks are taken and their counterparts are built to perform decoding and up-sampling to form the final Unet.

There has been much research on making these networks better by using backbones which are capable of extracting features better.

Using ResNet50 on top of U-Net is a common approach in some computer vision tasks to further enhance the performance of U-Net in capturing high-level features. This combination leverages the strengths of both architectures. Here's how it can be done:

1. **U-Net as Encoder-Decoder:** Use the U-Net architecture as the initial part of the network. The contracting path of U-Net acts as the encoder, capturing low-level and mid-level features. The expanding path serves as the decoder, recovering spatial information using skip connections.

2. **ResNet50 as Feature Extractor:** Replace the decoder part of U-Net with ResNet50. ResNet50 is a deeper architecture capable of capturing more complex and abstract features. By replacing the decoder with ResNet50, we can leverage its strong feature extraction capabilities to capture high-level information.

3. **Fine-tuning:** Optionally, we can choose to fine-tune the combined U-Net and

ResNet50 model. Fine-tuning allows the network to adapt the learned weights to the specific task and dataset you are working with. It helps to improve the model's performance by updating the parameters based on the specific requirements of the task.

4. **Output Layer:** Add an appropriate output layer to produce the desired output based on your task. For instance, if we are performing image segmentation, the output layer could be a convolutional layer with the number of output channels equal to the number of classes you want to segment.

This combination of U-Net and ResNet50 allows the model to capture both low-level and high-level features, taking advantage of the strengths of each architecture. The U-Net part helps in preserving spatial information and capturing local features, while ResNet50 provides deeper feature extraction capabilities for capturing global and abstract features.

Chapter 4

Experimental Result

Our proposed method is extensively studied on brain BRATS 2020 brain MR image volumes. We have taken both Training and validation data of Brain MR image dataset from BRATS website. The MICCAI BRATS 2020 Dataset used here contains 371 training files which were taken from Kaggle. We have done segmentation on the BraTS 2020 dataset and gotten great results. The challenge of identifying a tumor is quite difficult and time consuming. The position, form and structure of a tumor differ greatly from one patient to another.

Firstly, we have printed each slice of 2D image from 3D MRI data. We have gotten the following results :

