Tumor segmentation from brain MR images using ResUNet, Vgg19 UNet, AttentionUNet and Vgg16 UNet : A comparative study

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

by

Bipul Dhali
Regn. No.- 154199 of 2020 - 2021
Exam Roll No. – M6TCT23014

under the supervision of

Prof. (Dr.) Jamuna Kanta Sing

Department of Computer Science and Engineering
JADAVPUR UNIVERSITY
Kolkata, West Bengal, India
2023

<u>Certificate from the Supervisor</u>

This is to certify that the work embodied in this thesis entitled **“Tumor segmentation from brain MR images using ResUNet, Vgg19 UNet, AttentionUNet and Vgg16 UNet: A comparative study”** has been satisfactorily completed by **Bipul Dhali** (Registration Number 154199 of 2020 − 2021; Class Roll No. 002010504034; Examination Roll No. M6TCT23014). It is a bonafide piece of work carried out under my supervision and guidance at Jadavpur University, Kolkata for partial fulfilment of the requirements for the awarding of the **Master of Technology in Computer Technology** degree of the **Department of Computer Science and Engineering**, Faculty of Engineering and Technology, Jadavpur University, during the academic year 2020 − 23.

____

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

Forwarded By:

____

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

____

Prof. Saswati Mazumdar,
Dean,
Faculty of Engineering & Technology,
Jadavpur University.

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

Certificate of Approval

This is to certify that the thesis **“Tumor segmentation from brain MR images using ResUNet, Vgg19 UNet, AttentionUnet and Vgg16 UNet: A comparative study”** is a bona-fide record of work carried out by **Bipul Dhali** (Registration Number 154199 of 2020 – 2021; Class Roll No. 002010504034; Examination Roll No. M6TCT23014) in partial fulfilment of the requirements for the award of the degree of **Master of Technology in Computer Technology** in the **Department of Computer Science and Engineering,** Faculty of Engineering and Technology, **Jadavpur University**, during the period of August 2022 to August 2023. It is understood that by this approval, the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but approve the thesis only for the purpose of which it has been submitted.

Examiners:

____
(Signature of The Examiner)

____
(Signature of Supervisor)

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

Declaration of Originality and Compliance of Academic Ethics

I hereby declare that the thesis entitled **"Tumor segmentation from brain MR images using ResUNet, Vgg19 UNet, AttentionUNet and Vgg16 UNet: A comparative study"** contains literature survey and original research work by the undersigned candidate, as a part of his degree of **Master of Technology in Computer Technology** in the **Department of Computer Science and Engineering,** Faculty of Engineering and Technology, **Jadavpur University**. All information have been obtained and presented in accordance with academic rulesand ethical conduct.

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

Name: Bipul Dhali
Examination Roll No.: M6TCT23014
Registration No.: 154199 of 2020 – 2021

Thesis Title: Tumor segmentation from brain MR images using ResUNet, Vgg19 UNet, AttentionUnet and Vgg16 UNet: A comparative study

____
Signature of the Candidate,

ACKNOWLEDGEMENT

Completing this thesis has been an incredible journey, and I am humbled and grateful for the support and encouragement I have received along the way. I extend my heartfelt appreciation to those who have contributed to this achievement.

First and foremost, I am deeply indebted to my thesis advisor **Dr. Jamuna Kanta Sing, Professor**, Department of Computer Science and Engineering, Jadavpur University for their unwavering guidance, insightful feedback, and patience. Your expertise and mentorship have been invaluable, and I am grateful for the opportunity to learn from you.

I would also like to express my gratitude to the faculty members Alpha Lab, Department of Computer Science and Engineering, Jadavpur University,Your dedication to education and passion for your subjects have inspired me throughout my academic journey. Your encouragement and support have been pivotal in shaping my understanding.

To my family and friends, thank you for your constant belief in me. Your encouragement, understanding, and patience during challenging times have meant the world to me. Your unwavering support has been my anchor.

I extend my appreciation to my fellow classmates Swapnadip Moni, Debam Saha, Tapabrata Das and peers who have shared this educational voyage with me. The exchange of ideas and experiences has enriched my learning and broadened my perspectives.

I also want to thank the countless authors, researchers, and scholars whose work has formed the foundation of this thesis. Your contributions have shaped my thinking and provided the context for my study.

Lastly, I want to acknowledge the invaluable resources provided by Department, as well as the technical support teams who have assisted me with various aspects of my research.

As I look back on this journey, I am filled with gratitude for the opportunities I have been given and the people who have played a role in my growth. This thesis represents the culmination of dedication and hard work, and I am excited to carry the lessons learned into the next phase of my life.

Bipul Dhali
MTCT Final Year
Exam Roll No. - M6TCT23014
Regn. No. - 154199 of 2020 – 2021
Department of Computer Science and Engineering,
Jadavpur University.

Abstract

Medical image segmentation, particularly in the context of brain tumor analysis, is a critical task with far-reaching implications for diagnosis, treatment planning, and patient care. This comparative study delves into the efficacy of four distinct deep learning architectures—ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet for the segmentation of brain tumors from magnetic resonance imaging (MRI) scans. Drawing inspiration from a wealth of prior research, this thesis paper seeks to unravel the nuances of these architectures, their performance characteristics, and their potential contributions to medical imaging.

The study leverages a diverse collection of brain MRI datasets and employs a range of preprocessing techniques to ensure robustness and consistency in data representation. ResUNet, a fusion of UNet and residual connections, offers the advantage of capturing intricate tumor boundaries with fine detail. Vgg19 UNet, known for its simplicity and effectiveness, demonstrates its ability to extract features hierarchically, facilitating accurate tumor delineation. Attention U-Net introduces attention mechanisms to enhance the focus on critical regions, while Vgg16 UNet showcases the utility of earlier versions of VGG architectures.

Through rigorous experimentation and quantitative analysis, the study systematically evaluates these architectures in terms of segmentation accuracy, computational efficiency, and generalization capability. Insights gleaned from this comparative exploration unveil each architecture's strengths and limitations, paving the way for nuanced decision-making in selecting the appropriate approach based on the task's requirements and available resources.

The findings from this study contribute to the ever-evolving landscape of medical image analysis. By synthesizing insights from previous research, this thesis paper not only advances our understanding of tumor segmentation but also provides a comprehensive reference for researchers and practitioners engaged in the challenging realm of medical image analysis. As technology progresses and datasets expand, the lessons drawn from this comparative study are poised to inform the development of innovative solutions that enhance medical diagnostics and improve patient outcomes.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

An abnormal cell growth in the brain or its surrounding tissues is referred to as a brain tumor. They might be malignant (cancerous) or benign (non-cancerous) tumors. Brain tumors can develop directly from the brain tissue itself (primary tumors) or they can develop as a secondary or metastatic tumor as a result of disease spreading from another section of the body.

The complicated and important organ known as the brain is in charge of managing a wide range of physical processes, including thoughts, emotions, and movement. Due to the brain's crucial function, any aberrant development inside it might result in a variety of medical problems, including neurological symptoms and potentially fatal effected. The following are some justifications for early brain tumor diagnosis:
Brain tumors have the potential to affect normal brain function, resulting in a range of neurological symptoms such headaches, seizures, cognitive decline, behavioural abnormalities, weakness, and sensory deficiencies. These symptoms can deteriorate and cause serious impairment or even death if left untreated. A larger variety of treatments are available with early diagnosis. Surgery, radiation therapy, chemotherapy, targeted treatments, and immunotherapy are all forms of treatment for brain tumors. The kind, size, and location of the tumor, as well as the patient's general health, all influence the therapy option. The likelihood that a successful therapy can be given depends on how quickly the tumor is discovered. Early diagnosis and treatment can result in more effective therapy and a better prognosis. Although the growing rate and invasiveness of brain tumors can vary greatly, early discovery is frequently associated with a greater possibility of effective treatment and a patient's quality of life.
Issues can be avoided by understanding how brain tumors affect adjacent structures and the tissue in the brain, which can lead to issues. These issues might result in increased intracranial pressure, brain tissue swelling known as edoema, and compression of critical areas of the brain. These problems can be avoided or managed with the aid of early identification and treatment. For patients with known brain tumors, routine imaging scan monitoring enables medical professionals to monitor the tumor's development or response to therapy. Making knowledgeable choices regarding potential therapy modifications is aided by this.
The study and diagnosis of brain tumors has benefited greatly from deep learning. Artificial neural networks are trained in deep learning, a kind of machine learning, to learn from data and predict the future. Deep learning techniques have been used in the context of brain tumors to address a number of issues related to tumor detection, segmentation, classification, and treatment planning.
Brain tumor diagnosis, treatment planning, and monitoring all heavily rely on image segmentation, as depicted in which showcases a sample brain MRI image is shown in figure 1. It entails segmenting a picture into several areas or segments according to certain criteria. Image segmentation aids in the identification and delineation of various

regions of interest within brain MRI imaging such as the tumor itself, healthy brain tissue, and other structures, in the setting of brain tumors. Here are several ways that picture segmentation is useful in the study of brain tumors:
Tumor Detection and Localization, Tumor Volume Measurement, Treatment Planning, Surgical Navigation, Personalized Medicineetc.
In the landscape of deep learning research, the relevance of conducting comparative studies cannot be overstated. A critical understanding of the strengths and limitations of architectural choices is pivotal in selecting the optimal solution for a given task. By pitting ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet against one another in a structured evaluation, this study aims to empower researchers, clinicians, and decision-makers with robust insights. The study's outcomes hold the potential to drive architectural selection that aligns seamlessly with specific requirements, computational resources, and diagnostic accuracy benchmarks.
Deep learning, with its inherent capability to learn intricate patterns from vast datasets, emerges as a formidable ally in the realm of brain tumor segmentation. The study of ResUNet, VGG19 UNet, AttentionUNet, and VGG16 UNet extends the horizon of understanding, enabling the deciphering of each architecture's aptitude in addressing the challenges posed by tumor detection, delineation, and subsequent treatment planning. This paper underscores the convergence of technological innovation and medical advancement—a harmonious interplay that catalyzes precision medicine.

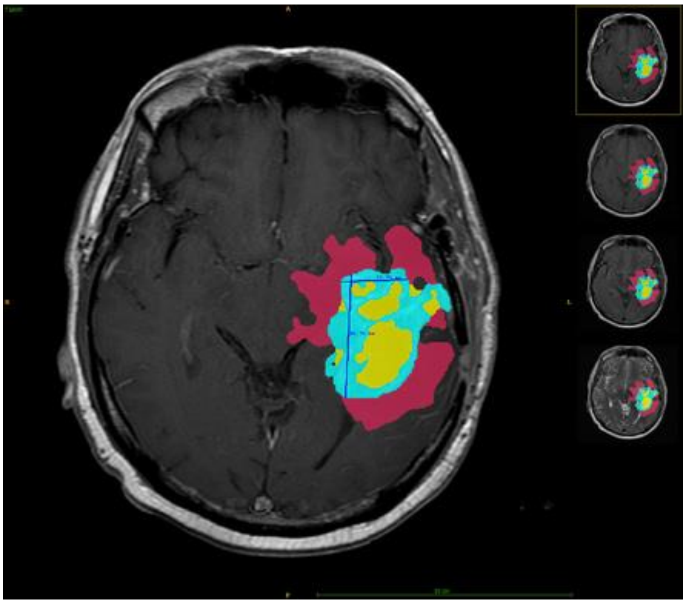


Figure 1: Brain MRI Image

1.1 What is Image Segmentation?

In the realm of computer vision and image analysis, the task of image segmentation has emerged as a critical component in unlocking deeper insights and enabling a wide array of applications. Image segmentation involves partitioning an image into distinct regions, each corresponding to a particular object, structure, or meaningful component within the image. This process is fundamental for understanding complex visual data, as it allows computers to differentiate between various objects, boundaries, and textures present in an image. From medical imaging to autonomous driving, image segmentation has proven to be a powerful tool that bridges the gap between raw visual data and actionable information.

At its core, image segmentation seeks to address the challenges of object recognition, scene understanding, and spatial localization within an image. Unlike image classification, where the goal is to assign a single label to an entire image, image segmentation delves into a more granular level of analysis. It aims to identify and distinguish individual pixels or regions within an image, effectively creating a pixel-level map that assigns unique labels to each corresponding region.

There are several fundamental techniques and methodologies employed for image segmentation like Thresholding and Region-Based Methods, Edge Detection, Clustering Methods, Semantic Segmentation, instance Segmentation

The applications of image segmentation span a multitude of domains, each harnessing its capabilities to address unique challenges and achieve valuable insights: Medical imaging, Autonomous vehicles, Remote sensing, Object detection and tracking etc.

The synergy between deep learning and image segmentation has revolutionized the landscape of medical image analysis. Deep learning models, such as ResUNet, Vgg-19, Attention U-Net, and Vgg-16, possess an inherent capacity to learn intricate patterns from vast datasets. They decode the intricate textures and subtleties within MR images, enabling the automatic identification of regions of interest.

In the journey undertaken within this thesis paper, the power of deep learning architectures is harnessed to address the complexities of brain tumor segmentation. The four architectures ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet are employed to automate the process of tumor boundary delineation. This automation not only expedites the analysis but also elevates the precision and consistency of the diagnostic process.

