

**NEAR EAST UNIVERSITY**

**INSTITUTE OF GRADUATE STUDIES**

**DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING**

**INCREASED DEPTH OF 3D RESIDUAL UNET NETWORK  
ARCHITECTURE WITH AN ATTENTION GATE FOR BRAIN TUMORS  
SEGMENTATION**

**MSc THESIS**

**MOHAMUD ABDULLAHI JAMA**

**NICOSIA – 2022**

**NEAR EAST UNIVERSITY**

**INSTITUTE OF GRADUATE STUDIES**

**DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING**

**INCREASED DEPTH OF 3D RESIDUAL UNET NETWORK  
ARCHITECTURE WITH AN ATTENTION GATE FOR BRAIN TUMORS  
SEGMENTATION**

**MSc THESIS**

**MOHAMUD ABDULLAHI JAMA**

**SUPERVISOR**

**Prof. Dr. FAHREDDIN SADIKOGLU**

**NICOSIA – 2022**

## APPROVAL

We certify that we have read the thesis submitted by MOHAMUD ABDULLAHI JAMA titled "INCREASED DEPTH OF 3D RESIDUAL UNET NETWORK ARCHITECTURE WITH AN ATTENTION GATE FOR BRAIN TUMORS SEGMENTATION" and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Electrical and Electronic Engineering.

Examining Committee Name-Surname

Signature

Head of the Committee: Assoc. Dr. Boran SEKEROGLU

Committee Member: Assoc. Prof. Dr. Kamil Dimililer

Supervisor: prof. Dr. Fahreddin M. SADIKOGLU

Approved by the Head of the Department





29/08/2022

prof. Dr. Bulent BILGEHAN

Head of Department

Approved by the Institute of Graduate Studies

.../.../20...

Prof. Dr. Kemal Hüsnü Can Başer



Head of the Institute

## DECLARATION

I hereby declare that all information in this document has 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 material and results that are not original to this work.

I hereby declare that the Cyprus International University, Institute of Graduate Studies and Research is allowed to store and make available electronically the present thesis.

Mohamud Abdullahi Jama

23/08/2022

Day/month/year

## **ACKNOWLEDGEMENT**

First and foremost, I would like to thank Almighty Allah for bestowing His blessings upon me and giving me the strength to carry out and complete this work.

second, I would like to express my special thanks to Prof. Dr. FAHREDDIN SADIKOGLU, my supervisor, who gave me the golden opportunity to do this, and the information and support essential to end my research. Which also helped me in gaining knowledge, He was replaced more like a mentor than a supervisor. I'm thankful for you, dear supervisor.

I welcome and am thankful to Near East University's lecturers and professors for sharing information and presenting honest and useful assistance during the course. Most highly, I want to express my gratitude to my parents for their reliable love,

Finally, I would like to use this chance to express my gratitude to everybody complicated for their choice knowledge and effort, which enabled me to finish this thesis. And, I would like to express my gratitude to everyone who has contributed to this project, whether directly or indirectly.

## **ABSTRACT**

### **INCREASED DEPTH OF 3D RESIDUAL UNET NETWORK ARCHITECTURE WITH AN ATTENTION GATE FOR BRAIN TUMORS SEGMENTATION**

**MOHAMUD ABDULLAHI JAMA**

**Prof. Dr. FAHREDDIN SADIKOGLU**

**M.S.c, Department of Electrical and Electronics**

**Engineering August, 2022, 64 pages**

Gliomas are one the deadliest forms of brain tumor, often resulting in a short life when they reach a high grade. Early detection of glioma is critical for patient survival. Magnetic resonance images (MRI) are frequently used to evaluate brain malignancies. Segmenting tumors from Magnetic resonance images of the brain is one of the highest priorities areas of medical science. Semantic segmentation gives detection information and helps doctors know the disease's early stage. A convolutional neural network is highly effective in segmenting medical images. This study presents a new deep learning method for accurate brain tumor segmentation that can be modified and expand the residual unet architecture. It increases the network's depth while keeping an extremely high level of accuracy. This study proposes a deep learning network architecture called increased depth of 3D Residual UNET Network Architecture with an attention gate for Brain Tumors Segmentation, which contains an attention gate and advanced 3D Residual UNET. The proposed architecture has increased the depth of the normal attention residual unet from four layers to six layers. However, the network loses a corresponding amount of spatial information, lowering segmentation performance. The 3D UNet transmits contextual and spatial information from the encoder part to the decoder by using skip links. Consequently, critical spatial information lost during down sampling can be recovered more effectively. By allowing only activations from important areas on the encoder side using attention gates and creating better feature mappings at the decoder, these modifications to the network enhanced the process of learning. Furthermore, the use of a combination of dice loss and focal loss helped the model in resolving class imbalance challenges where brain

tumors have a significant imbalance between foreground and background classes. Because of this, the model has improved and got a better segmentation achievement. The model outperformed baseline models such as UNet, Residual UNet, and attention gates with Residual UNet. Three separate datasets are evaluated to demonstrate that the presented model is superior to its baseline models and the existing state-of-the-art segmentation approaches. The suggested model was tested on BraTS 2020, BraTS 2019, and BraTS 2018 datasets. The model achieved the dice coefficient scores for WT, TC, and ET of 93.91%, 93.01%, and 89.21% on the BraTS 2020 dataset, 88.44%, 75.11%, and 79.87% on the BraTS 2019 dataset, and 88.36%, 83.17%, and 78.19% on the BraTS 2018 dataset, respectively.

**Keywords**

Semantic segmentation; Brain tumor segmentation; Increased depth; Res-UNet; Attention gates; BraTS 2020; BraTS 2019; BraTS 2018.

## ÖZ

### INCREASED DEPTH OF 3D RESIDUAL UNET NETWORK ARCHITECTURE WITH AN ATTENTION GATE FOR BRAIN TUMORS SEGMENTATION

**MOHAMUD ABDULLAHI JAMA**

**Prof. Dr. FAHREDDIN SADIKOGLU**

**M.S.c, Department of Electrical and Electronics**

**Engineering August, 2022, 64 pages**

Yüz İfadesi Tanıma (FER), bilgisayarlı görü alanında iyi çalışılmış bir konu olduğu kadar zor bir konu olmuştur. Fac, insan vücudunun duyguları hakkında bilgi aktaran en dinamik parçasıdır. Yüz geometrisi ve yüz görünümündeki çeşitlilik düzeyi, farklı insan ifadelerinin tespit edilmesini mümkün kılar. Çok sayıda yüz ifadesi arasında ayırım yapabilmek için, yüz ifadelerinin sınıflarını belirlemek çok önemlidir.

Bu makalede kullanılan metodoloji, Makine öğrenmesinde Konvolüsyonel Sinir Ağı (CNN)'e dayanmaktadır. Tezin ana katkısı, Alex Net mimarilerini ve derin öğrenme evrişimli sinir ağları incelemektir. Öğrenme aktarım yaklaşımı uygulanarak ve tam bağlantılı katman (SVM) sınıflandırıcı ile değiştirilerek iyileştirmeler sağlandı. Sistem, icv-MEFED veri tabanında incelenerek elde edilen sonuçlar başarılı oldu. İyileştirilmiş modeller, seçilen ifadelerin sınıflandırılması için yaklaşık %64,29 tanıma oranlarına ulaştı. Elde edilen sonuçlar kabul edilebilir ve literatürdeki ilgili sistemlerle karşılaştırılabilir niteliktedir ve daha fazla iyileştirme için bir arka plan sağlar.

Anahtar Kelimeler: Yüz ifadeleri, yüz ifadesi analizi, yüz ifadesi tanıma, icv-MEFED, derin öğrenme, Evrişimsel Sinir Ağları, Alex net, SoftMax, SVM.

## TABLE OF CONTENTS

ACKNOWLEDGEMENT .....	i
ABSTRACT.....	iii
TABLE OF CONTENTS .....	v
LIST OF TABLES .....	ix
LIST OF FIGURES .....	x
ABBREVIATIONS .....	xi
CHAPTER ONE .....	1
INTRODUCTION .....	1
1.1 Background Study .....	1
1.2 Problem Statement .....	3
1.3 Objectives .....	3
1.4 Organizations of the Thesis .....	4
CHAPTER TWO .....	5
LITERATURE REVIEW.....	5
2.1 Overview .....	5
2.2 Brain Tumor .....	5
2.2.1 Low-Grade Glioma .....	6
2.2.2 High-Grade Glioma .....	7
2.3 Magnetic Resonance Images .....	8
2.3.1 T1-weighted .....	9
2.3.2 Contrast-enhanced T1-weighted .....	9
2.3.3 T2-weighted .....	10
2.3.4 FLAIR.....	10
2.3.5 Ground Truth .....	11
2.4 Artificial Neural Network Model .....	12

2.4.1 Basics in Artificial Neural Networks .....	13
2.4.2 Convolutional Neural Networks .....	17
2.5 Previous Work .....	23
2.6 Semantic Segmentation .....	27
CHAPTER THREE.....	29
METHODOLOGY.....	29
3.1 Introduction .....	29
3.2 Network Architecture .....	29
3.2.1 Attention Gates .....	29
3.2.2 Residual Unet.....	30
3.3 Increased depth of 3D Attention Residual UNet.....	31
3.3.1 Encoder Path.....	32
3.3.2 Bottleneck.....	33
3.3.3 Decoder Path.....	33
3.4 Loss Function .....	34
3.4.1 Focal Loss .....	34
3.4.2 Dice Loss .....	35
3.5 Implementation Details .....	36
3.6 Data Pre-Processing .....	36
3.7 Model Training.....	37
CHAPTER FOUR.....	38
RESULTS AND DISCUSSION .....	38
4.1 Introduction .....	38
4.2 Segmentation Evaluation Metrics .....	38
4.3 Data Sets.....	41
4.3.1 BraTS2018.....	41
4.3.2 BraTS2019.....	42

4.3.3 BraTS2020 .....	42
4.4 Pre-Processing .....	42
4.5 Comparison Of The Proposed Models .....	43
4.5.1 Increased Depth Of 3D Unet Model .....	43
4.5.2 Increased Depth Of 3D Residual-Unet Model.....	44
4.5.3 Increased Depth Of 3D Attention Residual Unet .....	45
4.6 Using Dropout Regularization.....	46
4.7 Predicted Image Results Of The Proposed Model .....	49
4.7.1 BraTS 2018 .....	49
4.7.2 BraTS 2019 .....	50
4.7.3 BraTS 2020 .....	51
4.8 Comparison Of Increased Depth Of 3D Attention Residual-Unet With Other Deep Learning Segmentation Models .....	52
CHAPTER FIVE.....	56
CONCLUSION AND RECOMMENDATIONS.....	56
5.1 Conclusion.....	56
5.2 Future Research .....	56
REFERENCE .....	58
CURRICULUM VITAE .....	<b>Error! Bookmark not defined.</b>

## LIST OF TABLES

<b>Table 4.1:</b> Dice coefficient score and Jaccard score for the Increased depth of 3D UNet model on the BraTS 2020, 2019, and 2018 datasets .....	43
<b>Table 4.2:</b> Sensitivity and specificity for Increased depth of 3D UNet on BraTS 2020, 2019, and 2018 datasets. ....	43
<b>Table 4.3:</b> Dice coefficient score and Jaccard score for the Increased depth of 3D Residual-UNet model on BraTS 2020, 2019, and 2018 datasets .....	44
<b>Table 4.4:</b> Sensitivity and specificity for the Increased depth of 3D Residual-UNet on BraTS 2020, 2019, and 2018 datasets. ....	45
<b>Table 4.5:</b> Dice coefficient score and Jaccard score for Increased depth of 3D Attention Residual-UNet on BraTS 2020, 2019, and 2018 datasets. ....	45
<b>Table 4.6:</b> Sensitivity and specificity for Increased depth of 3D Attention Residual-UNet on BraTS 2020, 2019, and 2018 datasets. ....	46
<b>Table 4.7:</b> Dice coefficient and Jaccard score for Increased depth of 3D Attention Residual-UNet with and without dropout regularization on BraTS 2020, 2019, and 2018 datasets. The best scores are in bold. ....	47
<b>Table 4.8:</b> Sensitivity and specificity for Increased depth of 3D Attention Residual-UNet with and without dropout regularization on BraTS 2020, 2019, and 2018 datasets. The best scores are in bold. ....	47
<b>Table 4.9:</b> Comparing the presented model and state-of-the-art segmentation models regarding model size and the number of parameters using BraTS2020. The best two scores are in bold. ....	52
<b>Table 4.10:</b> Comparing the presented model and state-of-the-art segmentation models regarding model size and the number of parameters using BraTS 2019. The best two scores are in bold. ....	53
<b>Table 4.11:</b> Comparing the presented model and state-of-the-art segmentation models regarding model size and the number of parameters using BraTS 2018. The best two scores are in bold. ....	54

## LIST OF FIGURES

<b>Figure 2.1:</b> Magnetic resonance imaging of low-grade gliomas: A) FLAIR, B) T1 image, C) T1-ce image, D) T2weighted, E) ground truth.....	7
<b>Figure 2.2:</b> Magnetic resonance imaging of high-grade glioma: A) FLAIR, B) T1 image, C) T1-ce image, D) T2weighted, E) ground truth.....	8
<b>Figure 2.3:</b> Magnetic resonance imaging of a brain tumor: (left to right) FLAIR, T1-weighted, T1-ce, T2- images.....	11
<b>Figure 2.4:</b> (a) Complete tumor visible in FLAIR; (b) tumor core visible in T2; (c) enhanced and necrotic tumor component structures visible in T1ce; (d) final labels of the observable tumor structures noticeable: edema (yellow), necrotic/cystic core (light blue), enhanced core (red).....	12
<b>Figure 2.5:</b> Feedforward neural network.....	16
<b>Figure 2.6:</b> 2D convolution .....	18
<b>Figure 2.7:</b> 3D representation of a convolution layer.....	19
<b>Figure 0.8:</b> 2×2 max-pooling.....	20
<b>Figure 2.9:</b> Transposed convolution.....	22
<b>Figure 2.10:</b> Dilated convolution .....	23
<b>Figure 2.11:</b> Example of semantic image segmentation.....	28
<b>Figure 3.1:</b> Attention gate representation.....	30
<b>Figure 3.2:</b> Residual block. ....	31
<b>Figure 3.3:</b> An increased depth of 3D Residual Unet Network Architecture with an attention gate.....	32
<b>Figure 4.1:</b> Demonstrates True Positive, False Positive, True Negative, and False Negative. The green circle shows the findings, whereas the yellow circle indicates the ground truth. .	39
<b>Figure 4.2:</b> Predictions of the proposed model on BraTS 2018.....	49
<b>Figure 4.3:</b> Predictions of the proposed model on BraTS 2019.....	50
<b>Figure 4.4:</b> Predictions of the proposed model on BraTS 2020.....	51

## ABBREVIATIONS

ANN	Artificial neural networks
BraTS	Brain Tumor Segmentation
CNN	Convolutional Neural Networks
CSF	Cerebrospinal fluid
DCNN	Deep convolutional neural network
DRN	Deep Residual Network
E.T	Enhancing Tumor
F.N.	False Negative
F.P.	False Positive
FCN	Fully Convolutional Networks
FLAIR	Fluid attenuated inversion recovery
HGM	High Grade Glioma
Leaky ReLU	Leaky Rectified Linear Unit
LGM	Low Grade Glioma
MRI	Magnetic resonance imaging
ReLU	Rectified linear unit
RESNET	Residual UNET

T.C	Tumor core
T.N.	True negative
T.P.	True positive
T1ce	T1-contrast enhanced
VSD	Virtual skeleton database
W.T.	Whole tumor
WHO	World Health Organization

## CHAPTER ONE

### INTRODUCTION

#### 1.1 Background Study

A tumor is simply an unusual and excessive development of body tissue, while a brain tumor is an uncontrolled development of cancer cells in the brain. A brain tumor can also be called a mass of uncontrolled cells in the brain. For example, your skull is extremely hard, enclosing your brain. Any development inside such a constrained region might produce life-threatening problems (Wadhwa, Bhardwaj, & Verma, 2019). A brain tumor could be divided into malignant and benign cells in the brain, which could also be cancerous and non-cancerous. Benign tumors do not include any cancerous cells; when they develop near essential organs, they push on the nerves, impede blood flow, and cause significant complications. Generally, benign tumors react positively to treatments. In contrast, malignant tumors contain a cancerous cell and have a non-uniform structure (Bahadure, Ray, & Thethi, 2017; Pietrangelo, 2019, October 23; Wadhwa et al., 2019).

There are two different kinds of brain tumors which are primary and secondary. Primary tumors are composed entirely of the brain cells, while secondary tumors are composed entirely of cancer cells that have migrated to the brain from another affected body region. Meningiomas are quickly segment-able; however, glioblastomas and gliomas are difficult to detect and locate because of their large contrast and diffusion. Furthermore, its size, structure, and form are variable, making it difficult to spot (Maji, Sigedar, & Singh, 2022). The most popular type of tumor is known as glioma. They can range in grade from high-grade (H.G.) tumors named glioblastoma multiform (GBM) to low-grade (L.G.) tumors similar to oligodendrogliomas or astrocytoma's. They are developing quickly for growth (Abd-Ellah, Awad, Khalaf, & Hamed, 2019; El-Dahshan, Mohsen, Revett, & Salem, 2014).

