Faculty of Media Engineering and Technology
German University in Cairo

Experimentation on Various Activation Functions for Brain Tumor Segmentation Using U-Net

A thesis submitted in partial fulfillment of the requirements for the degree of
Bachelor of Science in Computer Science and Engineering

By

Amr Yasser Elsaid Galal

Supervised by

Dr. Mohamed Karam

January 23, 2025

This is to certify that:

(i) the thesis comprises only my original work toward the Bachelor Degree of Science (B.Sc.) at the German University in Cairo (GUC),

(ii) due acknowledgment has been made in the text to all other material used

<div align="right">

_____

Amr Yasser Elsaid Galal

January 23, 2025

</div>

Acknowledgments

I would like to thank the following for their efforts with me in my study and for tolerating me during this journey...

Abstract

Brain tumor segmentation is a critical challenge in medical imaging, pivotal for effective diagnosis, treatment planning, and monitoring disease progression. With the advent of deep learning, U-Net has emerged as a robust architecture for biomedical image segmentation due to its encoder-decoder structure and skip connections, which enable precise localization and contextual understanding [1]. However, challenges in brain tumor segmentation persist due to the high variability in tumor morphology, imaging artifacts, class imbalance, and the need for precise boundary delineation. This study investigates the influence of various activation functions on the performance of a U-Net-based model for brain tumor segmentation, aiming to enhance segmentation accuracy and computational efficiency. Using the BraTS 2020 dataset [2], which includes multimodal MRI scans (T1, T2, T1ce, and FLAIR) and expert-annotated segmentation masks, this study evaluates eight activation functions: Rectified Linear Unit (ReLU), ReLU6, Leaky Rectified Linear Unit (Leaky ReLU), Parametric Rectified Linear Unit (PReLU), Sigmoid, Tangent Hyperbolic Function (Tanh), Exponential Linear Unit (ELU), and Continuously Differentiable Exponential Linear Unit (CELU). The dataset is preprocessed to standardize orientation, remove irrelevant background regions, and focus on regions of interest, ensuring consistency and relevance of input data. The U-Net model architecture employed is optimized for 3D medical imaging, leveraging skip connections and a Dice Loss function to improve segmentation accuracy. Training and validation were conducted in a controlled environment, holding all other parameters constant to isolate the impact of the activation functions.

Experimental results reveal that ReLU consistently outperforms other activation functions across key metrics, including training loss and Dice similarity coefficients for the whole tumor, enhancing tumor, and tumor core regions. Its ability to prevent vanishing gradients while maintaining computational simplicity makes it particularly effective for segmentation tasks. Leaky ReLU follows closely, offering competitive performance by addressing the dead neuron problem but demonstrating slight variability in certain tumor regions. In contrast, PReLU and ReLU6 show slower convergence and higher loss values, limiting their efficacy for this application. Sigmoid and Tanh, while less effective initially, show potential for improvement with extended training, particularly for smaller or less defined tumor regions. Advanced functions like ELU and CELU demonstrate smoother gradient transitions, but their computational costs offset potential gains in accuracy.

This research underscores the critical role of activation functions in optimizing U-Net models for brain tumor segmentation. By systematically comparing the performance of these functions, the study provides valuable insights into their impact on model learning dynamics, segmentation accuracy, and computational efficiency. The findings highlight the necessity of selecting appropriate activation functions to address the unique challenges of medical image segmentation, contributing to the development of more reliable and efficient diagnostic tools. Future work may explore hybrid architectures, adaptive activation functions, and integration with attention mechanisms to further enhance segmentation performance and clinical applicability.

Contents

List of Abbreviations

MRI	Magnetic Resonance Imaging
CNN	Convolutional Neural Network
CLAHE	Contrast-Limited Adaptive Histogram Equalization
ResNets	Residual networks
ROI	Region of Interest
RES	Residual Extended Skip
WC	Wide Context
ReLU	Rectified Linear Unit
E-Tanh	Exponential Tangent Hyperbolic Function
Tanh	Tangent Hyperbolic Function
PReLU	Parametric Rectified Linear Unit
Leaky ReLU	Leaky Rectified Linear Unit
RAS	Right Anterior Superior
DSC	Dice Similarity Coefficient
ELU	Exponential Linear Unit
CELU	Continuously Differentiable Exponential Linear Unit

List of Figures

List of Tables

Chapter 1

Introduction

Brain tumor segmentation is a pivotal task in the domain of medical imaging, especially with the field of automated diagnostic tools that assist clinicians in planning effective treatment strategies. Accurate segmentation of brain tumors is critical not only for diagnosis but also for treatment planning, surgical interventions, and the evaluation of therapeutic outcomes. The necessity of precise segmentation becomes more apparent in conditions like gliomas, which are among the most aggressive and common types of brain tumors. These tumors exhibit a high degree of heterogeneity, both in morphology and growth patterns, making the task of segmentation inherently complex. Manual segmentation, while considered the gold standard, is labor-intensive, time-consuming, and susceptible to inter-observer variability. With the growing prevalence of brain tumor cases globally, automated methods for segmentation have become indispensable. Automated approaches not only offer consistency and reproducibility but also significantly reduce the workload of medical professionals, allowing them to focus on critical clinical decisions [8].

The use of deep learning techniques has revolutionized medical image analysis, with U-Net emerging as a state-of-the-art architecture for segmentation tasks. Originally proposed by Ronneberger et al [9]. for biomedical image segmentation, U-Net is a fully convolutional neural network that excels in capturing both local and global features through its encoder-decoder structure. The integration of skip connections in U-Net ensures the retention of fine-grained details, a feature crucial for medical images where subtle differences can have significant implications.

Despite the advancements brought about by U-Net, challenges remain, particularly in the domain of brain tumor segmentation. These challenges stem from the high variability in tumor appearance, the presence of imaging artifacts, and the class imbalance in medical datasets. The choice of activation function, a critical component of neural networks, directly influences the model's ability to learn and generalize from data. Activation functions introduce non-linearity

[8], enabling the network to capture complex patterns that linear models fail to address. While ReLU has been the default choice for many deep learning models, its limitations such as the dying neuron problem highlight the need for alternative functions that can enhance performance.

This study delves into the impact of various activation functions on the performance of the U-Net architecture for brain tumor segmentation. By systematically experimenting with functions like ReLU, ReLU6, Leaky ReLU, PReLU, Sigmoid, Tanh, ELU, and CELU, this research aims to identify the optimal configuration that balances segmentation accuracy, computational efficiency, and model stability. The BraTS 2020 dataset[2], known for its multimodal MRI scans and comprehensive annotations, serves as the basis for this analysis, providing a robust framework for evaluating model performance.

1.1   Problem Statement

Brain tumor segmentation remains a challenging task due to the high variability in tumor morphology, imaging artifacts, and class imbalance in datasets. Manual methods are time-consuming and inconsistent, while existing automated models struggle with precision and computational efficiency. This study addresses the need for optimized activation functions within U-Net architectures to improve segmentation accuracy, efficiency, and reliability in clinical applications.

Chapter 2

Literature Review

2.1 Significance of Brain Tumor Segmentation in Medical Diagnostics

Brain tumor segmentation is a vital component of diagnostic radiology, particularly for conditions such as gliomas, which are among the most common and aggressive types of brain tumors. Accurate segmentation plays a central role in identifying tumor location, structure, and spread, which directly impacts diagnosis, treatment planning, and patient outcomes. The segmentation process helps to differentiate between healthy and tumorous tissues. This segmentation also assists clinicians in identifying critical tumor subregions, such as necrotic tissue, enhancing regions, and surrounding edema, which are crucial for diagnosis and treatment decisions.

Manual segmentation of brain tumors performed by radiologists, is a labor-intensive and time-consuming task that is prone to variability between observers. The intricacies of brain structures and the diverse morphology of tumors further complicate the segmentation task. Tumors vary in appearance depending on type, malignancy, and growth patterns. For instance, malignant tumors such as glioblastomas often infiltrate surrounding tissues, making complete removal challenging and increasing the need for precise boundary delineation. In this context automated segmentation methods have become essential to ensure consistency and reproducibility in results.

MRI is the primary imaging modality used for brain tumor diagnosis due to its high-resolution soft tissue contrast, which effectively highlights abnormalities within brain tissue [10]. MRI enables non-invasive imaging, highlighting potential abnormalities without the need for surgical intervention. However, despite its advantages, MRI-based segmentation remains a complex task when performed manually. The rigid structure of the skull and the limited space mean that even minor tumor growth can disrupt essential neurological functions, emphasizing the need for timely and accurate detection through effective segmentation.

Automated segmentation, particularly using deep learning models like U-Net, offers a robust solution to these challenges. By learning complex spatial and contextual features, U-Net architectures can handle the diverse nature of tumors [10]. Automated methods not only streamline the segmentation process but also reduce dependency on highly skilled radiologists, who are often required to delineate tumor boundaries manually. This is particularly advantageous in regions with limited access to specialized healthcare providers. Furthermore, automated segmentation minimizes inter-observer variability, thus offering more consistent and reliable diagnostic outcomes [11].

The high incidence and mortality of brain tumors highlight the importance of early detection. With over 300,000 cases reported in 2020 on a worldwide basis [12], an early and precise segmentation in MRI images is key to improving their treatment outcomes. Assisting appropriate treatment planning that includes surgery, chemotherapy, and radiation therapy. Additionally, in follow-up care, accurate segmentation allows for consistent monitoring of tumor growth or response to therapy, providing clinicians with actionable insights for adjusting treatment regimens [13].

2.2 A Brief Introduction to the U-Net Model and Its Architecture

The U-Net model, introduced by Ronneberger et al. [9], has become a cornerstone in medical image segmentation, particularly for tasks involving complex and detailed structures like brain tumors. Originally developed for biomedical applications, U-Net is a fully Convolutional Neural Network (CNN) that employs an encoder-decoder structure with skip connections. This design allows the model to capture both local and contextual information at multiple scales, making it particularly effective for segmenting intricate structures in MRI data, where precise boundary delineation is critical.

In U-Net's architecture, the encoder path, or contracting path, progressively downsamples the input image [1], capturing contextual information while reducing spatial dimensions. This path is essential for learning high-level, abstract features within the image. Meanwhile, the decoder path, or expansive path, upsamples the encoded data, gradually reconstructing spatial resolution to produce a pixel-level segmentation map. Each level of the encoder corresponds to a similar level in the decoder, connected by skip connections, which transfer feature maps from the encoder to the decoder. These skip connections are fundamental in maintaining high-resolution features, as they combine coarse information from the encoder with finer details in the decoder, achieving accurate segmentation even in complex medical images. Tasks that require precise pixel-wise classification, such as brain tumor segmentation, benefit significantly from U-Net's architectural design. By retaining detailed information throughout the encoding and decoding process, U-Net ensures that each pixel in an image is assigned a label, which is

critical for segmenting heterogeneous and irregular structures like brain tumors [1].

Given its success in 2D segmentation, U-Net has been adapted to 3D versions, such as 3D U-Net, for volumetric data [1]. These adaptations apply the model's principles to 3D data, enabling segmentation across MRI volumes rather than individual 2D slices. This is particularly valuable for brain tumor segmentation, where understanding tumor depth and spread across multiple MRI slices can be crucial for diagnosis and treatment planning.

Many studies [10][1] highlight U-Net's encoder-decoder structure and skip connections as foundational elements for achieving precise, pixel-level segmentation. The model's effectiveness in MRI segmentation is enhanced by these architectural features, which allow it to capture both low-level spatial details and high-level contextual information. By addressing the need for both fine details and broad spatial context, U-Net provides an optimal balance for tasks requiring high spatial accuracy. Consequently, it has become a standard for segmenting MRI images in medical applications, particularly in identifying detailed tumor regions within brain scans.

2.3 U-Net Adaptations for Brain Tumor Segmentation

The U-Net architecture, originally developed for biomedical image segmentation, is particularly well-suited for brain tumor segmentation due to its encoder-decoder structure and skip connections [1]. This structure enables U-Net to capture spatial hierarchies within MRI images, preserving both high-level contextual information and fine-grained details essential for identifying intricate tumor boundaries. Standard U-Net has, however, shown limitations in handling complex tumor structures in MRI data [13], leading to numerous adaptations aimed at enhancing segmentation accuracy and efficiency in clinical settings.

One adaptation of U-Net is the application of 3D convolutions [1], which extend U-Net's capability from 2D to volumetric data, making it more effective for MRI scans. Models like 3D U-Net employ 3D operations that better capture the continuity across MRI slices, which is essential for volumetric segmentation in brain imaging. This modification allows U-Net to retain detailed spatial information across all three planes, ensuring consistent segmentation throughout the tumor's volume.

Other adaptations focus on enhancing U-Net's feature extraction capabilities by incorporating attention mechanisms and residual connections. For example, the AResU-Net model [14] combines attention modules and residual blocks, allowing the model to prioritize significant spatial features while maintaining feature continuity across layers. Attention mechanisms help U-Net selectively focus on tumor-relevant areas, which is particularly valuable in complex brain MRIs where tumors vary in shape, size, and contrast. Residual connections further support the network by alleviating the vanishing gradient problem, making it easier for deeper layers to retain essential information.

Attention-based U-Net architectures, such as the Attention 3D U-Net [15], extend this concept by adding multiple skip connections and integrating efficient computation modules, such as MobileNetV2 blocks, to reduce computational costs. This approach enables the model to capture hierarchical features across different spatial scales while maintaining computational efficiency, a critical requirement in real-time clinical applications. By incorporating MobileNetV2, the model becomes lightweight, making it suitable for high-throughput settings where quick, reliable segmentation is essential.

Advanced U-Net variants also utilize feature fusion techniques, which combine U-Net with other CNNs to enhance segmentation capabilities. For instance, FE-HU-NET models [16] employ feature enhancement techniques, such as Contrast-Limited Adaptive Histogram Equalization (CLAHE), to improve image contrast before segmentation. This preprocessing step helps increase visibility in complex brain tumor regions, leading to more accurate segmentation results. Similarly, hybrid approaches like HTTU-Net [17] use a dual-track architecture with different kernel sizes to capture both fine and large-scale tumor features, enhancing the model's ability to generalize across various tumor types.

Residual networks (ResNets) have been integrated into U-Net to further improve feature retention and alleviate vanishing gradient issues, particularly in deeper models. ResU-Net [1], for example, employs residual blocks within the encoder to allow uninterrupted gradient flow, enhancing the model's ability to learn complex features. This architecture performs well on segmentation tasks involving heterogeneous tumor structures, as seen in comparative studies on the BraTS dataset. Additionally, other U-Net variants like U-Net++ [10] have introduced nested skip pathways, which retain multi-scale features, further improving segmentation accuracy and boundary delineation.

Handling missing data modalities has also been a focus of U-Net adaptations, as incomplete MRI sequences can limit model performance. Models like U-HeMIS and HVED attempted to address this by employing shared latent spaces [18], although they encountered limitations in modeling modality-specific features. Building on these efforts, SMU-Net [18] introduced a style and content matching approach, which allowed the model to reconstruct missing modalities by learning from available full-modality data. This approach has been shown to improve segmentation robustness in clinical settings where complete MRI modalities may not always be available.

Region of Interest (ROI)-aided U-Net models have introduced another layer of refinement by focusing the segmentation task on a smaller, more relevant region [19], effectively excluding irrelevant background information. These models typically use a 2D U-Net for localizing the tumor and a 3D U-Net for detailed segmentation within this specified ROI. Comparative evaluations on benchmark datasets indicate that this approach significantly enhances segmentation accuracy, as measured by metrics like Dice similarity and Hausdorff distance, by concentrating the model's resources on relevant areas, thereby improving tumor boundary detection.

To further optimize U-Net for diverse MRI datasets, some studies have introduced multi-encoder frameworks that utilize pre-trained models, such as VGG-19, ResNet50, and MobileNetV2, within the U-Net architecture [12]. These multi-encoder designs capture complementary features, which are then processed by attention-based decoders, refining segmentation accuracy and adapting the model for different tumor morphologies and imaging conditions. For example, a U-Net model with a ResNet50 encoder benefits from enhanced feature extraction due to ResNet's depth and residual blocks, which improve segmentation accuracy, especially for capturing subtle tumor boundaries.

Another approach involves the BU-Net model [20], which introducesResidual Extended Skip (RES) connections and Wide Context (WC) blocks, drawing from deep lab structures to enhance multi-scale feature capture. TheRES blocks expand the receptive field, allowing the model to aggregate spatial information from a wider context, while WC blocks gather multi-scale features. This structure is beneficial for segmenting tumors with diverse morphologies and irregular boundaries, as it allows the model to retain both global and local information, thereby improving segmentation quality across different tumor types and sizes.

Other innovative adaptations leverage additional processing layers to optimize performance for specific MRI modalities or scenarios. For example, SMU-Net [18] introduces a dual-path architecture that handles both full and missing MRI modalities. This design incorporates style and content matching to reconstruct missing information by learning from complete modality data, ensuring consistent segmentation even when imaging conditions are suboptimal or incomplete.

Lastly, advancements such as the introduction of specialized activation functions, including Exponential Tangent Hyperbolic Function (E-Tanh) [8], have contributed to improved boundary detection in U-Net-based segmentation. E-Tanh, by preserving very negative values, enhances the model's sensitivity to boundary nuances, which traditional functions like ReLU may overlook.

The application of U-Net adaptations in brain tumor segmentation continues to grow, with each variation tailored to address specific segmentation challenges in MRI data. Whether through 3D convolutions, attention mechanisms, multi-encoder designs, specialized activation functions, or advanced preprocessing techniques, these adaptations aim to enhance U-Net's accuracy and adaptability in handling complex and diverse tumor presentations. Consequently, U-Net remains a highly flexible and effective model for brain tumor segmentation, supporting the development of robust, automated diagnostic tools for clinical use.

2.4 Overview of Some Activation Functions and Their Properties

Activation functions play a crucial role in neural networks by introducing non-linearity, which enables models to learn complex patterns and representations within data [8]. Without activation functions, neural networks would function as linear regressions, incapable of handling the intricate patterns required for tasks like medical image segmentation. The choice of activation function directly affects key performance metrics, including segmentation accuracy, model stability, and convergence speed.

One of the most widely used activation functions in deep learning is the ReLU, which is particularly common in CNNs and U-Net architectures. ReLU operates by zeroing out negative values allows models to ignore less relevant features while focusing computational resources on positive values [4]. This characteristic is advantageous in segmentation tasks, as it reduces computational overhead and supports stable learning. Which is an ideal choice for deep networks that require fast convergence. ReLU's computational efficiency helps prevent the vanishing gradient problem, a phenomenon in which gradients diminish as they propagate through layers, making it challenging for the model to learn in deeper architectures. However, ReLU's tendency to completely discard negative values can lead to the loss of subtle information in high-resolution data like MRI images, where small gradients might indicate critical details at tumor boundaries [8].