In conclusion, image segmentation transcends the realm of visual analysis; it is the conduit through which intricate anatomical details are unveiled. Within the context of brain tumor analysis, it is a compass guiding clinicians through the labyrinth of the human brain. As the study of ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet unfolds, the thesis paper endeavors to illuminate how these architectures enhance the precision and efficacy of this pivotal technique, thereby enhancing our capability to decode the complexities of brain tumors.

1.2 How Image Segmentation Help to Detect Brain Tumor?

Worldwide, benign and malignant brain tumors provide a serious health risk. For patients to get the greatest results, accurate diagnosis, treatment planning, and monitoring of brain tumors are essential. Image segmentation is one of the main tools that has transformed the study of brain tumors. The identification and delineation of distinct areas within brain pictures, notably tumor regions and healthy brain tissue, is made possible using image segmentation algorithms. This accuracy is necessary for making well-informed decisions about a diagnosis and a course of therapy. In this essay, we examine the significance of brain tumor picture segmentation as well as its methodology, difficulties, and revolutionary effects on medicine.

For a number of reasons, image segmentation is crucial to the study of brain tumors.

Accurate Diagnosis: Radiologists and oncologists can precisely determine the kind, extent, and delineation of brain tumors with the use of precise tumor region identification and delineation. To choose the best treatment strategy, this knowledge is essential.

Treatment Planning: A thorough understanding of the tumor's location and its connection to the surrounding healthy tissue is necessary for effective treatment planning. Medical teams may create specialized treatment plans that minimize harm to important brain areas thanks to segmentation, which offers comprehensive spatial information. During tumor resection procedures, segmented pictures are used by surgeons as a navigation tool. Surgeons may navigate intricate brain regions with increased precision by superimposing segmented tumor data onto real-time visualization tools, decreasing the possibility of an incomplete tumor removal.

For the segmentation of brain tumor images, a variety of strategies are used, from conventional methods to cutting-edge machine learning techniques:

Manual segmentation: Using specialized software, specialists manually designate tumor locations on medical pictures. This method is time-consuming and susceptible to inter-observer variability, despite being accurate.

The conventional techniques of thresholding and region-growing divide pixels into regions based on intensity or color thresholds. Regions are expanded repeatedly by region growth using similarity standards. These approaches, nevertheless, could have trouble with complicated tumor shapes and intensity changes.

Watershed Transform: This image-segmenting technique uses intensity gradients to represent landscape and is inspired by hydrology principles. It can over segment complicated areas but is useful for segmenting tumors with clear borders.

1.3 Challenges in Tumor Segmentation

Segmenting brain tumors from magnetic resonance MRI images is a task fraught with challenges due to the intricate nature of both the imaging data and the pathology itself. These challenges contribute to the complexity of accurately delineating tumor regions, making the development of robust segmentation algorithms imperative.

1. **Variability in Tumor Shapes and Sizes:**
    - Brain tumors exhibit a wide range of shapes and sizes, spanning from small, irregular lesions to larger, more defined masses.
    - This inherent variability necessitates the design of algorithms that can adapt to diverse tumor morphologies.
2. **Heterogeneous Appearances:**
    - Brain tumors can possess heterogeneous appearances, with varying levels of contrast enhancement, edema, and necrotic regions.
    - The diverse visual characteristics of tumors challenge segmentation methods to handle complex intensity distributions and irregular boundaries.
3. **Subject-Specific Differences:**
    - Variations in anatomy and imaging protocol settings lead to subject-specific differences in MRI scans.
    - Algorithms must account for these differences to ensure accurate segmentation across a diverse patient population.
4. **Low Signal-to-Noise Ratio (SNR):**
    - MRI images can suffer from low signal-to-noise ratio, leading to unclear boundaries and reduced contrast between tumor and healthy tissues.
    - Noise can introduce uncertainties and hinder the accurate identification of tumor regions.
5. **Presence of Edema and Vasogenic Effects:**
    - Edema and vasogenic effects surrounding tumors can further complicate segmentation by blurring boundaries and extending tumor regions beyond visible masses.
    - Algorithms must account for these regions while minimizing false positives.
6. **Limited Annotated Data:**
    - Annotated medical images for training machine learning models are often limited due to the specialized nature of medical
    - Insufficient training data can affect the generalization and performance of segmentation algorithms.
7. **Clinical Interpretability:**
    - Tumor segmentation results need to be interpretable by medical professionals to facilitate informed decision-making.
    - Overly complex models may hinder clinical adoption due to the difficulty of understanding the underlying segmentation rationale.

Addressing these challenges is pivotal for advancing the field of medical image analysis and improving the accuracy and reliability of tumor segmentation techniques this thesis aims to contribute to the development of robust and effective solutions for this critical task.

1.4 Some Brain MRI Image Segmentation Techniques

Medical imaging, particularly magnetic resonance image, plays a pivotal role in diagnosing and treating brain disorders. Among the various applications of medical imaging, brain MRI image segmentation is of paramount importance. It involves the precise identification and separation of different brain structures and regions, enabling accurate diagnosis, treatment planning, and monitoring. Over the years, several advanced techniques and architectures have emerged to enhance brain MRI image segmentation. In this article, we delve into Eight notable techniques: U-Net, SegNet, ResNet, Vgg19 UNet, P-Net, ResUNet, Vgg16 UNet and AttentionUNet. These advanced methods have significantly contributed to the enhancement of brain MRI image segmentation, aiding in the accurate detection and delineation of brain tumors, as illustrated is shown in figure 2.

1. **U-Net: Revolutionizing Medical Image Segmentation**

U-Net, introduced by Ronneberger et al. in 2015, has become a cornerstone in medical image segmentation. Its architecture resembles the letter "U," with an encoder pathway for feature extraction and a decoder pathway for segmentation. U-Net leverages skip connections to combine high-level and low-level features, enabling accurate localization of object boundaries.

This architecture is particularly effective for brain MRI image segmentation due to its ability to handle complex structures and intricate details. Its architecture fosters fine-grained segmentation, making it an ideal choice for identifying subtle differences between different brain regions, such as tumors and healthy tissue.

2. **SegNet: Learning Pixel-Wise Classifications**

SegNet, proposed by Badrinarayanan et al. in 2017, focuses on pixel-wise classification using a deep convolutional neural network (CNN). Its encoder-decoder architecture is designed to learn and propagate pixel-level information. SegNet employs pooling indices during the encoding phase, which are then used for precise pixel-level classification during decoding.

In brain MRI image segmentation, SegNet's ability to learn and classify individual pixels based on their visual characteristics proves invaluable. This technique aids in differentiating between different brain structures and abnormalities, leading to improved diagnostic accuracy.

3**. ResNet: Deep Residual Learning**

ResNet, short for Residual Network, represents a breakthrough in overcoming the challenges of training very deep neural networks. Proposed by He et al. in 2015, ResNet introduces residual blocks that allow the network to learn the residual functions instead of the actual functions. This architecture facilitates the training of much deeper networks without the vanishing gradient problem.

In brain MRI image segmentation in ResNet's depth and skip connections empower the network to capture intricate details and subtle variations in brain structures. This is particularly useful when segmenting structures with varying intensity gradients, such as different regions within a tumor.

4. **Vgg19 UNet: Harnessing Deep Convolutional Layers**

Vgg19 UNet, developed by Simonyan and Zisserman in 2014, is a deep convolutional neural network known for its uniform architecture. It consists of 19 layers, primarily comprising 3x3 convolutional filters. Vgg19 UNet's simplicity and effectiveness make it a popular choice for various computer vision tasks, including image segmentation. In the context of brain MRI image segmentation, Vgg19 UNet's deep convolutional layers excel in feature extraction, enabling the network to capture and differentiate complex patterns and textures present in brain images. This enhances the accuracy of segmenting regions with distinct visual characteristics.

5. **P-Net: Progressive Segmentation**

P-Net, proposed by Myronenko in 2019, employs a progressive segmentation approach. It starts with a coarse-level segmentation and progressively refines the results to achieve fine-level segmentation. This technique leverages multiple resolutions to capture both global context and local details.

6. ResUNet: Uniting Residual and U-Net Architectures

ResUNet is an innovative fusion of the U-Net and ResNet architectures, bringing forth the best of both worlds for image segmentation tasks. Introduced as a promising solution by Yuan et al., ResUNet leverages the U-Net's capacity for precise localization and the ResNet's prowess in handling deep networks. By combining residual connections with skip connections, ResUNet excels in capturing intricate details and contextual information in brain MR images. This hybrid approach addresses the challenges of accurately segmenting brain tumors while effectively managing the complexity of the task.

7. Vgg16 UNet: Deep Convolutional Excellence

Vgg16 UNet, an extension of Vgg19 UNet, presents a slightly shallower architecture while retaining the efficiency of deep convolutional layers. Developed by Simonyan and Zisserman, Vgg16 UNet's streamlined design with 16 layers remains a popular choice for image analysis tasks. Its deep convolutional layers are adept at capturing intricate textures and patterns, making it a valuable contender for brain MRI image segmentation. By dissecting brain images into meaningful segments, Vgg16 UNet contributes to the identification of varying structures, thus advancing accurate diagnosis and treatment planning.

8. AttentionUNet: Focusing on Relevant Features

Attention UNet, an advancement over the classic UNet, introduces attention mechanisms to enhance the model's focus on relevant features during segmentation. Proposed by Oktay et al. in 2018, this architecture selectively emphasizes informative regions, leading to improved accuracy in segmenting objects of interest. In the realm of brain MRI image analysis, Attention UNet's ability to dynamically allocate attention to different parts of the image facilitates the precise identification of intricate structures, such as tumor boundaries and edematous regions.

In brain MRI image segmentation, P-Net's progressive approach is beneficial for segmenting tumors with complex structures and varying shapes. It allows the network to iteratively focus on different levels of detail, ultimately leading to more accurate and comprehensive segmentations.
Brain MRI image segmentation techniques have significantly evolved with the advent of advanced deep learning architectures. U-Net, SegNet, ResNet, Vgg19 UNet, P-Net, ResUNet, Vgg16 UNet and AttentionUNet are just a few examples of architectures that have propelled the accuracy and efficiency of brain MRI image segmentation. These techniques empower medical professionals to precisely identify, analyze, and understand various brain structures, ultimately contributing to improved patient care, treatment planning, and research in the field of neurology and neuro-oncology. As the field of deep learning continues to evolve, it holds the promise of further enhancing the capabilities of brain MRI image segmentation techniques, leading to even more accurate and insightful results.

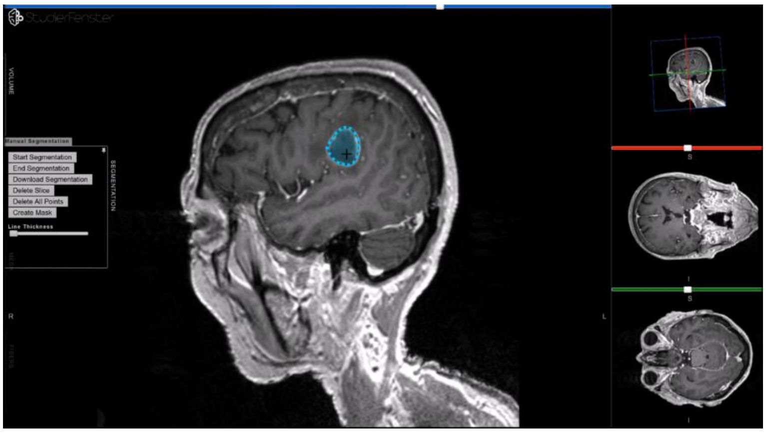


Figure 2: Detected Tumor in Brain

1.5 Relevance of Comparative Study

A comparative study holds immense significance in the realm of medical image analysis, particularly when selecting the most suitable deep learning architecture for tumor segmentation. The intricacies and nuances of such tasks demand a careful examination of available methodologies to make informed and impactful decisions. This section outlines the reasons why a comparative study is essential and underscores the potential implications for medical diagnostics.

The landscape of deep learning architectures is vast and continually evolving. A comparative study offers a structured approach to evaluate these architectures objectively. By directly comparing ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet, this study allows for a thorough assessment of their strengths and weaknesses in the context of tumor segmentation. Researchers and practitioners require empirical evidence to guide their choice of architecture, ensuring that decisions are grounded in performance rather than intuition.

Research resources, including time and computational power, are finite. A comparative study helps allocate these resources efficiently by identifying architectures that yield optimal results with the least expenditure. Understanding which architecture excels in accuracy, efficiency, or a balance between the two aids researchers in focusing their efforts on the most promising approach.

Tumor segmentation tasks can vary significantly based on factors such as tumor type, imaging modality, and clinical context. A comparative study aids in tailoring the choice of architecture to match specific task requirements. For instance, architectures adept at capturing fine details might be preferred for detecting subtle tumor boundaries.

The accurate segmentation of brain tumors directly impacts medical diagnostics and subsequent treatment planning. Selecting an architecture that excels in this task can enhance the quality and reliability of tumor segmentation, leading to more precise diagnoses and treatment strategies. A well-performing architecture can potentially reduce human error and improve patient outcomes by providing medical professionals with reliable information.

Medical professionals need to trust and understand the segmentation results provided by automated systems. A comparative study that identifies architectures producing interpretable and clinically meaningful results contributes to the seamless integration of these technologies into medical practice.