Diseases were diagnosed using medical imaging such as X-rays, Computed Tomography scans, and Medical Resonance Imaging (MRI) (Saeed et al., 2021;

Wadhwa et al., 2019). MRI has been extensively employed to identify and cure brain cancers. It delivers high-quality brain pictures that contribute to identifying and treating tumors. Utilizing MRI to segment a brain tumor significantly improves detection, medication, and growth rate (Maji et al., 2022). It is crucial to correctly differentiate tumor areas from MRI pictures. The complicated form of the brain tumor makes it difficult to distinguish malignancy from the brain. Therefore, an automatic segmentation approach is necessary to precisely and efficiently identify and segment tumor locations. Segmenting brain tumors manually takes time, relying on the clinician's experience and skills, yet it can occasionally be inaccurate (Maji et al., 2022; Saeed et al., 2021). This causes segmenting tumor cells to be a very difficult process. In the past, a variety of segmentation approaches for brain tumors in MRI images have been established (Angulakshmi & Lakshmi Priya, 2017; Wadhwa et al., 2019). Some of the most popular segmentation techniques include region-growing algorithms (Chakraborty, Chatterjee, Das, & Mali, 2020), clustering-based algorithms (C. Singh & Bala, 2019; M. Singh, Venkatesh, Verma, & Sharma, 2020), and watershed methods (Rajinikanth, Palani Thanaraj, Satapathy, Fernandes, & Dey, 2019; Sivakumar & Janakiraman, 2020). But, due to their susceptibility to noise, the efficiency of these methods is restricted. Recent Discoveries in Deep learning Networks have made significant progress in pixel-level semantic segmentation problems with rich hierarchical functions and comprehensive learning structures (Minaee et al., 2001). Convolutional Neural Networks (CNNs) (He, Zhang, Ren, & Sun, 2016; Krizhevsky, Sutskever, & Hinton, 2012b; Simonyan & Zisserman, 2014) and Fully Convolutional Networks (FCNs) inspired Deep Learning models such as SegNet (Badrinarayanan, Kendall, & Cipolla, 2017), Deep Neural Network (Havaei et al., 2017), Unet (Ronneberger, Fischer, & Brox, 2015), QuickNAT (Roy, Conjeti, Navab, Wachinger, & Initiative, 2019), DenseNet (Huang, Liu, Van Der Maaten, & Weinberger, 2017) and its variations (Alqazzaz, Sun, Yang, & Nokes, 2019), (Kermi, Mahmoudi, & Khadir, 2018), (Noori, Bahri, & Mohammadi, 2019), (Xiao, Lian, Luo, & Li, 2018) have demonstrated impressive segmentation ability.

This thesis presents a new deep learning design that increases the depth of the normal Residual Unet (Maji et al., 2022; Yang et al., 2019) and attention gates (Maji et al.,

2022; Oktay et al., 2018) from four layers into six layers. Additionally, the present work uses a combination of dice loss and focal loss to assist deep learning models in resolving class imbalance challenges where there is a significant unbalance between foreground and background categories during model training of brain tumor segmentation.

## **1.2 Problem Statement**

Semantic Segmentation gives information to detect and helps doctors to know the disease's early stage. When segmenting medical images on brain magnetic resonance imaging, A convolutional neural network (CNN) is highly effective in segmenting medical images. Due to its capacity for delivering high-accuracy of segmentations. Deep neural networks have a substantial influence on predicting the presence of tumors. Segmenting brain tumors manually takes time, relying on the clinician's experience and skills. As a result, using convolutional neural networks, image segmenting of various brain tumor sub-groups is essential for patient medication. This thesis proposes a convolutional neural network for segmenting brain tumor sub-areas into enhancing tumor (E.T.), whole tumor (W.T.), and tumor core (T.C.).

## **1.3 Objectives**

One of the key objectives of the thesis is to study and analyze the relevant problems and evaluate residual unet network performances to meet the purposes of Semantic Segmentation implementations.

The principal goals for this thesis can be concluded as follows: -

- I. To build a system capable of detecting, segmenting, and classifying mass areas in M.R. images for the identification of brain tumors. In addition, predict the molecular subtypes of masses based on segmented form characteristics.
- II. Modify and expand the residual net architecture such that it may be used with a small amount of training data and yet provide accurate segmentation.
- III. To study the performance of semantic segmentation utilizing Python based on TensorFlow.

The contribution of the proposed work is given below:

- To build a model that increases the network's depth while keeping an extremely high level of accuracy.
- Instead of 2D input, the provided model employs 3D input of MRI data, which provides more contextual and spatial information.
- The present work uses a combination of dice loss and focal loss to assist deep learning models in resolving class imbalance challenges.
- Three separate datasets are evaluated to demonstrate that the presented model is superior to its baseline models and many of the existing state-of-the-art segmentation approaches.

#### **1.4 Organizations Of The Thesis**

This thesis includes these chapters

- The first chapter is an introduction to the thesis. It consists of several sections, including the overview of the study, problem statements, objective of the project, and lastly, the organization of the study.
- Chapter two gives an overview of this project and the literature available on the evolutions of the arising semantic segmentation technology and residual unet network architecture. It also overviews the convolutional neural network architecture. The significance of semantic segmentation technology is briefly discussed as well.
- Chapter three contains the methods for implementing semantic segmentation and system designs. A detailed 3D residual unet design architecture will be used in this study.
- Chapter four focuses on and shows the results of the segmented image and will be discussed.
- Lastly, chapter five will establish the conclusions depending on the previous chapter's findings and provide recommendations for future work.

## Chapter 2

### LITERATURE REVIEW

#### 2.1 Overview

The human brain is a critical organ of the body as it regulates all the activities of the human body. Numerous disorders, including infections, strokes, and tumors, can affect the brain. Additionally, a brain tumor can be a malignant or non-cancerous mass of aberrant cell proliferation in the brain. The most critical technology for detecting a brain tumor is MRI. Recently, MRI medical image analysis has drawn attention to the need to efficiently and objectively evaluate huge amounts of data (Bauer et al., 2013). Magnetic resonance imaging (MRI) is critical for patients' medical care and human brain research since it allows for the detection of brain cancers and the automated categorization of brain tissue. Segmentation, which splits and isolates the image's objects for processing, is the most vital step in the medical imaging processing of a magnetic resonance image.

This chapter gives basic information on brain tumors and the procedures used to detect tumors. Additionally, this chapter addresses the fundamentals of MRI. The Convolutional neural networks are described. In addition, this chapter also summarizes several related studies about this thesis.

#### 2.2 Brain Tumor

The brain is a sophisticated organ that manages every physical activity, such as intellect, creativity, feelings, and memory. The cerebrum, cerebellum, and brainstem are the three parts of the brain encased inside the skull. In addition, the brain is linked to the spinal cord through the brainstem (Al-Qazzaz, 2020). The brain's largest structure is the cerebrum, which comprises the right and left hemispheres. It is in charge of vital functions, including touching, vision, hearing interpretations, voice production, thinking, feelings, learning, and fine motor control. In addition, muscle control, posture, and balancing are controlled by the cerebellum, positioned

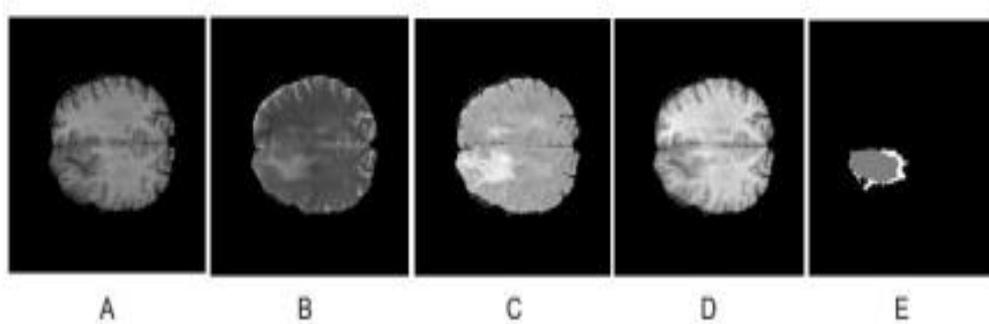
underneath the cerebrum. The midbrain, pons, and medulla make up the brainstem, a relay center for communication between the cerebrum and cerebellum, and the spinal cord (Al-Qazzaz, 2020). The cerebral cortex is the surface tissue of the cerebrum's outer layer. The cerebral cortex is a greyish, foldable region in the brain that contains over 70% of the mind's 100 billion nerve cells (Al-Qazzaz, 2020; Miller, Wittek, & Joldes, 2010). Axons, lengthy nerve fibers that link neurons and extend under the cortex to create the white matter, are also found in the brain (Miller et al., 2010). The central nervous system is split into four hollow parts in the brain, referred to as ventricles. These ventricles carry cerebrospinal fluid (CSF), responsible for fluid circulation throughout the central nervous system (Al-Qazzaz, 2020). When looking at the photos, the most noticeable and recognized sections of the brain are the grey matter (G.M.), white matter (W.M.), and cerebral spinal fluid (CSF).

In the brain or surrounding the brain, a brain tumor is a cluster of aberrant cells that have grown out of control. Because of their structure, these cells may be recognized from those in the neighboring tissue (Al-Qazzaz, 2020; Azhari, Hatta, Htike, & Win, 2014; thebraintumourcharity). Brain tumors are malignant tumors that begin in the brain and then migrate to other parts of the body. In other words, when tumor cells from another part of the body migrate via the bloodstream to the brain, a process known as metastasis. Secondary or spreading brain tumors are more common in older people (thebraintumourcharity). The four types of brain tumors include gliomas, meningiomas, pituitary adenomas, and nerve sheath tumors. Tumors are graded on a scale of 1 (least progressed) to 4 (most advanced) by the World Health Organization (WHO) (Louis et al., 2007). The World Health Organization (WHO) published a revised tumor categorization depending on histology and molecular criteria in 2016 (Louis et al., 2016).

### **2.2.1 Low Grade Glioma**

LGGs are the main tumor that starts in the brain or spinal cord's glial (supporting) cells. Even though gliomas of grades 1 and 2 were formerly classed as low tumors, the World Health Organization reclassified this kind of tumor as a high tumor in 2016, according to the World Health Organization. Several researchers (Louis et al., 2016). Even

though low-grade gliomas may develop over the years, they often progress to high-grade gliomas throughout the period. In 2007, the World Health Organization (WHO) classified gliomas according to their histological subtypes (Louis et al., 2007). The 2016 categorization includes astrocytic and oligodendroglial tumors, which were separated into two categories based on the presence or absence of IDH mutations and the 1p/19q codeletion (Stieber, 2001). LGG tumors have a variety of imaging characteristics that may be used to distinguish them from their histological subtypes. For example, LGG shows hyperintense on T2 images and may have diffuse, undetectable edges or focused forms with clear margins. This view might reveal cysts, often associated with moderate edema (swelling). It is possible to enhance the appearance of LGG tumors by adjusting the contrast of the images. LGG tumors are shown in Figure 2.1 using a variety of MRI methods.

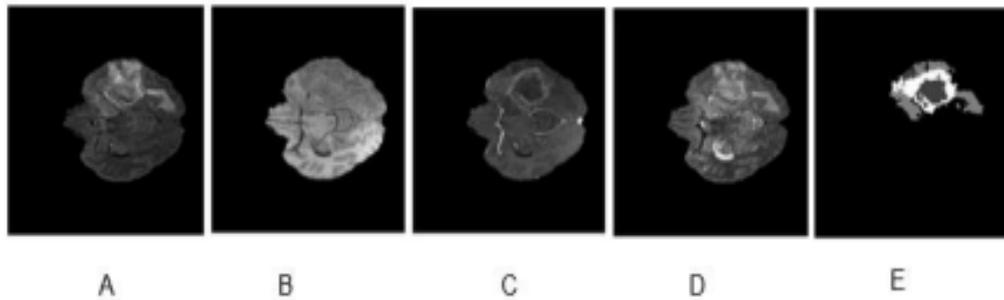


**Figure 2.1:** Magnetic resonance imaging of low-grade gliomas: A) FLAIR, B) T1 image, C) T1-ce image, D) T2weighted, E) ground truth.

### 2.2.2 High Grade Glioma

The fatal intra-axial glial tumor of the central nervous system is high-grade glioma. HGG is in charge of the great number of malignant primary brain tumors that fall into the WHO grade 3 and 4 classifications according to the World Health Organization (Louis et al., 2007). It is estimated that glioblastomas (GBMs) make up around 60 and 70 percent of all HGG, with anaplastic astrocytomas and anaplastic oligodendrogliomas coming in second and third place, respectively (Eftimov, Ivanov, Petkov, & Nakov, 2007). The outlook for people who have been diagnosed with HGG

remains bad. On magnetic resonance imaging (MRI), the HGG appears as a heterogeneous tumor with neighboring edema. The core portion frequently has necrosis and significant edema in the surrounding area (see Figure 2.2).



**Figure 2.2:** Magnetic resonance imaging of high-grade glioma: A) FLAIR, B) T1 image, C) T1-ce image, D) T2weighted, E) ground truth

### 2.3 Magnetic Resonance Images

MRI is an advanced medical imaging method for detecting and visualizing structural and functional aspects of the body's internal structure (Azhari et al., 2014; Hassan & Aboshgifa, 2015). The essential ideas of Nuclear Magnetic Resonance (NMR), a field of science that has been around and dates back to the early part of the twentieth century, are discussed here. Magnetic resonance imaging does not utilize X-rays to create images of the human body; rather, it employs a strong magnetic field and radio waves. After that, the images may be viewed on a computer's monitor or printed on photographic films. Because MRI gives substantially better brightness than C.T., it is particularly beneficial in diagnosing and treating neurological disorders. In addition, there are various benefits of MRI over other medical imaging modalities, including the fact that it is non-invasive and flexible, provides great tissue contrast, and is sensitive to flow and diffusion.

Several modalities are often acquired during a clinical M.R. examination. It is usual in clinical practice to acquiring the following M.R. modalities for patients with brain tumors: 1) a T1 picture, 2) a T1 image with contrast agent (T1c), 3) a FLAIR image, and 4) a T2 image. These four modalities (T1, T1c, T2, and FLAIR) are a globally

acknowledged imaging strategy for acquiring brain tumor MRI and are particularly prevalent in brain tumor segmentation (Menze et al., 2014). It would be highlighted that some organizations may acquire and further study an additional set of M.R. modalities for brain tumor patients, for example, as part of medical research initiatives or to get more insight into tumor metabolism. (See Figure 2.3)

### **2.3.1 T1-Weighted**

In situations where both the T.E. and the T.R. are short, a T1-weighted protocol is obtained. This has therapeutic relevance since tissues with low T1 values look bright on T1-weighted imaging, whereas tissues with high T1 values seem dark. The contrast between fluids, water-based tissues, and fat-based tissues is frequently superior to other methods when using this process. This method creates a great contrast between grey matter and white matter in brain images. To use a low T.R. number of around 500 ms, 3 unique tissue kinds (i.e., CSF, brain tissue, and fat) may be easily distinguishable. When viewing a T1-weighted image, the Cerebrospinal fluid seems to be hypo-intense, whilst the brain tissues (G.M. and W.M.) appear to be moderately intense, and fat appears to be significantly hyper-intense. As an alternative to contrast agents, patients can be given injections of contrast agents to help in enhancing specificity by delivering numerous images with different contrast levels (Al-Qazzaz, 2020).

### **2.3.2 Contrast-Enhanced T1-Weighted**

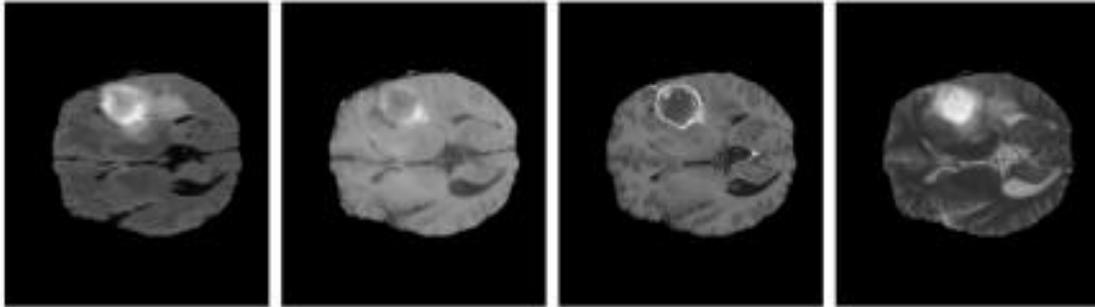
It is possible to increase the contrast of a T1-weighted image (T1ce) by employing a contrast agent derived from low-molecular-weight compounds like gadolinium. Contrast-enhancing drugs benefit brain tumor images due to their molecules do not cross through the blood vessel walls found in normal brain tissues and stay within the vessels. The blood-brain barrier is breached in the case of malignant primary brain tumors, allowing the contrast agent to exit the arteries and enter the intracranial area. Damaged tissue will have a shorter T1 on T1-weighted imaging, creating a strange appearance of hyperintensity in the surrounding healthy tissue (Bauer et al., 2013).

### 2.3.3 T2-Weighted

When both the T.E. and the T.R. are long enough, a T2-weighted procedure may be achieved. This approach makes it possible to distinguish between different types of brain cells (G.M., W.M., CSF) and scalp fat when analyzing the T2 values of brain tissues (Roberts & Mikulis, 2007). Specifically, spin interactions occur more quickly in more mobile tissue fluid areas, with a slower loss of transverse coherence. This results in a longer T2 duration in fluid regions of more mobile tissues. Furthermore, more restricted structures, such as those containing a dense population of cells, have longer spin interactions and faster transverse coherence loss, resulting in a shorter T2 time than more expansive ones. Fluids look bright on T2-weighted scans, but water- and fat-based tissues appear mid-grey on the same scans, respectively. Tumors often cause harm to the brain's microstructures, causing the T2 levels in the damaged tissue to rise as a result.