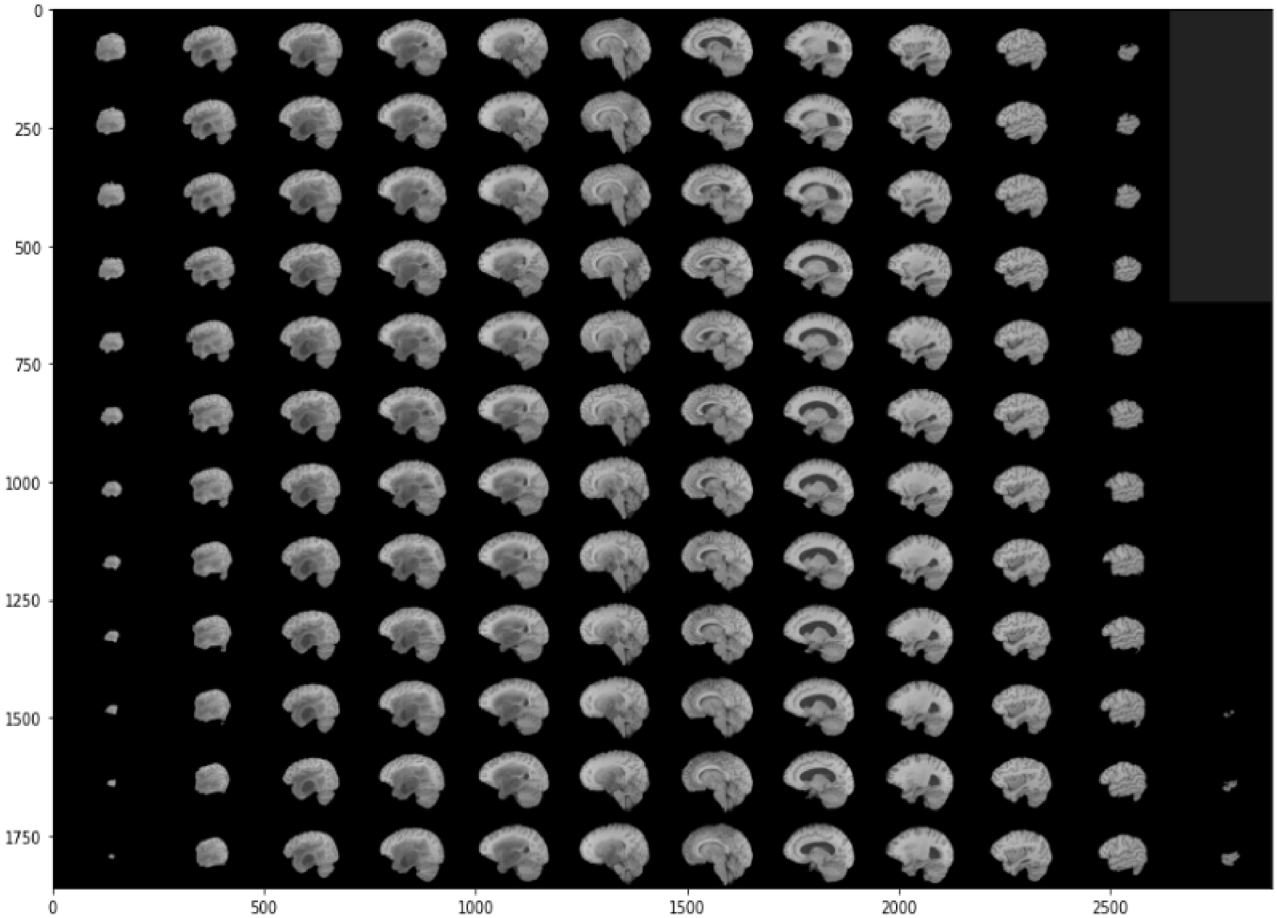


Figure 4.1: Each slice of MRI image from 3d data

Now we have done segmentation on those images and got the following results. Showing segment of tumor from each slice