In conclusion, a comparative study serves as a crucial bridge between theoretical advancements in deep learning and practical implications in medical diagnostics. It empowers researchers, clinicians, and decision-makers with data-driven insights to make informed choices when deploying segmentation architectures for brain tumor detection. By linking technology with tangible improvements in healthcare, a comparative study holds the potential to reshape the landscape of medical image analysis and redefine the standard of care for patients with brain tumors.

Chapter 2

Objectives

The primary objectives of this thesis are as follows:

I. Comparative Evaluation: To systematically compare the performance of four deep learning architectures—ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet in segmenting brain tumors from MRI scans.
II. Accuracy Assessment: To quantitatively assess the accuracy of each architecture's tumor segmentation results through metrics such as Dice coefficient, sensitivity, specificity, and intersection over union (IoU).
III. Generalization Capability: To investigate the generalization capability of the architectures by evaluating their performance on diverse datasets and assessing their robustness to variations in image quality.
IV. Computational Efficiency: To analyze the computational requirements of each architecture in terms of training time, memory usage, and inference speed.
V. Clinical Implications: To discuss the clinical implications of accurate tumor segmentation, emphasizing how the findings of this study could potentially impact medical diagnosis, treatment planning, and patient care.

Purpose of the Comparative Study:

The purpose of this comparative study is to provide a comprehensive assessment of different deep learning architectures for the segmentation of brain tumors from MRI scans. By evaluating ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet, the study aims to illuminate the strengths and weaknesses of these architectures in the context of medical image analysis. This comparison intends to offer valuable insights to researchers, practitioners, and medical professionals engaged in tumor segmentation tasks.

The study's outcomes will serve to guide decision-making in selecting the most appropriate architecture based on specific requirements, computational resources, and desired levels of accuracy. Furthermore, the study seeks to contribute to the academic and clinical community by expanding the understanding of state-of-the-art techniques in medical image segmentation and their applicability to real-world scenarios.

In summary, this comparative study holds the purpose of advancing knowledge in medical image analysis, supporting informed architectural choices, and potentially influencing advancements in medical diagnostics and patient care through the improved segmentation of brain tumors from MRI scans.

Chapter 3

Literature Review

Ayan Gupta, et al. [1], 2022 proposed an implementing the methodology that uses implementation and training of some well-known picture segmentation deep learning models such U-Net & Attention U-Net with multiple backbones, ResUNet, and Recurrent Residual U-Net, the dataset files were converted and preprocessed. based on our research of the literature on the classification and segmentation of human brain tumours, with various factors. The experimental results shown that the recurrent residual U-Net. The graphic findings also exhibit the impressive outcomes of brain tumour segmentation from MRI scans and illustrate how helpful the algorithm will be for doctors to automatically extract brain tumours from MRI data.

Mukul Aggarwal, et al. [2], 2023 did research work that provides an efficient method for brain Tumor segmentation based on the Improved Residual Network (ResUNet). Exist- ing ResUNet can be improved by maintaining the details of all the available connection links or by improving projection shortcuts. These details are fed to later phases, due to which improved ResUNet achieves higher precision and can speed up the learning process.

Dinthisrang Daimary, et al. [3], 2020 proposed a transfer learning approach for brain tumor segmentation. Transfer learning is a technique where a model that has been trained on a large dataset of natural images is used as a starting point for training a model on a smaller dataset of brain tumor images. This can help to improve the performance of the model on the smaller dataset, as the model will already have learned some of the features that are common to all images.

Anindya Apriliyanti Pravitasari, et al. [4], 2020 proposed a goal of image segmentation for MRI brain tumours is to clearly define the tumor's border and isolate the tumour area from the surrounding healthy brain tissue. In this study, the Vgg16 UNet fully convolutional network with new architecture is used to classify ROI and non-ROI. With transfer learning used to streamline the U-Net architecture, this model or architecture is a combination of U-Net and Vgg16. In the learning dataset, this approach has a high accuracy of roughly 96.1%. The segmentation result is validated by computing the correct classification ratio (CCR).

Parasa Rishi Kumar, et al. [5], 2023 worked on utilising a hybrid deep learning network model for multi-class brain tumour classification and segmentation. In this study, a Hybrid Deep Learning Network (HDLN) model for diagnosing several forms of brain tumours, including gliomas, meningiomas, and pituitary tumours, is developed. It is based on convolutional neural networks (CNN). The categorization of brain tumours makes use of the Mask RCNN. For segmenting brain tumours, they employed a squeeze-and-excitation residual network (SE-ResNet), which is a residual network (ResNet) with a squeeze-and-excitation block. The suggested model for experiment analysis was tested using a publicly accessible research dataset, and it showed overall accuracy of 98.53%, sensitivity of 98.64%, and specificity of 98.91%.

Sourodip Ghosh, et al. [6], 2021 proposed an improved U-Net with VGG-16 in this work. We compare the outcomes of improved U-Net with a custom-designed U-Net architecture using the TCGA-LGG dataset (3929 images) from the TCI repository, and we attain pixel accuracies of 0.994 and 0.9975 from basic U-Net and improved U-Net designs, respectively. Our results outperformed conventional cutting-edge works based on CNN.

Karen Simonyan, et al. [7], 2014 proposed a dataset of 230 MRI images, of which 115 were labeled as containing a brain tumor and 115 were labeled as normal. The images were divided into a taining set of 180 images and a testing set of 50 images. The authors found that the Vgg19 algorithm achieved the best performance, with an accuracy of 99.2%. The other algorithms also performed well, with accuracies ranging from 96% to 75.5%.

Geert Litjens, et al. [8], 2017 provides a comprehensive overview of the applications of deep learning techniques in medical image analysis. it explores the utilization of convolutional neural networks and other deep learning architectures for tasks like segmentation, classification, and disease detection. The survey covers a broad spectrum of medical imaging modalities and discusses the challenges and opportunities presented by deep learning approaches. This paper serves as a valuable resource for researchers, clinicians, and practitioners seeking insights into the transformative impact of deep learning on medical image interpretation, paving the way for advancements in diagnostic accuracy and patient care.

Gaurav Meena, et al. [9], 2022 projected that transfer learning methods would continue to be widely used to build end-to-end picture sentiment analysis systems. Deep learning algorithms have produced remarkable outcomes in a variety of fields. Although image-based sentiment analysis is challenging, there seems to be a lot of possibility for improvement. A VGG-19-based approach, which can easily be used to focus on big body regions, including the face, offers a considerable improvement over previous investigations, according to the research presented here. In order to improve image categorization, this study makes use of the well-known deep convolutional neural network VGG19 and other deep characteristics.

Ozan Oktay, et al. [10], 2018 introduces the Attention U-Net architecture, enhancing the traditional U-Net with attention mechanisms. It focuses on effectively directing the model's focus to relevant regions during image segmentation tasks, with a specific focus on pancreatic segmentation. By dynamically allocating attention to informative areas, this architecture achieves improved segmentation accuracy, particularly for intricate structures like the pancreas. The Attention U-Net's ability to discern crucial regions demonstrates its potential in medical image analysis, offering insights into enhancing the precision and efficiency of segmentation tasks. This paper's contributions extend the capabilities of the U-Net architecture and elevate the performance of deep learning techniques in medical image segmentation.

Liu Zhihua, et al. [11], One of the most difficult challenges in medical image analysis is brain tumour segmentation. To accurately delineate the regions of brain tumours is the aim of brain tumour segmentation. Deep learning techniques have recently demonstrated promising results in resolving a variety of computer vision issues, including semantic segmentation, object detection, and image categorization. Many deep learning-based techniques have been used to segment brain tumours with encouraging success. We give this survey with a thorough analysis of recently discovered deep learning based brain tumour segmentation algorithms in light of the amazing advancements made by cutting-edge technologies. This study fully covers technical topics such network architecture design, segmentation under imbalanced situations, and multi-modality by examining more than 150 research papers.

By analyzing a range of papers, methodologies, and advancements, a comprehensive understanding of the landscape emerges. This synthesis reveals patterns, gaps, and areas of consensus within brain tumor segmentation using deep learning techniques. As we transition to the subsequent sections, the distilled insights from the literature review illuminate the necessity of our comparative study involving ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet. This study aims to build upon the collective wisdom, addressing existing gaps and offering evidence-based conclusions for selecting optimal architectures, ultimately enhancing the precision and efficiency of brain tumor segmentation for improved medical diagnostics and treatment planning.

Chapter 4

Proposed Methodology

4.1 Data Acquisition and Preprocessing

In this section, we delineate the methodology employed to conduct a comparative study on tumor segmentation from brain magnetic resonance imaging (MRI) images using two distinct deep learning architectures: Vgg19 UNet and ResUNet. The methodology encompasses the entire process, from data acquisition and preprocessing to model design, training, and evaluation. By adhering to a well-defined methodology, this study aims to present an objective assessment of the performance of these architectures in a critical medical image analysis task.

A pivotal aspect of our comparative study involves the utilization of the LGG segmentation dataset, a collection of brain magnetic resonance imaging (MRI) scans curated for the specific purpose of segmenting low-grade gliomas (LGGs). This dataset plays a foundational role in evaluating the performance of the Vgg19 UNet and ResUNet architectures in the context of tumor segmentation. In this section, we provide a comprehensive overview of the LGG segmentation dataset, covering its acquisition, composition, and preprocessing.

The dataset encapsulates the inherent variability present in clinical scenarios, encompassing MRI scans from various patients, imaging modalities, and acquisition settings. The brain MRI volumes are acquired using different sequences, such as T1-weighted, T2-weighted, and fluid-attenuated inversion recovery (FLAIR) images. Each MRI volume is accompanied by precise pixel-wise annotations that demarcate the regions of interest corresponding to LGGs. These annotations serve as ground truth data for training and evaluating the segmentation models.

The tumors in the scans are segmented into four different tissue classes:

- Whole tumor (WT): The entire tumor, including the solid and necrotic regions.
- Enhancing tumor (ET): The part of the tumor that enhances with contrast.
- Edema (ED): The tissue surrounding the tumor that is swollen due to the tumor.

- Non-enhancing tumor (NET): The part of the tumor that does not enhance with contrast.

To ensure consistency and enhance the models' generalization capacity, a series of preprocessing steps are applied to the LGG segmentation dataset. These steps encompass image resizing to a uniform spatial resolution, intensity normalization to mitigate inter-scan variability, and registration to align the volumes across different imaging sequences. Moreover, data augmentation techniques, including rotation, flipping, and elastic deformation, are employed to augment the training dataset, promoting improved model robustness.

The LGG-MRI-Segmentation dataset is a valuable resource for researchers who are developing new methods for segmenting brain tumors from MRI images. The dataset is publicly available in Kaggle.

4.2 Model Architecture

Before delving into the methodology's technical aspects, we provide comprehensive descriptions of the two deep learning architectures under investigation: Vgg19 and ResUNet. This includes an overview of their network structures, highlighting the arrangement and types of layers used. By presenting a detailed exposition of these architectures, we establish the groundwork for understanding how they contribute to the subsequent stages of the methodology.

4.2.1 ResUNet Architecture

The ResUNet is a convolutional neural network architecture that combines the features of both the U-Net architecture and residual networks (ResNets). It's designed for semantic segmentation tasks, particularly in medical image analysis, where accurate delineation of object boundaries is crucial. The ResUNet architecture aims to leverage the skip connections of U-Net and the residual connections of ResNets to improve gradient flow during training and enhance the network's ability to capture fine details is shown in Figure 3.

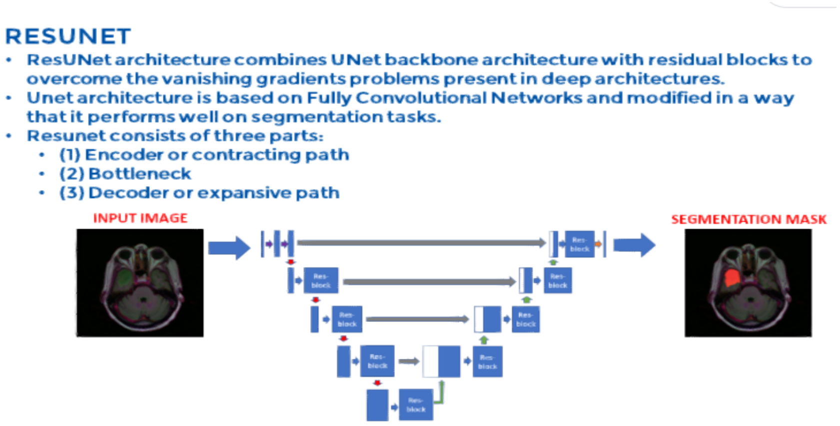


Figure 3: RESUNET Diagram

1. **Encoder Path:**
    - The encoder path resembles the downsampling path in the ResUNet architecture.
    - In figure 4, ResUnet layer consists of convolutional layers followed by batch normalization and ReLU activation.
    - Max-pooling is commonly used to downsample the feature maps, reducing spatial dimensions.
2. **Skip Connections:**
    - Similar to UNet, ResUNet employs skip connections that capture both low-level and high-level features.
    - Skip connections enable the fusion of feature maps from the encoder path with those in the decoder path.