### 2.3.4 FLAIR

The FLAIR image is created by using a  $180^\circ$  R.F. pulse rather than a  $90^\circ$  R.F. pulse to generate a fluid-attenuated inversion recovery (FLAIR) image; this contributes to the expansion of the dynamic range in T1-weighted pictures. When an R.F. pulse is used to completely invert the longitudinal magnetization of all tissue under  $B_0$ , it is referred to as an "inverting pulse." Consequently, the magnetization starts with a negative number and decreases until it reaches zero. It is the most extensively used method for identifying brain cancers with non-enhancing lesions. It gives a T2-weighted image with a diminished Cerebrospinal fluid signal, which is then compared to other procedures. (See Figure 2.3)



**Figure 2.3:** Magnetic resonance imaging of a brain tumor: (left to right) FLAIR, T1-weighted, T1-ce, T2- images

### 2.3.5 Ground Truth

The MR pictures in the Virtual Skeleton Database (VSD) system are augmented with handwritten annotations, referred to as ground truths. It is possible to obtain information on the locations of different tumor kinds from the ground truth. Intratumoral developments can be classified into edema, necrosis, and a larger tumor. These tumor types were identified by a group of radiologists and 7 radiographers who had received specialized training. The tumor structures were identified and tagged on each third axial slice of MRI. Additionally, morphological operators (region growth) were employed to interpolate the segmentation findings, and the outcomes were visually inspected to determine if any human changes were required (Menze et al., 2014).

The following approach was used to manually annotate tumor structures in both low- and high-grade instances in the BRATS challenge (Menze et al., 2014) (see Figure 2.4):

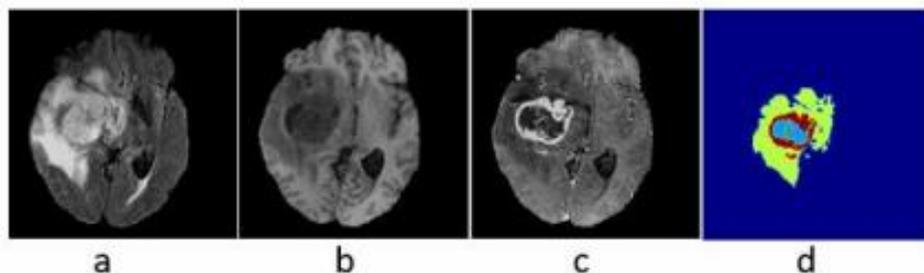
- 1- Segmentation of edema was performed using T2 and FLAIR images. The initial segmentation was performed using T2. Then, FLAIR was utilized to determine the extent of edema and differentiate it from other tissues such as necrosis and ventricles.

2- The whole tumor core, which contained all three tumor structures, was segmented utilizing the hyper-intense areas in T1-ce and the inhomogenous area of hyper-intense and hypointense lesions in T1.

3- The tumor's enhanced core was segmented utilizing T1-ce by thresholding on the whole tumor core region, which resulted in the gadolinium being retained and the necrotic tissue is removed. Visual inspection was used to determine the threshold levels for each case.

4- T1-ce was defined as necrosis or a fluid-filled core within the increased tumor's low-intensity structures. For the rare hemorrhages, the same designation was suggested.

5- The remaining portion of the whole tumor core was the non-enhanced core. They were derived by removing the corresponding areas from the whole tumor, as they were not detected in lists 3 and 4.



**Figure 2.4:** (a) Complete tumor visible in FLAIR; (b) tumor core visible in T2; (c) enhanced and necrotic tumor component structures visible in T1ce; (d) final labels of the observable tumor structures noticeable: edema (yellow), necrotic/cystic core (light blue), enhanced core (red).

## 2.4 Artificial Neural Network Model

A neural network is a consider obtained of how the human brain processes data. It simulates a huge number of linked processing units that simulate abstract representations of neurons in order to function. The processing units are organized

hierarchically. In a neural network, there are generally three layers: an input layer with units representing the input fields, one or more hidden layers, and an output layer with units representing the target field (Modeler, 2021).

### 2.4.1 Basics In Artificial Neural Networks

Artificial neural networks (ANNs) are mathematical or computational models that may be used to approximate or estimate functions in various situations. ANNs are artificial neural networks intended to imitate the biological neural networks that make up animal brains. Like biological neural networks, artificial neural networks (ANNs) are composed of many nodes (artificial neurons) that may be used to mimic complex data interactions. Modern artificial neural networks (ANNs) are a nonlinear dynamic system that is often enhanced through a mathematical statistics-based learning procedure. We may acquire many local structural spaces that functions can characterize by employing standard statistical mathematics techniques. However, in the realm of artificial perception, automatic decision-making may be accomplished by using quantitative statistics. In other words, artificial neural networks (ANNs) can make fundamental decisions and judgments, much like humans, utilizing statistical methodologies. This method outperforms formal logical deductions in most situations (Sze, Chen, Yang, & Emer, 2017; J. Wang, 2019). ANNs, along with other machine learning methodologies, have been used for a variety of problems that are difficult to solve with standard rule-based algorithms, such as computer vision, natural language processing, and recommendation systems.

#### 2.4.1.1 Artificial Neurons

Artificial neurons are the basic unit of ANNs. They mainly simulate the structures and characteristics of biological neurons by accepting a set of input signals and producing outputs. Assume a neuron accepts  $d$  inputs  $x_1, x_2, \dots, x_d$ , we use vector  $\mathbf{x} = [x_1; x_2; \dots; x_d]$  To denote this group of input, and the net input  $z \in \mathbb{R}$  to denote the weighted sum of  $\mathbf{x}$ .

$$z = \sum_{i=1}^d w_i x_i + b = \mathbf{w}^T \mathbf{x} + b \dots \dots \dots (2.1)$$



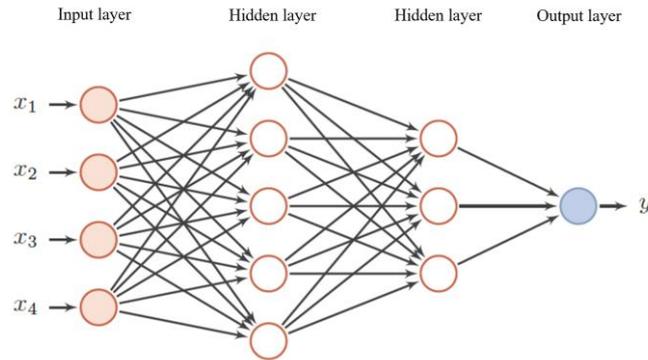
$$\text{ReLU}(x) = \max(0, x) \dots \dots \dots (2.3)$$

ReLU requires only additions, multiplications, and comparing operations from neurons, which is more computationally efficient. In addition, compared to other activation functions mentioned in the last paragraph that squashes the output range, ReLU is only left saturated, which alleviates the vanishing gradient problem and accelerates the learning process to some extent. One disadvantage of ReLU is that sometimes a neuron may never get activated during training if its output is zero for all data, which leads to the “dying ReLU” problem (Chen, Sathe, Aggarwal, & Turaga, 2017). To address this issue, several variants of ReLU have been proposed and widely used in practice, such as leaky ReLU (Maas, Hannun, & Ng, 2013), parametric ReLU (PReLU) (He, Zhang, Ren, & Sun, 2015), and exponential linear unit (ELU) (Clevert, Unterthiner, & Hochreiter, 2015).

#### 2.4.1.2 Feedforward Neural Network

When neurons receive information in a feedforward neural network (FNN), they are divided into groups based on the sequence in which they receive it. Each group may be regarded as a layer of the brain on which to operate. Each layer's neurons receive the previous layer's output and calculate outputs used as inputs by the following neurons. The input layer is the first layer, the output layer is the last, and the hidden layers are the intermediate levels between these two levels of representation. The data is only delivered in one direction, and there is no way to get feedback. To explain the FNN model, a directed acyclic graph might be employed. Figure. 2.5 displays the construction of a fully connected FNN (Nielsen, 2015), which has no intra-layer connections and links every pair of neurons in nearby layers. This FNN has no intra-layer connections and connects every pair of neurons in neighboring layers.

The feedforward network as a whole can be perceived as a function. Multiple compounds of simple nonlinear functions accomplish the complicated mapping from the input to the output space. This network topology is simple and quick to construct.



**Figure 2.5:** Feedforward neural network

**Source:** (Nielsen, 2015)

We use the following notations to describe an FNN:

- $L$ : the number of layers in the network
- $M^{(l)}$ : the number of neurons in layer  $l$
- $f(\cdot)$ : activation function of the neurons in layer  $l$
- $\mathbf{W}^{(l)} \in \mathbb{R}^{m^{(l)} \times m^{(l-1)}}$ : weight matrix from layer  $l - 1$  to layer  $l$
- $\mathbf{b}^{(l)} \in \mathbb{R}^{m^{(l)}}$ : bias vector in layer  $l$
- $\mathbf{z}^{(l)} \in \mathbb{R}^{m^{(l)}}$ : net inputs of neurons in layer  $l$
- $\mathbf{a}^{(l)} \in \mathbb{R}^{m^{(l)}}$ : outputs of neurons in layer  $l$

Signals are determined according to these equations:

$$\begin{aligned} \mathbf{z}^{(l)} &= \mathbf{W}^{(l)} \cdot \mathbf{a}^{(l-1)} + \mathbf{b}^{(l)} \\ \mathbf{a}^{(l)} &= f_l(\mathbf{z}^{(l)}) \end{aligned} \quad \dots \quad (2.4)$$

In this way, information is transmitted layer by layer until the final output  $\mathbf{a}^{(L)}$  is computed. The whole network can be expressed as a compound function  $\varphi(\mathbf{x}; \mathbf{W}, \mathbf{b})$ :

$$\mathbf{x} = \mathbf{a}^{(0)} \rightarrow \mathbf{z}^{(1)} \rightarrow \mathbf{a}^{(1)} \rightarrow \mathbf{z}^{(2)} \rightarrow \dots \rightarrow \mathbf{z}^{(L)} \rightarrow \mathbf{a}^{(L)} = \varphi(\mathbf{x}; \mathbf{W}, \mathbf{b}) \dots \quad (2.5)$$

Where  $\mathbf{W}$  and  $\mathbf{b}$  represent the weights and biases of all layers, respectively.

For simplicity, in the following discussion, we omit the parameters  $W$  and  $b$  in the notation of the compound function. In classification problems, the output of the network  $j(x)$  is passed into a classifier  $g(\cdot)$ , then the output of the classifier is  $\hat{y} = g(\varphi(\mathbf{x}))$ . Depending on the specific task, one can choose an appropriate loss function  $\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}})$ , where  $y$  is the label corresponding to the input  $x$ . The objective of the network is to learn a parametric function that can fit the data well, which is achieved by minimizing the loss function during the training process. Therefore, after being trained, the network can be used for performing predictions with input data that have not been shown to it. The parameters in the network are updated by stochastic gradient descent (Bottou, 2010), which calculates the partial derivatives of the loss function with respect to the parameters. In neural networks, this is approached by backpropagation (Chauvin & Rumelhart, 2013). I will not discuss these algorithms in this thesis. For more details, please refer to the references.

#### **2.4.2 Convolutional Neural Networks**

CNNs are a deep FNN model with local connections and weight sharing. CNNs are mainly used in many tasks related to image or video processing (J. Wang, 2019). Compared to the fully connected network discussed previously, CNNs have two major advantages:

1. The size of convolution kernels in each layer is much smaller than the inputs, and the weights can be shared in the same layer. At the same time, in fully connected networks, the number of connections can be huge, especially when the input signal is high-dimensional and the network becomes deeper. Therefore, the number of parameters in CNNs is much smaller, which boosts the training efficiency and reduces the chance of overfitting.
2. The architecture of CNNs allows them to capture the local invariance of objects in images. For example, zooming, translating, or rotating images should not change the semantic information. It is very difficult for fully connected networks to achieve this, as spatial information is absent.

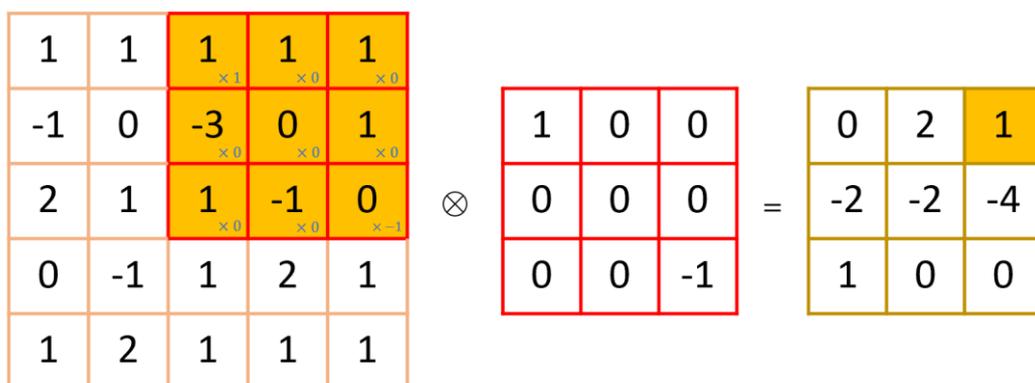
Generally, convolutional blocks in modern CNNs are formed by stacking convolutional layers and pooling layers, in which high-level features are extracted. Depending on the inference objective, fully connected or transpose convolutional layers are used after convolutional blocks. The entire network can be trained end-to-end through backpropagation.

### 2.4.2.1 2D Convolution

The 2D convolution is equivalent to the cross-correlation operation in deep learning, which does not involve flipping the kernel. Given an image  $X \in \mathbb{R}^{M \times N}$  And a filter  $W \in \mathbb{R}^{(m \times n)}$  (usually  $m \ll M, n \ll N$ ), the convolution is computed by

$$y_{ij} = \sum_{w=1}^m \sum_{v=1}^n w_{wv} \cdot x_{i+w-1, j+v-1} \dots \dots \dots \quad (2.6)$$

The spatial step that the convolution kernel jumps during sliding is called the stride. The size of the output decreases as the stride increases. We use the notation to denote the convolution operation in the following discussion. Figure 2.6 shows an example of 2D convolution with a stride of 1.



**Figure 2.6:** 2D convolution

**Source:** (Wang, 2019)

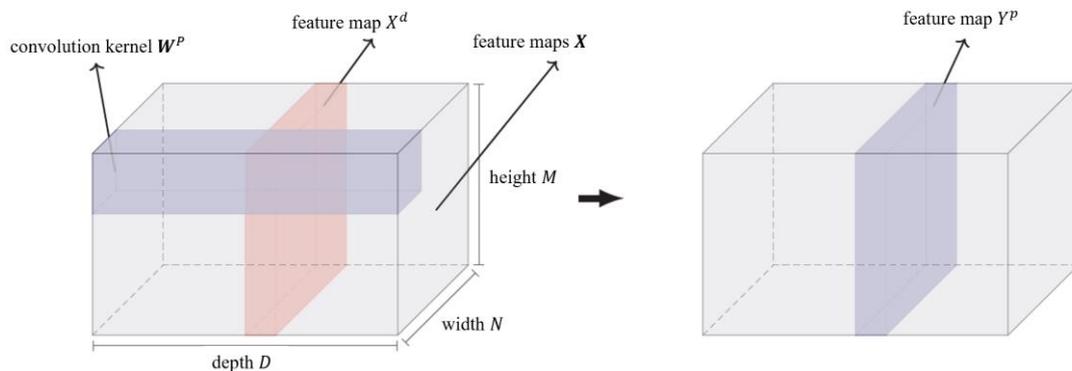
As shown in Figure 2.6, the convolution operation causes the size of the output to shrink. To keep the output size the same as the input, zero padding is often applied in

CNNs. In a one-dimensional example, assume the input size is  $M$ , the size of the kernel is  $m$  ( $m$  is usually an odd number for simplicity in CNNs), and we pad  $p$  zeros to both ends of the input. Then, in the scenario of unit stride, set  $p = (m - 1)/2$  will maintain the output size equal to  $M$ . This is called the *equal-width convolution*, which is the default type of convolution used in CNNs.

### 2.4.2.2 Convolution Layer

The purpose of the convolutional layer is to obtain information from a local region. Different kernels of convolution relate to distinct feature extractors. Normally, the input  $X$  of a convolution layer is organized into a 3D tensor with shape  $M \times N \times D$ , which can be interpreted as  $D$  feature maps, each with a size of height  $M \times$  width  $N$ . In the input layer, the feature map is the input image itself; therefore  $D = 1$  for grayscale pictures, and  $D = 3$  for natural Color-coded images. Generally, the structure of a convolution layer can be decomposed as follows:

The 3D representation of a convolution layer is shown in Figure 2.7.



**Figure 2.7:** 3D representation of a convolution layer.

**Source:** (Nielsen, 2015)

To compute feature map  $Y^p$ , we use kernel  $W^{p,1}, W^{p,2}, \dots, W^{p,D}$  to convolve with feature map  $X^1, X^2, \dots, X^D$  respectively. The outputs of the convolutions are summed

up and then added with a scalar  $b^p$  to get the net input  $Z^p$ . The final output feature map  $Y^p$  is produced by passing  $Z^p$  into an activation function.