To address some of the limitations of ReLU, several variants have been developed, including Leaky ReLU and PReLU. Leaky ReLU [8] introduces a small slope for negative values, allowing gradients to flow even when inputs are negative. This feature mitigates the problem of 'dead neurons' (units that no longer activate and thus stop learning) that can occur with standard ReLU. By maintaining a small gradient for negative values, Leaky ReLU helps retain more information during training, which can be particularly beneficial in medical imaging tasks where negative values might represent essential contrasts in MRI data. PReLU, an extension of Leaky ReLU, further refines this approach by making the slope of the negative part learnable [14], allowing the model to adaptively tune itself to the data. Studies incorporating PReLU in models like AResU-Net demonstrate improved convergence rates and enhanced detail retention in feature maps, particularly in complex tumor boundaries where subtle gradients play a significant role.

Sigmoid and Tanh, are other common activation functions, though less frequently used in modern deep networks due to their limitations [8]. Sigmoid compresses the input values to a range between 0 and 1, making it useful for binary classification tasks. However, it is prone to saturating gradients, which can slow down learning in deep layers. Tanh, on the other hand, scales values between -1 and 1, effectively centering the data and making it more suitable for inputs with both positive and negative values. While Tanh, is less, some U-Net adaptations

employ it for specific tasks where symmetrical data representation is advantageous.

Recent research has introduced more advanced activation functions, such as Swish and Mish, which aim to combine the simplicity of ReLU with more effective gradient propagation [8]. It retains a small gradient even for negative values, providing a smoother transition than ReLU. This smooth gradient flow can enhance model stability and learning, making Swish a promising choice for deep networks like U-Net, where capturing subtle boundary details is essential. Mish, another recent function, operates similarly to Swish and has shown promise in tasks requiring fine-grained detail retention, such as medical image segmentation. Both Swish and Mish have been found to enhance model accuracy in applications where precise boundary detection is critical, although they come with higher computational costs.

In addition, the E-Tanh function has been proposed for U-Net adaptations [8]. Unlike standard Tanh, which scales values between -1 and 1, E-Tanh retains a small slope for very negative values, allowing the model to capture subtle features that would otherwise be discarded. This is useful in MRI segmentation, where negative values can hold diagnostic relevance. By preserving more information at the boundary regions, E-Tanh enhances the accuracy of segmentation in complex medical images.

Despite the strong performance of ReLU in U-Net architectures, studies suggest that alternative activation functions could enhance U-Net's segmentation capabilities, especially in distinguishing boundaries with subtle contrasts [8]. The choice of activation function affects not only segmentation precision, but also computational efficiency. For example, PReLU's adaptive learning of the negative slope offers a balance between detail retention and computational cost, making it advantageous for applications that require both accuracy and efficiency [14]. Exploring alternatives such as Swish, Mish and E-Tanh could further improve boundary delineation, especially for brain tumors, where accurate segmentation impacts clinical outcomes.

2.5 Studies Focusing on Activation Functions within U-Net for Segmentation Tasks

Activation functions play a fundamental role in neural network architectures such as U-Net, impacting both learning efficiency and segmentation accuracy. Despite the pivotal nature of these functions, studies specifically focused on the effects of different activation functions within U-Net for medical image segmentation, particularly brain tumor segmentation, remain limited. The majority of studies employ standard activation functions like ReLU due to its computational simplicity and effectiveness in mitigating the vanishing gradient problem [8]. However, recent research suggests that exploring alternative activation functions within U-Net models could potentially enhance segmentation performance, especially when high boundary precision and sensitivity to subtle image features is needed.

Many U-Net-based segmentation studies rely on ReLU as a default activation function. For example, studies on U-Net adaptations like BU-Net [20] and HTTU-Net [17] employ ReLU extensively across convolutional layers, emphasizing its effectiveness in maintaining training stability and enhancing convergence rates. However, while ReLU has proven effective, it is argued that it may not be the optimal choice for segmentation tasks that involve capturing fine-grained details [8], particularly at tumor boundaries. The primary limitation with ReLU lies in its zeroing of negative values, which can lead to the loss of subtle boundary information essential for precise segmentation.

To address this gap, some U-Net adaptations have experimented with alternative functions. For instance, the AResU-Net model [14] employs PReLU in its architecture, leveraging PReLU's learnable parameter for the negative slope to capture a broader range of information. Studies using AResU-Net found that PReLU not only improved convergence rates but also enhanced the model's ability to retain finer details in boundary regions. This adaptation illustrates that incorporating alternative activation functions can positively impact model performance, especially in complex segmentation tasks where tumor boundaries are irregular and require careful delineation.

Other studies have investigated ReLU6 and Leaky ReLU as activation function choices within U-Net architectures. ReLU6, a capped version of ReLU that limits output to a maximum value of six, has shown promise in balancing computational efficiency with segmentation accuracy. For example, in Attention 3D U-Net with MobileNetV2 blocks, ReLU6 is applied to enhance computational efficiency without sacrificing detail, achieving a practical balance that allows for effective boundary detection in high-dimensional MRI data [15]. Similarly, Leaky ReLU, with its small slope for negative values, has been used in models like HTTU-Net, where the ability to retain subtle negative information has proven beneficial for segmenting brain tumors with intricate and ambiguous boundaries [17]. Both ReLU6 and Leaky ReLU have demonstrated effectiveness in segmentation, although direct comparative studies remain scarce.

In experiments on brain MRI segmentation, recent studies have introduced activation functions tailored for brain MRI segmentation. One such function, E-Tanh [8], has been incorporated into U-Net variants to overcome the limitations of ReLU in capturing subtle boundary details. E-Tanh-based models achieved higher Dice scores, particularly in capturing core and enhancing tumor regions, compared to models using ReLU. This improvement highlights the potential of specialized activation functions in boosting segmentation performance for medical imaging tasks that require high precision.

Despite these advancements, comprehensive studies comparing the effects of different activation functions within U-Net for brain tumor segmentation remain limited. The SMU-Net [18] and FE-HU-NET [16] models, for example, primarily emphasize architectural innovations like multi-path and hybrid structures, with less focus on experimenting with activation functions.

However, researchers acknowledge that further exploration into adaptive activation functions like Swish and Mish could provide additional benefits for segmentation accuracy. Swish and Mish offer smoother gradient flow and could enhance the network's ability to capture subtle variations across boundaries, which is especially relevant for complex segmentation tasks like brain tumors where feature differentiation is crucial [8].

Another promising area of study involves attention-based U-Net architectures. Some studies have integrated attention mechanisms with alternative activation functions to optimize feature retention and boundary precision. For instance, Attention U-Net models combined with PReLU have demonstrated improved segmentation outcomes by selectively focusing on tumor-relevant features [15]. This combination allows the model to enhance critical boundaries while reducing irrelevant background noise, thereby increasing segmentation accuracy in intricate areas like the tumor core and enhancing regions.

Models like RMU-Net [21], which experiment with adaptive activation functions, indicate that exploring alternatives to ReLU could yield significant performance improvements. Researchers suggest that future work might benefit from systematically comparing activation functions like Swish, PReLU, and E-Tanh across U-Net architectures, examining their impacts on segmentation metrics such as Dice scores, boundary precision, and convergence stability [8].

These preliminary findings suggest that a more targeted investigation into activation functions within U-Net could unlock further advancements in segmentation quality, particularly for tasks as demanding as brain tumor delineation in MRI data

2.6 Challenges in Achieving Precise Segmentation for Brain Tumors

Brain tumor segmentation presents numerous challenges due to the complex nature of brain anatomy, the variability in tumor characteristics, and the high volume of data associated with MRI imaging. Unlike other types of medical imaging, brain MRIs must capture and differentiate intricate structures, often with minimal contrast between tumor tissue and surrounding healthy brain regions. Tumors vary widely in size, shape, and location, and their boundaries are often irregular, making it difficult to accurately segment these regions with standard methods. Additionally, manual segmentation by clinicians is both time-intensive and prone to variability, highlighting the necessity for reliable automated segmentation tools that can consistently handle the nuances of brain MRI data [8].

Another primary difficulty in brain tumor segmentation arises from MRI artifacts and the diverse imaging conditions across different institutions [1]. Variability in MRI protocols, machine types, and patient positioning can result in inconsistencies in image quality and intensity, complicating the training process for automated segmentation models. To combat these issues,

many U-Net adaptations apply extensive data augmentation to account for variations in rotation, scale, and intensity across images, thus making models more robust. However, while data augmentation can help, it cannot fully resolve the inherent variations introduced by different MRI protocols, which remain a challenge for consistent segmentation accuracy across diverse datasets.

Class imbalance is another key challenge in brain tumor segmentation. Tumor pixels in MRI scans are significantly outnumbered by healthy brain tissue pixels, which can lead models to favor predictions for the dominant class (healthy tissue) and overlook smaller tumor regions. This imbalance reduces segmentation accuracy, particularly for small tumor areas that are clinically significant but represent a small fraction of the overall image. Addressing this, models like AResU-Net [14] employ custom loss functions, such as weighted cross-entropy combined with Dice loss, to emphasize tumor regions and improve boundary delineation, even in cases of highly imbalanced classes.

Segmentation accuracy is further hindered by the high computational demands of processing 3D MRI data, which can consist of hundreds of slices per patient scan. Processing entire MRI volumes is challenging for standard U-Net models, which were originally designed for 2D image segmentation. To address this, some models, such as the 3D U-Net, use three-dimensional convolutions to capture volumetric data, making it easier to recognize and segment tumors in all three spatial dimensions. However, these 3D models require substantial computational resources, which can limit their applicability in clinical settings where quick processing times are essential. Other approaches attempt to optimize efficiency by using techniques like ROI filtering to focus computational power on tumor-specific regions, thus improving segmentation accuracy without significantly increasing processing time.

Finally, achieving high segmentation accuracy is challenging due to the difficulty of precisely delineating tumor boundaries [8], particularly in cases where the tumor infiltrates normal brain tissue. Low contrast between the tumor and adjacent brain structures often results in blurred boundaries, making it hard to differentiate the two regions. Manual segmentation by radiologists is inherently subjective, leading to inter-observer variability, while automated models can struggle to consistently recognize these subtle contrasts. This issue has prompted some studies to explore alternative activation functions within U-Net, such as E-Tanh, which helps retain negative pixel values and finer detail at boundary regions, thus enabling the model to capture subtle contrasts that standard functions like ReLU may overlook.

In summary, precise segmentation of brain tumors in MRI images remains challenging due to the high variability in tumor morphology, imaging inconsistencies, computational demands, and class imbalances. Each of these factors complicates the segmentation task, making it difficult for models to achieve high accuracy consistently. Despite the progress made through U-Net adaptations, including 3D convolutions, attention mechanisms, and specialized loss functions, further innovations are required to fully address these challenges. Future developments may

benefit from exploring tailored activation functions, optimized architectures, and adaptive data preprocessing techniques to improve segmentation accuracy in real-world clinical applications.

2.7 Research Gaps and How Different Activation Functions Might Address Segmentation Challenges

While U-Net has become a cornerstone in brain tumor segmentation, research has primarily focused on architectural modifications, such as attention mechanisms and multi-scale feature extraction, rather than optimizing activation functions specifically for medical imaging. Most U-Net models rely on standard activation functions like ReLU, which are popular for their simplicity and computational efficiency. However, ReLU's tendency to zero out negative values could limit its effectiveness in applications where subtle contrasts, such as the edges of a tumor, are crucial. This reliance on ReLU reveals a notable research gap in exploring alternative activation functions that might better handle the intricate requirements of medical image segmentation.

while there have been promising results with newer activation functions in other domains, such as Swish and Mish, which smooth gradients and prevent abrupt zeroing out, their application within U-Net for medical image segmentation remains largely unexplored. These functions have demonstrated potential in enhancing gradient flow and retaining finer details, attributes that are critical in medical imaging tasks like brain tumor segmentation. Systematic studies that examine Swish, Mish, and other novel functions in U-Net-based models could reveal configurations that not only preserve complex boundary details but also improve model stability across diverse datasets.

Another underexplored area is the impact of adaptive activation functions, such as PReLU and Swish, within 3D U-Net architectures. 3D U-Net models, often used to capture volumetric features in MRI data, could benefit from adaptive functions that adjust their activation based on data characteristics. For instance, PReLU, with its learnable parameters for negative slopes, can dynamically adjust the flow of information through layers, potentially enhancing segmentation accuracy at tumor boundaries. However, research has yet to fully explore how adaptive functions might further improve 3D U-Net's capability to manage complex tumor shapes and textures across MRI slices, where preserving gradient flow and preventing vanishing gradients is crucial.

Despite the success of attention mechanisms in U-Net adaptations, research on integrating alternative activation functions within these attention-enhanced architectures is sparse. Attention modules in models like Attention U-Net and AResU-Net [14] have demonstrated improvements in focusing on tumor-relevant regions. However, further gains might be achieved by using activation functions that are optimized to work with attention mechanisms. For instance, Swish,

which provides smoother gradients and enhanced gradient flow, could work synergistically with attention modules to improve feature extraction and gradient flow within U-Net. Testing Swish or similar functions in conjunction with attention-based U-Net models could yield insights into how these functions enhance spatial awareness and feature sensitivity, particularly in brain tumor segmentation where details are crucial.

Another promising but underexplored function is E-Tanh [8], a modified hyperbolic tangent function designed to preserve negative values and maintain small gradients even for very low inputs. E-Tanh has shown potential for capturing subtle contrasts at tumor boundaries, a task where ReLU's properties may fall short. U-Net models for brain tumor segmentation could benefit from E-Tanh's ability to enhance edge detection and boundary precision without losing critical information in the negative range. However, there is little research on the application of E-Tanh in brain tumor segmentation.

Hybrid U-Net models, which combine the strengths of CNN-based encoders with specialized U-Net decoders, also represent a potential area for exploration with alternative activation functions. Models like BU-Net [20] and RMU-Net[21], which use ResNet50 or MobileNet encoders, perform well in handling the high-dimensional data of MRI images. However, research on how activation functions tailored to medical imaging, such as adaptive functions or functions specifically designed for MRI data, could impact the segmentation quality in these hybrid architectures is still limited. By experimenting with functions like PReLU, researchers might enhance these hybrid models segmentation performance, particularly in differentiating subtle tumor boundaries.

While U-Net-based models achieve strong performance in brain tumor segmentation, there remains a significant gap in research on alternative activation functions optimized for this task. Investigating functions like Leaky ReLU, PReLU, Swish, and E-Tanh in U-Net could reveal new ways to improve boundary delineation, gradient flow, and overall segmentation accuracy. Addressing these gaps may lead to models that not only perform better in brain tumor segmentation but also enhance clinical applicability through increased reliability in handling complex medical imaging challenges.

2.8 Justification for focusing on activation functions within U-Net for improved brain tumor segmentation

Focusing on optimizing activation functions within the U-Net architecture holds considerable promise for enhancing brain tumor segmentation accuracy and computational efficiency. Activation functions play a fundamental role in defining how neural networks learn complex patterns, especially in medical image segmentation where fine-grained detail retention and boundary precision are critical. By investigating alternative activation functions specifically tailored

to medical imaging, this research aims to improve segmentation outcomes in both accuracy and stability, addressing key clinical needs for precise and reliable diagnostic tools.

The literature shows that standard activation functions, such as ReLU, have been effective in U-Net architectures, but limitations arise in their ability to capture subtle contrasts and preserve boundary features essential for accurate tumor delineation. For instance, ReLU's tendency to zero out negative inputs can lead to the loss of critical information in regions where fine detail matters, potentially impacting diagnostic accuracy. By exploring alternative functions like Leaky ReLU, PReLU, and E-Tanh, this research aims to address these limitations and develop an enhanced approach that better preserves feature details at tumor boundaries. E-Tanh, in particular, has shown potential in retaining more boundary information through non-linear adjustments, which could improve model sensitivity to critical structures within MRI data, advancing segmentation capabilities for clinical applications.

Moreover, as brain tumor segmentation increasingly relies on automated systems, a high level of accuracy and model reliability is essential for practical deployment. Current U-Net adaptations often emphasize architectural changes, such as multi-scale feature extraction or attention mechanisms, but the impact of different activation functions has been underexplored. This research gap presents an opportunity to optimize U-Net's segmentation performance by focusing on alternative activation functions that balance gradient stability with feature retention, a critical factor in accurately distinguishing between tumor and healthy tissue. This focus aligns with clinical goals, where enhancing segmentation reliability can streamline the diagnostic process, reduce manual intervention, and ultimately support more accurate treatment planning for patients.

The integration of adaptive activation functions within U-Net could also benefit computational efficiency, a significant consideration given the large datasets typical in MRI-based segmentation tasks. Activation functions like Swish and PReLU, known for their smooth gradient flow and adaptability, could potentially improve convergence rates without compromising accuracy. Faster convergence is particularly relevant in real-time clinical settings, where processing speed is as crucial as segmentation quality. This research intends to determine whether such activation functions could reduce computational costs, thereby making U-Net-based models more accessible for deployment in varied healthcare environments, including resource-limited settings where processing power may be constrained.

Additionally, focusing on activation functions allows for fine-tuning of U-Net models to specific imaging characteristics, such as those unique to brain MRI. Tumors in these scans exhibit highly variable shapes, sizes, and textures, which demand precise boundary detection and adaptability from segmentation models. The use of customized activation functions within U-Net could enhance the model's ability to adapt to these variations [8], providing consistent performance across different imaging conditions and tumor morphologies. This adaptability is crucial for clinical application, where robust segmentation models must handle diverse patient

data reliably.

This research further builds on emerging findings in the field, suggesting that activation function optimization could enhance segmentation detail and robustness in brain tumor segmentation. While attention mechanisms and feature-matching techniques have shown promise, their integration with optimized activation functions may yield even better outcomes. For example, models like SMU-Net [18], which leverage style and content matching for missing modality scenarios, could further benefit from an activation function that preserves fine details in low-contrast regions, thereby improving segmentation accuracy under challenging conditions.

Chapter 3

Methodology

The primary objective of this study is to investigate the impact of various activation functions on the performance of the U-Net model for brain tumor segmentation. Accurate segmentation of brain tumors in medical imaging is crucial for diagnosis, treatment planning, and monitoring disease progression. Activation functions play a pivotal role in deep learning models by introducing non-linearity, which enables the model to learn complex patterns within the data [8]. This research aims to determine which activation functions enhance the U-Net model's ability to accurately segment brain tumors from MRI scans.

The U-Net architecture was selected for this study due to its proven efficacy in biomedical image segmentation tasks. Its encoder-decoder structure, combined with skip connections, allows for precise localization and contextual understanding, making it highly suitable for segmenting intricate structures like brain tumors [3]. Focusing on activation functions is essential because they significantly influence the model's learning dynamics, including gradient flow and convergence behavior [8]. By experimenting with different activation functions, this study seeks to optimize the U-Net model's performance specifically for brain tumor segmentation.