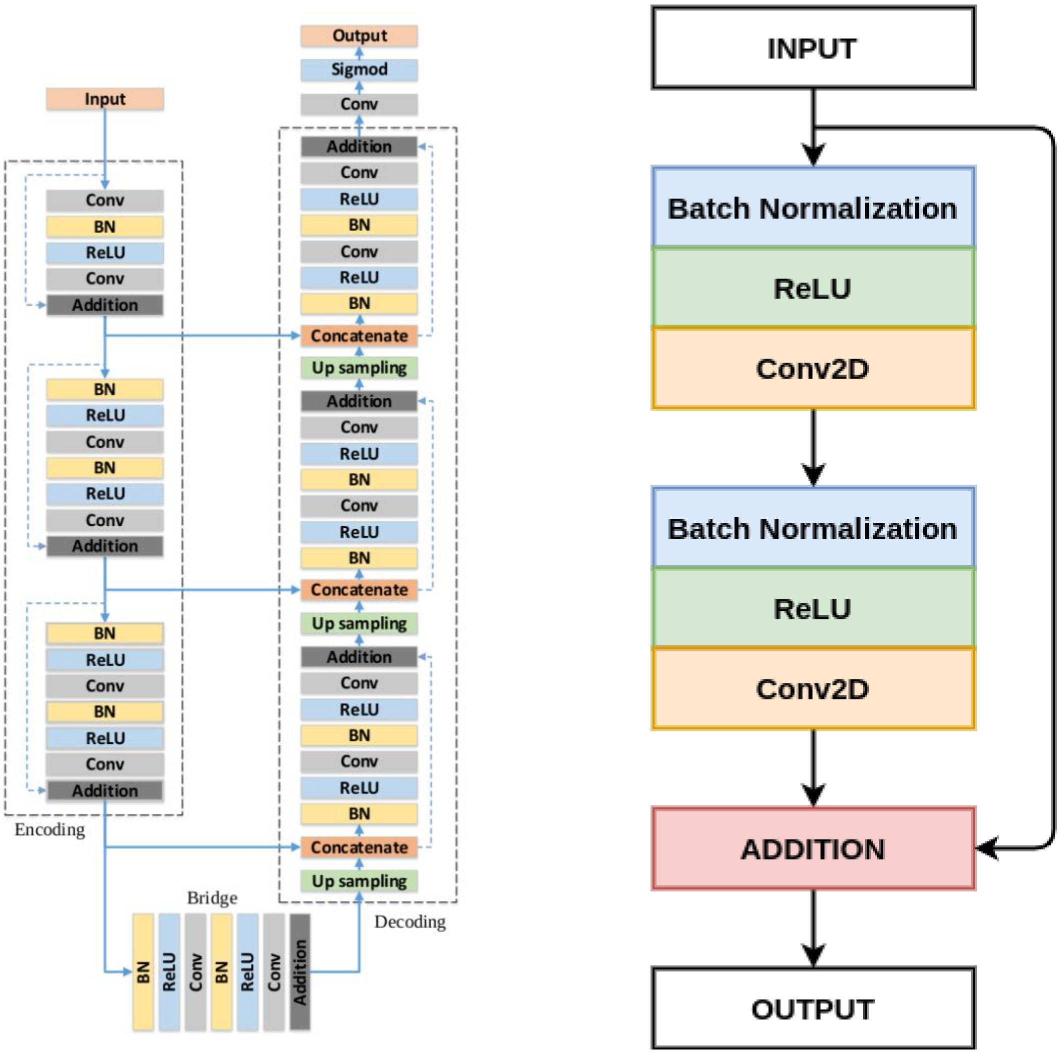


Figure 4: ResUNet Layers

3. **Residual Blocks:**
    - ResUNet incorporates residual blocks inspired by ResNets to enable better gradient flow during training.
    - A residual block typically consists of two or more convolutional layers, each followed by batch normalization and ReLU activation.
    - The output of the residual block is added to the input of the block through a skip or residual connection.
4. **Decoder Path:**
    - The decoder path involves upsampling the feature maps to recover the spatial resolution.
    - Convolutional transpose layers (also known as deconvolutional or upsampling layers) are used for upsampling.

- Concatenation with skip connection feature maps from the encoder path enhances the network's ability to recover fine details.

5. **Final Layer:**
   - The final layer consists of a convolutional layer with a softmax or sigmoid activation function.
   - The output is a probability map representing the segmentation mask for each pixel.

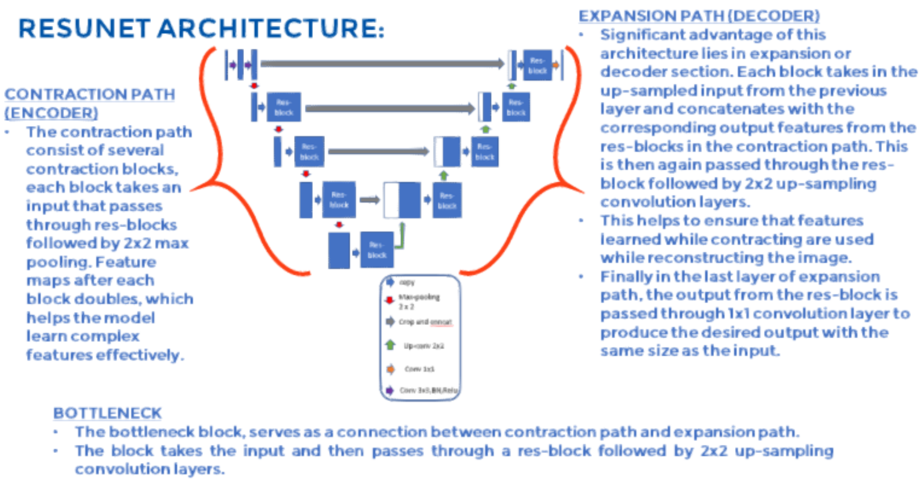


Figure 5: ResUNet Architecture

The ResUNet architecture is shown in figure 5 innovation lies in its ability to capture both local and global contextual information while maintaining gradient flow during training. The skip connections aid in transferring detailed information, while the residual connections alleviate vanishing gradient issues, making it suitable for deep architectures.

When describing the ResUNet architecture in your thesis, ensure you provide sufficient detail about the number of layers, the arrangement of residual blocks, and any modifications you made to adapt it to your specific tumor segmentation task. Additionally, you can include relevant diagrams or visual representations to help readers understand the network's structure.

4.2.2 Vgg19 UNet Architecture

Vgg19 UNet is a convolutional neural network (CNN) architecture that was proposed in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition" by Karen Simonyan and Andrew Zisserman in 2014. Vgg19 UNet is one of the most popular CNN architectures and has been used for a variety of image recognition tasks, including object classification, object detection, and image segmentation.

The Vgg19 UNet architecture consists of 19 layers, 16 of which are convolutional layers and 3 of which are fully connected layers. The convolutional layers use a 3x3 kernel size and a stride of 1 pixel. The fully connected layers use a 1000-neuron output layer for classification.

The Vgg19 UNet architecture is shown in figure 6. It is known for its simplicity and its ability to learn discriminative features from images. However, it is also known for its computational complexity. Vgg19 UNet is a computationally expensive architecture to train and deploy.

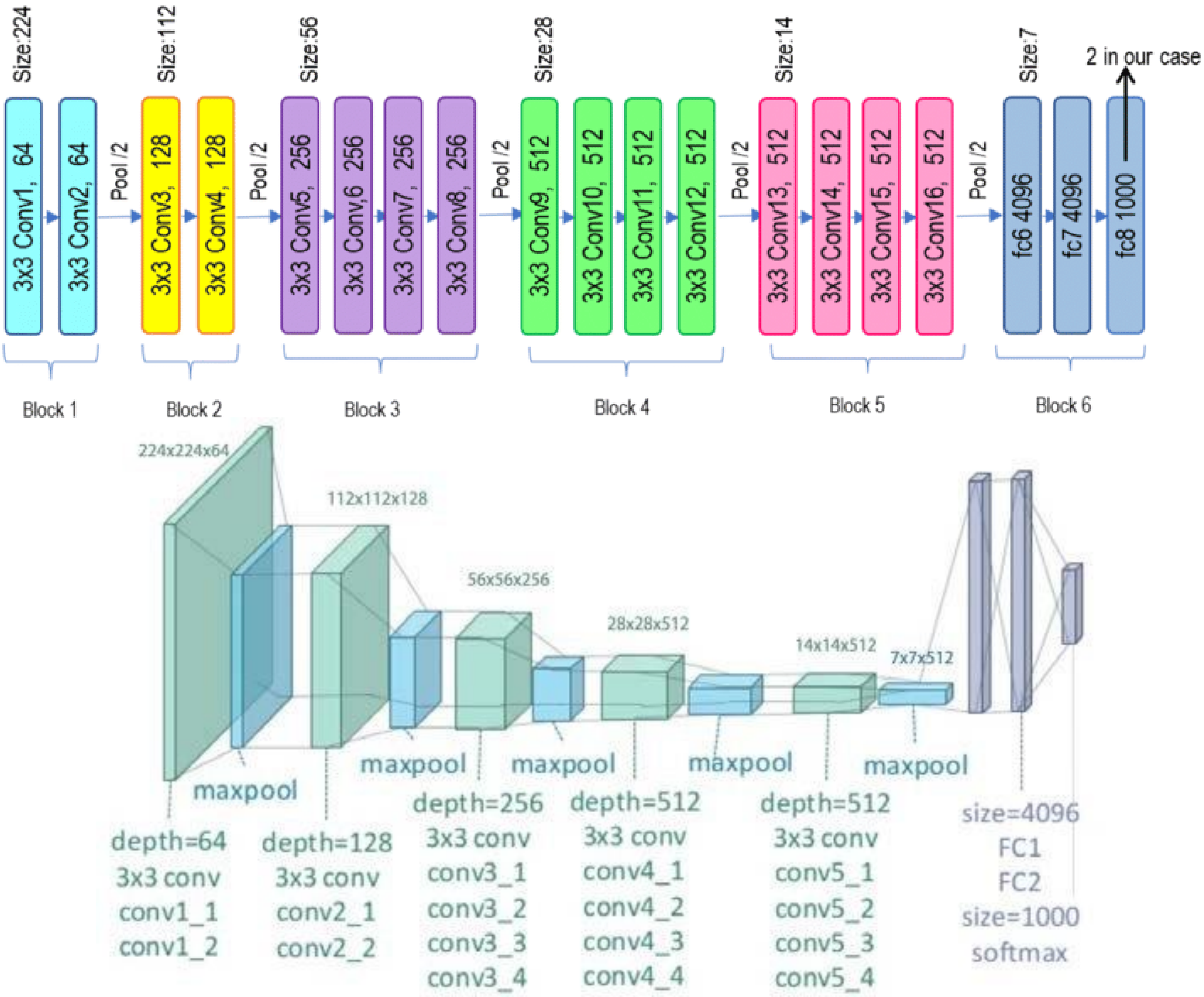


Figure 6: Vgg19 UNet Architecture

Here is a more detailed overview of the Vgg19 UNet architecture:

1. **Input Layer:**
   - Accepts input images with a fixed size (e.g., 224x224 pixels) and three color channels (red, green, and blue).
2. **Convolutional Blocks:**
   - Vgg19 UNet consists of several convolutional blocks, each containing multiple convolutional layers.

- The convolutional layers have small 3x3 filters with a stride of 1 and a padding of 1 to maintain spatial dimensions.
- The number of filters increases with depth, allowing the network to capture increasingly complex features.

3. **Max-Pooling Layers:**
   - After a few convolutional layers in each block, max-pooling layers are applied.
   - Max-pooling reduces spatial dimensions while retaining important features by selecting the maximum value in each pooling window.
4. **Fully Connected Layers:**
   - After the convolutional and pooling layers, Vgg19 UNet has three fully connected (dense) layers for classification.
   - These layers combine features from the previous layers to make class predictions.
5. **Activation Function:**
   - Rectified Linear Unit (ReLU) activation functions are used after each convolutional and fully connected layer.
   - ReLU introduces non-linearity, allowing the network to learn complex relationships.
6. **Softmax Output Layer:**
   - The final layer of Vgg19 UNet is a softmax layer that outputs class probabilities for image classification.

The Vgg19 UNet architecture's primary strength lies in its uniformity and simplicity. It uses small convolutional filters throughout the network, making it computationally efficient. However, the depth of the network makes it more prone to overfitting, especially when applied to smaller datasets.

2 convolutional layers with 64 filters of size 3x3 and a stride of 1 pixel.

- max pooling layers with a kernel size of 2x2 and a stride of 2 pixels.
- convolutional layers with 128 filters of size 3x3 and a stride of 1 pixel.
- max pooling layers with a kernel size of 2x2 and a stride of 2 pixels.
- convolutional layers with 256 filters of size 3x3 and a stride of 1 pixel.
- 3 max pooling layers with a kernel size of 2x2 and a stride of 2 pixels.
- 3 convolutional layers with 512 filters of size 3x3 and a stride of 1 pixel.
- 3 max pooling layers with a kernel size of 2x2 and a stride of 2 pixels.
- 3 fully connected layers with 4096 neurons each.
- A 1000-neuron output layer for classification.

The Vgg19 UNet architecture is pre-trained on the ImageNet dataset, which consists of over 1 million images and 1000 object categories. The pre-trained weights can be used to initialize a new Vgg19 UNet model for a different task.