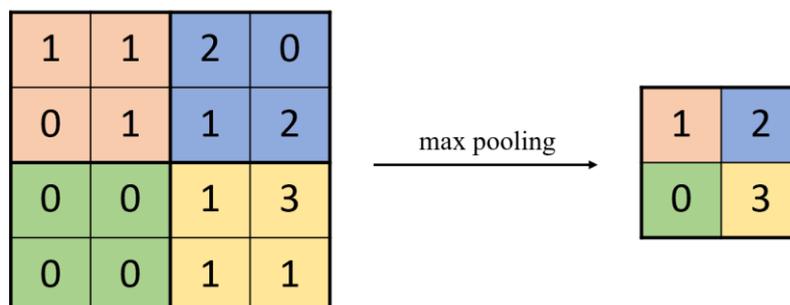
$$\begin{aligned} Z^p &= \mathbf{W}^p \otimes \mathbf{X} + b^p = \sum_{d=1}^D W^{p-d} \otimes \mathbf{X}^d + b^p \dots\dots\dots (2.7) \\ Y^p &= f(Z^p) \end{aligned}$$

To get  $P$  output feature maps, the operation above is repeated  $P$  times using  $P$  distinct kernels  $\mathbf{W}^p \in \mathbb{R}^{m \times n \times D}$ . A convolution layer requires  $P \times D \times m \times n + P$  parameters if both the weights and biases are counted.

### 2.4.2.3 Pooling Layer

Pooling layers are also called subsampling layers, whose role is to perform feature selection. This operation reduces the number of features; thus, the network's number of parameters is reduced to minimize overfitting. Each feature map in the inputs to a pooling layer can be parsed into many regions (non-overlapping in most CNN applications). The pooling operation refers to down sampling the region to a single value, which is the abstract of this region. Max pooling is the most common type of pooling operation in CNNs. It should be noted that pooling also makes the network invariant to some small morphological changes and enlarges the receptive field.

Figure 2.8 shows an example of a max-pooling operation using a sampling window size of  $2 \times 2$ . The input 2D feature map is downsampled by a factor of 2 both horizontally and vertically. The sampling window size is usually small; otherwise, it will cause excessive loss of information.



**Figure 0.8:**  $2 \times 2$  max-pooling

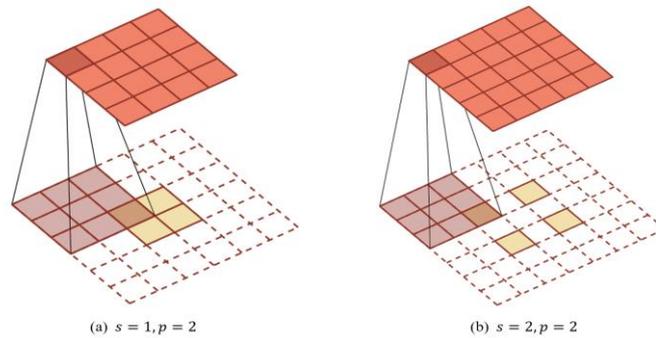
**Source:** (Wang, 2019)

#### 2.4.2.4 Transposed Convolution

We can generally convert high dimensional features to low dimensional features through convolutions. However, in some tasks, it is demanded to map low dimensional features to high dimensional features, and we still want to achieve this through convolution operations. This type of convolution is called the *transposed convolution*.

Given a high dimensional vector  $\mathbf{x} \in \mathbb{R}^d$  And a low dimensional vector  $\mathbf{z} \in \mathbb{R}^p, p < d$ . The mapping from the high dimension to the low dimension can be expressed as an affine transformation  $\mathbf{z} = W\mathbf{x}$ , where  $W \in \mathbb{R}^{p \times d}$  The inverse mapping can be simply accomplished by transposing  $W$ , i.e.,  $\mathbf{x} = W^T\mathbf{z}$ . Similarly, in CNNs, the 2D convolution operation can be viewed as an affine transformation; therefore, the mapping from low dimensional feature maps to high dimensional feature maps is named the transposed convolution. Note that the purpose of the transposed convolution is not to restore the previous high dimensional feature maps but up sample the features back to higher dimensions. Therefore, the weights in transposed convolutional layers can also be learned.

The operation of transposed convolutions is the same as convolutions. Assume the size of the input is  $p$ , and the size of the convolution kernel is  $m$ . If we pad  $m-1$  zeros to both input ends, we can get an output with size  $p + m - 1$ . A more efficient upsampling approach is to increase the stride  $s$ , which pads zeros between the elements in the input feature maps. Figure 2.9 shows the effect of using different strides (Nielsen, 2015). For a  $2 \times 2$  input feature map and a  $3 \times 3$  kernel, by increasing  $s$  from 1 to 2, the size of the output feature map expands from  $4 \times 4$  to  $5 \times 5$ .



**Figure 2.9:** Transposed convolution

**Source:** (Nielsen, 2015)

#### 2.4.2.5 Dilated Convolution

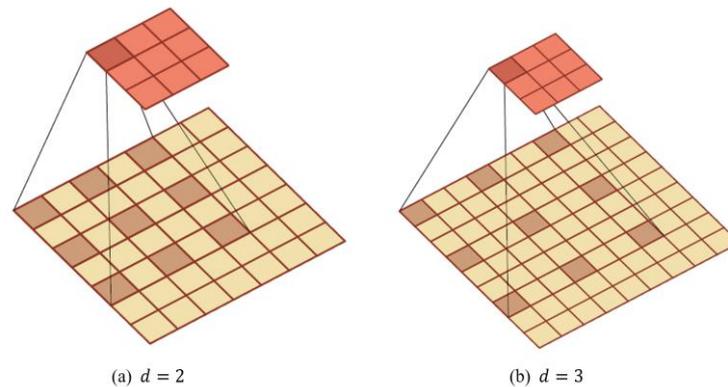
Many CNN architectures adopt pooling layers to enlarge the receptive field of each unit in the output feature maps without introducing excessive parameters (Krizhevsky, Sutskever, & Hinton, 2012a; Long, Shelhamer, & Darrell, 2015; Simonyan & Zisserman, 2014). However, this engenders a loss of information, which is negative for some tasks requiring dense spatial information. The dilated convolution reduces this loss while expanding the receptive field with the same number of parameters.

The dilated convolution operation inserts “holes” between the adjacent entries in the convolution kernels. Assume the size of the kernel is  $m$ . If  $d - 1$  holes are inserted, then the effective size of the kernel  $m'$  is given by

$$m' = m + (m - 1) \times (d - 1) \dots \dots \dots (2.8)$$

Where  $d$  is the dilation rate, note that if  $d = 1$ , the operation is a regular convolution.

Figure 2.10. shows examples of the dilated convolution. The original kernel size is  $3 \times 3$ . The effective kernel size is  $5 \times 5$  and  $7 \times 7$ , respectively, with  $d = 2$  and  $d = 3$ .



**Figure 2.10:** Dilated convolution

**Source:** (Nielsen, 2015)

## 2.5 Previous Work

There are various computer-aided methods for medical imaging segmentations, including threshold-based, region-based, model-based, and artificial neural network (ANN) approaches (Fu et al., 2018; Wenlu Zhang et al., 2015), which are discussed below. However, using a convolutional neural network (CNN) for medical image segmentation has been particularly effective in recent studies (Fu et al., 2018; Wenlu Zhang et al., 2015). Deep learning may be accomplished by using a convolutional multi-layer neural network with a large number of hidden layers and free parameters. The ultimate decision-making stages are reached after each Magnetic resonance imaging input image has been processed through several convolutional layers and a pooling layer, filters, and fully linked layers (Patterson & Gibson, 2017).

Several designs for deep learning networks, such as CNN, are utilized, including the deep residual Network (DRN), the deep feed-forward network, the deep convolutional neural network (DCNN), and the Unet. Deep learning networks, such as CNN, are used in several applications. When comparing these deep learning approaches, CNN has proven to be the most commonly used in image processing due to its distinctive framework, which contains an input layer, extracting features using convolutional layers. This activation layer uses a rectified linear unit (ReLU), different pooling

layers, and classifying layers (Patterson & Gibson, 2017). As a result, CNN has lately acquired prominence as a resource for evaluating medical images.

Because of the convolutional layers of CNNs, they can capture a considerable degree of non-linear matching between inputs and outputs (Maharjan, Alsadoon, Prasad, Al-Dalain, & Alsadoon, 2020; Mlynarski, Delingette, Criminisi, & Ayache, 2019; Qamar, Jin, Zheng, Ahmad, & Usama, 2020). As a result, CNN can extract complex characteristics from pictures of the brain. First, MRI brain images are recovered and fed into the CNN as inputs, with patches of information extracted from each image. Next, trainable convolutional filters and local subsampling extract the difficult features (Işın, Direkoğlu, & Şah, 2016; Mohan & Subashini, 2018).

Semantic segmentation is, without a doubt, among the most important medical image processing techniques since it provides extensive information on essential diseases such as brain tumors. The extraction of a precise tumor structure has proven to be a tough task, despite the various studies that have been undertaken on brain tumor segmentation. In this context, it presents a concise summary of numerous important results on brain tumor segmentation: -

Sérgio Pereira et al. (Pereira, Pinto, Alves, & Silva, 2016). Proposed an automatic segmentation system based on Convolutional Neural Networks (CNN) that analyzed tiny  $3 \times 3$  kernels of information. It was shown that using intensity normalization as a pre-processing step for brain tumor segmentation in MRI images, which is unusual in CNN-based segmentation algorithms, was extremely successful when paired with data augmentation. The suggested approach was tested on the BRATS 2013 and 2015 databases, and both were successful. The BRATS 2013 data set was able to acquire 0.88, 0.83, and 0.77 for the full region, the core area, and the enhancing region, respectively. The BRATS 2015 data set achieved 0.78, 0.65, and 0.75 for the complete, core, and enhancing regions.

Mohammad Havaei et al. (Havaei et al., 2017) presented a fully-automatic brain tumor segmentation approach based on Deep Neural Networks developed in their laboratory (DNNs). The networks that have been proposed are customized to high and low grades

that have been identified using magnetic resonance imaging (MRI). Brain tumors can occur anywhere and can be of any size, form, and contrast. Their CNN uses both local features and more global contextual elements simultaneously, which is unique. The last layer of their networks is a convolutional version of a fully connected layer, resulting in a 40-fold increase in runtime compared to most standard CNN applications. BRATS 2013 online assessment system results suggest that their best model outperformed the presently published state-of-the-art technique in terms of accuracy and speed, as demonstrated in MICCAI 2013 and BRATS 2013.(Havaei et al., 2017)

Peter D. Chang created and developed an (FCR-NN) utilizing convolutional residuals for medical image segmentation using linear identity mappings. FCR-NN is an image segmentation architecture that mixes optimization improvements from residual identification maps with a fully convolutional model that successfully accommodates low & high-level picture properties. When two different models were trained, one for complete tumor classification and the other for cells within a cell sub-region classifying, the serial FCR-NN architecture outperforms the state-of-the-art with Dice coefficients of 0.87, 0.81, and 0.72 for complete tumor classification, core tumor classification, and enhancing tumor validation, respectively.(Havaei et al., 2017)

Dinthisrang Daimary et al. (Daimary, Bora, Amitab, & Kandar, 2020) conducted hybridized convolutional neural networks using a mixture of U-SegNet, Seg-UNet, and Res-SegNet methods. To capture fine and detailed information, the Unet architecture uses skip connections. However, it requires more processing time for training than the SegNet design. As a result, the proposed hybrid models offer more effective outputs than other CNN models that are already accessible. The U-SegNet, Res-SegNet, and Seg-UNet each achieved a mean accuracy of 91.6 percent, 93.3 percent, and 93.1 percent, respectively, compared to the other networks. Because hybrid architectures include more layers and trainable parameters than traditional designs, training is no longer required once the model has been trained. However, the system can separate brain tumors from MRI images in seconds.

Salma Alqazzaz et al. (Alqazzaz et al., 2019) suggested a SegNet for automatic brain tumor segmentation, with each modality being trained separately and the SegNet output integrated during post-processing. Pre-processing begins with the inputs being pre-processed to remove unwanted artifacts while also improving segmentation outcomes via normalization and bias field correction. The SegNet network was then used to train each of the four MRI modalities on its own, independently of the others (Alqazzaz et al., 2019). The architecture is comprised of two components: an encoder (for downsampling) and a decoder (for up-sampling) (upsampling). To achieve this, the encoder includes 13 convolutional layers with three-by-three filters, batches of normalization layers and ReLU, and, lastly, max-pooling layers with two-by-two filters. In addition, the decoder has 13 convolutional layers that are used to correlate with the encoder's output. The enhanced characteristics of the decoder are passed to SoftMax layers, which classify each class pixel individually based on its location in the image. The segmentation approach had an accuracy of 85 percent for the overall tumor, an accuracy of 81 percent for the core tumor, and an accuracy of 79 percent for the expanding tumor (Alqazzaz et al., 2019).

Muhammad Saeed et al. (Saeed et al., 2021). Recommended a New Residual Mobile Unet model for BraTS from Magnetic Resonance Images; RMUnet is a hybrid deep learning model capable of segmenting tumors with high accuracy. The MobileNetV2, ResNet, and Unet networks are all included in the design. The inclusion of residual blocks to MobileNetV2 has allowed it to learn more in-depth features, which has improved its performance. The proposed network uses a modified Mobile Net V2 as an encoder and up-sampling layers of Unet s as a decoder, with the modified Mobile Net V2 serving as the encoder. In the BraTS 2020 dataset, the RMUnet accomplished dice coefficient scores of 0.91, 0.88, and 0.83 for complete, core, and enhancing tumors, respectively. On the BraTS 2019 dataset, the RMUnet achieved dice coefficient scores of 0.91, 0.91, and 0.83; on the BraTS 2018 dataset, the RMUnet achieved dice coefficients of 0.90, 0.86, and 0.79.

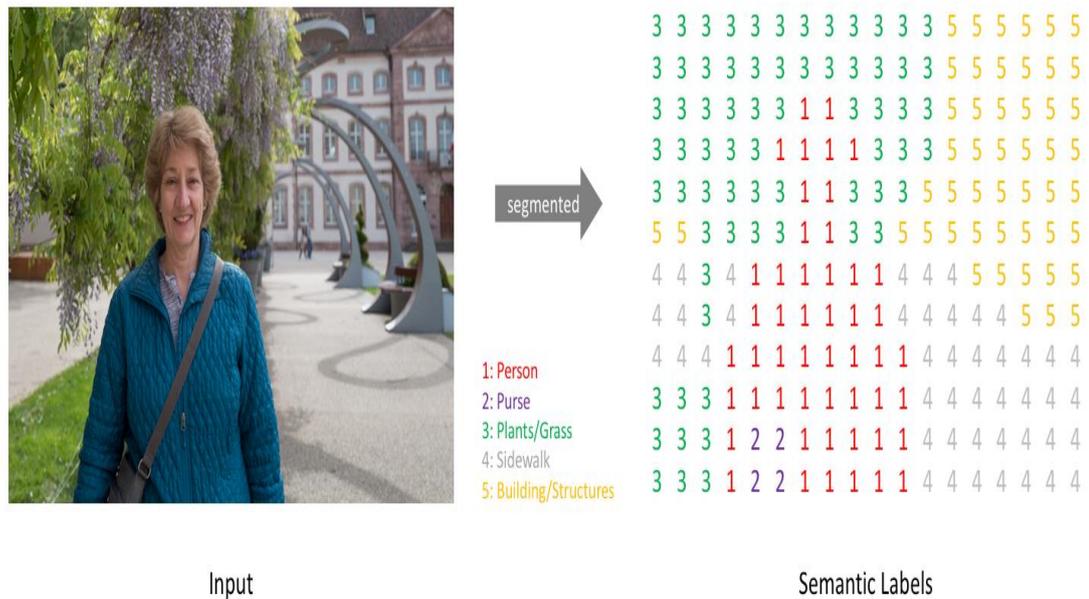
## 2.6 Semantic Segmentation

Image segmentation is currently the most widely used technology in predictive images. Indeed, most computer vision tasks above need intelligent image segmentation to completely comprehend the image's content and facilitate easier analysis between different parts of the image. The image segmentation technology based on deep learning employs convolutional neural networks to understand the real-world objects represented by each pixel in the image, which was unimaginable before. Image segmentation techniques based on deep learning are mainly divided into semantic segmentation and instance segmentation.

Semantic segmentation of images is an essential and difficult computer vision topic that researchers have been tackling for a long period. Essentially, it is the challenge of comprehending the contents of an image at the pixel level. Semantic segmentation is a classic computer vision problem, which takes some raw data as input (such as a two-dimensional (2D), Three-dimensional (3D)) and converts it into a mask defining an area of interest. More precisely, semantic image segmentation categorizes each pixel in an input image into a predetermined category, with each category corresponding to a particular object or portion of the image. This problem of per-pixel prediction is also known as a dense prediction. Semantic segmentation enables the identification of the categories of items contained in an image and their shapes and contours. But objects in the same category will not be distinguished. This provides an in-depth analysis of the problem. As a result, developing an automated semantic segmentation system is a high-priority topic of research in robotic vision, autonomous driving, and medical image processing, among other fields.

Figure 2.11 illustrates a semantic segmentation example. Each pixel in the input image is labeled with one of the labels from the collection (human, Purse, Plants, sidewalk, building) (Lamba, 2019, Feb 17; Mody, 2018, August 7; Mwiti, 2019, July 29; J. Wang, 2019). The objective is to take an image and generate an output that contains a segmented image in which each number of pixels (between 0 and 255) in the input image is converted to a class label value (0, 1, 2,..., n) (Jordan, 2018). Semantic segmentation in medical imaging is frequently associated with the label imbalance

problem. For instance, tumors can be extremely small in tasks involving tumor segmentation, and abnormal cases may account for only a small portion of the dataset. There are significantly fewer tumor samples than non-tumor samples in this scenario, which complicates capturing tumor characteristics.



**Figure 2.11:** Example of semantic image segmentation

**Source:** (Jordan, 2018)

The purpose of semantic segmentation technology is to understand the image content from the pixel level and to output a segmentation map of an RGB image or a single-channel grayscale image, which tells the category of each pixel in the image in a class label.