3.1 Dataset

The BRATS 2020 dataset, sourced from Kaggle [2], dataset is a valuable collection of multimodal MRI, designed specifically for the segmentation of brain tumors, with a focus on gliomas. The dataset contains 3D MRI volumes , each comprising 155 2D slices, with each slice measuring 240×240 pixels. These volumes are stored in NIfTI format as single-channel grayscale images.It includes 473 patients [3], each with four MRI modalities and an expert-annotated segmentation file which has five labels segmenting the tumor.
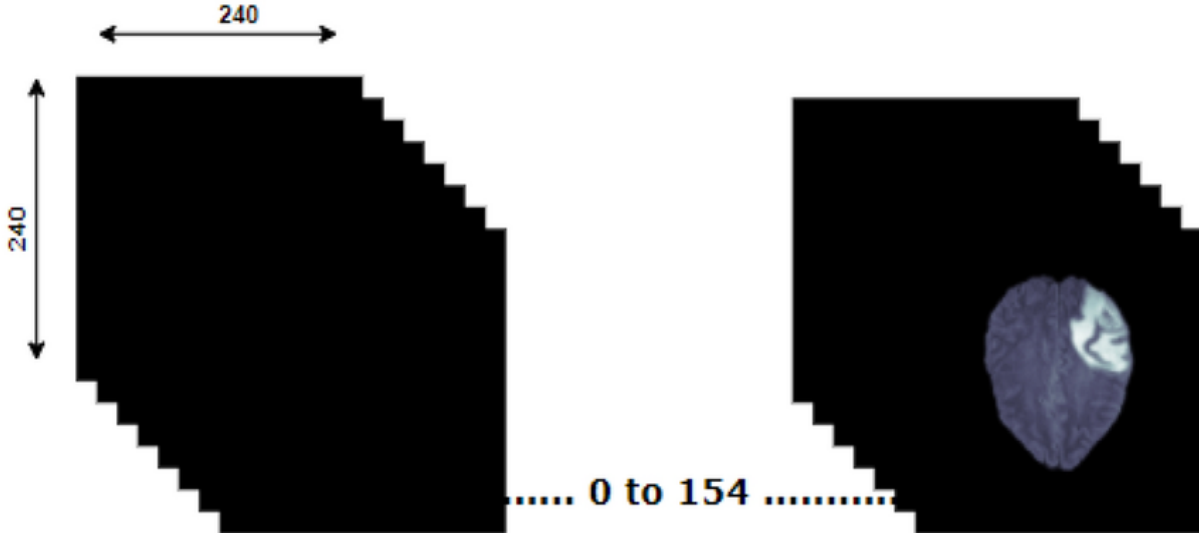
Figure 3.1: Visualization of 3D brain MRI [3]

The modalities are Native T1-weighted (T1), a standard MRI scan without contrast; post-contrast T1-weighted (T1ce), which highlights abnormal tissues by using contrast agents; T2-weighted (T2), which emphasizes tissue water content and helps identify swelling or edema; and T2-FLAIR (Fluid Attenuated Inversion Recovery), a specialized scan that suppresses cerebrospinal fluid signals to better visualize brain lesions. These four different MRI volumes provide a comprehensive view of the tumor and surrounding tissues, enhancing the accuracy of tumor detection and segmentation.

Accompanying the MRI scans are expert-annotated segmentation masks that divide the tumors into several sub-regions. These annotations are crucial for training and evaluating machine learning. The dataset categorizes the tumor regions into five labels: Label 0, which indicates areas that do not contain tumor tissue (Not Tumor, NT); Label 1, which corresponds to the necrotic and non-enhancing tumor core (NCR/NET), representing the dead or non-viable center of the tumor; Label 2, the peritumoral edema (ED), which is the surrounding area of swelling or fluid buildup caused by the tumor's presence; Label 3, Missing (No pixels in all the volumes contain label 3); and Label 4, which represents the GD-enhancing tumor (ET), the active, highly vascularized part of the tumor that absorbs the contrast agent and is indicative of the growing and aggressive nature of the tumor.
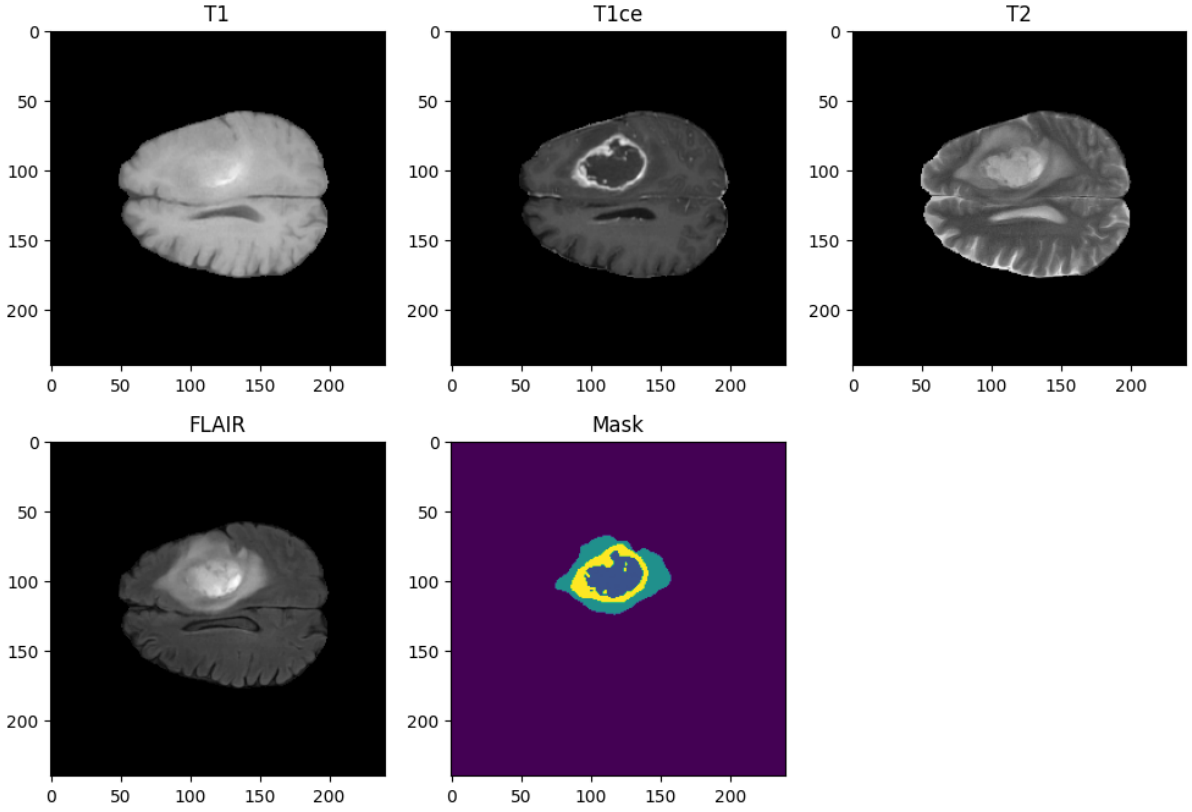
Figure 3.2: Different modalities of an MRI image and the mask

3.2 Preprocessing

Preprocessing plays a critical role in ensuring that the input data is clean, standardized, and appropriately formatted for the training of the U-Net model. The preprocessing pipeline employed in this work focuses on preparing the BraTS2020 dataset [2], which comprises multimodal MRI scans (T1, T2, T1ce, and FLAIR) and segmentation masks. Each preprocessing step is carefully designed to enhance the model's performance and ensure compatibility with the training framework.

The preprocessing begins with the use of the 'LoadImaged' transformation from MONAI [22], which facilitates efficient reading of the multimodal MRI data. This step ensures that the scans and segmentation masks are loaded into memory and associated with their respective patient IDs. The data is stored in a structured dictionary format, with keys corresponding to the modalities (t1, t2, t1ce, flair) and the ground truth segmentation mask (seg).

To adapt the dataset to the U-Net architecture, each MRI volume is processed using a custom 'AddChannelTransform'. This transformation adds a new channel dimension to each 3D scan and mask, as U-Net requires channel-specific inputs. Next, the images are standardized to a

common orientation using the Orientationd transformation with the Right Anterior Superior (RAS) coordinate system. This step ensures consistency in spatial orientation across all scans [22].

The 'CropForegroundd' transformation is applied to eliminate irrelevant background regions from the MRI volumes. By identifying the ROI based on the intensity values of the t1 modality, this step reduces the input size and ensures that the model focuses on meaningful areas.Given the variability in MRI scan dimensions, the dataset is standardized to a fixed spatial size of 160×160×80 using the 'SpatialPadd' transformation. Padding is applied symmetrically with a constant value, ensuring uniformity in input dimensions while preserving the original anatomical structures. To further refine the input size and focus on the central ROI, the 'CenterSpatial-Cropd' transformation is used. This operation extracts a 160×160×80 region from the center of each volume, maintaining the integrity of the target areas while discarding peripheral data [22].

The dataset is divided into training and validation sets, with 60 samples allocated for validation and 309 for training. Preprocessing is applied independently to both sets, ensuring that the validation data remains untouched by the training. This separation guarantees unbiased evaluation of the model's performance.
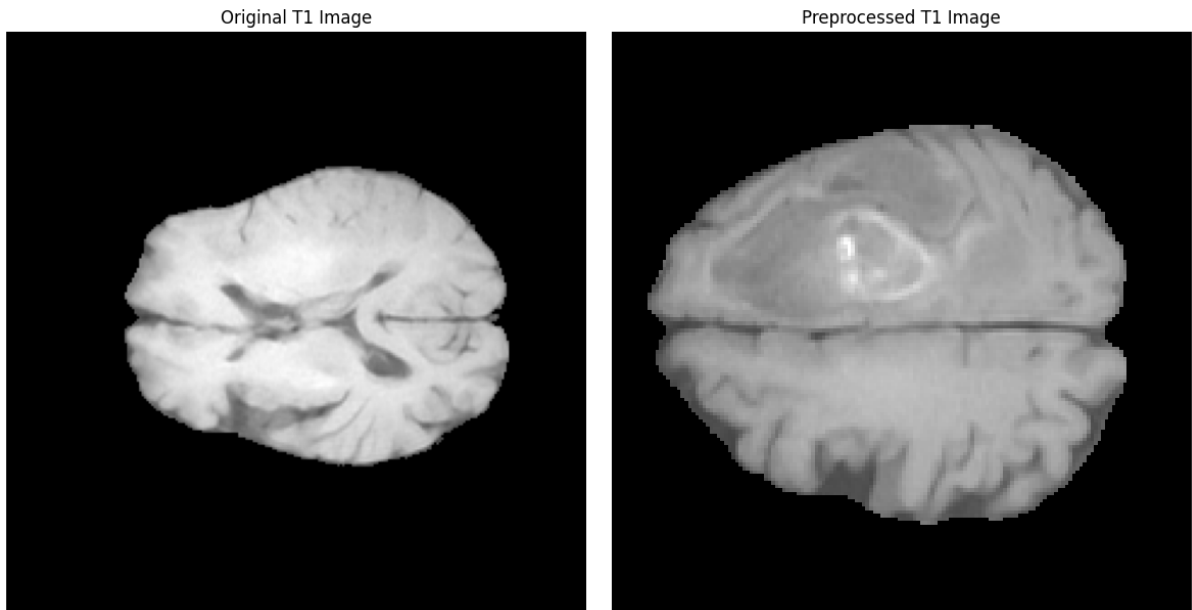
Figure 3.3: MRI image fot T1 before and after preprocessing

3.3 U-Net Model Architecture
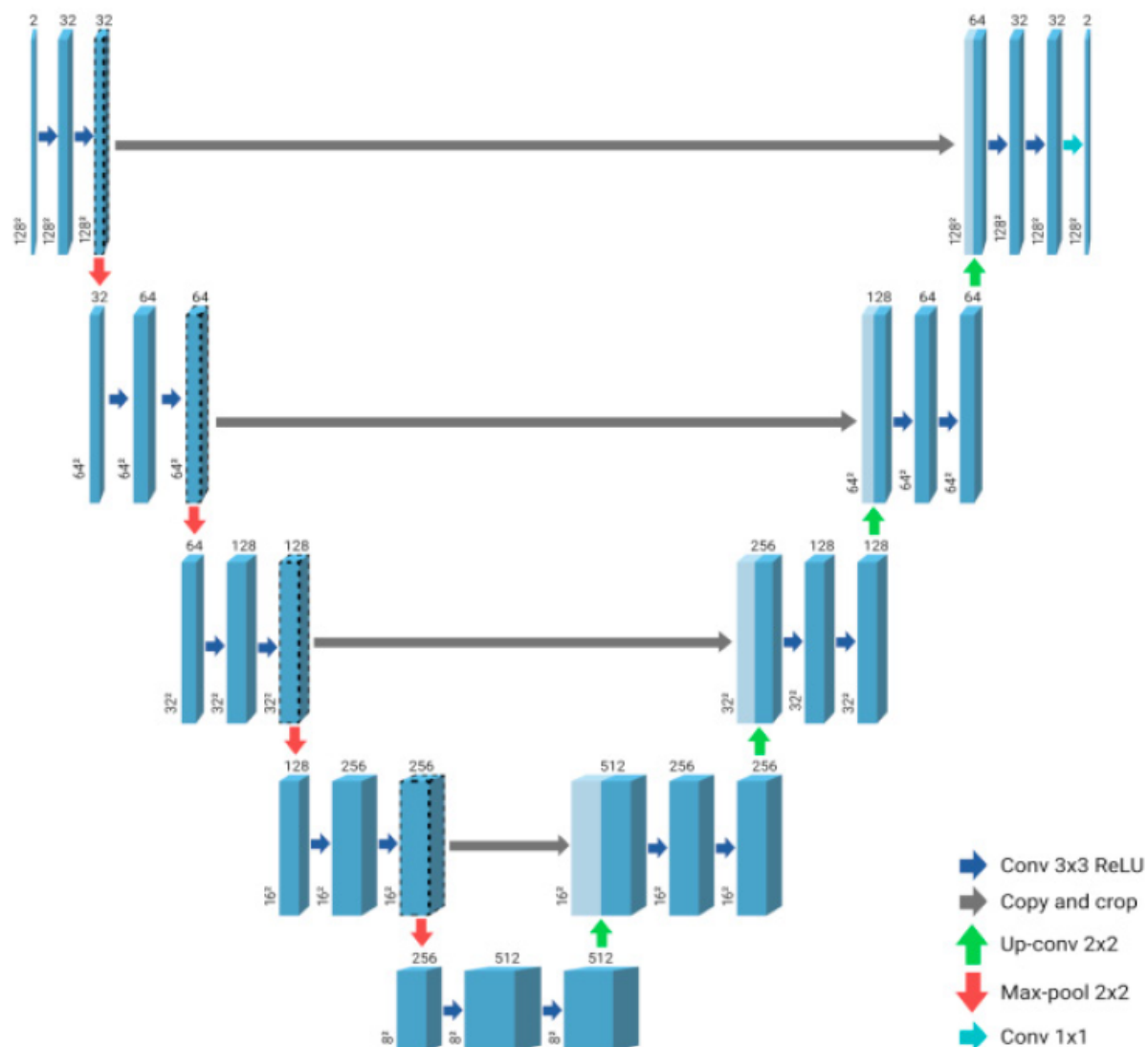
Figure 3.4: U-net architecture for the 3D image segmentation model [4]

For brain tumor segmentation, a 3D U-Net model was used for my test case. This choice was motivated by U-Net's demonstrated effectiveness in medical imaging tasks that require precise localization and segmentation of regions of interest.

The model was designed to process 3D MRI data and accommodate four input channels corresponding to the four MRI modalities: T1, T2, T1ce, and FLAIR. The architecture's spatial dimensions and channel configurations were specifically tailored to handle the complex patterns present in multi-modal MRI datasets. The network includes the following key components:

3.3.1 Input and Output Configuration

Input Channels where the model takes four input channels, each corresponding to an MRI modality (T1, T2, T1ce, and FLAIR). Output Channels where the model predicts four classes, corresponding to the background and three tumor regions: whole tumor, tumor core, and enhancing tumor.

3.3.2 Encoder-Decoder Structure

he U-Net architecture follows an encoder-decoder structure designed for efficient and precise segmentation tasks. The encoder progressively reduces spatial dimensions while capturing high-level feature representations through a series of convolutional blocks. Each block comprises 3D convolutional layers with a kernel size of 3 and an activation function. Downsampling is performed using strided convolutions, which halve the spatial dimensions at each stage. The feature maps in the encoder increase in depth, progressing through channel sizes of 16, 32, 64, and 128, enabling the network to learn increasingly abstract and complex features [22].

The decoder mirrors the encoder by reconstructing spatial dimensions step-by-step through upsampling layers. Skip connections between corresponding encoder and decoder layers are incorporated to preserve spatial details and ensure precise localization of features. This architecture effectively combines high-level feature extraction with fine-grained spatial information, making it well-suited for tasks like tumor segmentation in medical imaging [22].

3.3.3 Skip Connections

Skip connections bridge corresponding layers in the encoder and decoder, ensuring that spatial information is preserved throughout the network. These connections are essential for recovering fine details and improving segmentation accuracy [22].

The model was implemented using the MONAI library [22], a specialized framework for medical imaging applications. MONAI's U-Net implementation provided a flexible and efficient starting point, which why it was used to compare between activation functions.

3.4 Loss Function

The Dice Loss was employed as the primary loss function for optimizing the segmentation performance of the U-Net model. Dice Loss is particularly suitable for medical image segmentation tasks, as it directly optimizes the Dice Similarity Coefficient (DSC), a metric widely used for evaluating the overlap between predicted and ground truth regions. This approach ensures that the model learns to focus on improving the accuracy of the predicted segmentation boundaries [3].

$$\text{Dice Loss} = 1 - \frac{2 \times |P \cap G|}{|P| + |G|} \tag{3.1}$$

where P and G denote the predicted segmentation and ground truth masks, respectively, and $|P \cap G|$ represents the intersection of the two sets. This formulation ensures that the model is penalized more heavily for errors in smaller regions, which is advantageous when segmenting minor or less dominant structures like brain tumors [3].

In my implementation, the Dice Loss function was imported from the MONAI library [22], which is a specialized framework for medical imaging. The use of this library provides a robust and efficient implementation of the Dice Loss, tailored for 3D medical image data.

3.5 Training Process

The training process was designed to evaluate the effect of different activation functions on the performance of the U-Net model for brain tumor segmentation. To ensure a controlled and fair comparison, all other variables—including the dataset, preprocessing steps, optimizer, learning rate, batch size, and loss function—were held constant across all experiments. This strategy was used to isolate the impact of the activation function on model performance.