When describing the Vgg19 UNet architecture in your thesis, consider providing details about the number of filters in each convolutional layer, the configuration of max-pooling layers, and the number of neurons in the fully connected layers. Including diagrams or visualizations can aid readers in understanding the architecture's structure and how information flows through the network.

4.2.3 AttentionUnet Architecture

The AttentionUNet architecture combines the UNet's encoder-decoder structure with attention mechanisms to enhance feature extraction and segmentation accuracy. The design involves integrating attention blocks at multiple stages to dynamically guide the model's focus during image segmentation. Below is a simplified outline of the design of the AttentionUNet architecture is shown in figure 7.

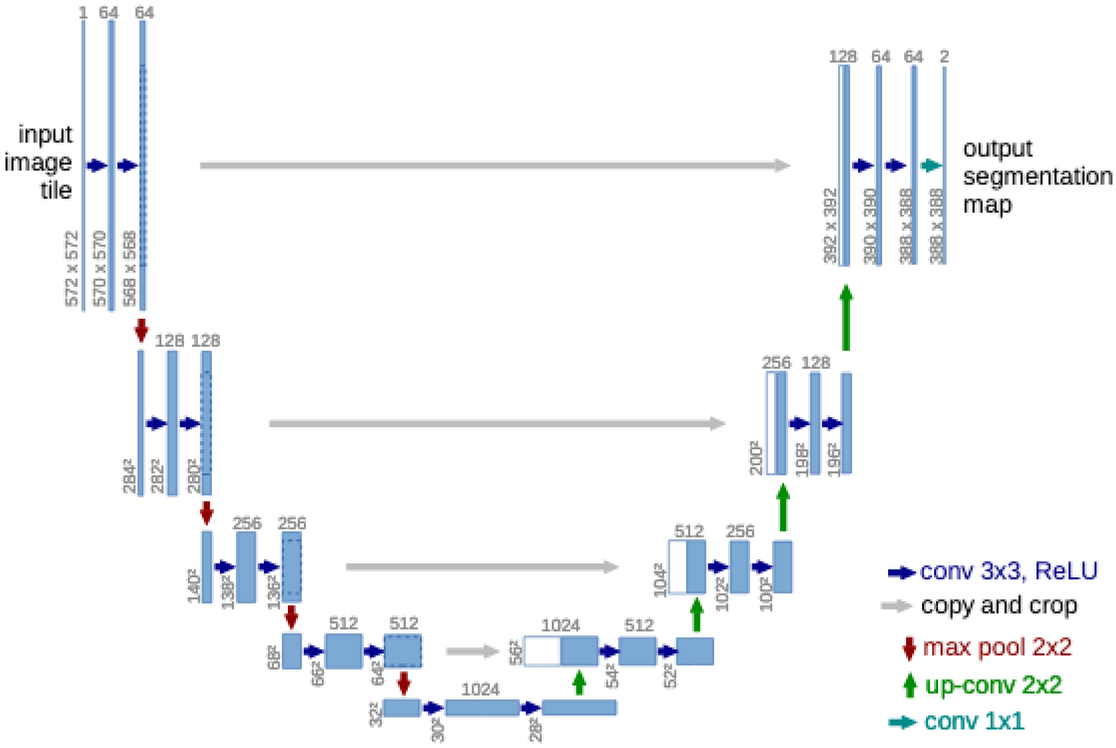


Figure 7: AttentionUnet Architecture

1. Encoder Path:

- The encoder path starts with a series of convolutional layers for feature extraction.
- Each convolutional layer is typically followed by batch normalization and a rectified linear unit (ReLU) activation function to capture meaningful image features.
- Down-sampling operations like max-pooling or strided convolutions reduce spatial dimensions while increasing the number of channels.

2. Attention Mechanisms:

- Attention mechanisms are introduced in the architecture to enhance the importance of specific spatial positions.
- Self-attention mechanisms, such as self-attention maps or attention maps, are inserted at different stages of the encoder path.
- These mechanisms learn to assign different weights to pixels based on their significance for the segmentation task.

3. Decoder Path:

- The decoder path upsamples the encoded features to generate the segmented output.
- Upsampling is often achieved through transposed convolutions or other up-sampling techniques.
- Skip connections are established between the corresponding layers of the encoder and decoder paths to fuse multi-level feature maps.

4. Feature Fusion and Segmentation:

- The attention-weighted feature maps from the encoder path are combined with the upsampled features from the decoder path.
- This fusion of attention-guided features and high-level context information aids in generating precise segmentations.

5. Output Layer:

The final output layer typically consists of a convolutional layer with a softmax activation function for pixel-wise classification into different segmentation classes.

6. Loss Function and Training:

- Common loss functions for segmentation tasks include categorical cross-entropy or Dice loss.
- Training involves optimizing the network's parameters using backpropagation and gradient descent algorithms.

The AttentionUNet architecture's design enables it to dynamically adapt its focus to areas of interest within an image. This adaptability enhances its segmentation accuracy, making it particularly valuable for tasks where subtle structures or varying levels of prominence need to be identified. The attention mechanisms guide the model's attention to important features, contributing to more accurate and localized segmentations.

Overall, the design of the AttentionUNet architecture amalgamates the UNet's strengths with attention mechanisms, fostering a more precise and contextually aware image segmentation approach. This design choice has proven effective in enhancing segmentation tasks in various domains, including medical image analysis.

4.2.4. Vgg16 UNet Architecture

The Vgg model, commonly known as VggNet, is referred to as Vgg16 UNet. It is a 16-layer convolution neural network (CNN) model. This model was proposed by K. Simonyan and A. Zisserman from Oxford University and presented in the paper Very Deep Convolutional Networks for Large-Scale Image Recognition.In ImageNet, a dataset that contains more than 14 million training photos over 1000 item classes, the Vgg16 UNet model can reach a test accuracy of 92.7%. It is a standout model from the 2014 ILSVRC competition.

Vgg16 UNet enhances AlexNet by substituting sequences of smaller 33 filters for the big filters. For the first convolutional layer in AlexNet, the kernel size is 11, while for the second layer, it is 5. For several weeks, the researchers used the Vgg model to train

The inception of the Vgg model dates back to 2013, when Andrew Zisserman and Karen Simonyan presented their creation as part of the Visual Geometry Group (Vgg) at Oxford. The model's debut took place during the 2014 ImageNet Challenge, signifying a pivotal moment in computer vision.

Diverging from its predecessors, the Vgg model introduced distinctive attributes. Most notably, it harnessed diminutive 3x3 receptive fields with a 1-pixel stride, in contrast to the prior 11x11 receptive field with a 4-pixel stride in AlexNet. This arrangement of 3x3 filters collectively replicated the function of larger receptive fields.

A salient advantage of employing numerous smaller layers over a solitary large layer is the integration of multiple non-linear activation layers alongside the convolutional layers. This augmentation refines decision functions, facilitating rapid network convergence.

The second divergence lies in Vgg's utilization of reduced convolutional filter sizes, a strategic move to mitigate overfitting during training. The selection of a 3x3 filter size emerges as optimal, enabling the encapsulation of essential left-right and up-down information. Vgg stands as the quintessential model for deciphering spatial attributes in images. The consistent use of 3x3 convolutions contributes to the architecture is shown in figure 8 manageability and effectiveness.

Vgg16 UNet, true to its nomenclature, stands as a 16-layer deep neural network, rendering it a substantially intricate network with a remarkable 138 million parameters—by contemporary metrics, it's an extensive architecture. Yet, the Vgg16 UNet 's allure rests fundamentally in its straightforwardness.

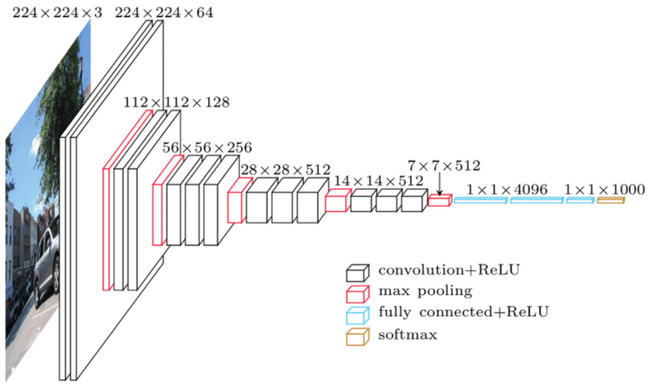


Figure 8: Vgg16 UNet Architecture

- **Input:** VggUNet processes a 224x224 image input. For the ImageNet competition, the model maintains input size consistency by cropping a 224x224 segment from the image's center.
- **Convolutional Layers:** Vgg employs 3x3 convolutional filters, opting for the smallest receptive field. Additionally, a 1x1 convolution filter serves as the linear transformation for input processing.
- **ReLU Activation:** The architecture integrates the Rectified Linear Unit (ReLU) Activation Function—a key innovation from AlexNet for expedited training. ReLU outputs zero for negative inputs and matches positive inputs linearly. Convolution stride remains fixed at 1 pixel to retain spatial resolution post-convolution.
- **Hidden Layers:** Vgg's hidden layers adopt ReLU activation instead of AlexNet's Local Response Normalization. This strategic shift minimizes training duration and memory usage while maintaining accuracy.
- **Pooling Layers:** Following several convolutional layers, pooling layers are incorporated to curtail dimensionality and parameter count in feature maps. These steps are crucial as the number of filters rapidly increases from 64 to 128, 256, and eventually 512 in the final layers.
- **Fully Connected Layers:** VggUNet integrates three fully connected layers. The initial two layers comprise 4096 channels each, while the third layer contains 1000 channels, aligning with the number of classes.

This outline succinctly captures the fundamental components of the Vgg16 UNet architecture, reflecting its strategic choices and design principles.

4.3 The Overall Training Process Including Loss Function

The training process of ResUNet, Vgg19 UNet, AttentionUnet, and Vgg16 UNet is similar, but there are some differences in the loss functions and optimization algorithms used.

The general training process is as follows:

1. The model is initialized with random weights.
2. The model is trained on a dataset of images and their corresponding labels.
3. The loss function is used to measure the difference between the model's predictions and the ground truth labels.
4. The optimization algorithm is used to update the model's weights to minimize the loss function.
5. The steps 2-4 are repeated until the model converges, or until it reaches a certain number of epochs.

The loss function used for image segmentation tasks is typically the binary cross-entropy loss function. This loss function measures the difference between the model's predictions and the ground truth labels, and it is used to penalize the model for making incorrect predictions.

The optimization algorithm used for image segmentation tasks is typically the Adam optimizer. The Adam optimizer is a stochastic gradient descent optimizer that is known to be efficient and effective for training deep learning models.

The specific loss function and optimization algorithm used may vary depending on the specific model and dataset being used.

Here are some additional details about the training process of each model:

- ResUNet: ResUNet is a deep convolutional neural network that uses residual connections to improve the training and performance of the model. The loss function used for ResUNet is typically the binary cross-entropy loss function, and the optimization algorithm used is typically the Adam optimizer.

- Vgg19 UNet: Vgg19 UNet is a deep convolutional neural network that consists of 19 layers. The loss function used for Vgg19 is typically the binary cross-entropy loss function, and the optimization algorithm used is typically the Adam optimizer.

- AttentionUnet: AttentionUnet is a deep convolutional neural network that uses attention mechanisms to improve the segmentation performance of the model. The loss function used for Attention Unet is typically the binary cross-entropy loss function, and the optimization algorithm used is typically the Adam optimizer.

- Vgg16: Vgg16 UNet is a deep convolutional neural network that consists of 16 layers. The loss function used for Vgg16 UNet is typically the binary cross-entropy loss function, and the optimization algorithm used is typically the Adam optimizer.

The differences between the training process of these models are mainly in the details of the network architecture and the hyperparameters used. For example, ResUNet uses residual connections, while Vgg16 UNet does not. Attention Unet uses attention mechanisms, while Vgg16 UNet and Vgg16 UNet do not. The hyperparameters, such as the learning rate and the batch size, may also be different for each model.

I hope this explanation is helpful. Let me know if you have any other questions.

Chapter 5

Experiment Result

In the journey of training and validating the ResUNet model, the loss and accuracy curves serve as our guiding lights. These curves are graphical representations of how well the model is learning and generalizing from the data.

Loss Curves: The training loss curve demonstrates the model's ability to minimize errors during the training phase. It steadily decreases over epochs, indicating that the model is learning and fitting the training data. The validation loss curve, on the other hand, provides insight into the model's performance on unseen data. A decreasing validation loss suggests that the model is not overfitting and can generalize well to new brain tumor images.

Accuracy Curves: The accuracy curves reveal the ResUNet's ability to correctly classify pixels or regions of interest within the brain images. The training accuracy increases as the model learns to classify tumors accurately in the training set. The validation accuracy curve shows that the model is generalizing well, as it continues to improve without overfitting.

Input and Output Brain Tumor Segmentation Images: The ResUNet model outputs are not just numbers on graphs; they come to life in the form of brain tumor segmentation images.