## Chapter 3

### METHODOLOGY

#### 3.1 Introduction

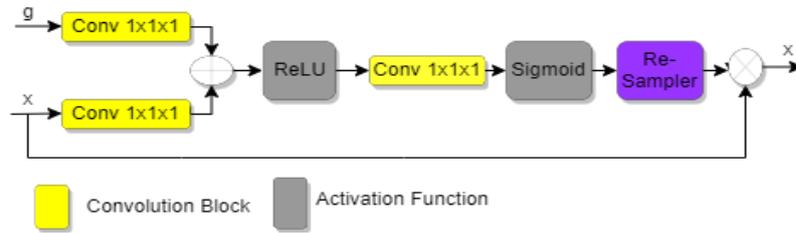
This section describes the network architecture specifics of the proposed model, which is an increased depth of 3D Residual Unet Network Architecture with an attention gate for brain tumor segmentation and the loss functions used in the training process. In addition, this part explains implementation details, and model training.

#### 3.2 Network Architecture

The design of the suggested architecture is shown in Figure 3.3. The various components of the model, especially the attention gate and residual block, are described below.

##### 3.2.1 Attention Gates

The attention mechanism concentrates on a certain portion of the image while paying less attention to the rest of the image. As with human visual attention, it can be trained to concentrate on a specified location or region while ignoring nearby areas in order to maximize performance. In addition, attention gates can help limit the number of false positives by reducing feature activation in unimportant image areas. As part of the current work, they support parameter updates in a model in geographical locations that are significant for tumor segmentation. As shown in Figure 3.1, the skip link utilizes an attention gate to establish a link between both the encoder and its equivalent decoder. The attention gate receives two parameters: The first input is from the appropriate encoder, which includes all of the contextual and spatial information for that layer, while the other comes from the below decoder layer, which acts as the gated signal. For concatenation purposes, the attention gate's output is also sent into the decoder through the attention gate's output.



**Figure 3.1:** Attention gate representation

Every attention stage receives inputs from the last attention layer and the equivalent resolution of the mainstream. For example, let  $s_l$  and  $m_l$  represent the attention and mainstream layers inputs at level  $l$ . First,  $s_l$  and  $m_l$  are connected, followed by a  $1 \times 1$  convolution layer  $C_{1 \times 1}$  and a sigmoid function  $\sigma$  to produce an attention map:

$$\alpha_l = \sigma(C_{1 \times 1}(s_l \parallel m_l)) \dots \dots \dots (3.1)$$

The outcome of the attention layer is then obtained by performing an element-wise multiplying on the input to the attention layer, written as:-

$$o_l = s_l \odot \alpha_l \dots \dots \dots (3.2)$$

**3.2.2 Residual Unet**

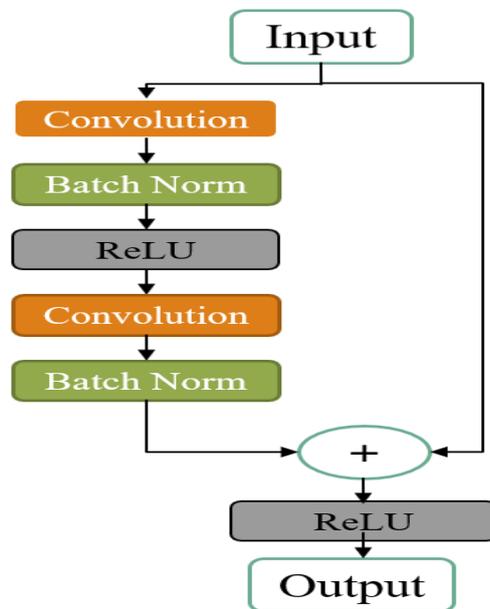
It is an encoder-decoder design for semantic segmentation created by Zhengxin Zhang et al. (S. Wang, Zhang, & Wu, 2019). In remote sensing, data analysis; was originally applied to the extraction of roads from high-resolution satellite photos. Afterward, researchers used it for numerous other applications, including brain tumor segmentation and human image segmentation. With a growing number of classification tasks and the complexity of recognizing, individuals have developed increasingly stringent criteria for the depth of convolutional neural network layers as convolutional neural networks have progressed through the development and study of the technology. It was introduced in 2015, and it connects two neural layers via a skip link, known as the residual network (Wilber, 2016). Figure. 3.2 shows how a skip link could perform this multi-layer identity function and learn the succeeding levels of a deep network's identity mapping. As a result, the network's ability to express features

can be maintained even as the depth of the network increases, gently resolving the issue of gradient disappearing or explosion caused by layer depth increase.

If one of the link layers is  $H(x)$ , the residual network block could be described as follows:

$$H(x) = F(x) + x \dots\dots\dots (3.3)$$

Where  $x$  represents the network-passed value,  $F(x)$  is the residual parameter, and  $H(x)$  is the predicted output.



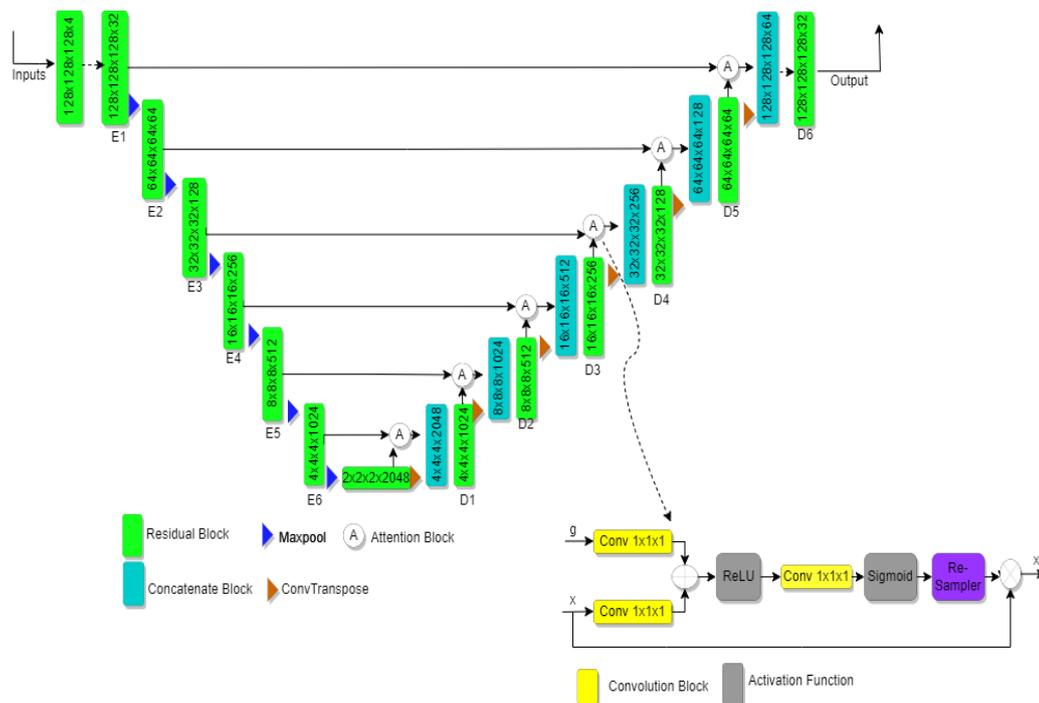
**Figure 3.2:** Residual block.

### 3.3 Increased Depth Of 3D Attention Residual Unet

Tumors must be segmented into distinct subregions to estimate their development and plan treatment, which is critical. On the other hand, Tumor classification is time-consuming and requires expertise. More complicated characteristics must be learned from data, and deeper networks must be used to accomplish this. However, with increasing depth, the network loses a corresponding amount of spatial information,

reducing segmentation accuracy. The UNet transmits contextual and spatial information between the encoder and the decoder using skip links.

Consequently, critical spatial information lost during down sampling can be recovered more effectively. However, all data is transmitted through skip links on the UNet network. Therefore, attention gates must be employed to combine only the decoder's relevant feature activations. Furthermore, each layer's characteristics are enhanced thanks to residual blocks, which help the UNet extract more information from each layer. Furthermore, to improve the accuracy of segmentation. This paper proposes the increased depth of 3D Residual Unet Network Architecture with an attention gate for brain tumor segmentation, as shown in Figure 3.3.



**Figure 3.3:** An increased depth of 3D Residual Unet Network Architecture with an attention gate.

### 3.3.1 Encoder Path

The encoder pathway is made up of six layers, each of which includes a Residual block and a  $2 \times 2 \times 2$  kernel size of a maximum pooling layer. There are four channels in the

proposed network, which correspond to the four modalities of the Magnetic resonance imaging datasets (Flair, T1, T1ce, and T2). The proposed network accepts an input of shape  $128 \times 128 \times 128 \times 4$ , where  $128 \times 128 \times 128$  is the image resolution of the input shape. The residual blocks are made up of  $3 \times 3 \times 3$  dilated convolutional layers, each of which follows a Batch Normalization (B.N.) layer and an activation function for the Leaky Rectified Linear Unit (Leaky ReLU). The first residual block contains 32 feature maps with a dimension of  $128 \times 128 \times 128$ . During the down-sampling block, the amount of kernels doubles, allowing the framework to learn complex structures efficiently, and the size of each feature map is reduced by half. As a result, the residual block at the sixth layer contains 1024 feature maps with a dimension of  $4 \times 4 \times 4$ .

### **3.3.2 Bottleneck**

The encoder and decoder of the suggested architecture are connected through the bottleneck layer. It possesses a residual block that has a dimension of  $2 \times 2 \times 2$ . And contains 2048 feature mappings. The convolution layers that make up the residual block each have a dilation rate of three. The output of the bottleneck layer is split between two distinct pathways. The first path goes into the decoder channel, followed by the convolutional transpose layer, while the second is the attention gate's gated signal.

### **3.3.3 Decoder Path**

Additionally, the decoder route comprises six layers, followed by a  $3 \times 3 \times 3$  convolutional transpose layer. Each decoder layer is linked to its matching encoder layer using a skip link and an attention gate.  $3 \times 3 \times 3$  convolutional transposition is used to achieve up sampling in this case. The outcome of the attention gate is combined with the up sampled output of the previous decoder layer. After that, the output is concatenated and distributed to the residual blocks. The exact residual blocks used in the encoder route are also used in this route. The number of feature maps in each decoder layer has been reduced by half, while the size of the feature maps has been increased. The output of the residual block is up-sampled to a resolution of  $128 \times 128 \times 128$  at the final layer, and the output is sent via a SoftMax activation

function. SoftMax activation transforms feature maps into probabilistic representations. The outputs are then evaluated to the ground truth to calculate the amount of loss.

### 3.4 Loss Function

In order to obtain a decent model, it is required to build a suitable loss function that measures the degree of similarity between the model's anticipated outcome and the actual value in the training data. The present work segments MRI scans using a mix of dice loss and focal loss.

#### 3.4.1 Focal Loss

Focal loss is a loss function suggested by Lin et al. (Lin, Goyal, Girshick, He, & Dollár, 2017) to assist deep learning models in resolving class imbalance challenges where there is a significant imbalance between foreground and background classes during model training (e.g., 1:1000).

The focused loss function was offered as a solution to the issue of a severe imbalance between the classes. It was achieved by adding weights to the majority and the minority samples. For frequent occurrences of majority data, the weight was lowered by a minor amount. For minority data, the weight was comparatively increased. By adjusting the value, it was possible to manage the contribution of the shared weight of different sample numbers to the overall loss. The addition of a modulating factor then regulated the weight of the easy-to-classify samples and the difficult-to-classify samples. This component decreased the weight of easy-to-classify data, allowing the model's training to place greater emphasis on difficult-to-classify samples.

The following are some descriptions of focus loss:

$$F. L. (p_t) = -\alpha_t(1 - p_t)^\gamma \log (p_t) \dots\dots\dots (3.4)$$

Where

$$p_t = \begin{cases} p & y = 1 \\ 1 - p & y = 0 \end{cases} \dots\dots\dots (3.5)$$

And

$$\alpha_t = \begin{cases} \alpha & y = 1 \\ 1 - \alpha & y = 0 \end{cases} \dots\dots\dots (3.6)$$

Equations. (3.4)-(3.6),  $y \in (0,1)$  is the real labeling of a patch, and  $p \in [0,1]$  is the model's predicted probabilities for the class labeled  $y = 1$ .

During recent decades, focus loss has been utilized extensively in image identification and categorization domains. For example, Lin et al. (Lin et al., 2017) used focus loss to develop a dense detector, RetinaNet, which produced impressive detecting performance. Furthermore, by implementing focus loss in their suggested neural network, Shu et al. (Shu et al., 2019) were able to tackle the problem of class imbalance during training and achieve improved classification performance for breast cancer. In light of the achievement of focused loss in the aforementioned disciplines, it feels that the potential of focal loss in clustering tasks is likewise deserving of inquiry and discussion; this is the driving force for this work.

### 3.4.2 Dice Loss

The dice coefficient is a commonly employed statistic for assessing segmentation output. It has also been adapted for use as a loss function, satisfying the segmentation objective's mathematical description. Later in 2016, it was also implemented as the Dice Loss function (Yeung, Sala, Schönlieb, & Rundo, 2022).

$$D = \frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \dots\dots\dots (3.7)$$

Demonstrates the Dice coefficient equation, where  $p_i$  and  $g_i$  indicate pairs of pixel numbers assigned to the predictions and ground truth, respectively. For example, in a boundary detection situation, the values of  $p_i$  and  $g_i$  are either 0 or 1, indicating whether or not the pixel is a boundary (value of 1) or not (value of 0). a consequence

of this, the number of total border pixels serves as the denominator for both ground truth and prediction. On the other hand, the numerator is the total of correctly predicted border pixels, and this amount can only grow when  $p_i$  and  $g_i$  are in a match.

$$D(x, y) = \frac{2|x \cap y|}{|x| + |y|} \dots\dots\dots (3.8)$$

Dice coefficient (DSC) quantifies the overlap of two different sets. For instance, if two sets A and B completely overlap, DSC's highest value is 1. Otherwise, DSC decreases until it reaches its minimum value of 0 if the two sets do not overlap. Therefore, the DSC range is between 0 and 1, and the higher it is, the better.

Designers are aware that the loss values require low values to fix the backpropagation weights. Therefore, 1-DSC may be used as the Dice loss to optimize the overlap between two sets.

$$Dice\ loss = 1 - DSC \dots\dots\dots (3.9)$$

This equation can determine that the dice loss decreases while the dice value increases. When the dice value reaches its maximum, the loss equals zero, indicating that the model is accurate.

### 3.5 Implementation Details

The suggested network goes through 100 training iterations with a batch size of 4. TensorFlow was utilized as the backend for this project. It utilized the Adam optimizer with a learning rate of 0.0001, batch normalization, and a learning rate of 0.0001 to boost the network's stability and normalize the model at each layer. Using a dilation rate of 1 reveals how much the kernel is expanded. The training for the network was done on the NVIDIA Tesla K80 GPU hosted in Google Colab.

### 3.6 Data Pre-Processing

Every modality of the Magnetic resonance imaging scans includes a three-dimensional, volumetric scan of the brain (240, 240, 155). As the 4 types of Magnetic

resonance images are utilized simultaneously to locate distinct parts of cancer, the study has integrated the slices from the four sequences into channels for each volume (128 slices). Training a model with an input image of (240,240,155,4) is computationally heavy and makes memory insufficient. So, to improve memory requirements. All the images were cropped (128,128,128,4). The input data size becomes (X,128,128,128, 4), where X is the number of images in each dataset,  $128 \times 128 \times 128$  is the dimension of every picture, and four are the channels. The dataset is then separated into a training data set consisting of 75% images, a validation set consisting of 15% images, and a testing set consisting of 10% images.

### **3.7 Model Training**

After normalizing, cropping, and resampling the pictures, the model was trained to automatically identify multiclass tumor segments. Due to the dimension of the data, samples were handled individually instead of in batches. The training dataset is separated into 75%, 15%, and 10% training, validation, and test samples. The training time for all three network models is 100 epochs long, and the learning rate is 0.0001. The Adam method [73], which is an adjustable first-order gradient optimizing methodology, is utilized when training the model. The model uses early stopping, which means the training process will be terminated if there is no improvement after twenty iterations on the validation data. The loss function for the project utilizes a combination of focused loss and dice loss.

## Chapter 4

### RESULTS AND DISCUSSION

#### 4.1 Introduction

This chapter discusses the performance of the proposed models. Several tests were undertaken to determine the model's improvements. In this chapter, a detailed description of experiments and an overview of the research's investigations are provided; after that, the selection of the optimal model configuration, BraTS 2020, BraTS 2019, and BraTS 2018 datasets were obtained.

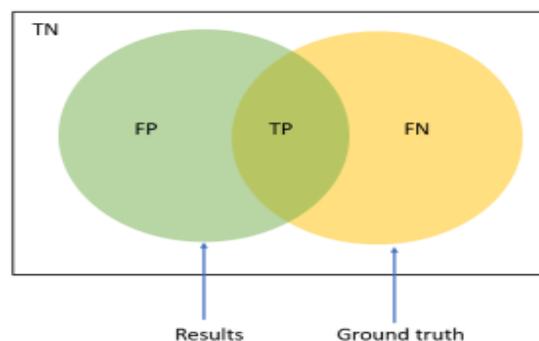
#### 4.2 Segmentation Evaluation Metrics

Data processing, pattern identification, and deep learning all require classification. Additionally, it permits rational decision-making (Norouzi et al., 2014). Classification is a method of building a network or classifiers in order to anticipate class labels. The objective of classifications is to properly anticipate the target class for each data instance (Kesavaraj & Sukumaran, 2013). The simplest kind of classification is binary classification, in which there are only two potential values for the desired characteristic. Multiclass targets include several parameters. In order to employ the automated classification approach, training data must be specified, which consists of a collection of observations for which the desired value has already been defined. This data must also include the inputs and outputs intended to be produced. In order for a diagnostic test to be accurate, it must be able to discriminate between individuals with and without illness (Kesavaraj & Sukumaran, 2013).