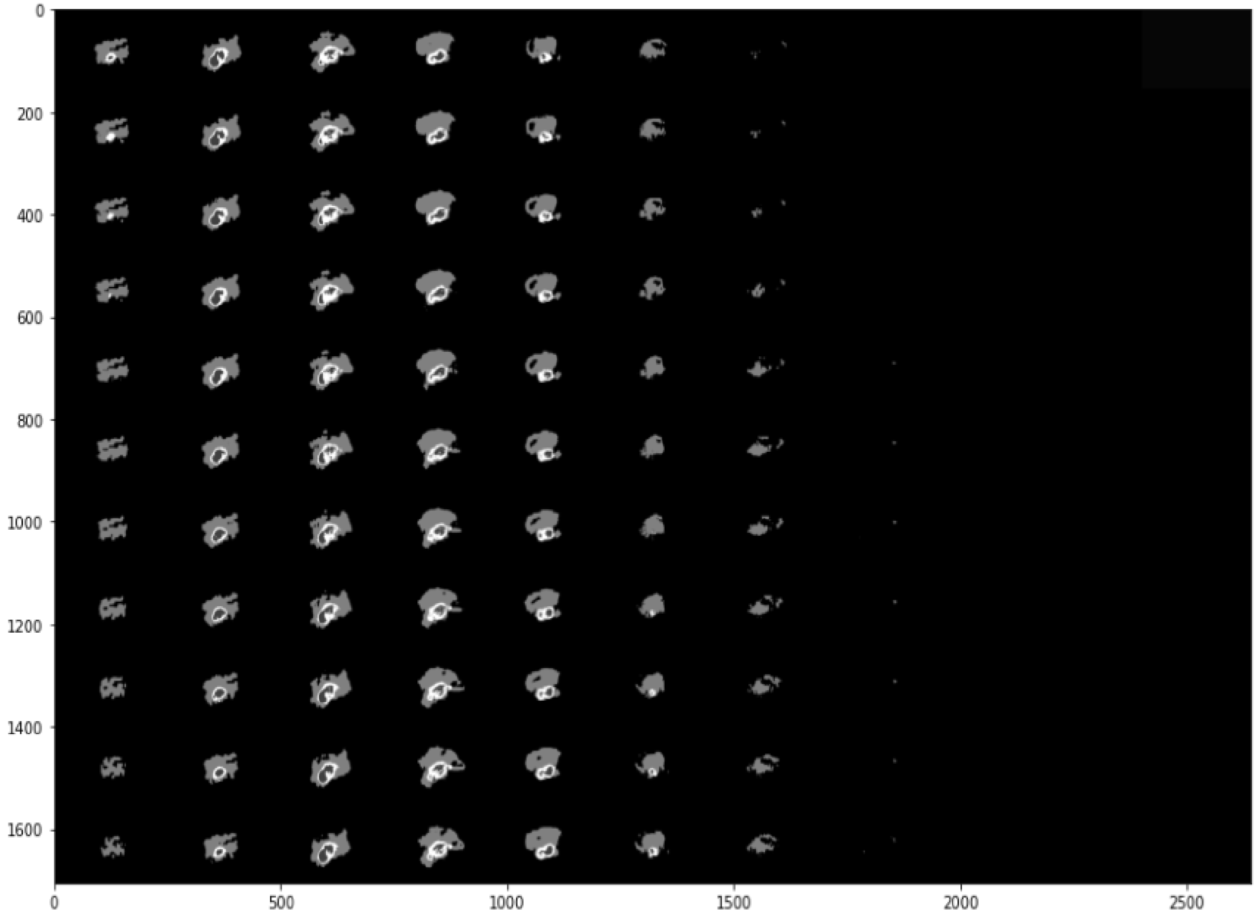


Figure 4.2: Segment of tumor from each slide

4.1 Evaluation Metrics

4.1.1 Accuracy Calculation

Accuracy is a metric that generally describes how the model performs across all classes. It is useful when all classes are of equal importance. It is calculated as the ratio between the number of correct predictions to the total number of predictions.

$$\textbf{Accuracy Score} = \textbf{(TP + TN) / (TP + TN + FP + FN)} \tag{4.1}$$

4.1.2 Precision Calculation

The precision is calculated as the ratio between the number of Positive samples correctly classified to the total number of samples classified as Positive (either correctly or incorrectly). The precision measures the model's accuracy in classifying a sample as positive.

$$\textbf{Precision} = \textbf{TP / (TP + FP)} \tag{4.2}$$

When the model makes many incorrect Positive classifications, or few correct Positive classifications, this increases the denominator and makes the precision small. On the other hand, the precision is high when:

1. The model makes many correct Positive classifications (maximize True Positive).
2. The model makes fewer incorrect Positive classifications (minimize False Positive).

4.1.3 Recall Calculation

The recall is calculated as the ratio between the number of Positive samples correctly classified as Positive to the total number of Positive samples. The recall measures the model's ability to detect Positive samples. The higher the recall, the more positive samples detected.

$$\textbf{Recall} = \textbf{TP / (TP + FN)} \tag{4.3}$$

The recall cares only about how the positive samples are classified. This is independent of how the negative samples are classified, e.g. for the precision. When the model classifies all the positive samples as Positive, then the recall will be 100% even if all the negative samples were incorrectly classified as Positive.

4.1.4 Dice Coefficient Calculation

Another popular loss function for image segmentation tasks is based on the Dice Coefficient, which is essentially a measure of overlap between two samples. This measure ranges from 0 to 1 where a Dice coefficient of 1 denotes perfect and complete overlap. The Dice coefficient was originally developed for binary data, and can be calculated as:

4.1.5 Imaging Data Description

All BraTS multimodal scans are available as NIfTI files (.nii.gz) and describe as follows: (a) native (T1) and (b) post−contrast T1−weighted (T1Gd), (c) T2−weighted (T2), and (d) T2 Fluid Attenuated Inversion Recovery (T2−FLAIR) volumes, and were acquired with different clinical protocols and various scanners from multiple (n=19) institutions, mentioned as data contributors here. All the imaging datasets have been segmented manually, by one to four raters, following the same annotation protocol, and their annotations were approved by experienced neuro−radiologists. Annotations comprise the GD−enhancing tumor (ET − label 4), the peritumoral edema (ED − label 2), and the necrotic and non−enhancing tumor core (NCR/NET − label 1), as described both in the BraTS 2012−2013 TMI paper and in the latest BraTS summarizing paper. The provided data are distributed after their pre−processing i.e., co−registered to the same anatomical template, interpolated to the same resolution (1 mm$\hat{3}$) and skull−stripped.