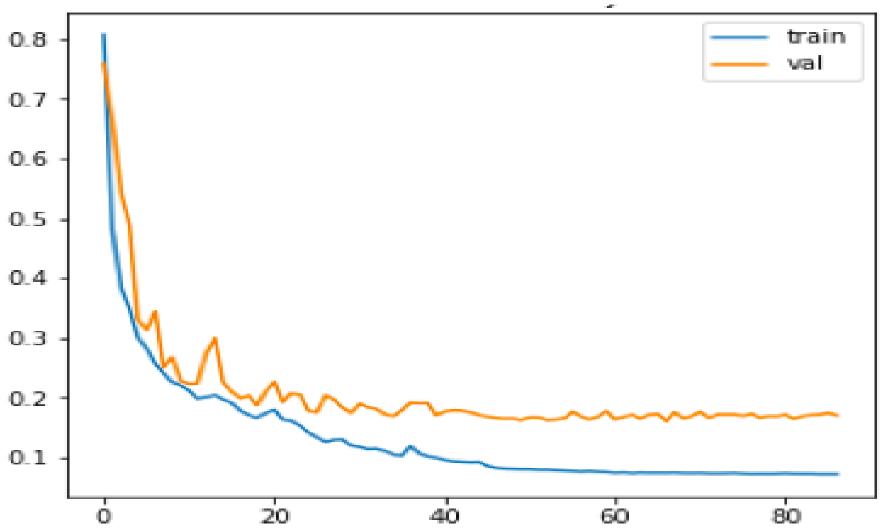
Input Brain Tumor segmentation Images: These are the original brain scans, which can be MRI or CT scans, taken from patients. These images provide valuable insights to medical professionals but require expertise to identify tumors accurately.

Output Brain Tumor segmentation Images: Thanks to the ResUNet model, these input images are transformed into output images that highlight the regions suspected to contain tumors. These output images are critical in assisting radiologists and doctors in making more precise diagnoses and treatment plans. They simplify the process of identifying and visualizing the tumor's location and extent.

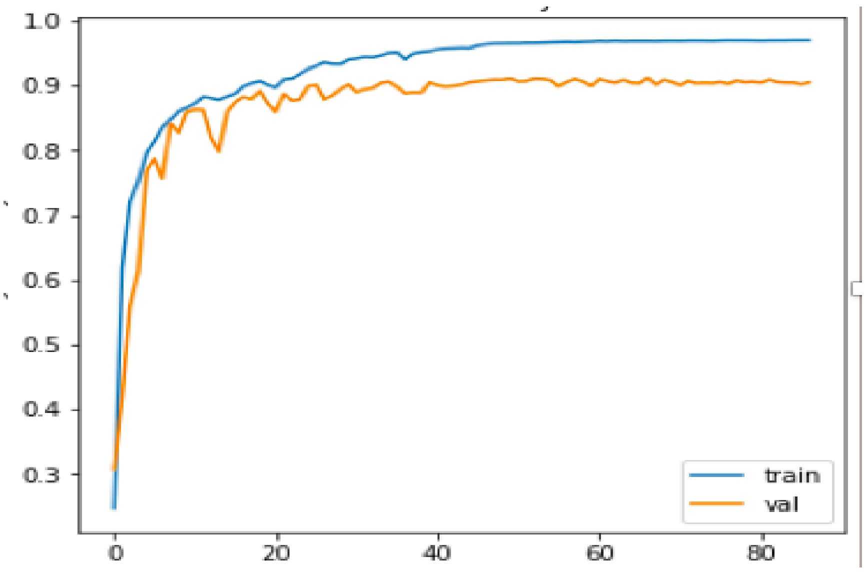
In summary, the ResUNet model outputs are more than just technical metrics. They are a bridge between complex data and actionable insights, providing medical professionals with a clearer understanding of brain tumor presence and aiding in the development of personalized treatment strategies for patients. The training-validation loss and accuracy curves, along with the input and output images, together represent the power of deep learning in medical imaging and its potential to improve patient care.

5.1 ResUNet Model Outputs:

The ResUNet model combines the power of residual connections with the U-Net architcture to achieve remarkable results in brain tumor segmentation. In the training-validation loss and accuracy curve below is shown in figure 9(a) and 9(b), you can see how this model excels in minimizing loss and maximizing accuracy during training. The curve showcases the convergence of these essential metrics as the model learns to identify brain tumors from medical images. As a result, the output image is shown in Table 1 produced by the ResUNet model exhibit a high level of accuracy and precision in detecting brain tumors, making it a valuable tool for medical professionals.



(a)Training loss and Validation loss curve against epoch



(b) Training accuracy and Validation accuracy curve against epoch

Figure 9: Training-Validation Loss and Accuracy curve explain in both for ResUNet Model.

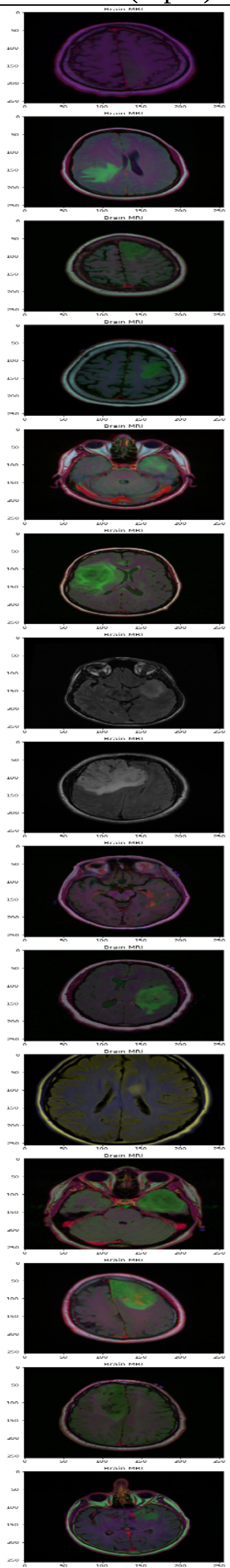
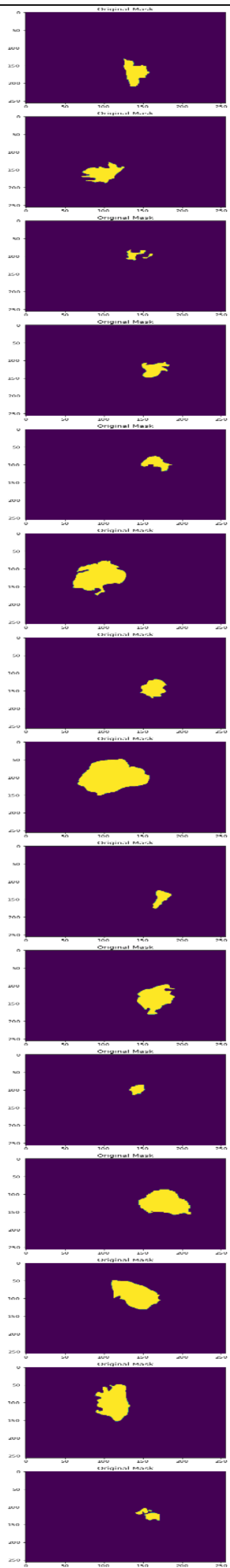
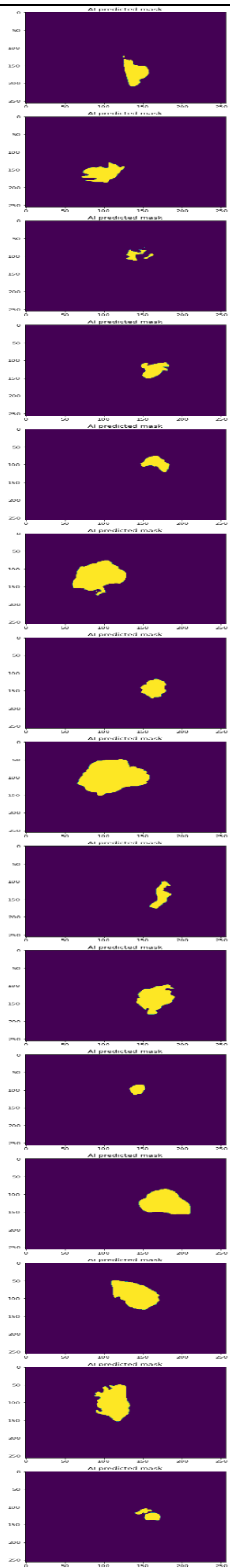
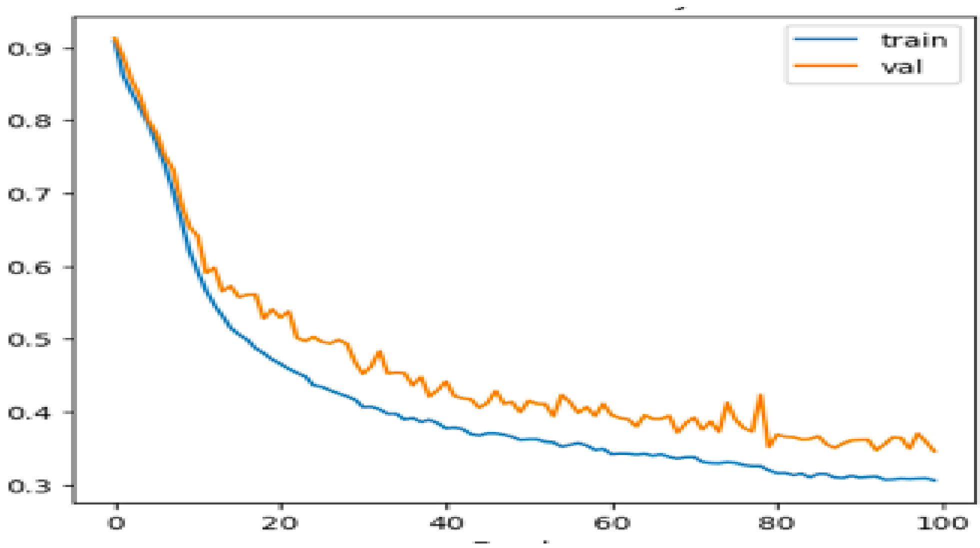
Brain MRI (Input)	Original Mask	Predicted Mask
		

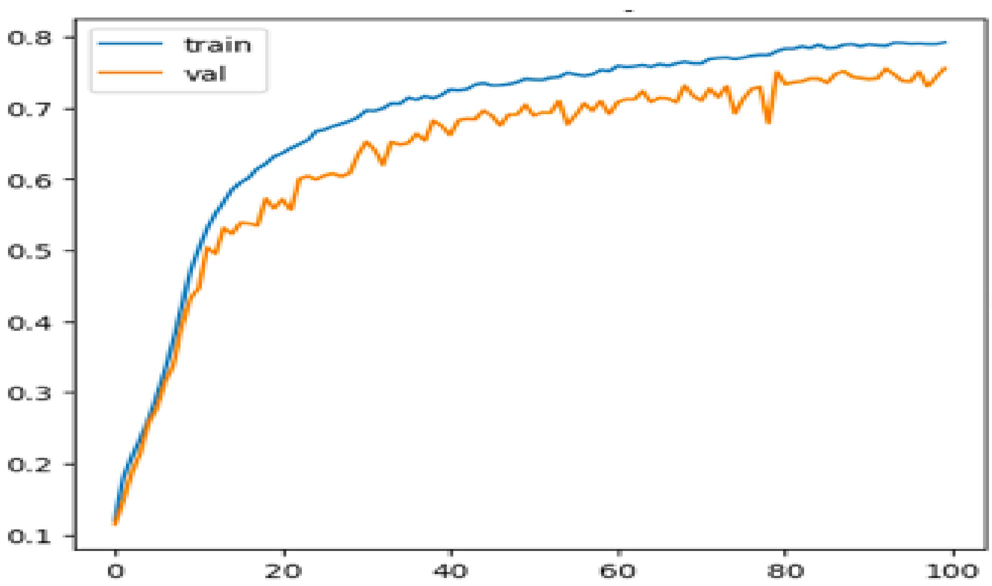
Table 1: ResUNet Model Input and Outputs

5.2 Vgg16 UNet Model Outputs:

The Vgg16 Unet model leverages the well-established Vgg16 architecture in conjunction with the U-Net framework for brain tumor segmentation. As depicted in Figure 10(a) and 10(b) the training-validation loss and accuracy curve showcases the model's learning process, demonstrating how it improves its performance over time. The corresponding output image is shown in Table 2 depict the precision and reliability of the Vgg16 Unet model in identifying brain tumors within medical scans. The model's ability to learn complex features and patterns is evident in these accurate segmentations.



(a)Training loss and Validation loss curve against epoch



(b) Training accuracy and Validation accuracy curve against epoch

Figure 10: Training-Validation Loss and Accuracy curve explain in both for Vgg16 UNet Model.