Generally, test findings are either positive or negative for a specific problem. These phrases of the terms positive and negative may be traced back to older medical uses in which persons who had a recognized health phenomenon (e.g., a sickness) were described as positive and all others as negative. A true positive (TP) occurs whenever a person has the disease and is classified as positive by the test. A false negative (F.N.) happens whenever the classifier projects a negative outcome. Likewise, a false positive

(F.P.) is identified if a person is negative for the disease. Still, the algorithm's forecast is positive; however, a true negative (T.N.) is established if both predictions are negative. Additionally, accuracy, sensitivity, and specificity are employed to evaluate clinical conditions. Sensitivity is the portion of true positives accurately detected. (Harefa, Alexander, & Pratiwi, 2017; Orru, Pettersson-Yeo, Marquand, Sartori, & Mechelli, 2012).

On the other hand, specificity relates to the fraction of real negatives that were accurately detected (Harefa et al., 2017; Orru et al., 2012). Often, the sensitivity of a binary test is stated as its precision (Kesavaraj & Sukumaran, 2013). Accuracy, precision, sensitivity, specificity, and the balanced error rate (BER) are the five measurements used to evaluate binary categorization (Majnik & Bosnić, 2013). Several researchers have employed varied combinations of tumor components to assess the approaches, making comparisons difficult. Therefore, the BRATS challenge organizers have supplied the most often mentioned tumor components. Several segmentation techniques include the Dice (DSC), sensitivity, and positive predictive value (PPV). Figure 4.1. Illustrates the segmented regions utilized to evaluate the segmentation technique in this study. According to Dvořák and Menze. (Dvořák & Menze, 2015), the tumor area was evaluated, which included the entire tumor, the core tumor excluding "edema," and the increasing tumor, which contained the increased tumor shape.



**Figure 4.1:** Demonstrates True Positive, False Positive, True Negative, and False Negative. The green circle shows the findings, whereas the yellow circle indicates the ground truth.

Comparing segmented pictures to measure accuracy is crucial for measuring the neural network's performance. The segmentation of brain tumors is severely affected by class imbalance difficulties. In addition to Intersection over Union (IoU), the Dice Score (D.S.) coefficient was used as the assessment measure for all the experiments since it addresses the problem of class imbalance.

The dice coefficient (DSC) (Yeghiazaryan & Voiculescu, 2015) is the largest frequent and widely used assessment metric for evaluating a segmented picture to its ground truth. It compares two sets,  $x$ , and  $y$ , by normalizing the sizes of their overlaps according to their average sizes. The following formula denotes the DSC formula:

$$D(x, y) = \frac{2|x \cap y|}{|x| + |y|} \dots\dots\dots (4.1)$$

Dice coefficient (DSC) (Yeghiazaryan & Voiculescu, 2015) quantifies the overlapping of two components. For instance, if two sets  $x$  and  $y$  fully overlapped, DSC's highest score is 1. Otherwise, DSC decreases until it reaches its minimum value of 0 if the two sets do not overlap. Hence, the DSC range is between 0 and 1, and the higher it is, the greater.

The Jaccard coefficient (Jaccard) (Yeghiazaryan & Voiculescu, 2015) is a segmentation technique assessment metric. For instance, Jaccard provides the below equation to compute the match between two  $x$  and  $y$  sets by normalizing the area of their intersection over their union:

$$\text{Jaccard} = \frac{|x \cap y|}{|x \cup y|} \dots\dots\dots (4.2)$$

Statistical decision concept metrics sensitivity and specificity are calculated using the following formulas:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \dots\dots\dots (4.10)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \dots\dots\dots (4.11)$$

The effectiveness of the proposed model was evaluated using the Jaccard score, dice coefficient score, sensitivity, and specificity.

### 4.3 Data Sets

The clinical magnetic resonance imaging scans image data represent low-grade gliomas and high-grade gliomas.

The dataset images utilized in the investigation all contain the 4 MRI scans listed below.

- T1 is a T1-weighted, native image that may be acquired in sagittal or axial 2D, and it has a cell size between 1 to 6 mm.
- T1-ce: The majority of patients will get a T1-weighted, contrast-enhanced (with gadolinium), three-dimensional image with a voxel size of 1 millimeter isotropic.
- T2: T2-weighted picture, acquired in axial 2D, having a 2- and 6-mm cell size.
- FLAIR: T2-weighted FLAIR images can be acquired in axial, coronal, or sagittal planes, with a cell size between 2 and 6 mm.

#### 4.3.1 BraTS 2018

BraTS 2018 [39, 80, 81] was utilized in this experiment, which delivers multimodal 3D brain MRIs and ground truth brain tumor segmentations labeled by clinicians, with 4 MRI modalities for each patient (T1, T1c, T2, and FLAIR). The MRI volumes were manually segmented by one to four reviewers, and professional neuroradiologists validated their labels. The annotations include three subregions of the tumor: the increasing tumor, the edema, and the necrotic and non-increasing tumor core. In addition, the annotations were grouped into three subgroups: the complete tumor, the core tumor, and the increasing tumor. MRI scanners from 19 organizations were used to acquire the data. The BraTS 2018 training data includes a total of 285 pictures, of which 210 are HGG, and 75 are LGG. The validation data includes 66 unique MRI

scans. Therefore,  $240 \times 240 \times 155$  is the volume dimension for each MRI in the BraTS 2018 dataset.

### **4.3.2 BraTS 2019**

The BraTS 2019 data set (Bakas et al., 2017b; Bakas et al., 2018; Menze et al., 2014) comprises a total of 259 high-grade gliomas (HGG) MRI scans and 76 low-grade gliomas (LGG) MRI images. This dataset was provided from the BraTS website and is utilized to analyze suggested designs. The MRI performed on each person includes the following order: FLAIR, T1, contrast-enhanced T1, and T2. The ground truth of each image was manually constructed using the same annotation technique. Professional neuroradiologists accepted the annotations (Bakas et al., 2017a), which include enhancing tumor (E.T. labeled 4), peritumoral edema (E.D. labeled 2), and necrotic or non-increasing tumor core (NCR/NET labeled 1). This research examines both HGG and LGG MRI images.

### **4.3.3 BraTS 2020**

This research uses the BraTS 2020 dataset (Bakas et al., 2017a; Bakas et al., 2017b; Bakas et al., 2018; Menze et al., 2014) to evaluate the effectiveness of the presented network. There are 125 for validating and 369 training research related to the multimodal brains of M.R. T1-image, T1ce, T2-image, and FLAIR sequences are included in every dataset. The dimension of every MRI picture is  $240 \times 240 \times 155$  pixels. Furthermore, professionals analyzed every study's increasing tumor (E.T.), peritumoral edema (E.D.), and necrotic and non-increasing tumor core (NET). Labels for training experiments are publicly available for online assessment, but validation and test trial annotations are kept private.

## **4.4 Pre-Processing**

The first step in any data-driven study is to pre-process the raw images. First, the images of all three datasets are resized to  $128 \times 128 \times 128$  for feeding as input to UNet. In every dataset, each subject contains four images with annotated masks. Next, all the

images are given to the networks by considering each image separately as W.T., E.T., and T.C. classes.

#### 4.5 Comparison Of The Proposed Models

This section evaluates the results of the proposed model against the baseline models to show the performance.

##### 4.5.1 Increased Depth Of 3D Unet Model

First, the BraTS datasets have been used for training the increased depth of 3D UNet architecture. Tables 4.1 and 4.2 provide the findings of the model, which could be found below. When compared to the performance of other models, UNet's performance is rated lower in terms of the dice coefficient value. Check out Tables 4.3 and 4.5.

**Table 4.1:** Dice coefficient score and Jaccard score for the Increased depth of 3D UNet model on the BraTS 2020, 2019, and 2018 datasets

Configuration	Dataset	Dice Coefficient Score			Jaccard Score		
		WT	ET	TC	WT	ET	TC
Increased Depth Of 3D Unet	BraTS 2020	0.9277	0.8792	0.9180	0.8652	0.7844	0.8484
	BraTS 2019	0.8280	0.7763	0.7993	0.7064	0.6344	0.6657
	BraTS 2018	0.8318	0.7720	0.8121	0.7121	0.6287	0.6836

**Table 4.2:** Sensitivity and specificity for Increased depth of 3D UNet on BraTS 2020, 2019, and 2018 datasets.

Configuration	Dataset	Sensitivity			Specificity		
		WT	ET	TC	WT	ET	TC
Increased Depth Of 3D Unet	BraTS 2020	0.9005	0.9356	0.9373	0.9980	0.9990	0.9990
	BraTS 2019	0.7650	0.7556	0.8124	0.9964	0.9980	0.9964
	BraTS 2018	0.8844	0.7280	0.7803	0.9873	0.9976	0.9968

#### 4.5.2 Increased Depth Of 3D Residual-Unet Model

To increase the dice score, a hybrid deep learning model Residual-UNet is used in which the Residual block is used as an encoder part for feature extraction. The utilization of residual blocks allows for building a bigger network to be accomplished, not need to be concerned with the issue of disappearing gradients or exploding gradients (Maji et al., 2022). It also helps in easy training of the network. In addition, the Residual UNet's skip links contribute to a more efficient transfer of information across the various layers, which also, in turn, contributes to a more effective transfer of gradients during training. RESNET (Maji et al., 2022) was developed to achieve great efficiency by having a reduced number of parameters. It is an advancement in UNET architecture. The increased depth of 3D Residual-UNet are presented in Table 4.3 and Table 4.4.

**Table 4.3:** Dice coefficient score and Jaccard score for the Increased depth of 3D Residual-UNet model on BraTS 2020, 2019, and 2018 datasets

Configuration	Dataset	Dice Coefficient Score			Jaccard Score		
		WT	ET	TC	WT	ET	TC
	BraTS 2020	0.9361	0.8919	0.8833	0.8977	0.8045	0.7910

Configuration	Dataset	Dice Coefficient Score			Jaccard Score		
		WT	ET	TC	WT	ET	TC
Increased Depth Of 3D Residual-UNet	BraTS 2019	0.8914	0.7578	0.7365	0.8040	0.6100	0.5828
	BraTS 2018	0.8484	0.7127	0.8033	0.7367	0.5537	0.6713

**Table 4.4:** Sensitivity and specificity for the Increased depth of 3D Residual-UNet on BraTS 2020, 2019, and 2018 datasets.

Configuration	Dataset	Sensitivity			Specificity		
		WT	ET	TC	WT	ET	TC
Increased Depth Of 3D Residual-UNet	BraTS 2020	0.9551	0.9392	0.9608	0.9964	0.9992	0.9979
	BraTS 2019	0.8426	0.6286	0.7046	0.9979	0.9997	0.9965
	BraTS 2018	0.8338	0.5703	0.7047	0.9931	0.9995	0.9989

#### 4.5.3 Increased Depth of 3D Attention Residual UNet

This model is a combination of attention mechanism and residual UNet. The attention function concentrates on a certain portion of the picture whilst overlooking the rest of the image. As with human visual attention, it may be trained to concentrate on a particular region while overlooking nearby areas in order to maximize performance (Maji et al., 2022). In addition, attention gates can help limit the number of false positives by reducing feature activation in unimportant image areas.

**Table 4.5:** Dice coefficient score and Jaccard score for Increased depth of 3D Attention Residual-UNet on BraTS 2020, 2019, and 2018 datasets.

Configuration	Dataset	Dice Coefficient Score			Jaccard Score		
		WT	ET	TC	WT	ET	TC
Increased Depth Of 3D Attention ResNet	BraTS 2020	0.9391	0.8921	0.9301	0.8851	0.8052	0.8694
	BraTS 2019	0.8844	0.7987	0.7511	0.7927	0.6649	0.6014
	BraTS 2018	0.8836	0.7819	0.8317	0.7914	0.6419	0.7119

**Table 4.6:** Sensitivity and specificity for Increased depth of 3D Attention Residual-UNet on BraTS 2020, 2019, and 2018 datasets.

Configuration	Dataset	Sensitivity			Specificity		
		WT	ET	TC	WT	ET	TC
Increased Depth Of 3D Attention ResNet	BraTS 2020	0.9484	0.9090	0.9375	0.9966	0.9993	0.9992
	BraTS 2019	0.8160	0.7752	0.7248	0.9987	0.9985	0.9966
	BraTS 2018	0.8569	0.7053	0.7758	0.9957	0.9985	0.9979

#### 4.6 Using Dropout Regularization

A training method known as a dropout involves ignoring a certain number of neurons at random. They simply "disappear" for no apparent reason. It suggests that their impact on the activity of neurons farther downstream is removed temporarily during the forward transit of the pathway. On the reverse pass, any weight modifications are not transmitted to the neuron (Saeed et al., 2021). This experiment uses a combination of dropout values of 0.1 and 0.2; the results are revealed in Table 4.7 and Table 4.8. The findings for this dropout number indicate that the model's predictive accuracy did not significantly increase as a consequence of the changes made.

**Table 4.7:** Dice coefficient and Jaccard score for Increased depth of 3D Attention Residual-UNet with and without dropout regularization on BraTS 2020, 2019, and 2018 datasets. The best scores are in bold.

Configuration	Dataset	Dice Coefficient			Jaccard Score		
		WT	ET	TC	WT	ET	TC
Increased Depth Of 3D A-ResNet (without dropout)	BraTS 2020	<b>0.9391</b>	<b>0.8921</b>	0.9301	<b>0.8851</b>	<b>0.8052</b>	0.8694
Increased Depth Of 3D A-ResNet (with dropout)		0.92	0.8899	<b>0.9539</b>	0.8518	0.8016	<b>0.9119</b>
<b>Average Improvement</b>		0.08%			0.29%		
Increased Depth Of 3D A-ResNet (without dropout)	BraTS 2019	0.8844	<b>0.7987</b>	0.7511	0.7927	<b>0.6649</b>	0.6014
Increased Depth Of 3D A-ResNet (with dropout)		<b>0.8942</b>	0.7714	<b>0.7970</b>	<b>0.8087</b>	0.6279	<b>0.6625</b>
<b>Average Improvement</b>		0.97%			1.34%		
Increased Depth Of 3D A-ResNet (without dropout)	BraTS 2018	<b>0.8836</b>	<b>0.7819</b>	<b>0.8317</b>	<b>0.7914</b>	<b>0.6419</b>	<b>0.7119</b>
Increased Depth Of 3D A-ResNet (with dropout)		0.8413	0.7357	0.8078	0.7261	0.5819	0.6775
<b>Average Improvement</b>		-3.75%			-5.32%		

**Table 4.8:** Sensitivity and specificity for Increased depth of 3D Attention Residual-UNet with and without dropout regularization on BraTS 2020, 2019, and 2018 datasets. The best scores are in bold.