The Adam optimizer was employed with an initial learning rate of 1×10^{-4}. This choice ensured efficient convergence while maintaining stability across different activation function experiments. The optimizer's settings were consistent for all runs to minimize confounding variables.

Due to the high-dimensional nature of 3D MRI data, a batch size of 1 was used to fit the model and data within the available GPU memory. The training was conducted for 50 epochs, ensuring sufficient iterations for convergence. Epoch loss was recorded for each training session to monitor learning progress.

The preprocessing method and Dice Loss described earlier were applied uniformly across all experiments to ensure consistency in input data and evaluation metrics. This approach prevented the introduction of variability unrelated to the activation function, enabling a fair and controlled comparison.

3.6 Activation Functions

Activation functions are a critical component of neural networks, introducing non-linearity to the model and enabling it to learn complex patterns [8]. They play a significant role in how information flows through the network and influence both the training dynamics and the final performance of the model. In this study, 8 activation functions were systematically tested in

a controlled environment to evaluate their impact on brain tumor segmentation using a U-Net architecture. All other variables, such as the dataset, preprocessing steps, optimizer, learning rate, and loss function, were held constant to isolate the effect of the activation functions. Each activation function and its mathematical formulation are described below.

3.6.1 ReLU

The ReLU activation function is one of the most commonly used activation functions in deep learning due to its simplicity and computational efficiency. It is defined as:

$$f(x) = \max(0, x) \tag{3.2}$$

ReLU outputs the input directly if it is positive and zero otherwise, introducing sparsity in the activations. This sparsity reduces the computational load and helps mitigate the vanishing gradient problem by allowing gradients to pass through for positive inputs. However, ReLU suffers from the dead neurons problem, where neurons can become inactive and stop learning if their input is always negative [5].
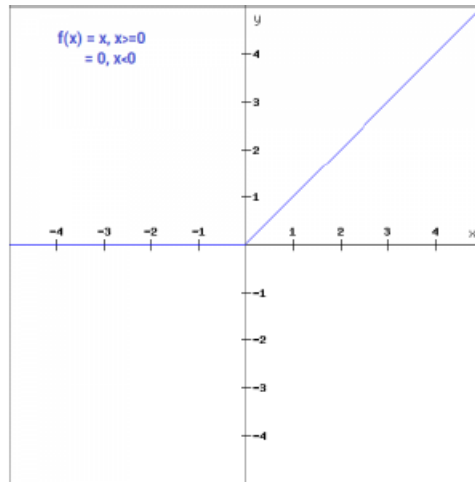
Figure 3.5: ReLU Activation Function plot [5]

3.6.2 ReLU6

ReLU6 is a variation of ReLU that limits the output to a maximum value of 6. It is defined as:

$$f(x) = \min(\max(0, x), 6) \tag{3.3}$$

By capping the output, ReLU6 helps prevent excessively large activations, which can stabilize training, especially in deep networks [6].
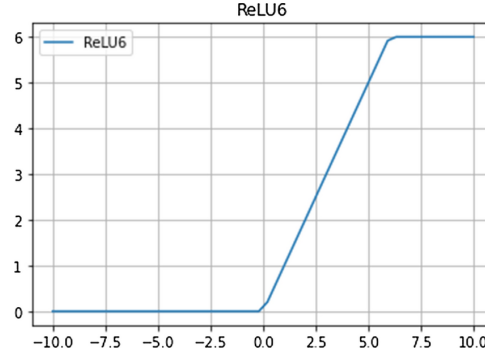
Figure 3.6: ReLU6 Activation Function plot [6]

3.6.3 Leaky ReLU

Leaky ReLU addresses the issue of dead neurons by allowing a small gradient for negative inputs.It is defined as:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{if } x \leq 0 \end{cases} \qquad (3.4)$$

where α is a small constant (e.g.,0.01). This ensures that no neuron becomes completely inactive, enhancing the model's ability to learn [5].
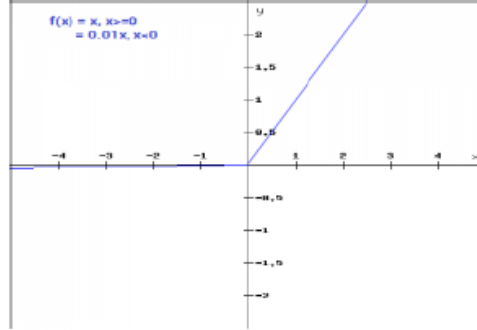
Figure 3.7: Leaky ReLU Activation Function plot [5]

3.6.4 PReLU

PReLU is an extension of Leaky ReLU where the slope of the negative part is learned during training. It is defined as:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{if } x \leq 0 \end{cases} \qquad (3.5)$$

where α is a learnable parameter. This added flexibility allows the network to adapt the activation function based on the data [5].
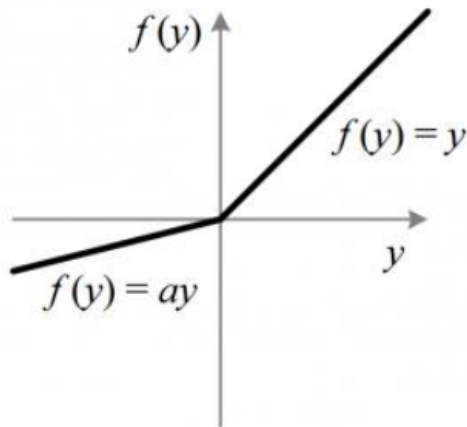
Figure 3.8: PReLU Activation Function plot [5]

3.6.5   sigmoid

The sigmoid function maps inputs to the range (0, 1), making it useful for binary classification tasks. It is defined as:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3.6}$$

While sigmoid provides a smooth gradient, it suffers from vanishing gradients for large positive or negative inputs, which can slow down learning in deep networks [5].
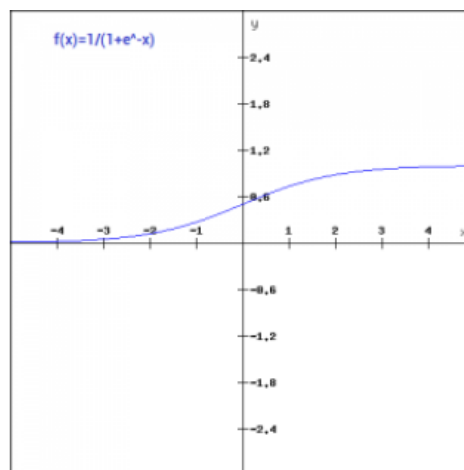
Figure 3.9: Sigmoid Activation Function plot [5]

3.6.6   Tanh

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.7}$$

Unlike sigmoid, Tanh centers the output around zero, making it more suitable for inputs with both positive and negative values. However, it is still prone to the vanishing gradient problem [5].
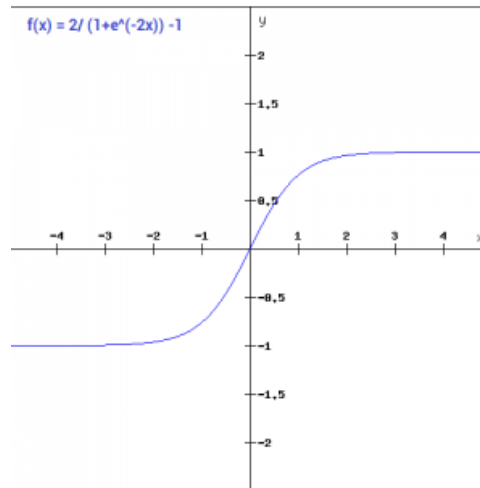
Figure 3.10: Tanh Activation Function plot [5]

3.6.7 ELU

ELU introduces smoothness for negative inputs and is defined as:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{if } x \le 0 \end{cases} \tag{3.8}$$

where α is typically set to 1. ELU avoids dead neurons and provides smoother transitions compared to ReLU, aiding optimization[5]. [5].
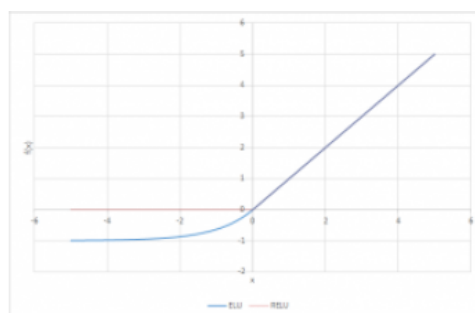
Figure 3.11: ELU Activation Function plot [5]

3.6.8 CELU

CELU is a variant of ELU that ensures continuous differentiability. It is given by:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha(\exp(x/\alpha) - 1) & \text{if } x \leq 0 \end{cases} \qquad (3.9)$$

Here, α controls the shape of the curve. CELU retains ELU's benefits while improving gradient behavior.
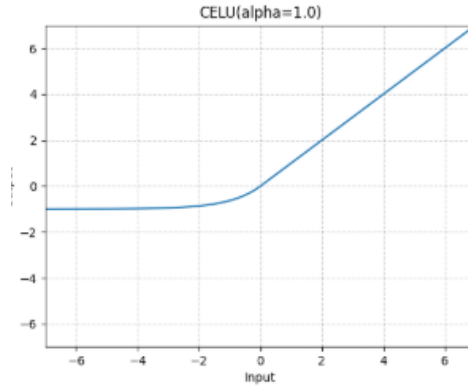
Figure 3.12: CELU Activation Function plot [7]

3.7 Computational Environment

Training was conducted on a high-performance computing setup equipped with an NVIDIA GeForce RTX 3070 GPU, providing sufficient computational power to handle the extensive training process. The system specifications included an Intel Core i7-10700K CPU, 32 GB DDR4 RAM, and the PyTorch framework with the MONAI library for model implementation. Data parallelism and GPU acceleration were leveraged to expedite the training process.The operating system used was Windows, and the experiments were implemented using Python 3.11.10.

Chapter 4

Results

The results section presents the performance analysis of different activation functions employed in a U-Net-based architecture for brain tumor segmentation. The focus lies on comparing the performance of each activation function in terms of model loss and segmentation accuracy, specifically evaluating the Dice similarity coefficients across whole tumor, enhancing tumor, and tumor core regions. By analyzing these metrics over training epochs, the strengths and weaknesses of each activation function are identified, enabling the selection of the most suitable activation function for this brain tumor segmentation.
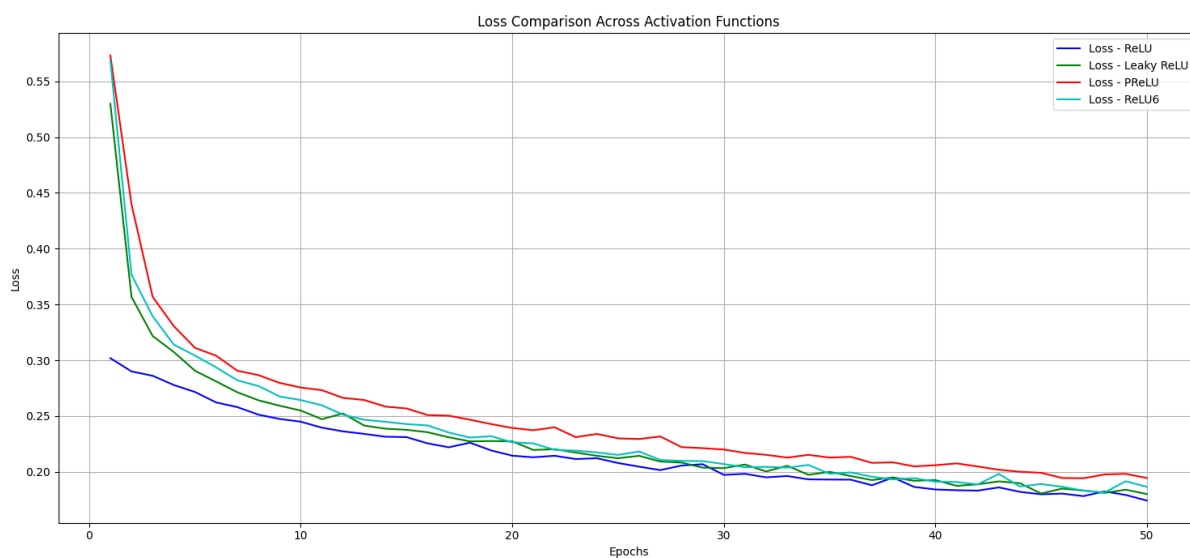
Figure 4.1

The graph presented showcases the training loss trends for four activation functions—ReLU, Leaky ReLU, PReLU, and ReLU6—across 50 epochs. The training loss metric quantifies the model's error during the training process, with lower values indicating better performance. The

trends observed in the graph highlight the relative strengths and weaknesses of these activation functions in facilitating efficient learning for brain tumor segmentation tasks.

ReLU demonstrates the most favorable performance, with the fastest convergence rate and the lowest final loss value. By around the 20th epoch, ReLU achieves a stable loss and continues to maintain its minimal value throughout the remaining epochs. This stability underscores its effectiveness in enabling the model to optimize efficiently during training. The superior performance of ReLU can be attributed to its simplicity and its ability to mitigate vanishing gradient issues, ensuring robust gradient flow and rapid error minimization.

Leaky ReLU, while competitive, converges at a slightly slower rate compared to ReLU during the initial epochs. Its final loss value stabilizes close to that of ReLU, but it remains marginally higher, indicating less efficient training. This slight inefficiency may stem from the small non-zero gradient it assigns to negative values, which, although beneficial in certain contexts, appears to introduce minor delays in convergence for this specific application.

PReLU, on the other hand, exhibits a slower convergence rate and consistently higher loss values throughout the training process. While it demonstrates steady improvement over epochs, its performance does not match that of ReLU or Leaky ReLU. Similarly, ReLU6 shows moderate convergence behavior, stabilizing at a loss level comparable to PReLU but higher than both ReLU and Leaky ReLU. ReLU6's performance suggests that its upper cap on activation values might constrain the model's ability to optimize effectively in this scenario.

The results indicate that ReLU performs the best among the tested activation functions, achieving the lowest and most stable training loss. Its effectiveness can be attributed to its simplicity and the absence of saturation issues, enabling faster and more efficient learning. Leaky ReLU follows closely but demonstrates slightly higher loss values, while PReLU and ReLU6 are less effective in this context.
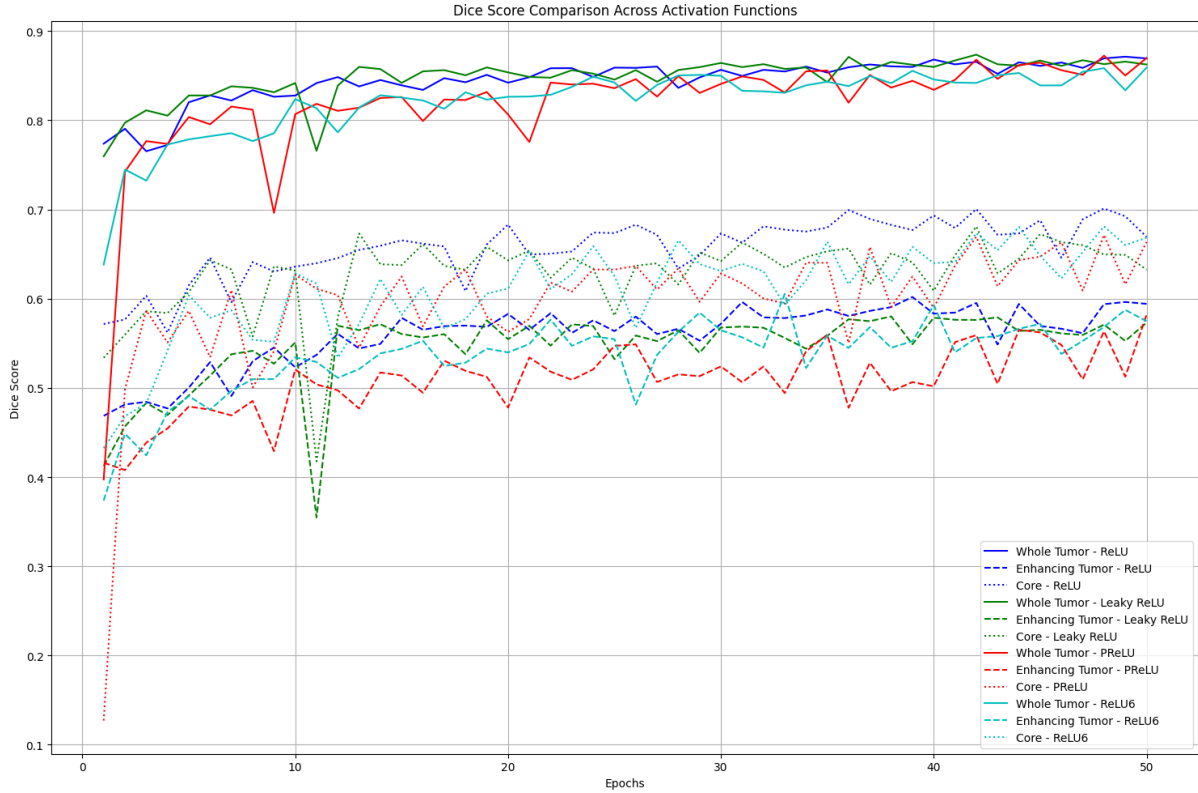
Figure 4.2

The graph illustrates the Dice score trends for different regions of brain tumors—Whole Tumor, Enhancing Tumor, and Tumor Core—across various activation functions (ReLU, Leaky ReLU, PReLU, and ReLU6) over 50 epochs. The Dice score is a performance metric commonly used in image segmentation tasks, measuring the overlap between the predicted segmentation and the ground truth, with higher values indicating better segmentation accuracy. The purpose of this comparison is to evaluate the effectiveness of each activation function in accurately segmenting the regions of interest.

ReLU demonstrates the most consistent performance across all tumor regions, maintaining high Dice scores with minimal fluctuations throughout the training process. For the Whole Tumor region, ReLU achieves Dice scores above 0.85 after the initial epochs and stabilizes at approximately 0.88, indicating superior segmentation accuracy. Similarly, for the Enhancing Tumor and Core regions, ReLU consistently outperforms the other activation functions, showcasing its robustness and efficiency in handling diverse tumor structures.

Leaky ReLU performs comparably to ReLU in the Whole Tumor segmentation, achieving similar Dice scores above 0.85. However, its performance in the Enhancing Tumor and Core regions shows more variability, with occasional dips in Dice scores, suggesting that it may be less stable compared to ReLU for these regions. While it still provides competitive performance,

the instability limits its reliability in achieving consistently high segmentation accuracy.