4.1.6 Training, validation and test set distribution

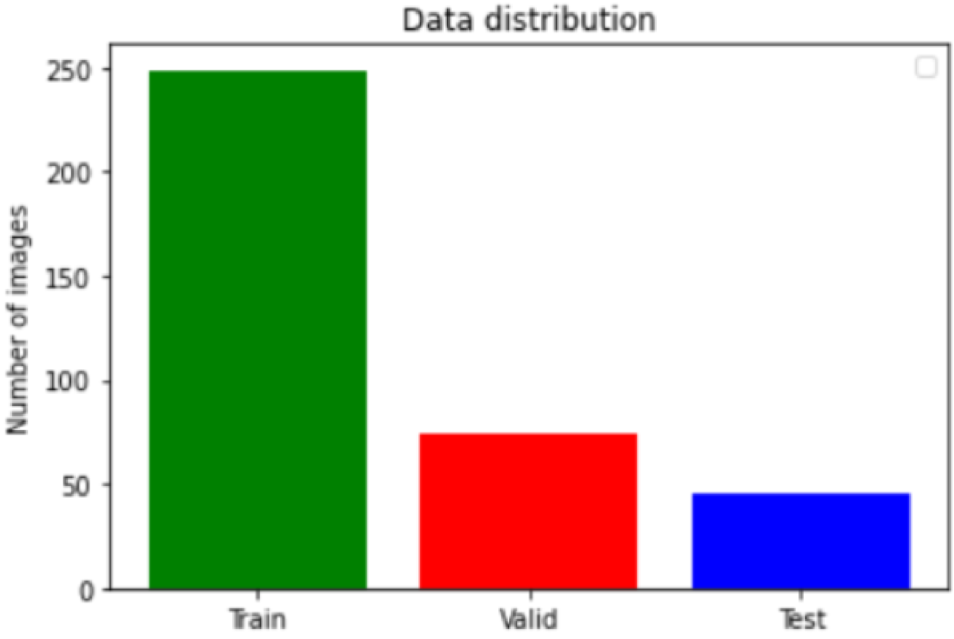
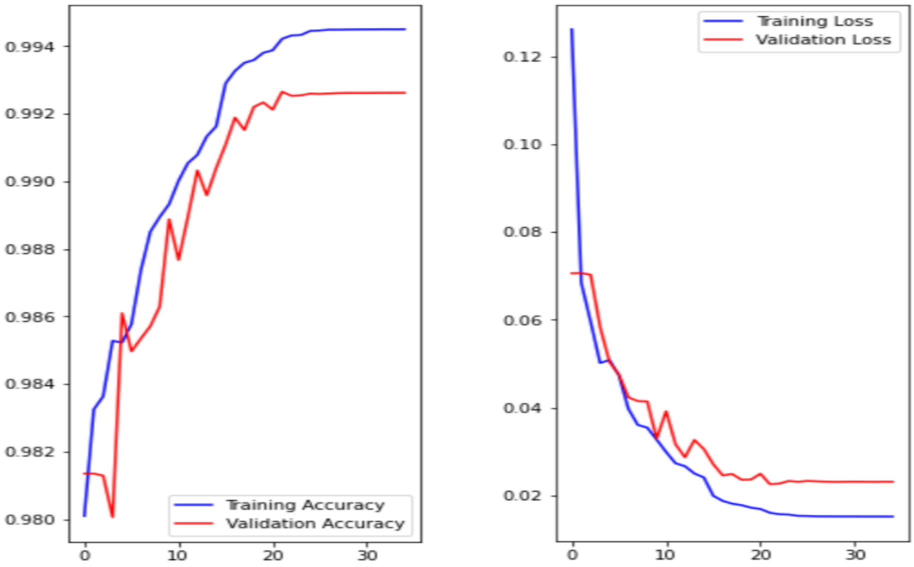