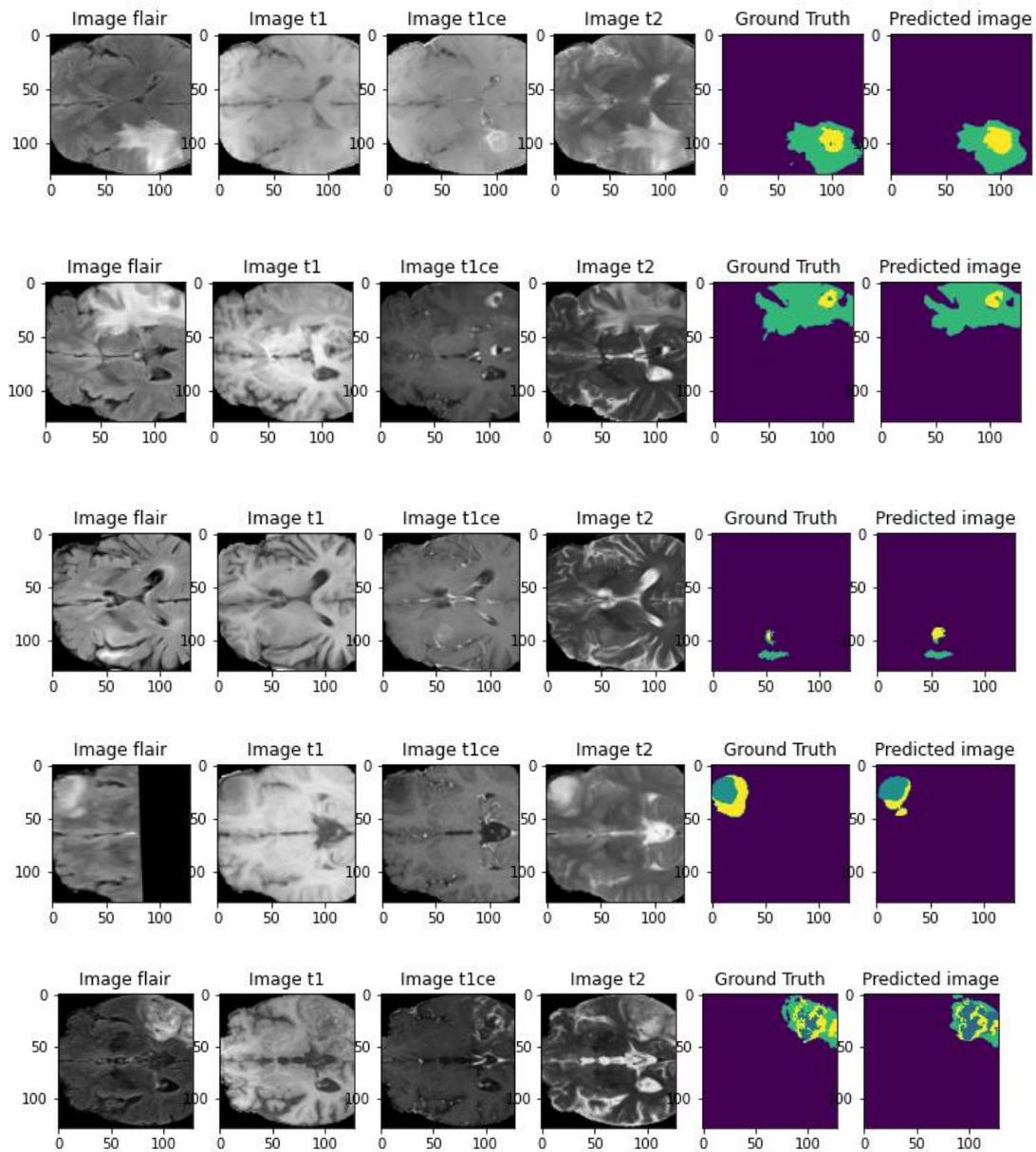
Configuration	Dataset	Sensitivity			Specificity		
		WT	ET	TC	WT	ET	TC
Increased Depth Of 3D A-ResNet (without dropout)	BraTS 2020	0.9484	<b>0.9090</b>	0.9375	<b>0.9966</b>	0.9993	0.9992
Increased Depth Of 3D A-ResNet (with dropout)		<b>0.9797</b>	0.894	<b>0.9478</b>	0.9928	<b>0.9996</b>	<b>0.9994</b>
Average Improvement			0.89%			-0.11%	
Increased Depth Of 3D A-ResNet (without dropout)	BraTS 2019	0.8160	<b>0.7752</b>	0.7248	<b>0.9987</b>	0.9985	0.9966
Increased Depth Of 3D A-ResNet (with dropout)		<b>0.8928</b>	0.6665	<b>0.7883</b>	0.9955	<b>0.9994</b>	<b>0.9969</b>
Average Improvement			1.05%			-0.067%	
Increased Depth Of 3D A-ResNet (without dropout)	BraTS 2018	0.8569	<b>0.7053</b>	<b>0.7758</b>	<b>0.9957</b>	0.9985	<b>0.9979</b>
Increased Depth Of 3D A-ResNet (with dropout)		<b>0.8831</b>	0.6169	0.7565	0.9886	<b>0.9991</b>	0.9973
Average Improvement			-2.72%			-0.237%	

## 4.7 Predicted Image Results of The Proposed Model

This section displays the predicted image results of the suggested model.

### 4.7.1 BraTS 2018

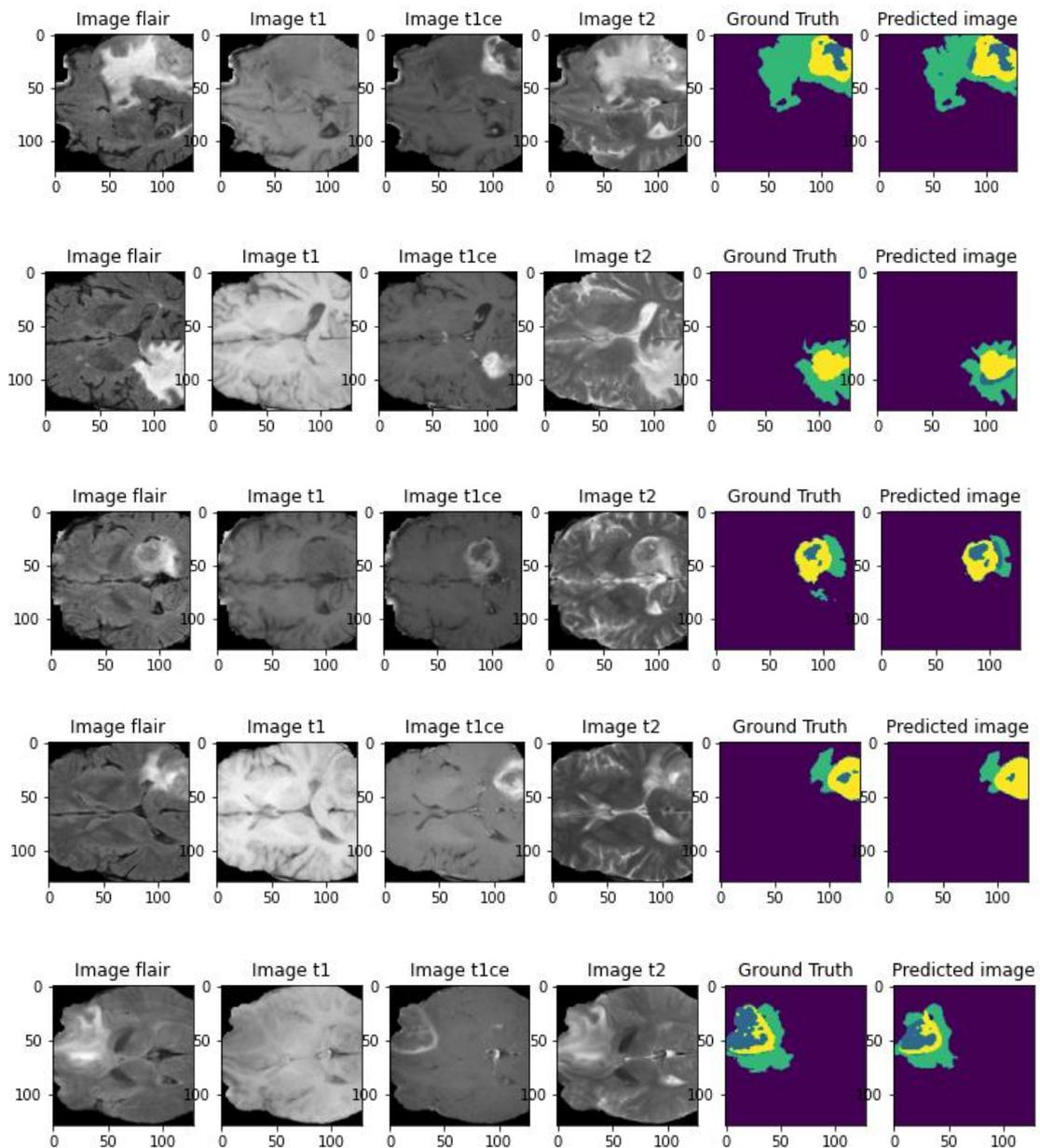
Figure 4.2 displays some sample images taken from the BraTS 2018 dataset. These images include the ground truth and the predicted image of the proposed model.



**Figure 4.2:** Predictions of the proposed model on BraTS 2018

### 4.7.2 BraTS 2019

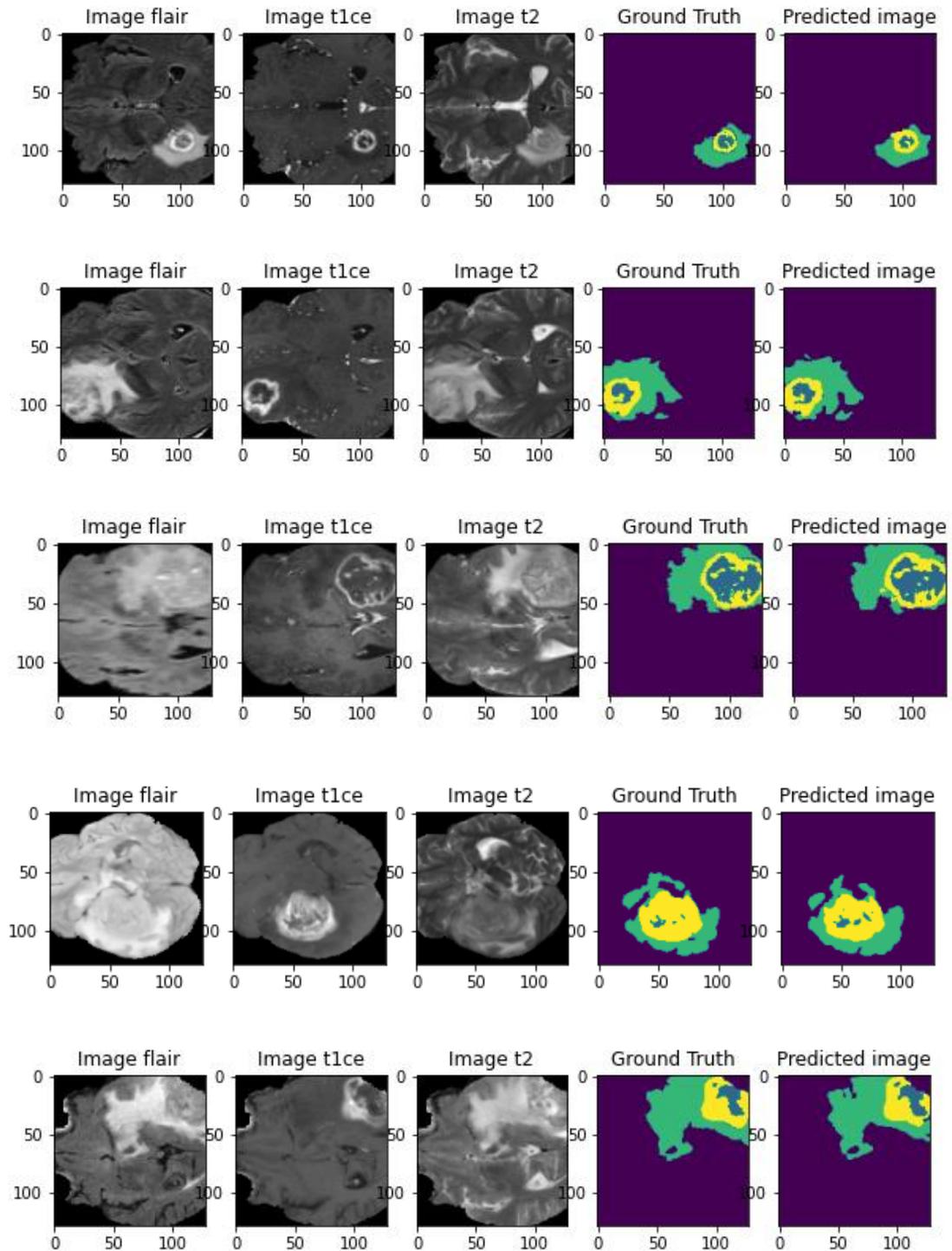
Figure 4.3 displays some sample images taken from the BraTS 2019 dataset. These images include the ground truth and the predicted image of the proposed model.



**Figure 4.3:** Predictions of the proposed model on BraTS 2019

### 4.7.3 BraTS 2020

Figure 4.4 displays some sample images taken from the BraTS 2020 dataset. These images include the ground truth and the predicted image of the proposed model.



**Figure 4.4:** Predictions of the proposed model on BraTS 2020

#### 4.8 Comparison Of Increased Depth Of 3D Attention Residual-Unet With Other Deep Learning Segmentation Models

The results obtained from different segmentation models are shown in Tables 4.9-4.11, including a brief explanation and the models' size.

**Table 4.9:** Comparing the presented model and state-of-the-art segmentation models regarding model size and the number of parameters using BraTS2020. The best two scores are in bold.

References	Dataset Used	Architecture Information	Dice Coefficient Score		
			Whole Tumor (WT)%	Enhanced Tumor (ET)%	Tumor Core (TC)%
(Fidon, Ourselin, & Vercauteren, 2020)	BraTS 2020	3D Unet architecture with additional layers	88.9	81.4	84.1
(Y. Wang et al., 2020)	BraTS 2020	Modality pairing architecture like 3D Unet architecture	89.1	81.6	84.2
(Jia, Cai, Huang, & Xia, 2020)	BraTS 2020	Single and cascaded HNF-Net	<b>91.29</b>	78.75	<b>85.46</b>
(Isensee, Jäger, Full, Vollmuth, & Maier-Hein, 2020)	BraTS 2020	nnUnet architecture with augmentation and modification	88.95	<b>82.03</b>	85.06
(Wenbo Zhang et al., 2021)	BraTS 2020	Multi-encoder architecture with Categorical dice score	70.24	73.86	88.26
(Silva, Pinto, Pereira, &	BraTS 2020	Three deep layer aggregation neural networks using previous outputs as input	88.58	79	82.97

References	Dataset Used	Architecture Information	Dice Coefficient Score		
			Whole Tumor (WT)%	Enhanced Tumor (ET)%	Tumor Core (TC)%
Lopes, 2020)					
(Ahmad, Qamar, Shen, & Saeed, 2020)	BraTS 2020	Modified Unet architecture with densely connected blocks	89.12	79.12	84.74
(Henry et al., 2020)	BraTS 2020	Ensemble model with multiple Unet networks	91	81	85
(Russo, Liu, & Ieva, 2020)	BraTS 2020	Lesion encoder framework with DCNN network	86.87	78.98	80.66
(Messaoudi et al., 2020)	BraTS 2020	Efficient network as an encoder with the three-dimensional network for segmentation	80.68	69.59	75.20
<b>Proposed Model</b>	BraTS 2020	<b>Increased depth of 3D Residual UNET Network Architecture with an attention gate</b>	<b>93.91</b>	<b>89.21</b>	<b>93.01</b>

**Table 4.10:** Comparing the presented model and state-of-the-art segmentation models regarding model size and the number of parameters using BraTS 2019. The best two scores are in bold.

References	Dataset Used	Architecture Information	Dice Coefficient Score		
			Whole Tumor (WT)%	Enhanced Tumor (ET)%	Tumor Core (TC)%
(Li, Luo, & Wang, 2019)	BraTS 2019	A multi-step cascaded model with hierarchical topology	<b>88.60</b>	77.10	<b>81.30</b>
(F. Wang, Jiang, Zheng, Meng, & Biswal, 2019)	BraTS 2019	3D Unet based deep learning model using brain-wise normalization and patching strategies	85.20	<b>77.80</b>	79.80
(Kim, Luna, Chikontwe, & Park, 2019)	BraTS 2019	Two-Step Unet for Brain Tumor Segmentation and Random Forest	85.80	74.30	<b>80.40</b>
(Amian & Soltaninejad, 2019)	BraTS 2019	Multi-resolution 3D CNN for MRI Brain Tumor Segmentation and Survival Prediction	84	71	74
(Islam et al., 2019)	BraTS 2019	Brain Tumor Segmentation and Survival Prediction Using 3D Attention Unet	86.89	77.80	77.71
<b>Proposed Model</b>	BraTS 2019	<b>Increased depth of 3D Residual UNET Network Architecture with an attention gate</b>	<b>88.44</b>	<b>79.87</b>	75.11

**Table 4.11:** Comparing the presented model and state-of-the-art segmentation models regarding model size and the number of parameters using BraTS 2018. The best two scores are in bold.

References	Dataset Used	Architecture Information	Dice Coefficient Score		
			Whole Tumor (WT)%	Enhanced Tumor (ET)%	Tumor Core (TC)%
(Weninger, Rippel, Koppers, & Merhof, 2018)	BraTS 2018	Two 3D Unet were used. One for finding tumor location and the second for detecting subtle tumor	84.40	62.10	72.80
(Kermi et al., 2018)	BraTS 2018	Modified Unet architecture based on 2D deep neural network	86.80	<b>78.30</b>	80.50
(Kong & Zhang, 2021)	BraTS 2018	Multi-modal Brain Tumor Segmentation Using Cascaded 3D Unet	88.30	74.60	<b>83.40</b>
(Puch, Sánchez, Hernández, Piella, & Prčkovska, 2018)	BraTS 2018	Global Planar Convolutions for improved context aggregation	<b>89.70</b>	75.20	79.07
(Myronenko, 2018)	BraTS 2018	Ensemble of ten deep learning models with auto-regularization	<b>88.39</b>	76.64	81.54
<b>Proposed Model</b>	BraTS 2018	<b>Increased depth of 3D Residual UNET Network Architecture with an attention gate</b>	88.36	<b>78.19</b>	<b>83.17</b>

## Chapter 5

### CONCLUSION AND RECOMMENDATIONS

#### 5.1 Conclusion

Segmentation of brain tumors is a key diagnostic technique. This research proposes the increased depth of 3D Residual UNET Network Architecture with an attention gate for brain tumor segmentation. The presented model has increased the depth of the normal attention residual unet from four layers into six layers. By allowing only activations from important areas on the encoder side using attention gates and creating better feature mappings at the decoder, these modifications to the basic network of attention Res-UNET enhanced the process of learning. The use of a combination of dice loss and focal loss helped the model in resolving class imbalance challenges where brain tumors have a significant imbalance between foreground and background classes. Because of this, the model has improved and got a better segmentation achievement. The datasets from the BraTS 2020, BraTS 2019, and BraTS 2018 competitions are used throughout the training and testing of the suggested model. The proposed model has gained dice coefficient scores for WT, TC, and ET as 93.91%, 93.01%, and 89.21% on BraTS 2020, 88.44%, 75.11%, and 79.87% on BraTS 2019, 88.36%, 83.17%, and 78.19% on BraTS 2018 datasets, respectively. As shown in tables 4.9,4.10, and 4.11, the results of the experiments revealed that the suggested design is superior to most of the existing models that are considered to be state-of-the-art.

#### 5.2 Future Research

The study described in this thesis and the results obtained enable the prediction of future research directions. Some research was excluded from the thesis because of the time limitations, while others are new forecasts based on concerns during the experiment's result. It is feasible that an increased depth of 3D Attention Residual UNET Network with a guided decoder could be done at some point in the future.