PReLU exhibits a slower improvement in Dice scores during the early epochs, particularly for the Enhancing Tumor and Core regions. While it eventually reaches reasonable performance, its Dice scores remain slightly lower and more volatile compared to ReLU and Leaky ReLU. This indicates that PReLU may require additional epochs to achieve comparable stability and performance, but its effectiveness is relatively constrained in this specific application.

ReLU6 shows reasonable performance for the Whole Tumor region, with Dice scores stabilizing around 0.85. However, its performance in the Enhancing Tumor and Core regions is less competitive, with lower Dice scores and significant fluctuations. This suggests that ReLU6 may be less suited for capturing the finer details required for accurate segmentation of these regions.

Overall, ReLU emerges as the best-performing activation function across all tumor regions—tumor core, enhancing tumor, and whole tumor—due to its consistent stability, superior segmentation accuracy, and minimal variability in Dice scores. Its ability to effectively handle the complexities of segmentation tasks makes it the most reliable choice for this application. Leaky ReLU serves as a viable alternative for Whole Tumor segmentation but falls short in stability for the other regions. PReLU and ReLU6, while effective to some extent, are less competitive in terms of both accuracy and consistency, highlighting the importance of selecting an activation function that aligns with the specific demands of medical image segmentation tasks.
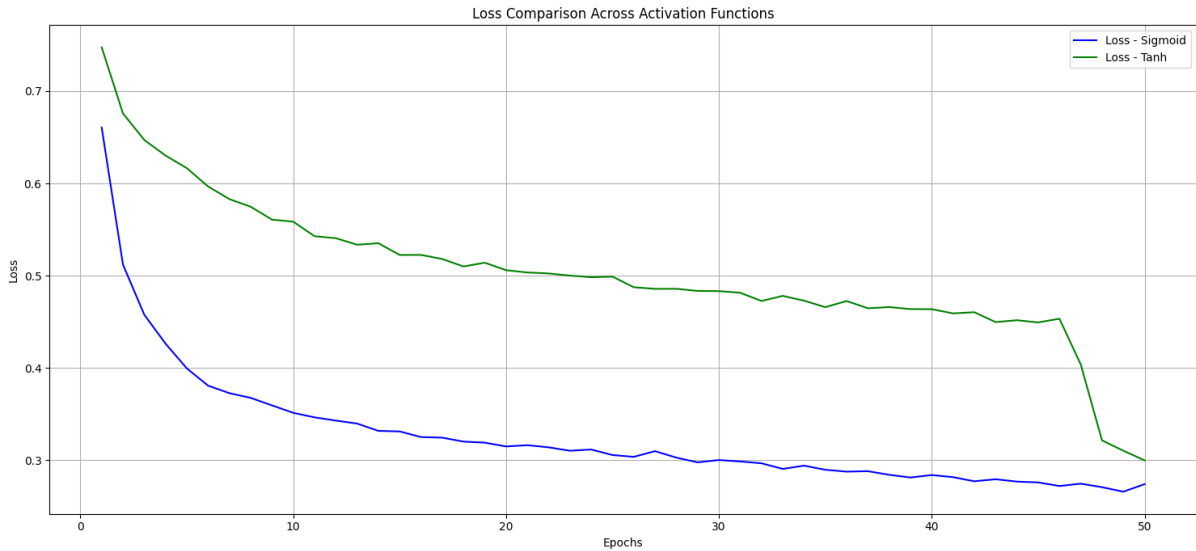
Figure 4.3

he graph compares the training loss trends of two activation functions, Sigmoid and Tanh, across 50 epochs during the training phase of the U-Net-based model for brain tumor segmen-

tation.The Sigmoid activation function demonstrates superior performance compared to Tanh. Sigmoid exhibits a faster convergence rate, with its loss decreasing rapidly within the initial epochs. By the 20th epoch, the loss stabilizes at a lower value, indicating effective optimization. The stability of Sigmoid's loss throughout the training phase underscores its ability to facilitate efficient learning and maintain consistent performance.

n contrast, Tanh exhibits slower convergence and higher loss values compared to Sigmoid. While Tanh's loss decreases steadily over the epochs, it stabilizes at a higher value for most of the training phase, suggesting that it is less effective in optimizing the model during earlier epochs. However, a notable improvement in Tanh's performance can be observed after the 48th epoch, where the loss begins to decrease significantly. This indicates that Tanh might perform better if the model is trained for additional epochs, as it shows potential for further optimization at later stages of training. The delayed convergence suggests that Tanh may require more iterations to adapt effectively to the patterns in the data compared to Sigmoid.
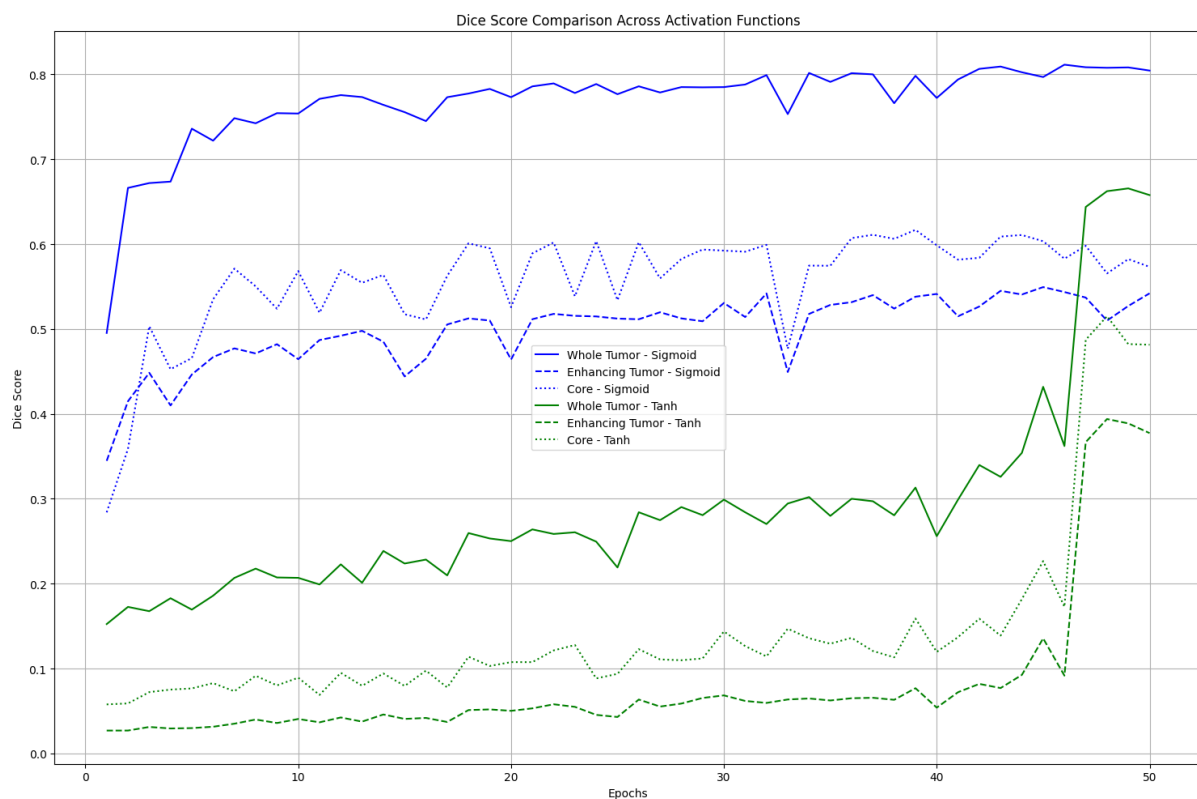
Figure 4.4

The graph compares the Dice score trends for the Sigmoid and Tanh activation functions across three regions of interest—Whole Tumor, Enhancing Tumor, and Tumor Core—over 50 training epochs. The Sigmoid activation function demonstrates superior performance overall. For the Whole Tumor region, Sigmoid achieves consistent Dice scores above 0.75 after the ini-

tial epochs, stabilizing near 0.8 as training progresses. Similarly, for the Enhancing Tumor and Tumor Core regions, Sigmoid maintains relatively higher Dice scores, although these regions exhibit more variability compared to the Whole Tumor segmentation. The steady improvement and stabilization of Sigmoid's Dice scores highlight its effectiveness in handling segmentation tasks across different tumor regions.

In contrast, the Tanh activation function shows significantly lower Dice scores across all regions for the majority of the training process. For the Whole Tumor region, Tanh starts with low scores and exhibits gradual improvement, stabilizing around 0.2 by epoch 30. However, a notable surge in Dice scores is observed for all regions after epoch 40, suggesting that Tanh may require more epochs to fully optimize and achieve competitive performance. This late improvement indicates that Tanh has the potential to perform better with extended training, particularly as its scores for the Enhancing Tumor and Tumor Core regions start to approach those of Sigmoid in the final epochs.

The superior performance of Sigmoid can be attributed to its ability to maintain stable gradients, ensuring effective learning during backpropagation. Its consistent and higher Dice scores across all tumor regions demonstrate its robustness and reliability for this application. Tanh, while showing promise in the later stages of training, struggles with slower initial convergence and lower stability, making it less effective within the current training timeframe. However, its late improvement suggests it could benefit from further exploration with longer training durations.
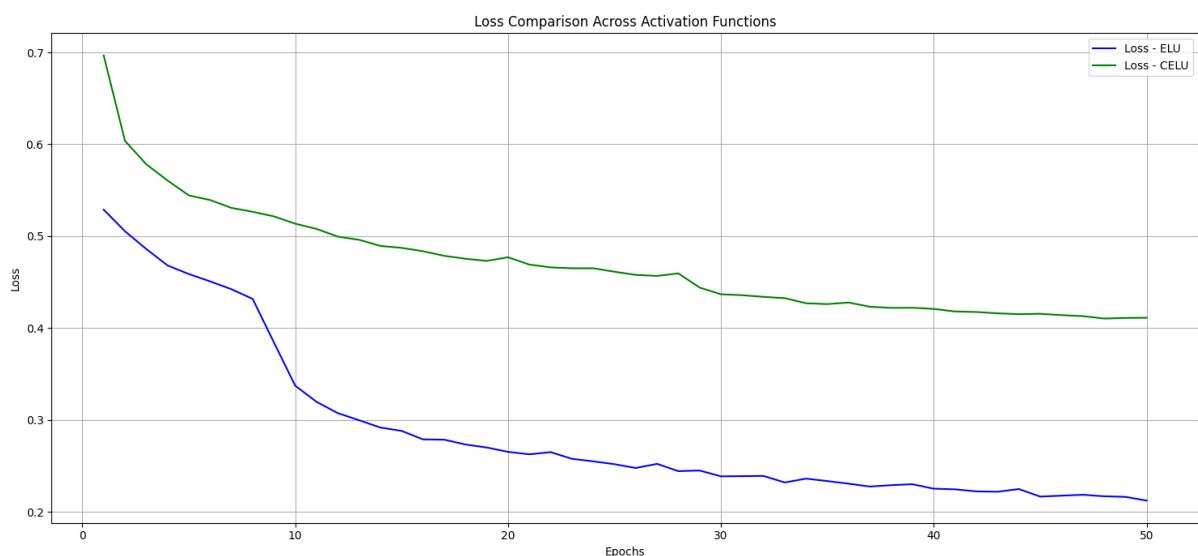
Figure 4.5

The graph illustrates the training loss trends of two activation functions, ELU and CELU, over 50 epochs.The ELU activation function demonstrates superior performance compared to

CELU in this case. ELU shows a faster convergence rate, with its training loss decreasing significantly within the first 10 epochs and stabilizing at a lower value around 0.25 as training progresses. This consistent decrease and early stabilization suggest that ELU provides a more effective gradient flow during training, enabling the model to optimize more efficiently and achieve better overall performance.

CELU, on the other hand, exhibits a slower convergence rate and maintains a higher loss value throughout the training process. While its loss decreases steadily over the epochs, it stabilizes at a value higher than that of ELU, indicating less effective optimization. The gradual decline in CELU's loss implies that it may require additional epochs to reach a more optimal loss value, but its overall performance remains inferior compared to ELU.
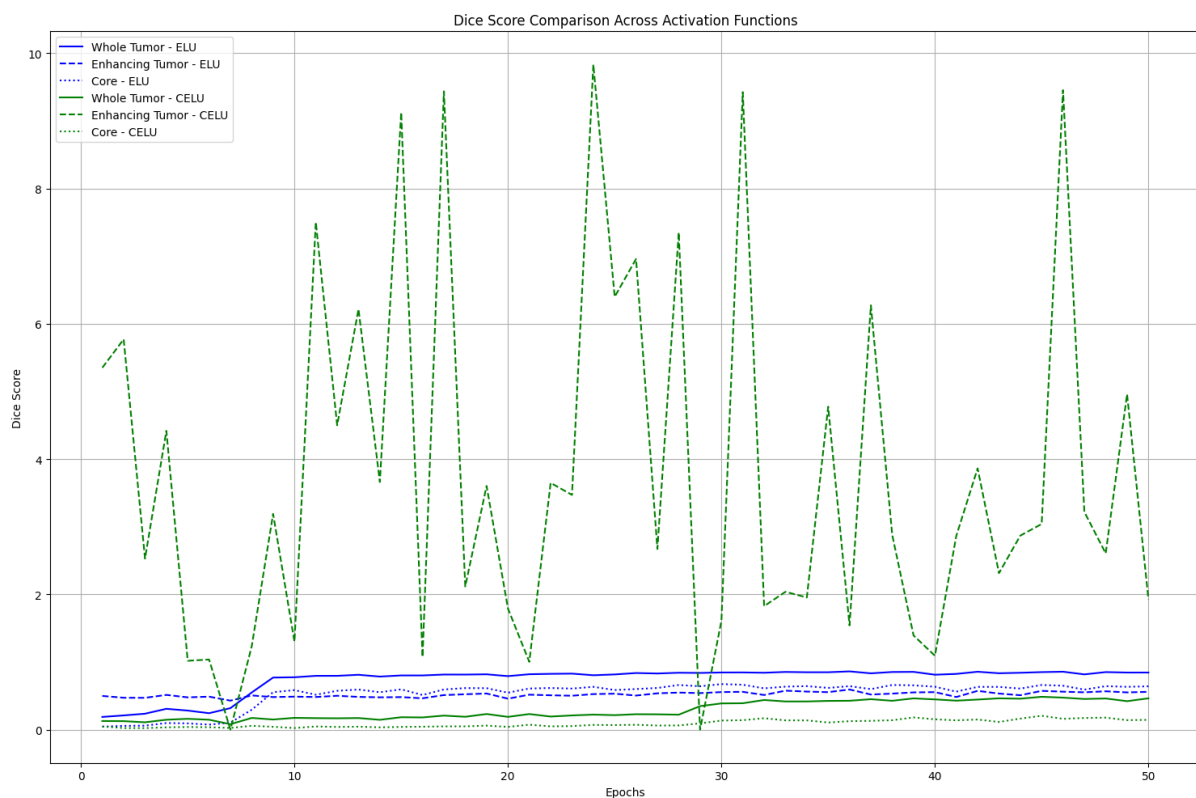
Figure 4.6

The graph presents a comparison of Dice scores for two activation functions, ELU and CELU across three tumor regions—Whole Tumor, Enhancing Tumor, and Tumor Core—over 50 epochs. For ELU, the Dice scores exhibit stability across all regions, with consistent values throughout the training process. The Whole Tumor Dice score is the highest among the three regions, stabilizing around 0.7 after the initial epochs. The Enhancing Tumor and Core regions also show consistent Dice scores but remain lower than the Whole Tumor region, reflecting the inherent complexity of accurately segmenting these smaller and more intricate regions.

In contrast, CELU demonstrates high variability in Dice scores across all regions, particularly for the Enhancing Tumor and Tumor Core. The Dice scores for CELU fluctuate significantly, often reaching extreme values that are not consistent with effective segmentation. This high degree of variability suggests instability in the optimization process, preventing CELU from achieving reliable segmentation performance. For the Whole Tumor region, CELU's performance is notably inconsistent, and the scores do not stabilize throughout the epochs.

ELU outperforms CELU by maintaining stable and higher Dice scores across all tumor regions. The stability of ELU can be attributed to its exponential property for negative values, which prevents vanishing gradients and ensures smooth gradient flow during optimization. This characteristic enables ELU to achieve consistent learning and reliable segmentation performance. On the other hand, the instability observed with CELU may result from its sensitivity to hyperparameter tuning or its inability to handle the complex patterns in the data effectively.
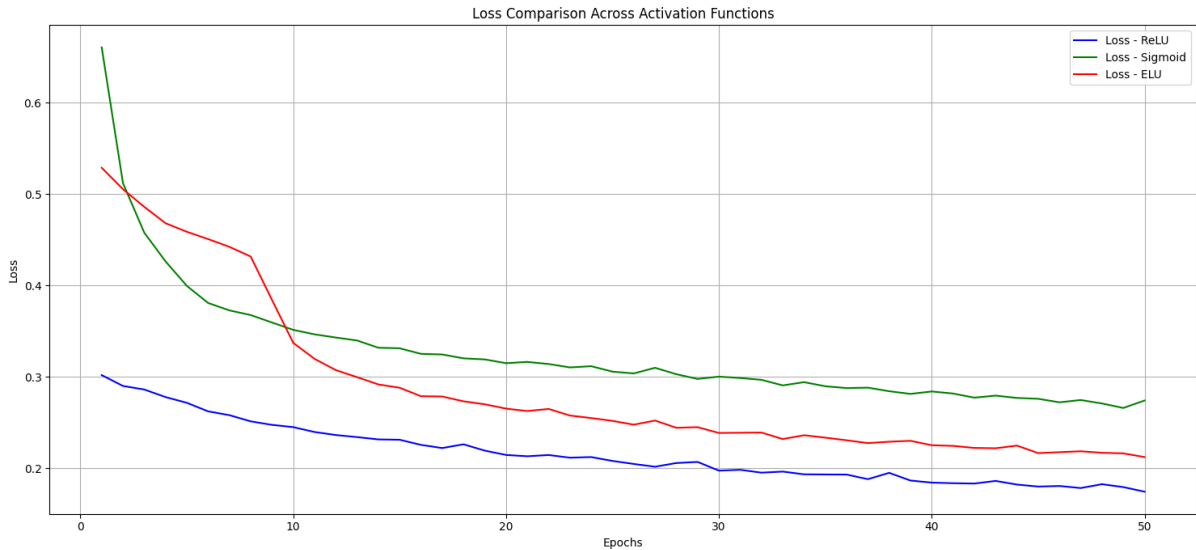
Figure 4.7

The graph compares the training loss trends for the three best preforming activation functions—ReLU, Sigmoid, and ELU—across 50 epochs,ReLU demonstrates the best overall performance, exhibiting the lowest and most stable loss values throughout the training process. From the initial epochs, ReLU achieves rapid convergence and maintains consistent performance, stabilizing at a minimal loss value of approximately 0.2. This superior performance can be attributed to RELU's simplicity and its ability to avoid vanishing gradients, ensuring efficient gradient flow during back propagation. Its effectiveness in minimizing error highlights its suitability for this application.