Figure 4.3: Data Distribution

4.2 Results

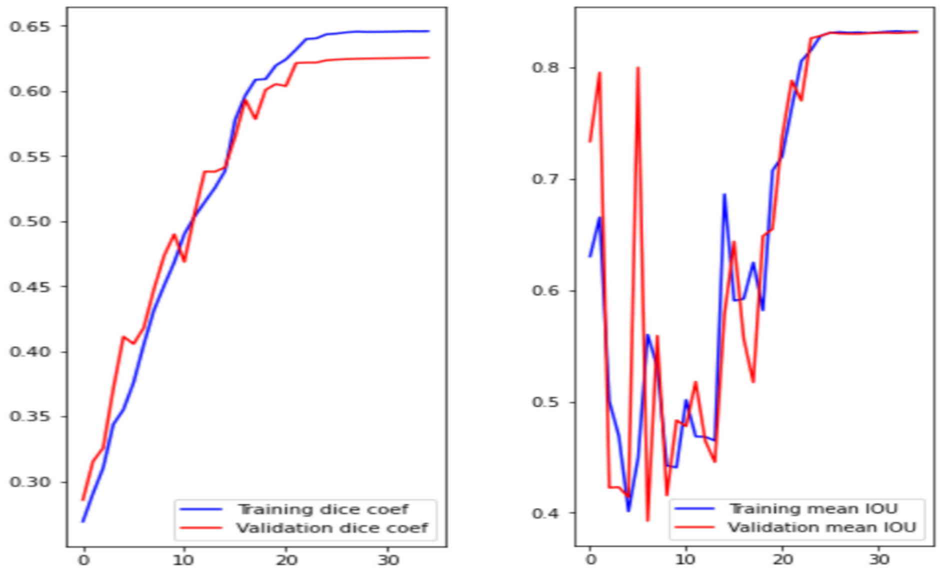
4.2.1 U-Net model outputs



(a) Training accuracy and Validation accuracy curve against epoch

(b) Training loss and Validation loss curve against epoch

Figure 4.4: Training-Validation Loss and Accuracy curve explain in both images for U-Net Model



(a) Training dice coef and Validation dice coef curve against epoch

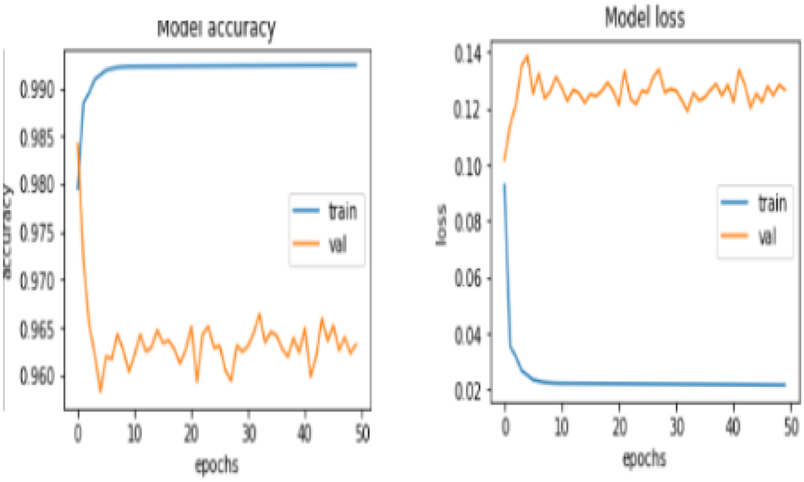
(b) Training mean iou and Validation mean iou curve against epoch

Figure 4.5: Training-Validation dice coef and mean iou curve explain in both images for U-Net Model

Accuracy	Loss	Precision	Dice-Coefficient	Sensitivity	Specificity
0.9944	0.0172	0.6512	0.9946	0.9932	0.9982

Table 4.1: Detailed U-Net model outputs

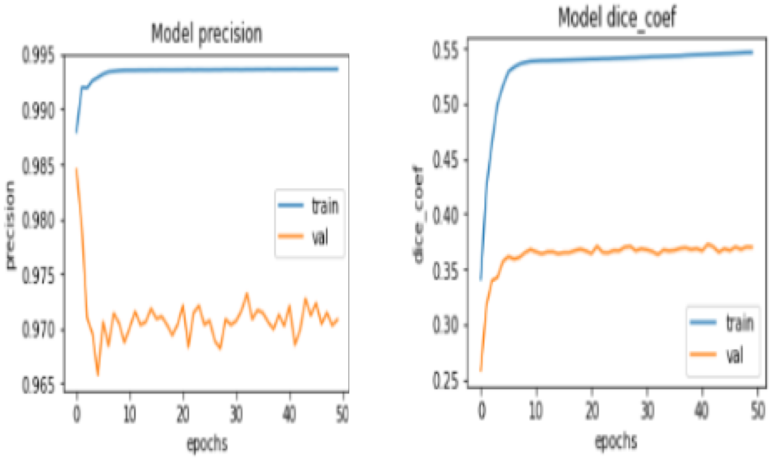
4.2.2 ResNet + U-Net model output



(a) Training-Validation accuracy curve against epoch

(b) Training-Validation loss curve against epoch

Figure 4.6: Training-Validation accuracy and loss curve explain in both images for U-Net Model



(a) Training-Validation precision curve against epoch

(b) Training-Validation dice coef curve against epoch

Figure 4.7: Training-Validation accuracy and dice coef curve explain in both images for U-Net Model

Accuracy	Loss	Precision	Dice-Coefficient	Sensitivity	Specificity
0.9887	0.0318	0.4196	0.9934	0.9853	0.9977

Table 4.2: Detailed ResNet + U-Net model outputs

4.2.3 Comparison of two models

Firstly we used only U-Net for Brain MRI image segmentation. Second time we used ResNet50 as backbone. We replaced the auto-encoder part of U-Net model with ResNet50.