On the other hand, the training of the recommended model is computationally intensive; as a result, more work can be done to cut down on the amount of memory and processing time required. In addition, however, many manually labeled data relating to brain tumors are required for the proposed model to be trained. As a result, future research may concentrate on developing unsupervised and semi-supervised methods for segmenting brain tumors.

## REFERENCE

- Abd-Ellah, M. K., Awad, A. I., Khalaf, A. A., & Hamed, H. F. (2019). A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. *Magnetic resonance imaging*, *61*, 300-318.
- Ahmad, P., Qamar, S., Shen, L., & Saeed, A. (2020). *Context aware 3D UNet for brain tumor segmentation*. Paper presented at the International MICCAI Brainlesion Workshop.
- Al-Qazzaz, S. (2020). *Deep learning-based brain tumour image segmentation and its extension to stroke lesion segmentation*. Cardiff University.
- Alqazzaz, S., Sun, X., Yang, X., & Nokes, L. (2019). Automated brain tumor segmentation on multi-modal MR image using SegNet. *Computational Visual Media*, *5*(2), 209-219.
- Amian, M., & Soltaninejad, M. (2019). *Multi-resolution 3D CNN for MRI brain tumor segmentation and survival prediction*. Paper presented at the International MICCAI Brainlesion Workshop.
- Angulakshmi, M., & Lakshmi Priya, G. (2017). Automated brain tumour segmentation techniques—a review. *International journal of imaging systems and technology*, *27*(1), 66-77.
- Azhari, E.-E. M., Hatta, M. M. M., Htike, Z. Z., & Win, S. L. (2014). Tumor detection in medical imaging: a survey. *International Journal of Advanced Information Technology*, *4*(1), 21.
- Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, *39*(12), 2481-2495.
- Bahadure, N. B., Ray, A. K., & Thethi, H. P. (2017). Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. *International journal of biomedical imaging*, 2017.
- Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., . . . Davatzikos, C. (2017a). Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection. *The cancer imaging archive*, *286*.
- Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J. S., . . . Davatzikos, C. (2017b). Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Scientific data*, *4*(1), 1-13.
- Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., . . . Rozycki, M. (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*.
- Bauer, S., Lu, H., May, C. P., Nolte, L. P., Büchler, P., & Reyes, M. (2013). Integrated segmentation of brain tumor images for radiotherapy and neurosurgery. *International journal of imaging systems and technology*, *23*(1), 59-63.
- Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent *Proceedings of COMPSTAT'2010* (pp. 177-186): Springer.
- Chakraborty, S., Chatterjee, S., Das, A., & Mali, K. (2020). Penalized fuzzy c-means enabled hybrid region growing in segmenting medical images *Hybrid Machine Intelligence for Medical Image Analysis* (pp. 41-65): Springer.
- Chauvin, Y., & Rumelhart, D. E. (2013). *Backpropagation: theory, architectures, and applications*: Psychology press.

- Chen, J., Sathe, S., Aggarwal, C., & Turaga, D. (2017). *Outlier detection with autoencoder ensembles*. Paper presented at the Proceedings of the 2017 SIAM international conference on data mining.
- Clevert, D.-A., Unterthiner, T., & Hochreiter, S. (2015). Fast and accurate deep network learning by exponential linear units (elus). *arXiv preprint arXiv:1511.07289*.
- Daimary, D., Bora, M. B., Amitab, K., & Kandar, D. (2020). Brain tumor segmentation from MRI images using hybrid convolutional neural networks. *Procedia Computer Science, 167*, 2419-2428.
- Dvořák, P., & Menze, B. (2015). *Local structure prediction with convolutional neural networks for multimodal brain tumor segmentation*. Paper presented at the International MICCAI workshop on medical computer vision.
- Eftimov, N., Ivanov, I., Petkov, A., & Nakov, E. (2007). Management of recurrent high-grade gliomas. *Cancer Ther, 5*, 243-252.
- El-Dahshan, E.-S. A., Mohsen, H. M., Revett, K., & Salem, A.-B. M. (2014). Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm. *Expert systems with Applications, 41*(11), 5526-5545.
- Fidon, L., Ourselin, S., & Vercauteren, T. (2020). *Generalized wasserstein dice score, distributionally robust deep learning, and ranger for brain tumor segmentation: Brats 2020 challenge*. Paper presented at the International MICCAI Brainlesion Workshop.
- Fu, X., Liu, T., Xiong, Z., Smail, B. H., Stiles, M. K., & Zhao, J. (2018). Segmentation of histological images and fibrosis identification with a convolutional neural network. *Computers in biology and medicine, 98*, 147-158.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*: MIT press.
- Harefa, J., Alexander, A., & Pratiwi, M. (2017). Comparison classifier: support vector machine (SVM) and K-nearest neighbor (K-NN) in digital mammogram images. *Jurnal Informatika dan Sistem Informasi, 2*(2), 35-40.
- Hassan, E., & Aboshgifa, A. (2015). Detecting brain tumour from MRI image using matlab gui programme. *International Journal of Computer Science & Engineering Survey (IJCES)*, 6(6).
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., . . . Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical image analysis, 35*, 18-31.
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Delving deep into rectifiers: Surpassing human-level performance on imagenet classification*. Paper presented at the Proceedings of the IEEE international conference on computer vision.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Henry, T., Carré, A., Lerousseau, M., Estienne, T., Robert, C., Paragios, N., & Deutsch, E. (2020). *Brain tumor segmentation with self-ensembled, deeply-supervised 3D U-net neural networks: a BraTS 2020 challenge solution*. Paper presented at the International MICCAI Brainlesion Workshop.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). *Densely connected convolutional networks*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Isensee, F., Jäger, P. F., Full, P. M., Vollmuth, P., & Maier-Hein, K. H. (2020). *nnU-net for brain tumor segmentation*. Paper presented at the International MICCAI Brainlesion Workshop.

- Işın, A., Direkoğlu, C., & Şah, M. (2016). Review of MRI-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, 102, 317-324.
- Islam, M., Vibashan, V., Jose, V., Wijethilake, N., Utkarsh, U., & Ren, H. (2019). *Brain tumor segmentation and survival prediction using 3D attention UNet*. Paper presented at the International MICCAI Brainlesion Workshop.
- Jia, H., Cai, W., Huang, H., & Xia, Y. (2020). H 2 NF-Net for Brain Tumor Segmentation Using Multimodal MR Imaging: 2nd Place Solution to BraTS Challenge 2020 Segmentation Task. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 58-68.
- Jordan, J. (2018). An overview of semantic image segmentation. Retrieved from <https://www.jeremyjordan.me/semantic-segmentation/>
- Kermi, A., Mahmoudi, I., & Khadir, M. T. (2018). *Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes*. Paper presented at the International MICCAI Brainlesion Workshop.
- Kesavaraj, G., & Sukumaran, S. (2013). *A study on classification techniques in data mining*. Paper presented at the 2013 fourth international conference on computing, communications and networking technologies (ICCCNT).
- Kim, S., Luna, M., Chikontwe, P., & Park, S. H. (2019). *Two-step U-Nets for brain tumor segmentation and random forest with radiomics for survival time prediction*. Paper presented at the International MICCAI Brainlesion Workshop.
- Kong, L., & Zhang, Y. (2021). *Multi-modal Brain Tumor Segmentation Using Cascaded 3D U-Net*. Paper presented at the 2021 6th International Conference on Image, Vision and Computing (ICIVC).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012a). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012b). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
- Lamba, H. (2019, Feb 17). Understanding Semantic Segmentation with UNET. Retrieved from <https://towardsdatascience.com/understanding-semantic-segmentation-with-unet-6be4f42d4b47>
- Li, X., Luo, G., & Wang, K. (2019). *Multi-step cascaded networks for brain tumor segmentation*. Paper presented at the International MICCAI Brainlesion Workshop.
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). *Focal loss for dense object detection*. Paper presented at the Proceedings of the IEEE international conference on computer vision.
- Long, J., Shelhamer, E., & Darrell, T. (2015). *Fully convolutional networks for semantic segmentation*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvett, A., . . . Kleihues, P. (2007). The 2007 WHO classification of tumours of the central nervous system. *Acta neuropathologica*, 114(2), 97-109.
- Louis, D. N., Perry, A., Reifenberger, G., Von Deimling, A., Figarella-Branger, D., Cavenee, W. K., . . . Ellison, D. W. (2016). The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica*, 131(6), 803-820.

- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). *Rectifier nonlinearities improve neural network acoustic models*. Paper presented at the Proc. icml.
- Maharjan, S., Alsadoon, A., Prasad, P., Al-Dalain, T., & Alsadoon, O. H. (2020). A novel enhanced softmax loss function for brain tumour detection using deep learning. *Journal of neuroscience methods*, 330, 108520.
- Maji, D., Sigedar, P., & Singh, M. (2022). Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors. *Biomedical Signal Processing and Control*, 71, 103077.
- Majnik, M., & Bosnić, Z. (2013). ROC analysis of classifiers in machine learning: A survey. *Intelligent data analysis*, 17(3), 531-558.
- Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., . . . Wiest, R. (2014). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE transactions on medical imaging*, 34(10), 1993-2024.
- Messaoudi, H., Belaid, A., Allaoui, M. L., Zetout, A., Allili, M. S., Tliba, S., . . . Conze, P.-H. (2020). *Efficient embedding network for 3D brain tumor segmentation*. Paper presented at the International MICCAI Brainlesion Workshop.
- Miller, K., Wittek, A., & Joldes, G. (2010). Biomechanics of the brain for computer-integrated surgery. *ACTA of Bioengineering and Biomechanics*, 12(2), 25-37.
- Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2001). Image segmentation using deep learning: A survey. arXiv 2020. *arXiv preprint arXiv:2001.05566*.
- Mlynarski, P., Delingette, H., Criminisi, A., & Ayache, N. (2019). 3D convolutional neural networks for tumor segmentation using long-range 2D context. *Computerized Medical Imaging and Graphics*, 73, 60-72.
- Modeler, S. (2021). The Neural Networks Model. Retrieved from <https://www.ibm.com/docs/en/spss-modeler/18.0.0?topic=networks-neural-model>
- Mody, P. (2018, August 7). Semantic Segmentation. Retrieved from <https://www.playment.io/blog/semantic-segmentation>
- Mohan, G., & Subashini, M. M. (2018). MRI based medical image analysis: Survey on brain tumor grade classification. *Biomedical Signal Processing and Control*, 39, 139-161.
- Mwiti, D. (2019, July 29). A Guide to Semantic Segmentation. Retrieved from <https://heartbeat.fritz.ai/a-2019-guide-to-semantic-segmentation-ca8242f5a7fc>
- Myronenko, A. (2018). *3D MRI brain tumor segmentation using autoencoder regularization*. Paper presented at the International MICCAI Brainlesion Workshop.
- Nielsen, M. A. (2015). *Neural networks and deep learning* (Vol. 25): Determination press San Francisco, CA, USA.
- Noori, M., Bahri, A., & Mohammadi, K. (2019). *Attention-guided version of 2D UNet for automatic brain tumor segmentation*. Paper presented at the 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE).
- Norouzi, A., Rahim, M. S. M., Altameem, A., Saba, T., Rad, A. E., Rehman, A., & Uddin, M. (2014). Medical image segmentation methods, algorithms, and applications. *IETE Technical Review*, 31(3), 199-213.
- Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., . . . Kainz, B. (2018). Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*.

- Oro, D., Fernández, C., Martorell, X., & Hernando, J. (2016). *Work-efficient parallel non-maximum suppression for embedded GPU architectures*. Paper presented at the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).
- Orru, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1140-1152.
- Patterson, J., & Gibson, A. (2017). *Deep learning: A practitioner's approach*: " O'Reilly Media, Inc."
- 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.
- Pietrangelo, A. (2019, October 23). Benign and Malignant Tumors: How Do They Differ? Retrieved from <https://www.healthline.com/health/cancer/difference-between-benign-and-malignant-tumors>
- Puch, S., Sánchez, I., Hernández, A., Piella, G., & Prčkovska, V. (2018). *Global planar convolutions for improved context aggregation in brain tumor segmentation*. Paper presented at the International MICCAI Brainlesion Workshop.
- Qamar, S., Jin, H., Zheng, R., Ahmad, P., & Usama, M. (2020). A variant form of 3D-UNet for infant brain segmentation. *Future Generation Computer Systems*, *108*, 613-623.
- Rajinikanth, V., Palani Thanaraj, K., Satapathy, S. C., Fernandes, S. L., & Dey, N. (2019). Shannon's entropy and watershed algorithm based technique to inspect ischemic stroke wound *Smart intelligent computing and applications* (pp. 23-31): Springer.
- Roberts, T. P., & Mikulis, D. (2007). Neuro MR: principles. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *26*(4), 823-837.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-net: Convolutional networks for biomedical image segmentation*. Paper presented at the International Conference on Medical image computing and computer-assisted intervention.
- Roy, A. G., Conjeti, S., Navab, N., Wachinger, C., & Initiative, A. s. D. N. (2019). QuickNAT: A fully convolutional network for quick and accurate segmentation of neuroanatomy. *NeuroImage*, *186*, 713-727.
- Russo, C., Liu, S., & Ieva, A. D. (2020). *Impact of spherical coordinates transformation pre-processing in deep convolution neural networks for brain tumor segmentation and survival prediction*. Paper presented at the International MICCAI Brainlesion Workshop.
- Saeed, M. U., Ali, G., Bin, W., Almotiri, S. H., AlGhamdi, M. A., Nagra, A. A., . . . Amin, R. u. (2021). RMU-net: a novel residual mobile U-net model for brain tumor segmentation from MR images. *Electronics*, *10*(16), 1962.
- Shu, W., Wang, S., Chen, Q., Hu, Y., Cai, Z., & Lin, R. (2019). *Pathological image classification of breast cancer based on residual network and focal loss*. Paper presented at the Proceedings of the 2019 3rd International Conference on Computer Science and Artificial Intelligence.
- Silva, C. A., Pinto, A., Pereira, S., & Lopes, A. (2020). *Multi-stage deep layer aggregation for brain tumor segmentation*. Paper presented at the International MICCAI Brainlesion Workshop.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

- Singh, C., & Bala, A. (2019). A transform-based fast fuzzy C-means approach for high brain MRI segmentation accuracy. *Applied Soft Computing*, 76, 156-173.
- Singh, M., Venkatesh, V., Verma, A., & Sharma, N. (2020). Segmentation of MRI data using multi-objective antlion based improved fuzzy c-means. *Biocybernetics and Biomedical Engineering*, 40(3), 1250-1266.
- Sivakumar, V., & Janakiraman, N. (2020). A novel method for segmenting brain tumor using modified watershed algorithm in MRI image with FPGA. *Biosystems*, 198, 104226.
- Stieber, V. W. (2001). Low-grade gliomas. *Current treatment options in oncology*, 2(6), 495-506.
- Sze, V., Chen, Y.-H., Yang, T.-J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, 105(12), 2295-2329.
- thebraintumourcharity. What is a brain tumour? Retrieved from <https://www.thebraintumourcharity.org/brain-tumour-diagnosis-treatment/how-brain-tumours-are-diagnosed/brain-tumour-biology/what-is-a-brain-tumour/>
- Wadhwa, A., Bhardwaj, A., & Verma, V. S. (2019). A review on brain tumor segmentation of MRI images. *Magnetic resonance imaging*, 61, 247-259.
- Wang, F., Jiang, R., Zheng, L., Meng, C., & Biswal, B. (2019). *3d u-net based brain tumor segmentation and survival days prediction*. Paper presented at the International MICCAI Brainlesion Workshop.
- Wang, J. (2019). *Medical Imaging Semantic Segmentation with Ising Conditional Random Field Post-processing and Lesion Localization*. University of Toronto (Canada).
- Wang, S., Zhang, C., & Wu, M. (2019). *Accurate Semantic Segmentation in Remote Sensing Image*. Paper presented at the Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition.
- Wang, Y., Zhang, Y., Hou, F., Liu, Y., Tian, J., Zhong, C., . . . He, Z. (2020). *Modality-pairing learning for brain tumor segmentation*. Paper presented at the International MICCAI Brainlesion Workshop.
- Weninger, L., Rippel, O., Koppers, S., & Merhof, D. (2018). *Segmentation of brain tumors and patient survival prediction: Methods for the brats 2018 challenge*. Paper presented at the International MICCAI brainlesion workshop.
- Wilber, S. G. a. M. (2016). Training and investigating Residual Nets. Retrieved from <http://torch.ch/blog/2016/02/04/resnets.html>
- Xiao, X., Lian, S., Luo, Z., & Li, S. (2018). *Weighted res-unet for high-quality retina vessel segmentation*. Paper presented at the 2018 9th international conference on information technology in medicine and education (ITME).
- Yang, C., Guo, X., Wang, T., Yang, Y., Ji, N., Li, D., . . . Ma, T. (2019). *Automatic brain tumor segmentation method based on modified convolutional neural network*. Paper presented at the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
- Yeghiazaryan, V., & Voiculescu, I. (2015). An overview of current evaluation methods used in medical image segmentation. *Department of Computer Science, University of Oxford*.
- Yeung, M., Sala, E., Schönlieb, C.-B., & Rundo, L. (2022). Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, 95, 102026.
- Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., & Shen, D. (2015). Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation. *NeuroImage*, 108, 214-224.

Zhang, W., Yang, G., Huang, H., Yang, W., Xu, X., Liu, Y., & Lai, X. (2021). ME-Net: Multi-encoder net framework for brain tumor segmentation. *International journal of imaging systems and technology*, 31(4), 1834-1848.