ELU initially shows a rapid decline in loss, performing comparably to Sigmoid in the early epochs. However, as training progresses, ELU surpasses Sigmoid and stabilizes at a loss value

lower than Sigmoid but higher than ReLU, approximately 0.25. The exponential smoothing of negative values by ELU contributes to its improved gradient flow and optimization, but it does not outperform ReLU in terms of final loss value and stability.

Sigmoid, while demonstrating a steady decline in loss, converges more slowly compared to ELU and ReLU. It stabilizes at the highest loss value among the three activation functions, approximately 0.3. Sigmoid's relatively slower convergence and higher final loss can be attributed to its tendency to saturate for extreme input values, which can hinder effective gradient updates and optimization in deeper layers.
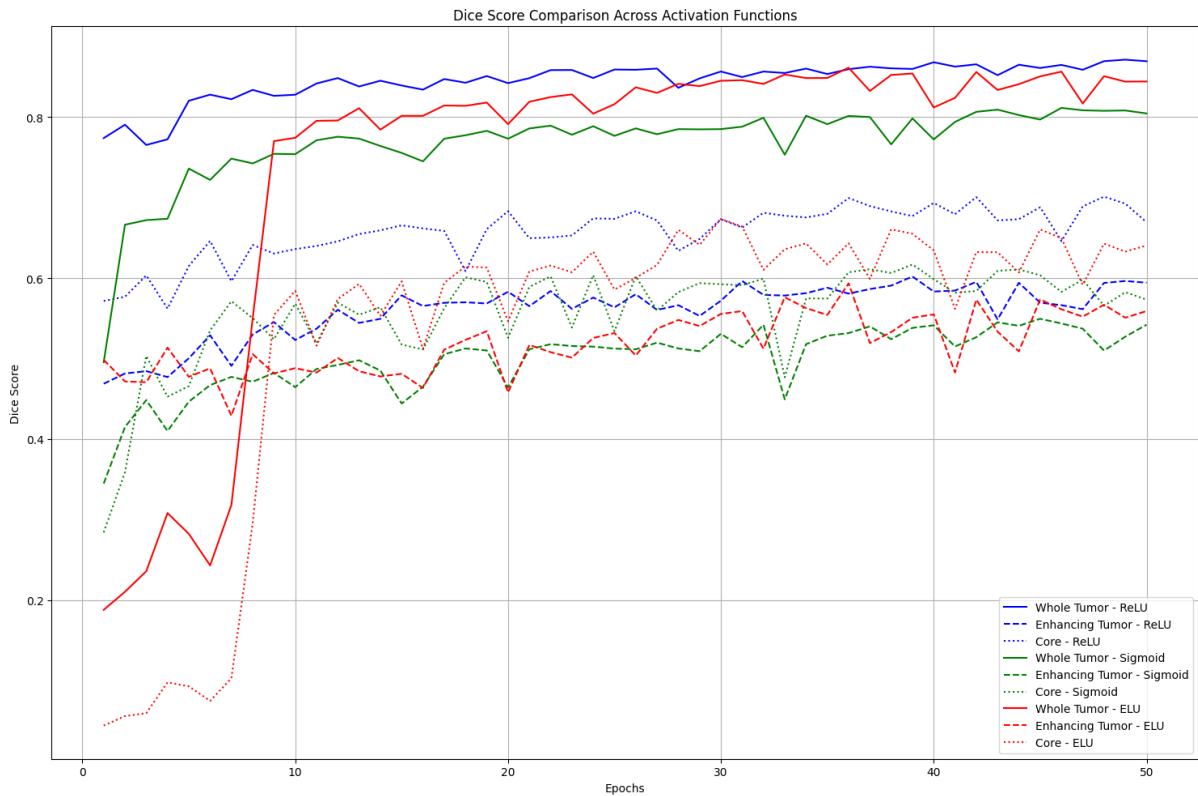
Figure 4.8

The graph provides a comparative analysis of Dice scores for the three best preforming activation functions—ReLU, Sigmoid, and ELU—across three tumor regions: Whole Tumor, Enhancing Tumor, and Tumor Core. ReLU exhibits the most consistent and robust performance across all tumor regions. For the Whole Tumor region, ReLU achieves the highest Dice scores, stabilizing above 0.85 early in the training and remaining steady until the final epochs. Similarly, in the Enhancing Tumor and Tumor Core regions, ReLU maintains competitive scores with minimal fluctuations, indicating its capability to adapt effectively to varying segmentation challenges. Its stability and high scores reflect its suitability for this segmentation task, likely due to its simplicity and efficient gradient flow, which avoid issues such as vanishing gradients.

Sigmoid performs comparably to ReLU for the Whole Tumor region, achieving Dice scores above 0.8 after the initial epochs. However, its performance for the Enhancing Tumor and Tumor Core regions is slightly less stable, with more noticeable fluctuations in Dice scores throughout the training. While Sigmoid demonstrates reasonable segmentation accuracy, its variability suggests that it may not generalize as well as ReLU to complex tumor regions.

ELU, on the other hand, shows slower initial convergence for all regions but demonstrates competitive performance as training progresses. For the Whole Tumor region, ELU reaches Dice scores above 0.8, approaching the performance of Sigmoid. However, for the Enhancing Tumor and Tumor Core regions, ELU exhibits more variability and slightly lower scores compared to ReLU and Sigmoid. Its exponential smoothing property for negative values aids optimization but does not surpass RELU's performance in this specific application.

ReLU outperforms all activation functions tested in this study, particularly in terms of stability and Dice scores for all tumor regions and achieving the lowest and most stable training loss. Its ability to maintain consistently high scores across epochs makes it the most reliable activation function for this segmentation task. Sigmoid, while effective for simpler regions like the Whole Tumor, shows more variability in complex regions. ELU, though competitive in later stages of training, struggles to achieve the same level of stability and accuracy as ReLU, emphasizing the importance of selecting an activation function that balances optimization efficiency and segmentation reliability.

The following figures present some of the segmentation results obtained from training the U-Net model with the activation functions across 50 epochs. As observed, Tanh and CELU emerged as the poorest-performing activation functions for the task of brain tumor segmentation, producing outputs with significant inaccuracies and deviations from the ground truth. In contrast, the remaining activation functions demonstrated relatively similar performance, with ReLU producing the most accurate segmentations, closely aligning with the ground truth. This highlights ReLU's effectiveness in retaining critical features and achieving precise boundary delineation compared to other functions.
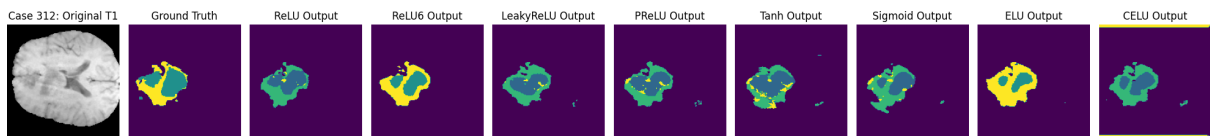
Figure 4.9
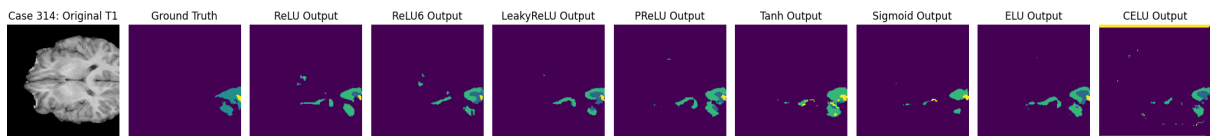
Figure 4.10
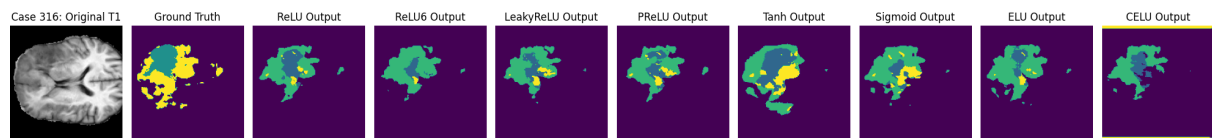
Figure 4.11
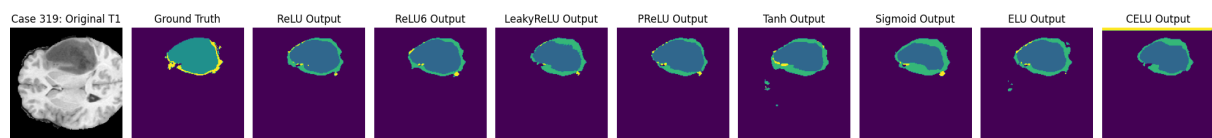
Figure 4.12
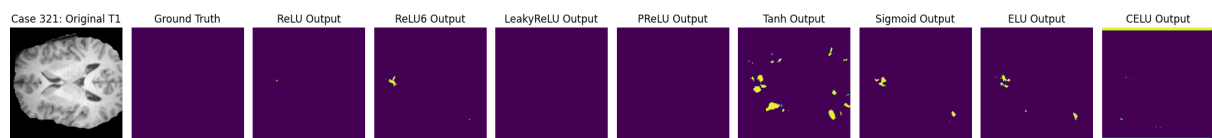
Figure 4.13
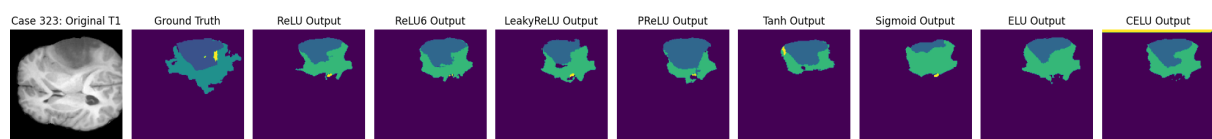
Figure 4.14
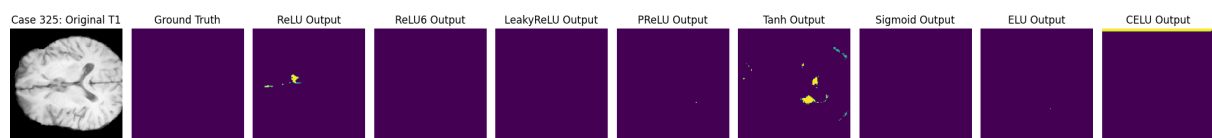
Figure 4.15
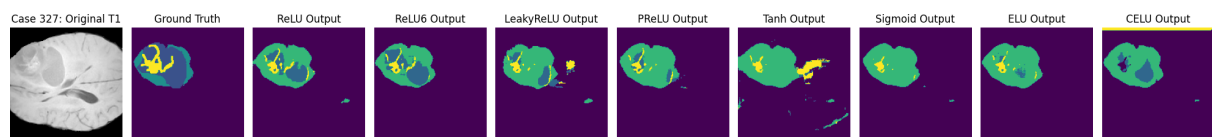
Figure 4.16
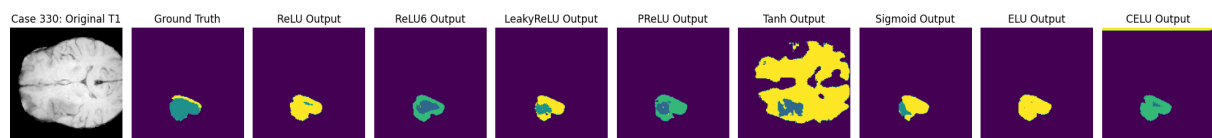
Figure 4.17
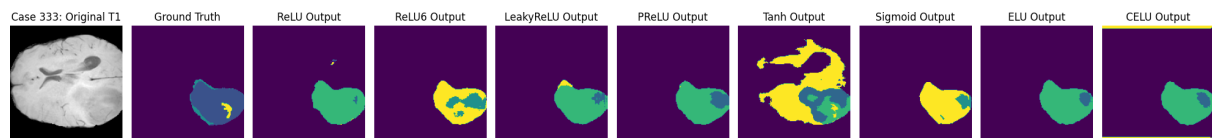
Figure 4.18
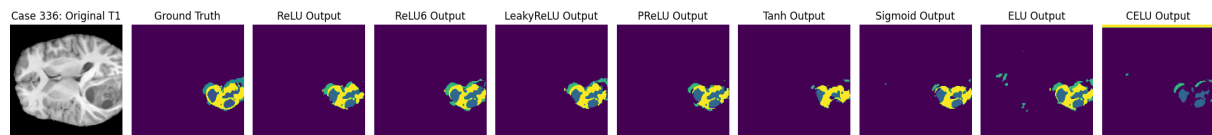
Figure 4.19
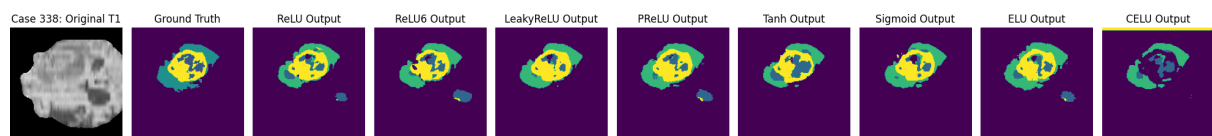
Figure 4.20
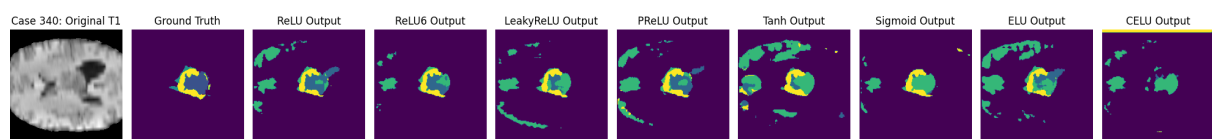
Figure 4.21
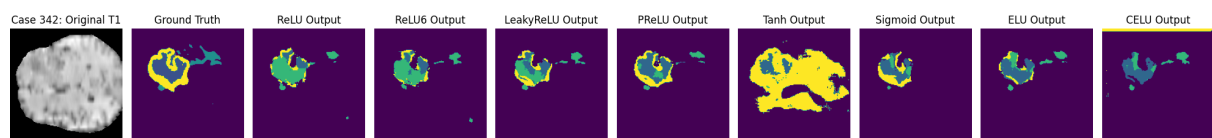
Figure 4.22
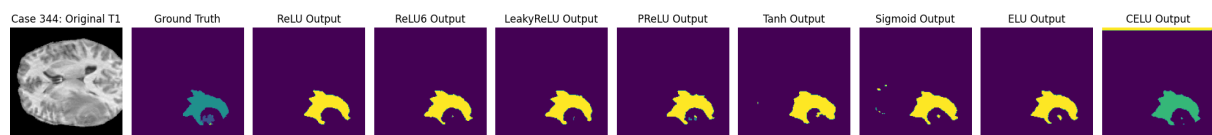
Figure 4.23
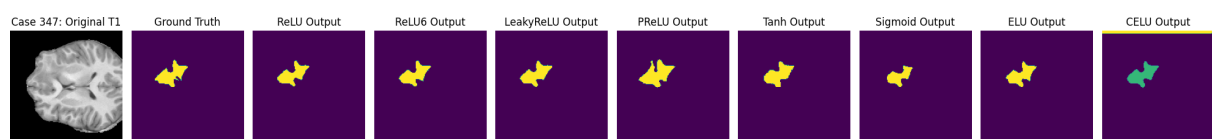
Figure 4.24
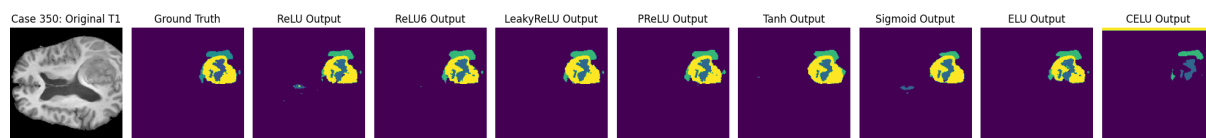
Figure 4.25
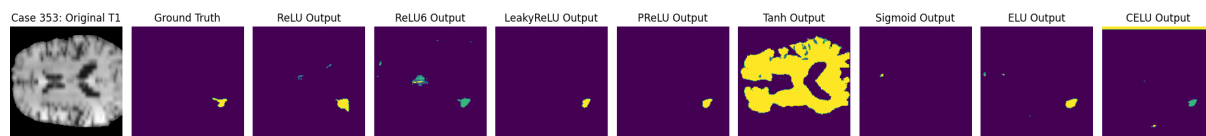
Figure 4.26
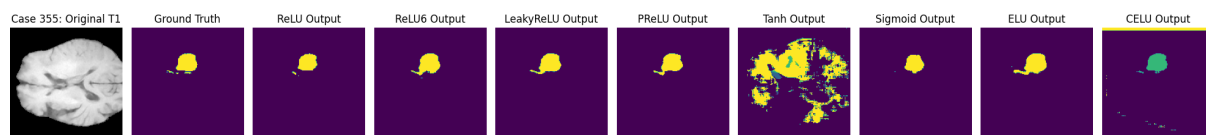
Figure 4.27
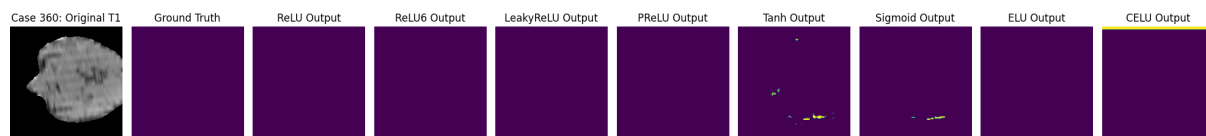
Figure 4.28

Next is the full results obtained from my experimentation.