Here are the experimental results of both models: Training Results:-

Model Name	Accuracy	Loss	Precision	Dice-Coefficient	Sensitivity	Specificity
U-Net	0.9944	0.0172	0.6512	0.9946	0.9932	0.9982
ResNet+U-Net	0.9887	0.0318	0.4196	0.9934	0.9853	0.9977

Table 4.3: Comparison between two models

Chapter 5

Conclusion and Future Work

5.1 Conclusion

In conclusion, this thesis conducted a comparative study between two popular deep learning architectures, U-Net and ResNet, for brain MRI image segmentation. The objective was to evaluate their performance and identify the strengths and weaknesses of each approach in this specific domain. The results of the comparative study indicate that both U-Net and ResNet are capable of achieving accurate segmentation results. However, certain differences emerged during the evaluation process. U-Net demonstrated excellent performance in segmenting brain structures, particularly when dealing with limited training data. Its architecture, with a contracting path for capturing context and an expanding path for precise localization, proved effective in capturing fine details and producing accurate segmentation. U-Net's skip connections also facilitated the integration of low-level and high-level features, leading to improved segmentation. On the other hand, ResNet, with its deep residual learning framework, showed robustness and resilience to vanishing gradients, making it suitable for training very deep networks. ResNet performed well in segmenting larger brain structures and exhibited good generalization capabilities. Its skip connections, although not as extensive as U-Net, allowed for effective feature reuse and improved gradient flow, contributing to accurate segmentation results. It is worth noting that the performance of U-Net and ResNet can be influenced by various factors, including the size and diversity of the training dataset, the complexity of the segmentation task, and the availability of computational resources. Therefore, the choice between

U-Net and ResNet should be made based on the specific requirements and constraints of the segmentation problem at hand. In conclusion, both U-Net and ResNet have demonstrated their effectiveness in brain MRI image segmentation, each with its own strengths and advantages.. Future research could explore hybrid architectures that combine the strengths of both U-Net and ResNet to achieve even better segmentation performance and interpretability.

5.2 Future Work

The future scope of brain MRI image segmentation using deep learning is promising and holds great potential for advancements in medical imaging analysis. Here are some key areas of future development:

- **Enhanced Segmentation Accuracy:** Deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), can be further optimized to improve the accuracy of brain MRI image segmentation.
- **Multi-Modal Segmentation:** Combining information from multiple imaging modalities can provide a more comprehensive understanding of brain structures. Future research can focus on developing deep learning models capable of integrating and utilizing multimodal data for improved segmentation accuracy and reliability.
- **Fine-Grained Segmentation:** Current segmentation methods often focus on segmenting major brain structures, such as gray matter, white matter, and cerebrospinal fluid. Future work can aim to achieve finer-grained segmentation by delineating substructures within these major regions, such as specific cortical regions or deep brain nuclei. This level of detail can assist in precise diagnosis and treatment planning for neurological disorders.
- **Real-Time Segmentation:** Real-time or near-real-time segmentation of brain MRI images can greatly benefit neurosurgery and intraoperative decision-making. Developing deep learning models that can provide accurate and fast segmentation results within a clinical time frame would be highly valuable. This could involve the optimization of existing architectures or the development of new network architectures that are computationally efficient.

- **Generalization Across Datasets and Pathologies:** Deep learning models trained on large and diverse datasets have shown excellent performance. Future research can focus on developing techniques that can transfer knowledge learned from one dataset or pathology to another, reducing the need for large labeled datasets for each specific case.
- **Inter pretability and Explainability:** Deep learning models are often considered black boxes, making it difficult to interpret their decisions. In the context of brain MRI image segmentation, there is a need for models that can provide inter pretable and explainable segmentation results. Researchers can explore methods that provide insights into which image features or regions contribute most to the segmentation output, aiding clinicians in understanding and trusting the results.
- **Clinical Integration:** Ultimately, the success of deep learning-based segmentation techniques depends on their seamless integration into clinical workflows. Future research should focus on bridging the gap between academic research and clinical adoption, ensuring that the developed models are user-friendly, robust, and easily accessible by healthcare professionals.

Overall, the future of brain MRI image segmentation using deep learning holds great potential for improving diagnostic accuracy, treatment planning, and patient outcomes in neuro imaging. Continued research and collaboration between deep learning experts and medical professionals will be crucial in realizing these advancements.