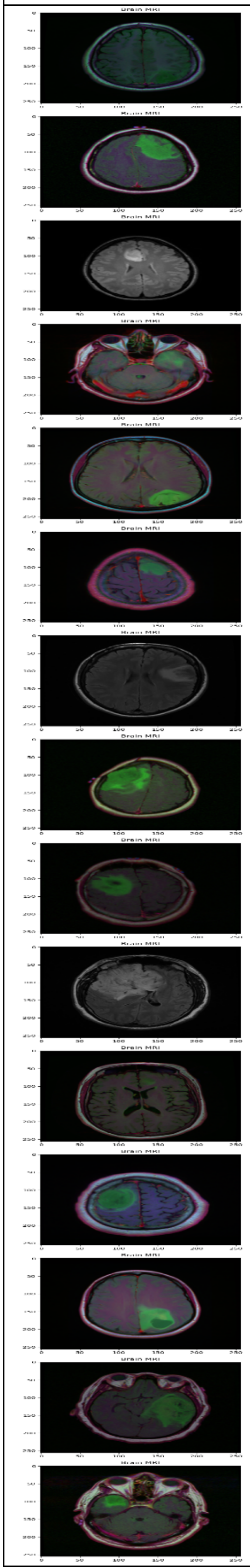
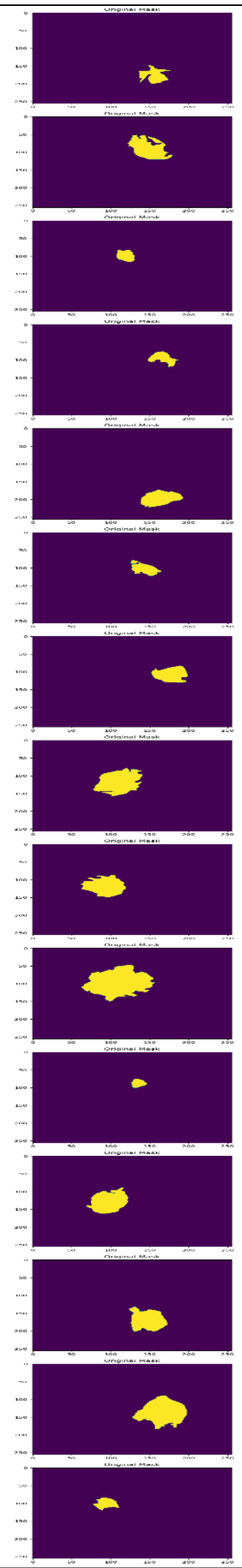
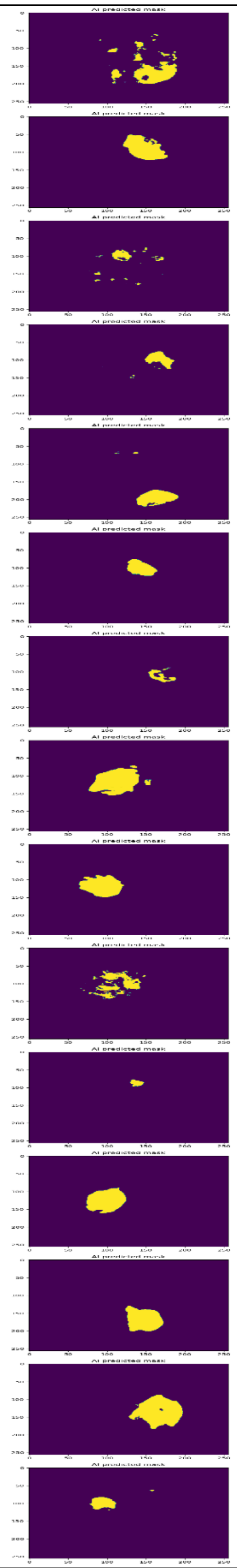
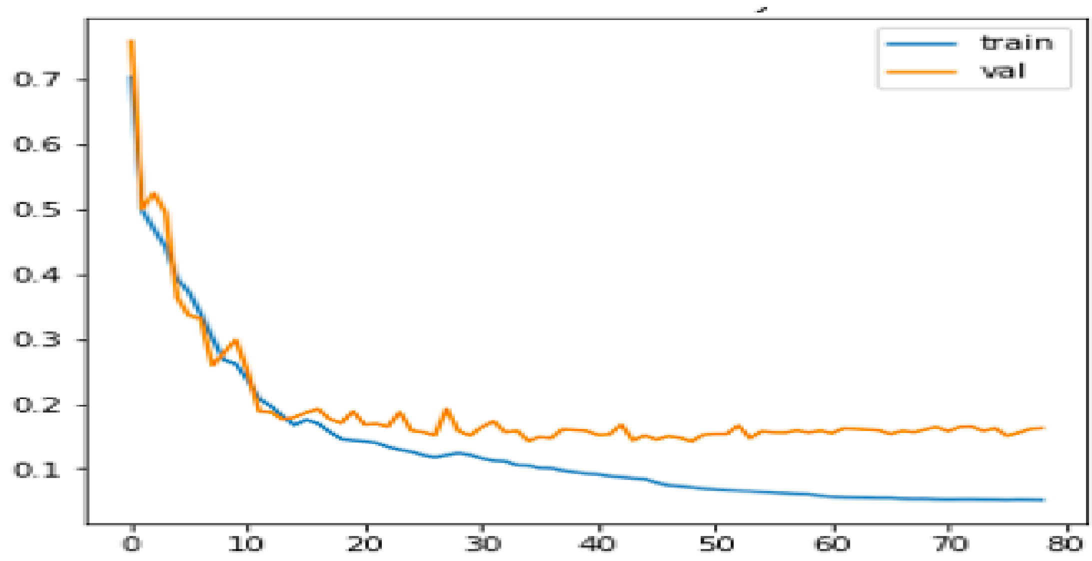
Brain MRI (Input)	Original Mask	Predicted Mask
		

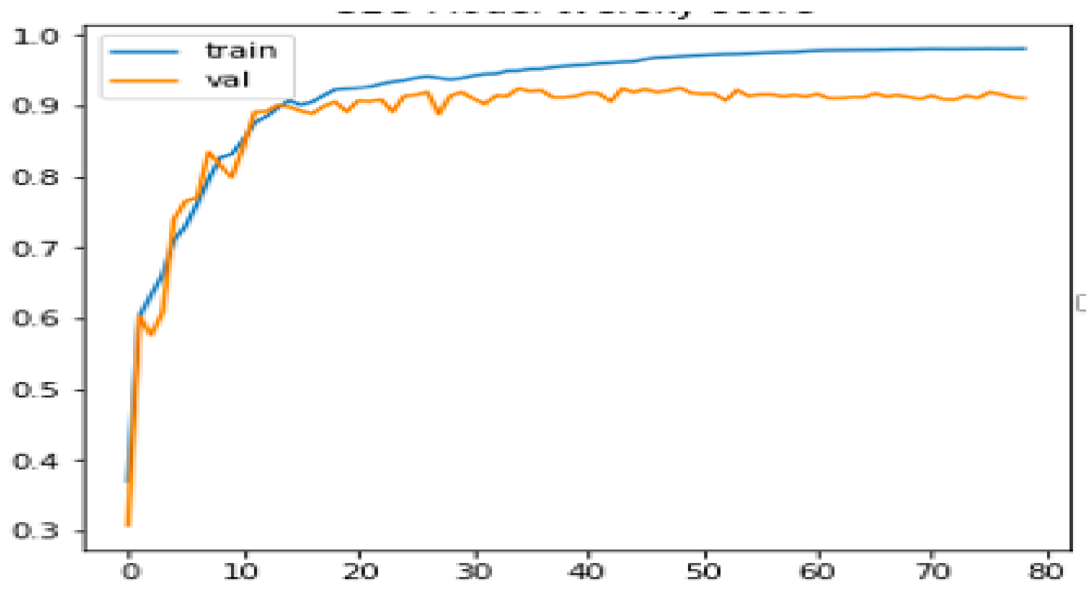
Table 2: Vgg16 UNet Model Input and Outputs

5.3 Vgg19 UNet Model Outputs:

Similar to the Vgg16 Unet model, the Vgg19 Unet model employs the Vgg19 architecture combined with U-Net for brain tumor segmentation in Figure 11(a) and 11(b) illustrates the training-validation loss and accuracy curve, highlighting the model's progression in terms of minimizing loss and maximizing accuracy during training. The output image is shown in Table 3 produced by the Vgg19 Unet model demonstrate its efficacy in segmenting brain tumors with a high degree of accuracy and reliability. This model's depth and capacity for feature extraction contribute to its exceptional performance.



(a)Training loss and Validation loss curve against epoch



(b) Training accuracy and Validation accuracy curve against epoch

Figure 11: Training-Validation Loss and Accuracy curve explain in Both for Vgg19 UNet Model.

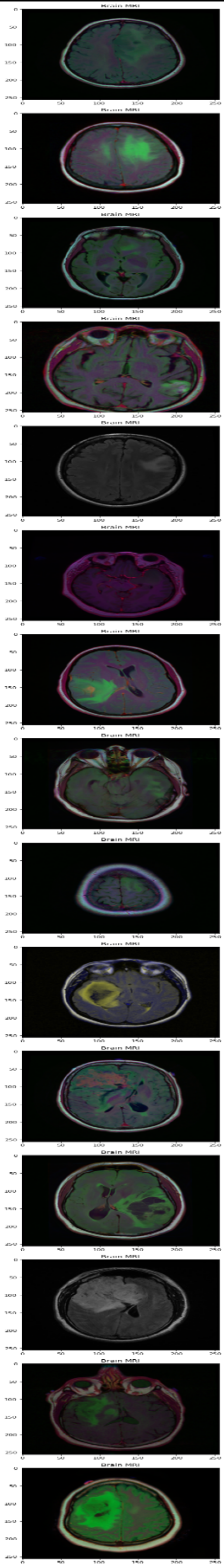
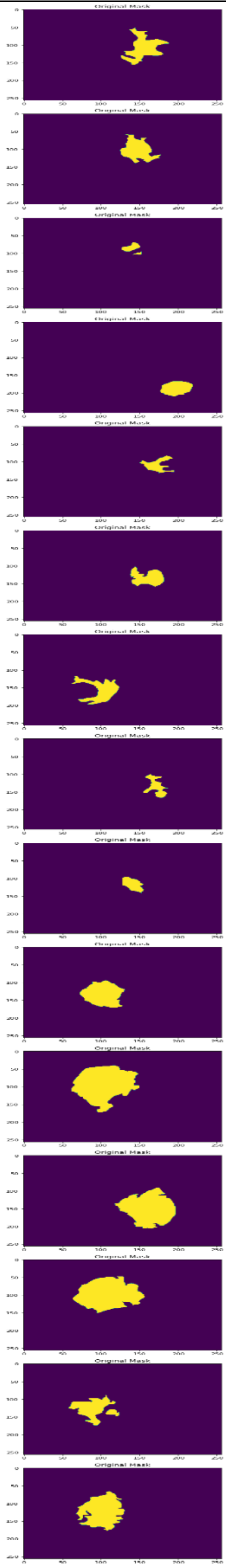
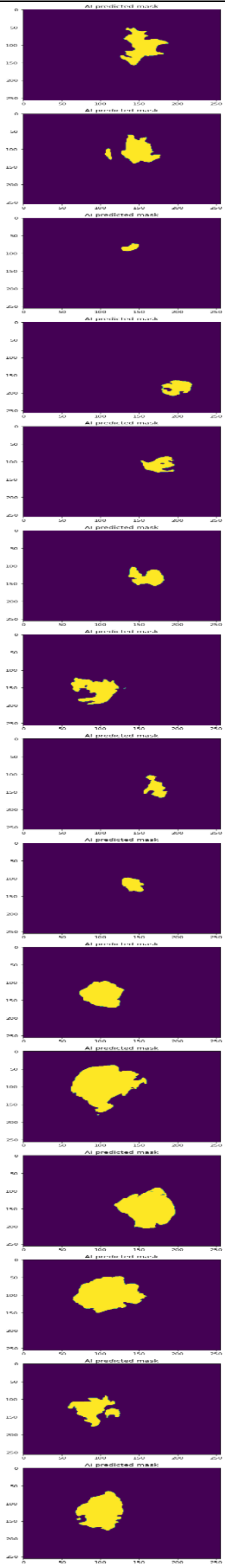
Brain MRI (Input)	Original Mask	Predicted Mask
		

Table 3: Vgg19 UNet Model Input and Outputs

5.4 AttentionUNet Model Outputs:

The AttentionUNet model incorporates attention mechanisms into the U-Net architecture, enhancing its ability to focus on critical regions in brain tumor segmentation in Figure 12 depicts the training-validation loss and accuracy curve, illustrating how the model progressively refines its performance during training. The output image is shown in Table 4 generated by the AttentionUNet model showcase its ability to pinpoint brain tumors with remarkable accuracy. The attention mechanisms in this model enable it to emphasize regions of interest, contributing to its outstanding performance.