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
1	0.9714	0.7738	0.4687	0.5716	0.3019
2	0.9713	0.7905	0.4815	0.5768	0.2900
3	0.9711	0.7653	0.4843	0.6033	0.2861
4	0.9723	0.7724	0.4770	0.5621	0.2778
5	0.9741	0.8203	0.5004	0.6150	0.2715
6	0.9751	0.8279	0.5290	0.6461	0.2622
7	0.9746	0.8222	0.4909	0.5960	0.2580
8	0.9758	0.8338	0.5299	0.6411	0.2512
9	0.9749	0.8264	0.5453	0.6305	0.2474
10	0.9758	0.8278	0.5230	0.6359	0.2449
11	0.9762	0.8417	0.5370	0.6398	0.2396
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
12	0.9770	0.8484	0.5606	0.6456	0.2363
13	0.9761	0.8380	0.5443	0.6547	0.2340
14	0.9773	0.8452	0.5493	0.6593	0.2315
15	0.9767	0.8392	0.5785	0.6655	0.2311
16	0.9761	0.8341	0.5652	0.6617	0.2255
17	0.9770	0.8472	0.5692	0.6585	0.2220
18	0.9754	0.8427	0.5699	0.6085	0.2261
19	0.9773	0.8510	0.5684	0.6608	0.2192
20	0.9777	0.8422	0.5830	0.6831	0.2145
21	0.9770	0.8483	0.5651	0.6495	0.2130
22	0.9776	0.8584	0.5838	0.6505	0.2144
23	0.9771	0.8585	0.5617	0.6529	0.2115
24	0.9778	0.8485	0.5758	0.6742	0.2122
25	0.9783	0.8591	0.5636	0.6736	0.2079
26	0.9785	0.8588	0.5799	0.6830	0.2046
27	0.9778	0.8602	0.5603	0.6715	0.2016
28	0.9772	0.8363	0.5663	0.6338	0.2056
29	0.9774	0.8481	0.5529	0.6486	0.2068
30	0.9779	0.8566	0.5719	0.6731	0.1973
31	0.9783	0.8497	0.5963	0.6627	0.1982
32	0.9789	0.8566	0.5791	0.6812	0.1951
33	0.9786	0.8547	0.5783	0.6774	0.1962
34	0.9783	0.8602	0.5812	0.6753	0.1933
35	0.9783	0.8535	0.5880	0.6798	0.1931
36	0.9790	0.8595	0.5807	0.6994	0.1930
37	0.9784	0.8626	0.5863	0.6894	0.1880
38	0.9784	0.8605	0.5906	0.6827	0.1949
39	0.9784	0.8598	0.6019	0.6770	0.1865
40	0.9795	0.8681	0.5833	0.6933	0.1842
41	0.9787	0.8627	0.5844	0.6793	0.1835
42	0.9794	0.8655	0.5952	0.7005	0.1832
43	0.9782	0.8521	0.5487	0.6716	0.1860
44	0.9785	0.8650	0.5941	0.6732	0.1821
45	0.9785	0.8610	0.5695	0.6879	0.1799
46	0.9774	0.8648	0.5663	0.6459	0.1805
47	0.9781	0.8588	0.5614	0.6890	0.1782
48	0.9794	0.8694	0.5939	0.7011	0.1824
49	0.9796	0.8713	0.5963	0.6924	0.1793
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
50	0.9790	0.8694	0.5942	0.6693	0.1742

Table 4.1: Summary of validation scores and losses across 50 epochs for ReLU

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
1	0.9629	0.6383	0.3738	0.4324	0.5686
2	0.9684	0.7448	0.4486	0.4682	0.3766
3	0.9693	0.7324	0.4244	0.4826	0.3394
4	0.9720	0.7729	0.4731	0.5418	0.3140
5	0.9714	0.7786	0.4908	0.6038	0.3041
6	0.9720	0.7822	0.4756	0.5786	0.2936
7	0.9732	0.7856	0.4967	0.5871	0.2820
8	0.9723	0.7767	0.5098	0.5538	0.2769
9	0.9731	0.7855	0.5099	0.5517	0.2676
10	0.9750	0.8238	0.5343	0.6286	0.2643
11	0.9753	0.8136	0.5291	0.6170	0.2596
12	0.9728	0.7865	0.5113	0.5341	0.2512
13	0.9741	0.8143	0.5211	0.5717	0.2466
14	0.9759	0.8281	0.5387	0.6222	0.2449
15	0.9751	0.8255	0.5438	0.5817	0.2428
16	0.9759	0.8223	0.5529	0.6134	0.2415
17	0.9747	0.8129	0.5246	0.5670	0.2352
18	0.9745	0.8314	0.5283	0.5771	0.2307
19	0.9760	0.8232	0.5441	0.6057	0.2319
20	0.9764	0.8264	0.5398	0.6114	0.2265
21	0.9761	0.8267	0.5492	0.6537	0.2255
22	0.9756	0.8286	0.5766	0.6107	0.2197
23	0.9765	0.8375	0.5472	0.6274	0.2190
24	0.9776	0.8489	0.5577	0.6589	0.2174
25	0.9765	0.8424	0.5546	0.6255	0.2152
26	0.9758	0.8218	0.4814	0.5672	0.2182
27	0.9763	0.8397	0.5371	0.6183	0.2106
28	0.9780	0.8503	0.5629	0.6655	0.2098
29	0.9773	0.8509	0.5844	0.6388	0.2096
30	0.9772	0.8499	0.5647	0.6310	0.2070
31	0.9765	0.8331	0.5567	0.6391	0.2041
32	0.9767	0.8325	0.5451	0.6306	0.2045
33	0.9755	0.8309	0.6055	0.5929	0.2038
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
34	0.9770	0.8392	0.5224	0.6207	0.2061
35	0.9772	0.8433	0.5585	0.6639	0.1983
36	0.9771	0.8383	0.5448	0.6162	0.1995
37	0.9778	0.8496	0.5682	0.6474	0.1956
38	0.9766	0.8415	0.5451	0.6188	0.1932
39	0.9782	0.8555	0.5524	0.6581	0.1944
40	0.9774	0.8458	0.5927	0.6400	0.1912
41	0.9767	0.8423	0.5399	0.6411	0.1909
42	0.9769	0.8419	0.5560	0.6735	0.1888
43	0.9777	0.8505	0.5581	0.6544	0.1981
44	0.9783	0.8530	0.5662	0.6805	0.1868
45	0.9770	0.8391	0.5718	0.6481	0.1891
46	0.9772	0.8391	0.5381	0.6229	0.1866
47	0.9779	0.8544	0.5527	0.6524	0.1831
48	0.9788	0.8585	0.5683	0.6810	0.1815
49	0.9772	0.8337	0.5874	0.6601	0.1916
50	0.9785	0.8591	0.5749	0.6682	0.1865

Table 4.2: Summary of validation scores and losses across 50 epochs for ReLU6

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
1	0.9655	0.7596	0.4126	0.5338	0.5299
2	0.9716	0.7973	0.4568	0.5598	0.3568
3	0.9728	0.8112	0.4832	0.5860	0.3217
4	0.9735	0.8052	0.4696	0.5839	0.3074
5	0.9747	0.8278	0.4916	0.6075	0.2906
6	0.9742	0.8278	0.5138	0.6431	0.2810
7	0.9758	0.8381	0.5378	0.6328	0.2713
8	0.9748	0.8364	0.5416	0.5577	0.2641
9	0.9744	0.8315	0.5271	0.6362	0.2593
10	0.9763	0.8418	0.5510	0.6314	0.2549
11	0.9718	0.7658	0.3548	0.4177	0.2472
12	0.9755	0.8389	0.5697	0.5790	0.2523
13	0.9783	0.8599	0.5647	0.6731	0.2415
14	0.9773	0.8575	0.5714	0.6386	0.2386
15	0.9760	0.8420	0.5604	0.6377	0.2377
16	0.9777	0.8549	0.5565	0.6611	0.2355
17	0.9774	0.8562	0.5603	0.6368	0.2310
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
18	0.9775	0.8505	0.5378	0.6325	0.2273
19	0.9781	0.8593	0.5758	0.6581	0.2276
20	0.9780	0.8537	0.5547	0.6430	0.2273
21	0.9782	0.8486	0.5696	0.6540	0.2196
22	0.9777	0.8478	0.5474	0.6234	0.2203
23	0.9778	0.8562	0.5713	0.6460	0.2173
24	0.9779	0.8523	0.5693	0.6337	0.2144
25	0.9767	0.8456	0.5322	0.5810	0.2122
26	0.9782	0.8563	0.5588	0.6369	0.2143
27	0.9779	0.8433	0.5524	0.6397	0.2092
28	0.9776	0.8564	0.5646	0.6159	0.2083
29	0.9779	0.8596	0.5393	0.6519	0.2038
30	0.9778	0.8643	0.5677	0.6422	0.2033
31	0.9785	0.8596	0.5687	0.6631	0.2065
32	0.9782	0.8629	0.5675	0.6502	0.2003
33	0.9780	0.8576	0.5560	0.6349	0.2055
34	0.9784	0.8592	0.5439	0.6467	0.1973
35	0.9761	0.8427	0.5586	0.6531	0.2000
36	0.9793	0.8711	0.5770	0.6563	0.1962
37	0.9776	0.8565	0.5748	0.6158	0.1926
38	0.9787	0.8654	0.5802	0.6512	0.1949
39	0.9779	0.8623	0.5490	0.6407	0.1920
40	0.9773	0.8598	0.5787	0.6089	0.1926
41	0.9784	0.8671	0.5764	0.6488	0.1875
42	0.9797	0.8736	0.5762	0.6810	0.1888
43	0.9784	0.8628	0.5791	0.6283	0.1914
44	0.9785	0.8613	0.5639	0.6448	0.1898
45	0.9791	0.8670	0.5651	0.6724	0.1806
46	0.9788	0.8610	0.5611	0.6634	0.1850
47	0.9791	0.8674	0.5594	0.6600	0.1832
48	0.9788	0.8630	0.5709	0.6496	0.1811
49	0.9788	0.8658	0.5527	0.6492	0.1840
50	0.9785	0.8626	0.5732	0.6325	0.1801

Table 4.3: Summary of validation scores and losses across 50 epochs for Leaky ReLU

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
1	0.9023	0.3974	0.4161	0.1271	0.5731
2	0.9667	0.7425	0.4079	0.4986	0.4398
3	0.9685	0.7767	0.4386	0.5867	0.3568
4	0.9704	0.7735	0.4545	0.5509	0.3306
5	0.9740	0.8037	0.4791	0.5864	0.3111
6	0.9727	0.7955	0.4756	0.5345	0.3041
7	0.9738	0.8155	0.4692	0.6097	0.2906
8	0.9718	0.8119	0.4854	0.5005	0.2867
9	0.9693	0.6962	0.4291	0.5412	0.2797
10	0.9725	0.8069	0.5207	0.6269	0.2756
11	0.9746	0.8185	0.5039	0.6112	0.2732
12	0.9752	0.8106	0.4973	0.6040	0.2663
13	0.9744	0.8141	0.4769	0.5452	0.2644
14	0.9747	0.8250	0.5173	0.5914	0.2585
15	0.9752	0.8261	0.5139	0.6249	0.2568
16	0.9747	0.7992	0.4945	0.5690	0.2508
17	0.9752	0.8233	0.5305	0.6136	0.2503
18	0.9752	0.8228	0.5191	0.6327	0.2467
19	0.9756	0.8317	0.5124	0.5800	0.2428
20	0.9747	0.8065	0.4781	0.5628	0.2393
21	0.9734	0.7757	0.5341	0.5772	0.2373
22	0.9765	0.8423	0.5182	0.6181	0.2399
23	0.9768	0.8401	0.5091	0.6079	0.2311
24	0.9770	0.8410	0.5208	0.6328	0.2340
25	0.9765	0.8360	0.5472	0.6327	0.2300
26	0.9776	0.8462	0.5488	0.6367	0.2294
27	0.9760	0.8266	0.5066	0.6103	0.2317
28	0.9776	0.8496	0.5152	0.6339	0.2221
29	0.9761	0.8307	0.5132	0.5965	0.2212
30	0.9770	0.8408	0.5241	0.6283	0.2200
31	0.9774	0.8492	0.5063	0.6171	0.2170
32	0.9766	0.8453	0.5240	0.6003	0.2152
33	0.9759	0.8311	0.4943	0.5946	0.2127
34	0.9781	0.8549	0.5411	0.6404	0.2152
35	0.9771	0.8558	0.5597	0.6402	0.2128
36	0.9750	0.8198	0.4777	0.5503	0.2134
37	0.9780	0.8508	0.5282	0.6580	0.2080

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
38	0.9765	0.8367	0.4963	0.5908	0.2085
39	0.9772	0.8443	0.5065	0.6249	0.2049
40	0.9765	0.8341	0.5018	0.5897	0.2059
41	0.9774	0.8454	0.5513	0.6372	0.2076
42	0.9789	0.8680	0.5592	0.6693	0.2048
43	0.9767	0.8465	0.5046	0.6140	0.2019
44	0.9784	0.8619	0.5652	0.6431	0.2001
45	0.9787	0.8646	0.5622	0.6472	0.1992
46	0.9781	0.8562	0.5481	0.6632	0.1945
47	0.9773	0.8510	0.5096	0.6090	0.1942
48	0.9791	0.8725	0.5635	0.6714	0.1976
49	0.9773	0.8503	0.5126	0.6162	0.1982
50	0.9790	0.8699	0.5819	0.6636	0.1945

Table 4.4: Summary of validation scores and losses across 50 epochs for PReLU

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
1	0.9425	0.4954	0.3447	0.2840	0.6606
2	0.9526	0.6663	0.4149	0.3588	0.5120
3	0.9606	0.6720	0.4484	0.5033	0.4578
4	0.9642	0.6737	0.4100	0.4526	0.4263
5	0.9651	0.7360	0.4464	0.4658	0.3995
6	0.9666	0.7219	0.4670	0.5350	0.3808
7	0.9675	0.7484	0.4772	0.5712	0.3727
8	0.9674	0.7424	0.4712	0.5503	0.3676
9	0.9687	0.7543	0.4821	0.5239	0.3594
10	0.9685	0.7539	0.4644	0.5683	0.3514
11	0.9688	0.7712	0.4870	0.5192	0.3465
12	0.9692	0.7756	0.4921	0.5694	0.3430
13	0.9699	0.7733	0.4979	0.5544	0.3397
14	0.9678	0.7641	0.4850	0.5638	0.3318
15	0.9694	0.7555	0.4441	0.5174	0.3312
16	0.9686	0.7450	0.4650	0.5113	0.3251
17	0.9706	0.7731	0.5054	0.5626	0.3245
18	0.9701	0.7774	0.5125	0.6010	0.3202
19	0.9705	0.7829	0.5100	0.5951	0.3190
20	0.9704	0.7731	0.4641	0.5258	0.3149
21	0.9714	0.7859	0.5116	0.5891	0.3163
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
22	0.9704	0.7894	0.5178	0.6020	0.3140
23	0.9711	0.7781	0.5155	0.5385	0.3103
24	0.9723	0.7886	0.5148	0.6036	0.3116
25	0.9711	0.7767	0.5123	0.5339	0.3056
26	0.9685	0.7859	0.5115	0.6023	0.3037
27	0.9717	0.7787	0.5198	0.5596	0.3099
28	0.9712	0.7850	0.5125	0.5825	0.3028
29	0.9719	0.7848	0.5092	0.5937	0.2977
30	0.9703	0.7850	0.5307	0.5924	0.3002
31	0.9710	0.7881	0.5141	0.5911	0.2987
32	0.9719	0.7991	0.5419	0.5992	0.2967
33	0.9693	0.7533	0.4494	0.4767	0.2906
34	0.9722	0.8017	0.5178	0.5747	0.2941
35	0.9703	0.7912	0.5284	0.5745	0.2897
36	0.9727	0.8014	0.5316	0.6071	0.2876
37	0.9729	0.8000	0.5400	0.6110	0.2881
38	0.9723	0.7661	0.5241	0.6062	0.2842
39	0.9720	0.7983	0.5381	0.6168	0.2813
40	0.9691	0.7723	0.5413	0.5986	0.2839
41	0.9724	0.7940	0.5150	0.5817	0.2817
42	0.9731	0.8065	0.5265	0.5839	0.2772
43	0.9735	0.8092	0.5450	0.6089	0.2794
44	0.9723	0.8025	0.5407	0.6108	0.2769
45	0.9727	0.7969	0.5494	0.6034	0.2759
46	0.9724	0.8114	0.5436	0.5828	0.2721
47	0.9726	0.8084	0.5371	0.5981	0.2746
48	0.9728	0.8078	0.5099	0.5655	0.2708
49	0.9721	0.8082	0.5272	0.5823	0.2659
50	0.9728	0.8045	0.5420	0.5732	0.2740

Table 4.5: Summary of validation scores and losses across 50 epochs for sigmoid

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
1	0.5271	0.1524	0.0269	0.0578	0.7474
2	0.5358	0.1726	0.0270	0.0589	0.6758
3	0.5778	0.1676	0.0312	0.0722	0.6471
4	0.5660	0.1828	0.0294	0.0752	0.6301
5	0.5948	0.1694	0.0298	0.0766	0.6166
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
6	0.6234	0.1858	0.0315	0.0828	0.5967
7	0.6375	0.2068	0.0350	0.0732	0.5830
8	0.6656	0.2177	0.0400	0.0915	0.5747
9	0.6399	0.2073	0.0358	0.0801	0.5607
10	0.6580	0.2069	0.0406	0.0892	0.5585
11	0.6760	0.1991	0.0367	0.0689	0.5428
12	0.6976	0.2228	0.0424	0.0950	0.5406
13	0.6389	0.2011	0.0375	0.0797	0.5335
14	0.6917	0.2385	0.0459	0.0942	0.5352
15	0.7001	0.2237	0.0407	0.0795	0.5224
16	0.7078	0.2284	0.0418	0.0975	0.5225
17	0.6364	0.2098	0.0371	0.0777	0.5181
18	0.7159	0.2596	0.0511	0.1139	0.5100
19	0.7129	0.2532	0.0519	0.1030	0.5140
20	0.7154	0.2501	0.0502	0.1076	0.5059
21	0.7317	0.2639	0.0531	0.1076	0.5035
22	0.7403	0.2586	0.0579	0.1214	0.5024
23	0.7177	0.2606	0.0549	0.1277	0.5001
24	0.7259	0.2495	0.0454	0.0884	0.4984
25	0.6581	0.2191	0.0429	0.0938	0.4990
26	0.7444	0.2841	0.0635	0.1228	0.4875
27	0.7396	0.2748	0.0552	0.1108	0.4857
28	0.7500	0.2903	0.0587	0.1098	0.4858
29	0.7536	0.2807	0.0654	0.1119	0.4835
30	0.7589	0.2990	0.0684	0.1437	0.4833
31	0.7522	0.2841	0.0618	0.1264	0.4815
32	0.7457	0.2703	0.0595	0.1143	0.4726
33	0.7679	0.2944	0.0635	0.1470	0.4781
34	0.7688	0.3020	0.0647	0.1357	0.4729
35	0.7439	0.2799	0.0624	0.1293	0.4659
36	0.7752	0.3000	0.0650	0.1362	0.4725
37	0.7673	0.2971	0.0656	0.1207	0.4647
38	0.7583	0.2805	0.0632	0.1132	0.4660
39	0.7955	0.3132	0.0768	0.1589	0.4638
40	0.7275	0.2559	0.0539	0.1196	0.4637
41	0.7683	0.2990	0.0721	0.1370	0.4591
42	0.8112	0.3398	0.0819	0.1587	0.4604
43	0.7991	0.3258	0.0770	0.1388	0.4497
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
44	0.8071	0.3541	0.0925	0.1818	0.4517
45	0.8550	0.4319	0.1356	0.2266	0.4493
46	0.8333	0.3622	0.0917	0.1728	0.4533
47	0.9016	0.6438	0.3667	0.4873	0.4039
48	0.9046	0.6624	0.3940	0.5151	0.3215
49	0.9028	0.6657	0.3888	0.4823	0.3103
50	0.9036	0.6578	0.3774	0.4815	0.2999