References

[1] Mukul Aggarwal et al. “An early detection and segmentation of Brain Tumor using Deep Neural Network”. In: *BMC Medical Informatics and Decision Making* 23.1 (2023), pp. 1–12.

[2] Salma Alqazzaz et al. “Automated brain tumor segmentation on multi-modal MR image using SegNet”. In: *Computational Visual Media* 5 (2019), pp. 209–219.

[3] Lamia H Shehab et al. “An efficient brain tumor image segmentation based on deep residual networks (ResNets)”. In: *Journal of King Saud University-Engineering Sciences* 33.6 (2021), pp. 404–412.

[4] Sakshi Ahuja, BK Panigrahi, and Tapan Gandhi. “Transfer learning based brain tumor detection and segmentation using superpixel technique”. In: *2020 international conference on contemporary computing and applications (IC3A)*. IEEE. 2020, pp. 244–249.

[5] Guotai Wang et al. “Interactive medical image segmentation using deep learning with image-specific fine tuning”. In: *IEEE transactions on medical imaging* 37.7 (2018), pp. 1562–1573.

[6] D Ramya et al. “Segmentation of Brain MRI Images Using U-Net”. In: *Recent Trends in Computational Intelligence and Its Application*. CRC Press, 2023, pp. 275–281.

[7] Wenbin Wu et al. “Inner Cascaded U^2-Net: An Improvement to Plain Cascaded U-Net.” In: *CMES-Computer Modeling in Engineering & Sciences* 134.2 (2023).

[8] Guibo Luo et al. “Influence of data distribution on federated learning performance in tumor segmentation”. In: *Radiology: Artificial Intelligence* 5.3 (2023), e220082.

[9] Qing Xu et al. “DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation”. In: *Computers in Biology and Medicine* 154 (2023), p. 106626.

[10] Bumshik Lee, Nagaraj Yamanakkanavar, and Jae Choi. “Automatic segmentation of brain MRI using a novel patch-wise U-net deep architecture”. In: *PLOS ONE* 15 (Aug. 2020), e0236493. DOI: `10.1371/journal.pone.0236493`.

[11] Akash Saxena et al. “Numerical Simulation and Design of Semantic Segmentation Using Improved Resnet-50 Based Deep Learning Techniques”. In: *International Journal of Intelligent Systems and Applications in Engineering* 11.6s (2023), pp. 804–815.

[12] Bjoern H Menze et al. "The multimodal brain tumor image segmentation benchmark (BRATS)". In: *IEEE transactions on medical imaging* 34.10 (2014), pp. 1993–2024.

[13] BK Tripathy et al. "Brain MRI segmentation techniques based on CNN and its variants". In: *Brain Tumor MRI Image Segmentation Using Deep Learning Techniques.* Elsevier, 2022, pp. 161–183.

[14] Parasa Rishi Kumar et al. "Multi-class Brain Tumor Classification and Segmentation using Hybrid Deep Learning Network Model". In: *Scalable Computing: Practice and Experience* 24.1 (2023), pp. 69–80.

[15] Mobeen Ur Rehman et al. "Bu-net: Brain tumor segmentation using modified u-net architecture". In: *Electronics* 9.12 (2020), p. 2203.

[16] Tiejun Yang and Jikun Song. "An automatic brain tumor image segmentation method based on the u-net". In: *2018 IEEE 4th International Conference on Computer and Communications (ICCC).* IEEE. 2018, pp. 1600–1604.

[17] Nahian Siddique et al. "U-net and its variants for medical image segmentation: A review of theory and applications". In: *Ieee Access* 9 (2021), pp. 82031–82057.

[18] Martin Kolařık et al. "Optimized high resolution 3d dense-u-net network for brain and spine segmentation". In: *Applied Sciences* 9.3 (2019), p. 404.

[19] Narinder Singh Punn and Sonali Agarwal. "Modality specific U-Net variants for biomedical image segmentation: a survey". In: *Artificial Intelligence Review* 55.7 (2022), pp. 5845–5889.

[20] Elias Hossain et al. "Brain Tumor Auto-Segmentation on Multimodal Imaging Modalities Using Deep Neural Network." In: *Computers, Materials & Continua* 72.3 (2022).

[21] Spyridon Bakas et al. "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge". In: *arXiv preprint arXiv:1811.02629* (2018).

[22] Spyridon Bakas et al. "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features". In: *Scientific data* 4.1 (2017), pp. 1–13.