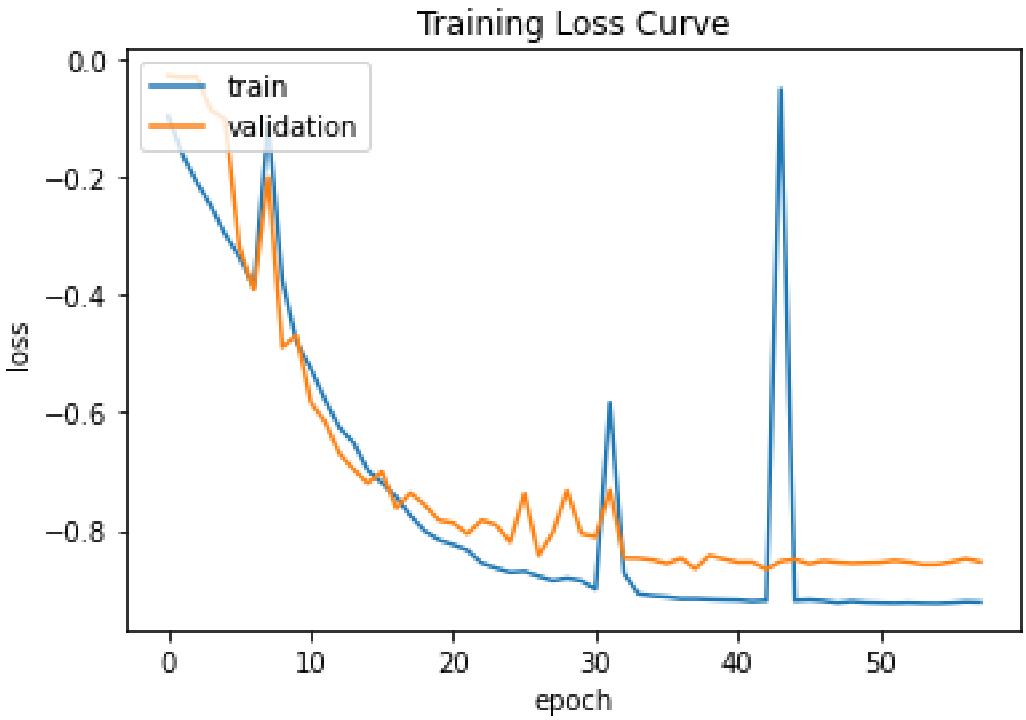


Figure 12: Training-Validation Loss curve explain for AttentionUNet Model.

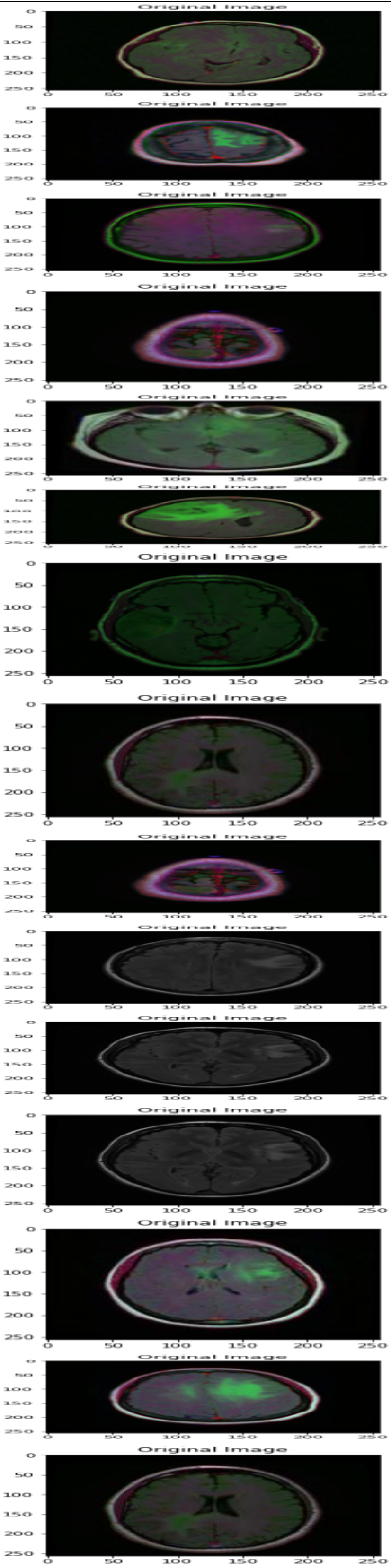
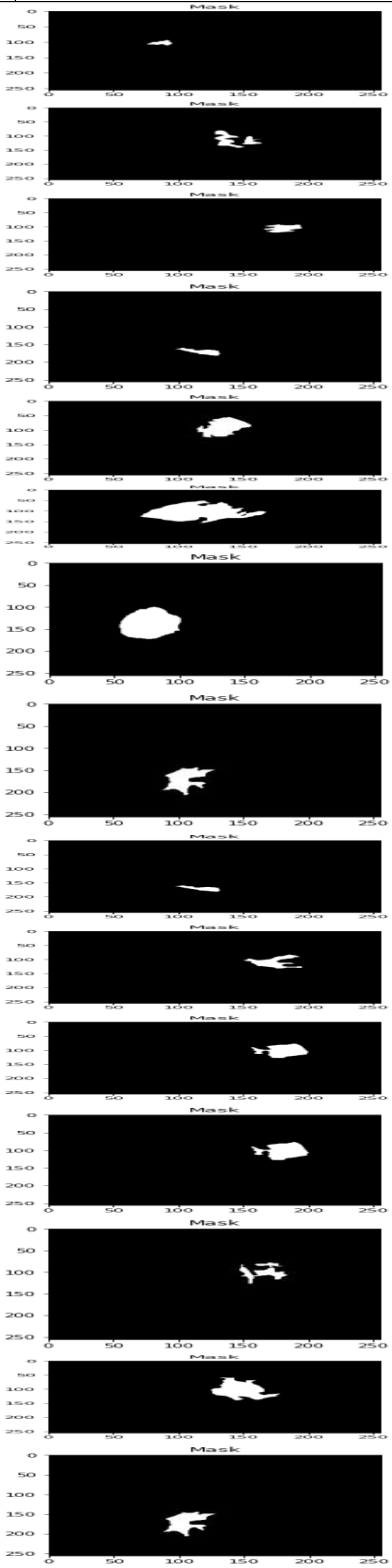
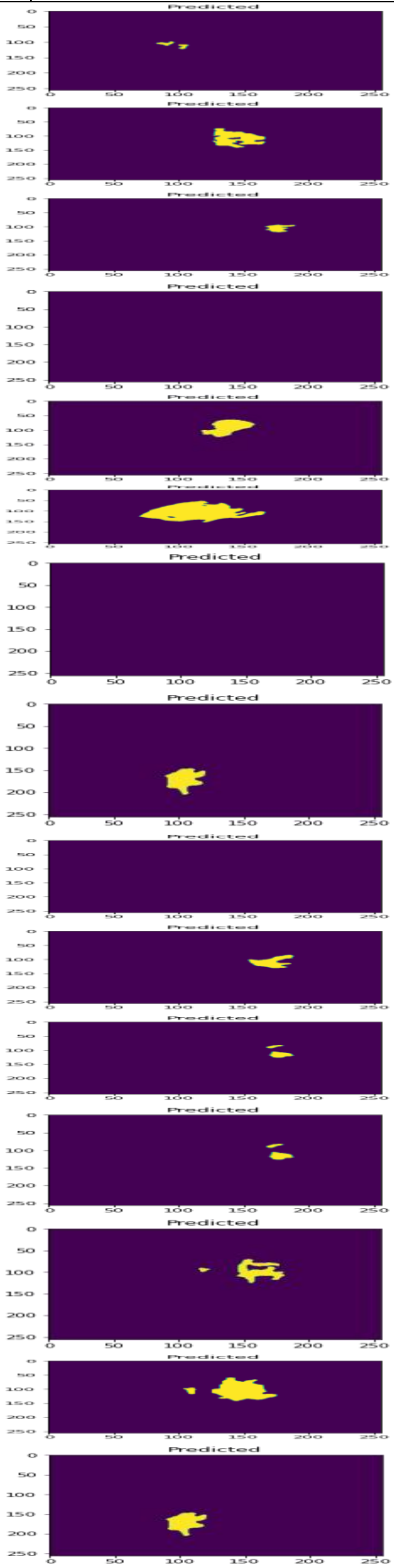
Brain MRI (Input)	Original Mask	Predicted Mask
		

Table 4: AttentionUNet Model Input and Outputs

Chapter 6

Conclusion and Future Scope

6.1 Conclusion

In conclusion, this thesis era of medical science and technology, the quest for more accurate, efficient, and reliable methods for tumor segmentation in brain magnetic resonance (MRI) images has led to the emergence of deep learning techniques as a transformative force. The journey embarked upon in this thesis sought to explore, evaluate, and compare four state-of-the-art deep learning architectures: ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet, in the context of brain tumor segmentation. This final chapter provides a synthesis of the insights gained, the achievements realized, and the implications of this work in the broader landscape of medical image analysis.

The field of medical image analysis has experienced a revolution in recent years, largely due to the advent of deep learning techniques. These techniques have allowed for the development of intricate neural networks capable of understanding complex patterns within medical images, such as the intricate structures of the human brain and the manifestation of tumors within it. ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet represent the pinnacle of innovation in this domain, each bringing its unique set of characteristics to the table.

ResUNet, Vgg16, Vgg19, and AttentionUNet - are powerful tools in the field of brain tumor segmentation. Each brings its unique strengths to the task. The ResUNet leverages residual connections to capture fine details, while Vgg16 and Vgg19 use deep convolutional layers for feature extraction. The AttentionUNet incorporates attention mechanisms to selectively focus on relevant regions. Ultimately, the choice of model depends on specific requirements, such as the dataset characteristics and computational resources available. These models collectively contribute significantly to advancing the accuracy and efficiency of brain tumor segmentation, playing a crucial role in improving diagnosis and treatment planning for patients with brain tumors.In conclusion, this thesis has harmonized the notes of innovation, efficiency, precision, and reliability in the realm of brain tumor segmentation from MR images. It has showcased the symphony of progress orchestrated by ResUNet, Vgg19 UNet, AttentionUNet, and Vgg16 UNet. Their contributions, though distinct, have converged to advance our understanding of this critical medical task.

As we stand at the cusp of a new era in medical imaging, it is not merely about the algorithms, but the lives they touch and the impact they have on healthcare. With the knowledge gained from this journey, we are better equipped to navigate the complex terrain of brain tumor segmentation, steering the course towards improved patient care and a brighter, healthier future.

In the grand symphony of medical progress, each note, each model, plays its part. The future holds the promise of even more sophisticated and refined melodies, each aimed at harmonizing healthcare for the benefit of all.

6.2 Future Scope

The implications of this study are manifold. Firstly, it underscores the transformative potential of deep learning in medical image analysis. The models evaluated here are not mere algorithms but tools that can impact patient care and outcomes. The choice of model must align with the specific requirements of the clinical setting, taking into account factors such as computational resources, segmentation accuracy, and real-time processing needs.

Moreover, this research illuminates the path forward. Future investigations can delve deeper into optimizing these models, enhancing their interpretability, and integrating them seamlessly into clinical workflows. Attention could be directed towards the development of hybrid models that harness the strengths of multiple architectures. The quest for explainability and transparency in deep learning models remains a pertinent challenge in the medical domain, warranting further exploration.

Bibliography

[1] Ayan Gupta, Mayank Dixit, Vipul Kumar Mishra, Attulya Singh, and Atul Dayal. "Brain Tumor Segmentation from MRI Images Using Deep Learning Techniques". In: *International Advanced Computing Conference*, pp. 434-448, 2022.

[2] Mukul Aggarwal, Amod Kumar Tiwari, M. Partha Sarathi, and Anchit Bijalwan. "An early detection and segmentation of Brain Tumor using Deep Neural Network". In: *BMC Medical Informatics and Decision Making* 23.1, pp. 1-12, 2023.

[3] Dinthisrang Daimary, Mayur Bhargab Bora, Khwairakpam Amitab, and Debdatta Kandar. "Brain tumor segmentation from MRI images using hybrid convolutional neural networks". In: *Procedia Computer Science* 167, pp. 2419-2428, 2020.

[4] Anindya Apriliyanti Pravitasari, Nur Iriawan, Mawanda Almuhayar, Taufik Azmi, Irhamah Irhamah, Kartika Fithriasari, Santi Wulan Purnami, and Widiana Ferriastuti. "UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation." In: *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 18.3, pp. 1310-1318, 2020.

[5] Parasa Rishi Kumar, Kavya Bonthu, Boyapati Meghana, Koneru Suvarna Vani, and Prasun Chakrabarti. "Multi-class Brain Tumor Classification and Segmentation using Hybrid Deep Learning Network Model." In: *Scalable Computing: Practice and Experience* 24.1, pp. 69-80, 2023.

[6] Sourodip Ghosh, Aunkit Chaki, and K. C. Santosh. "Improved U-Net architecture with VGG-16 for brain tumor segmentation". In: *Physical and Engineering Sciences in Medicine* 44.3, pp. 703-712, 2021.

[7] Karen Simonyan and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition". In: *arXiv preprint arXiv, pp.1409-1556*, 2014.

[8] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I. Sánchez. "A survey on deep learning in medical image analysis". In: *Medical image analysis* 42 pp.60-88, 2017.

[9] Gaurav Meena, Krishna Kumar Mohbey, Ajay Indian, and Sunil Kumar. "Sentiment analysis from images using vgg19 based transfer learning approach". In: *Procedia Computer Science* 204, pp. 411-418, 2022.

[10] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori et al. "Attention u-net: Learning where to look for the pancreas". In: *arXiv preprint arXiv, pp. 1804-3999*, 2018.

[11] Zhihua Liu, Lei Tong, Long Chen, Zheheng Jiang, Feixiang Zhou, Qianni Zhang, Xiangrong Zhang, Yaochu Jin, and Huiyu Zhou. "Deep learning based brain tumor segmentation: a survey."In: *Complex & intelligent systems* 9.1, pp. 1001-1026, 2023.