[23] Raphael Meier et al. "A hybrid model for multimodal brain tumor segmentation". In: *Multimodal Brain Tumor Segmentation* 31 (2013), pp. 31–37.

[24] Mina Ghaffari, Arcot Sowmya, and Ruth Oliver. "Automated brain tumour segmentation using cascaded 3d densely-connected u-net". In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part I 6.* Springer. 2021, pp. 481–491.

[25] Maodan Zhang et al. "Improved u-net with multi-scale cross connection and dilated convolution for brain tumor segmentation". In: *2019 International Conference on Medical Imaging Physics and Engineering (ICMIPE).* IEEE. 2019, pp. 1–5.

[26] Mingquan Lin et al. "Fully automated segmentation of brain tumor from multiparametric MRI using 3D context deep supervised U-Net". In: *Medical Physics* 48.8 (2021), pp. 4365–4374.

[27] Getao Du et al. "Medical image segmentation based on u-net: A review." In: *Journal of Imaging Science & Technology* 64.2 (2020).

[28] Wentao Wu et al. "An intelligent diagnosis method of brain MRI tumor segmentation using deep convolutional neural network and SVM algorithm". In: *Computational and Mathematical Methods in Medicine* 2020 (2020).

[29] Pradeep Kumar Mallick et al. "Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network". In: *IEEE Access* 7 (2019), pp. 46278–46287.

[30] Adel Kermi, Issam Mahmoudi, and Mohamed Tarek Khadir. "Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes". In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II 4*. Springer. 2019, pp. 37–48.

[31] Michał Futrega et al. "Optimized U-Net for brain tumor segmentation". In: *International MICCAI Brainlesion Workshop*. Springer. 2021, pp. 15–29.

[32] Nahian Siddique et al. "U-Net and its variants for medical image segmentation: theory and applications". In: *arXiv preprint arXiv:2011.01118* (2020).

[33] Feifan Wang et al. "3d u-net based brain tumor segmentation and survival days prediction". In: *International MICCAI Brainlesion Workshop*. Springer. 2019, pp. 131–141.

[34] Jiong Wu et al. "Skip connection U-Net for white matter hyperintensities segmentation from MRI". In: *IEEE Access* 7 (2019), pp. 155194–155202.

[35] M Mohammed Thaha et al. "Brain tumor segmentation using convolutional neural networks in MRI images". In: *Journal of medical systems* 43 (2019), pp. 1–10.

[36] M Malathi and P Sinthia. "Brain tumour segmentation using convolutional neural network with tensor flow". In: *Asian Pacific journal of cancer prevention: APJCP* 20.7 (2019), p. 2095.

[37] Mina Jafari et al. "DRU-Net: an efficient deep convolutional neural network for medical image segmentation". In: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2020, pp. 1144–1148.

[38] Zaiwang Gu et al. "Ce-net: Context encoder network for 2d medical image segmentation". In: *IEEE transactions on medical imaging* 38.10 (2019), pp. 2281–2292.

[39] Mohamed Arbane et al. "Transfer learning for automatic brain tumor classification using MRI images". In: *2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH)*. IEEE. 2021, pp. 210–214.

[40] Jianxin Zhang et al. "Attention gate resU-Net for automatic MRI brain tumor segmentation". In: *IEEE Access* 8 (2020), pp. 58533–58545.

[41] Weihao Weng and Xin Zhu. "INet: convolutional networks for biomedical image segmentation". In: *Ieee Access* 9 (2021), pp. 16591–16603.

[42] Ning Sheng et al. "Second-order ResU-Net for automatic MRI brain tumor segmentation". In: *Mathematical biosciences and engineering* 18.5 (2021), pp. 4943–4960.

[43] Hyunho Hwang, Hafiz Zia Ur Rehman, and Sungon Lee. "3D U-Net for skull stripping in brain MRI". In: *Applied Sciences* 9.3 (2019), p. 569.

[44] Dmitry Lachinov, Evgeny Vasiliev, and Vadim Turlapov. "Glioma segmentation with cascaded UNet". In: *International MICCAI Brainlesion Workshop*. Springer. 2018, pp. 189–198.

[45] Liyan Sun et al. "A 3D spatially weighted network for segmentation of brain tissue from MRI". In: *IEEE transactions on medical imaging* 39.4 (2019), pp. 898–909.

[46] Motahareh Aghalari, Ali Aghagolzadeh, and Mehdi Ezoji. "Brain tumor image segmentation via asymmetric/symmetric UNet based on two-pathway-residual blocks". In: *Biomedical signal processing and control* 69 (2021), p. 102841.