Table 4.6: Summary of validation scores and losses across 50 epochs for Tanh

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
1	0.5946	0.1878	0.4983	0.0440	0.5288
2	0.6473	0.2102	0.4715	0.0558	0.5053
3	0.6977	0.2357	0.4706	0.0594	0.4860
4	0.7803	0.3080	0.5136	0.0977	0.4681
5	0.7644	0.2822	0.4773	0.0928	0.4587
6	0.7331	0.2430	0.4878	0.0748	0.4507
7	0.8261	0.3181	0.4289	0.1028	0.4422
8	0.9145	0.5497	0.5055	0.2951	0.4317
9	0.9640	0.7701	0.4816	0.5541	0.3842
10	0.9679	0.7742	0.4880	0.5842	0.3371
11	0.9712	0.7953	0.4826	0.5159	0.3197
12	0.9702	0.7958	0.5007	0.5739	0.3073
13	0.9736	0.8110	0.4842	0.5928	0.2996
14	0.9723	0.7844	0.4778	0.5532	0.2917
15	0.9723	0.8015	0.4810	0.5959	0.2880
16	0.9727	0.8015	0.4634	0.5119	0.2788
17	0.9748	0.8144	0.5110	0.5942	0.2784
18	0.9749	0.8140	0.5232	0.6144	0.2732
19	0.9739	0.8180	0.5340	0.6130	0.2699
20	0.9734	0.7913	0.4579	0.5463	0.2652
21	0.9752	0.8190	0.5169	0.6081	0.2625
22	0.9753	0.8248	0.5079	0.6156	0.2649
23	0.9752	0.8282	0.5011	0.6072	0.2576
24	0.9720	0.8042	0.5256	0.6328	0.2548
25	0.9744	0.8160	0.5317	0.5857	0.2518
26	0.9752	0.8370	0.5037	0.6003	0.2476
27	0.9762	0.8300	0.5371	0.6162	0.2521
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
28	0.9763	0.8413	0.5480	0.6599	0.2442
29	0.9764	0.8385	0.5404	0.6412	0.2449
30	0.9768	0.8451	0.5554	0.6734	0.2385
31	0.9770	0.8458	0.5590	0.6639	0.2388
32	0.9764	0.8412	0.5123	0.6102	0.2390
33	0.9768	0.8529	0.5762	0.6358	0.2319
34	0.9775	0.8485	0.5627	0.6430	0.2360
35	0.9770	0.8486	0.5542	0.6169	0.2334
36	0.9778	0.8614	0.5934	0.6432	0.2305
37	0.9760	0.8325	0.5192	0.5990	0.2274
38	0.9776	0.8523	0.5328	0.6605	0.2289
39	0.9778	0.8541	0.5507	0.6550	0.2299
40	0.9728	0.8119	0.5549	0.6344	0.2251
41	0.9750	0.8239	0.4826	0.5618	0.2244
42	0.9773	0.8560	0.5732	0.6323	0.2222
43	0.9767	0.8336	0.5335	0.6321	0.2218
44	0.9768	0.8409	0.5089	0.6070	0.2247
45	0.9769	0.8506	0.5735	0.6604	0.2165
46	0.9782	0.8565	0.5612	0.6497	0.2175
47	0.9737	0.8168	0.5521	0.5915	0.2185
48	0.9775	0.8508	0.5668	0.6428	0.2168
49	0.9764	0.8441	0.5508	0.6328	0.2162
50	0.9765	0.8442	0.5590	0.6405	0.2121

Table 4.7: Summary of validation scores and losses across 50 epochs for ELU

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
1	0.5364	0.1274	5.3497	6.0000	0.6967
2	0.5527	0.1258	5.7698	6.0000	0.6037
3	0.5707	0.1088	2.5284	6.0000	0.5782
4	0.6011	0.1466	4.4167	6.0000	0.5606
5	0.6217	0.1603	1.0179	5.0000	0.5443
6	0.6299	0.1468	1.0365	5.0000	0.5393
7	0.6283	0.0808	0.0000	0.0252	0.5308
8	0.6431	0.1730	1.2351	5.0000	0.5265
9	0.6574	0.1500	3.1889	6.0000	0.5215
10	0.6664	0.1743	1.2999	5.0000	0.5135
11	0.6827	0.1701	7.5004	6.0000	0.5079
					Continued on next page

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
12	0.7002	0.1678	4.4995	6.0000	0.4995
13	0.7063	0.1715	6.2151	6.0000	0.4960
14	0.7085	0.1454	3.6610	6.0000	0.4894
15	0.7214	0.1839	9.1233	7.0000	0.4872
16	0.7254	0.1808	1.0737	6.0000	0.4835
17	0.7389	0.2077	9.4344	7.0000	0.4786
18	0.7336	0.1901	2.1168	6.0000	0.4754
19	0.7416	0.2310	3.6026	6.0000	0.4730
20	0.7335	0.1888	1.7911	6.0000	0.4770
21	0.7532	0.2306	1.0018	6.0000	0.4690
22	0.7520	0.1936	3.6474	7.0000	0.4660
23	0.7520	0.2102	3.4696	6.0000	0.4650
24	0.7597	0.2222	9.8300	7.0000	0.4650
25	0.7579	0.2143	6.3981	7.0000	0.4613
26	0.7564	0.2280	6.9566	7.0000	0.4578
27	0.7731	0.2264	2.6705	7.0000	0.4567
28	0.7740	0.2212	7.3492	8.0000	0.4595
29	0.8886	0.3475	0.0000	0.0928	0.4440
30	0.8997	0.3877	1.6105	7.0000	0.4367
31	0.9003	0.3903	9.4227	7.0000	0.4357
32	0.9088	0.4384	1.8258	5.0000	0.4339
33	0.9119	0.4166	2.0380	6.0000	0.4325
34	0.9178	0.4167	1.9532	7.0000	0.4269
35	0.9150	0.4249	4.7730	7.0000	0.4260
36	0.9097	0.4273	1.5409	6.0000	0.4277
37	0.9176	0.4500	6.2719	6.0000	0.4232
38	0.9189	0.4269	2.8778	7.0000	0.4219
39	0.9184	0.4622	1.3917	5.0000	0.4220
40	0.9183	0.4471	1.0924	5.0000	0.4208
41	0.9187	0.4285	2.8632	7.0000	0.4180
42	0.9193	0.4446	3.8620	5.0000	0.4174
43	0.9189	0.4630	2.3130	5.0000	0.4160
44	0.9207	0.4576	2.8681	7.0000	0.4150
45	0.9232	0.4857	3.0398	5.0000	0.4155
46	0.9233	0.4733	9.4524	7.0000	0.4141
47	0.9256	0.4533	3.2175	7.0000	0.4129
48	0.9254	0.4606	2.6053	7.0000	0.4103
49	0.9245	0.4209	4.9589	7.0000	0.4109
Continued on next page					

Epoch	Mean Dice	Whole Tumor	Enhancing	Core	Loss
50	0.9257	0.4630	1.9276	6.0000	0.4111

Table 4.8: Summary of validation scores and losses across 50 epochs for CELU
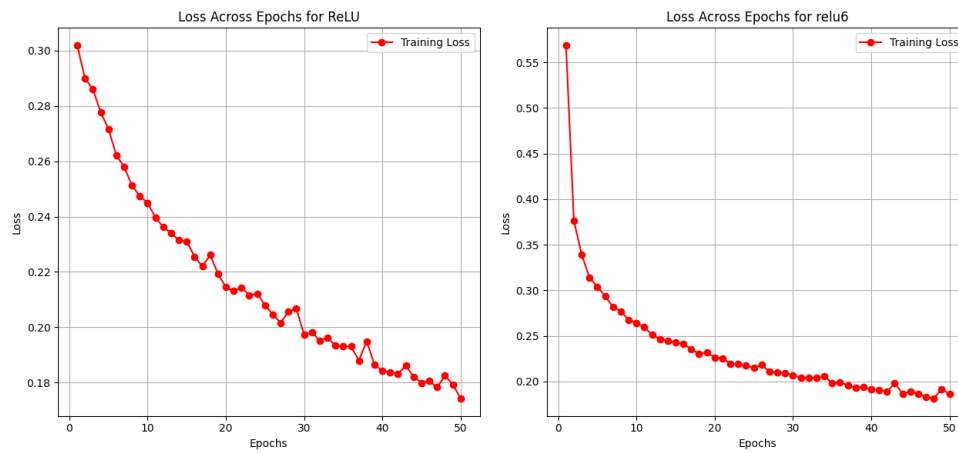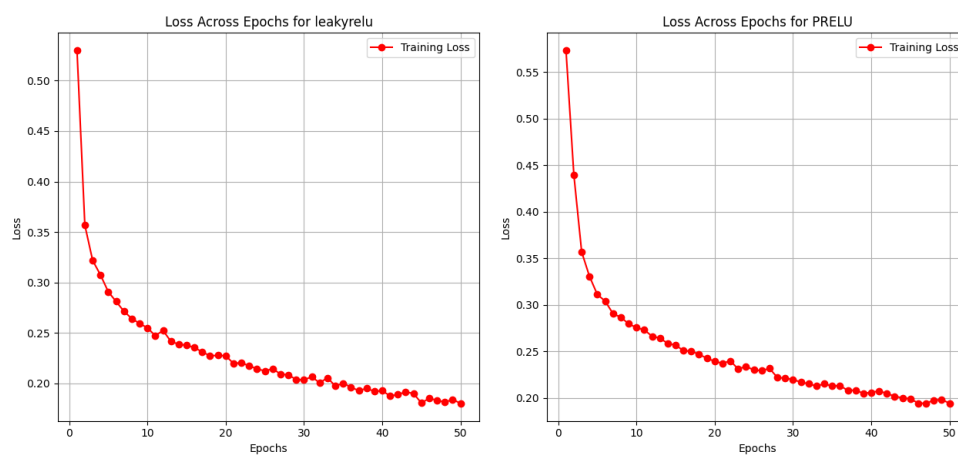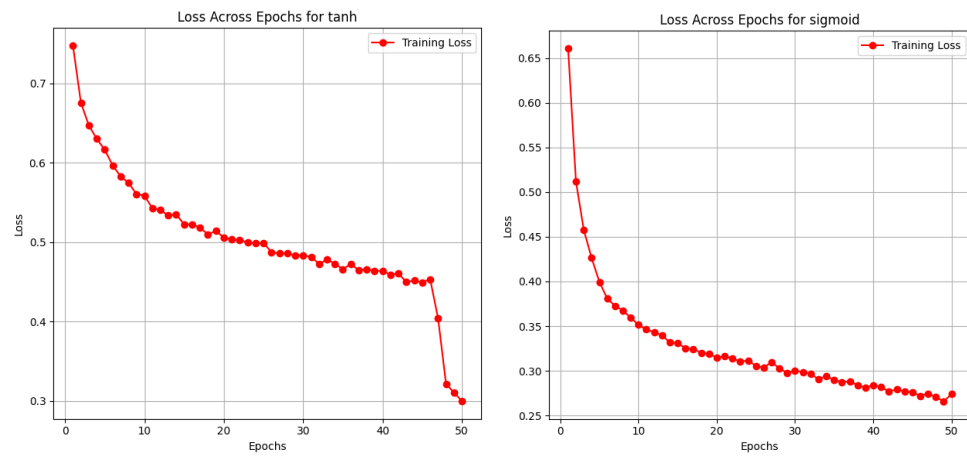

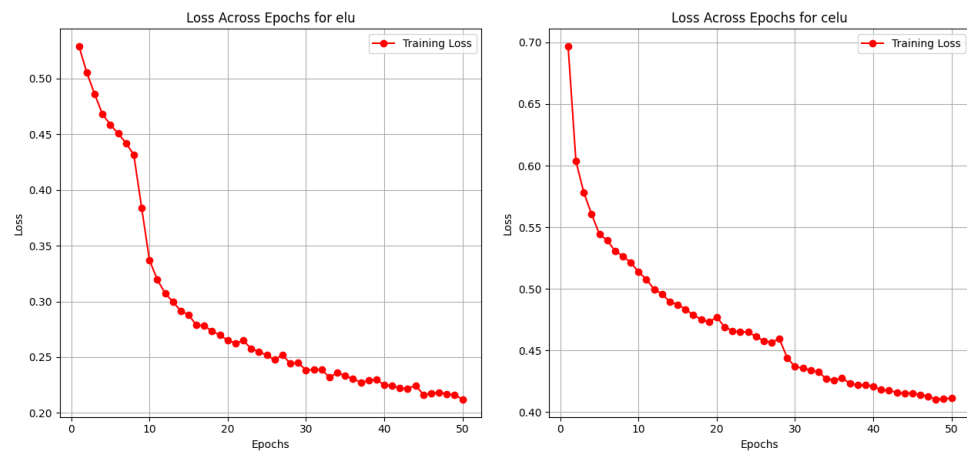Figure 4.29

Figure 4.30

Figure 4.31

Figure 4.32

Chapter 5

Conclusion

The advancements in brain tumor segmentation utilizing U-Net architectures underscore the significant strides made in medical image analysis. The intricate nature of brain tumors, combined with the vital role of accurate segmentation in diagnosis and treatment planning, emphasizes the necessity for reliable, efficient, and precise automated tools. This study's focus on optimizing activation functions within the U-Net framework highlights a key area of improvement in medical image segmentation, particularly for brain tumors.

5.1   Summary of Findings

This thesis presents a comprehensive evaluation of various activation functions within the U-Net architecture for the task of brain tumor segmentation. By leveraging the BraTS2020 dataset, which includes multimodal MRI scans and expert-annotated segmentation masks, the study aimed to address key challenges in achieving precise segmentation, such as class imbalance, imaging inconsistencies, and tumor boundary delineation.

The results indicate that the choice of activation function has a profound impact on the performance of U-Net models in terms of segmentation accuracy, computational efficiency, and stability during training. Among the tested activation functions, ReLU consistently outperformed its counterparts across all evaluation metrics, demonstrating the highest Dice similarity scores and the most stable training loss trends. ReLU's ability to effectively mitigate the vanishing gradient problem and focus computational resources on significant features made it particularly well-suited for this application. ELU and Sigmoid also showed competitive performance, particularly in capturing finer tumor details, but were slightly less efficient compared to ReLU.

In conclusion, based on the experimental parameters and results obtained, ReLU emerged as the best-performing activation function for brain tumor segmentation. It consistently achieved

the highest Dice scores across the core, enhancing, and whole tumor regions, demonstrating superior accuracy and reliability. These results highlight ReLU's ability to effectively handle the complexities of the segmentation task within the controlled experimental setup, making it the most suitable choice among the activation functions tested in this study.

5.2   Contributions to the Field

The findings of this study contribute to the growing body of knowledge on U-Net adaptations for medical imaging, emphasizing the pivotal role of activation functions in achieving precise segmentation. The research addressed several challenges inherent in brain tumor segmentation, such as class imbalance, computational demands, and variability in MRI data. By integrating optimized activation functions, the study showcased improved segmentation accuracy and model robustness. These advancements hold significant implications for clinical applications, where reliable and efficient segmentation tools can enhance diagnostic accuracy and streamline treatment planning.

5.3   Practical Implications

The implementation of optimized activation functions within U-Net architectures offers several practical benefits for clinical use. Improved boundary detection and segmentation accuracy can lead to more precise tumor localization, enabling tailored treatment strategies that minimize damage to healthy brain tissue.

The integration of computationally efficient activation functions also addresses the challenge of processing high-dimensional MRI data in real-time. By balancing segmentation precision with computational demands, these advancements pave the way for the development of lightweight, deployable models suitable for resource-constrained environments. This is particularly critical in regions with limited access to advanced medical imaging facilities, where automated tools can bridge the gap in diagnostic capabilities.

5.4   Limitations

While this study provides valuable insights into the impact of activation functions on brain tumor segmentation, it is important to acknowledge certain limitations. A key aspect of machine learning experimentation is running multiple training trials for each configuration to account for variability introduced by factors such as random weight initialization, data shuffling, and stochastic optimization processes. By repeating the experiments and reporting metrics such as the mean and standard deviation (e.g., Dice coefficient ± standard deviation), a more robust and reliable evaluation of model performance can be achieved.

However, due to the sheer number of activation functions tested in this study (eight in total) and the limited time available, conducting multiple trials for each activation function was not feasible. As a result, the reported results are based on single training runs for each activation function, which may not fully capture the variability inherent in the training process. Future work could address this limitation by performing multiple trials to ensure more statistically robust conclusions.

Chapter 6

Future Work

While this study has advanced our understanding of the role of activation functions in brain tumor segmentation, several avenues for future research could further enhance the effectiveness and applicability of U-Net models in medical imaging. The study primarily focused on established activation functions such as ReLU, Sigmoid, and ELU. Future work could explore newer functions like Swish, Mish, and E-Tanh, which have shown promise in retaining finer details and providing smoother gradient flow in other domains. Investigating their performance in U-Net architectures, particularly for brain MRI segmentation, could yield insights into achieving even higher segmentation accuracy.

In real-world clinical settings, complete multimodal MRI datasets are not always available. Models like SMU-Net [18] have introduced techniques to handle missing modalities, but their integration with alternative activation functions remains underexplored. Research could focus on developing robust methods to reconstruct missing modalities while optimizing activation functions for consistent segmentation accuracy.

Given the high computational demands of 3D U-Net models, optimizing activation functions for faster convergence and lower resource consumption is a critical area for future research. Functions like Swish and PReLU, known for their adaptability and smooth gradient flow, could be explored in this context.

While the experimental results provide a strong foundation, clinical validation on larger and more diverse datasets is essential to assess the generalizability and reliability of the proposed methods. Collaborating with medical professionals to evaluate model performance in real-world diagnostic settings would provide valuable insights.

As medical AI applications become more widespread, ensuring ethical use and minimizing bias in segmentation models is paramount. Future studies should focus on developing transparent and interpretable models that perform equitably across diverse patient demographics.

By addressing these areas, future research can build on the findings of this study to develop more robust, efficient, and clinically applicable U-Net models for brain tumor segmentation and beyond.

Appendix

Bibliography

[1] Rammah Yousef, Shakir Khan, Gaurav Gupta, Tamanna Siddiqui, Bader M. Albahlal, Saad Abdullah Alajlan, and Mohd Anul Haq. U-net-based models towards optimal mr brain image segmentation. *Diagnostics*, 13(9), 2023.

[2] Kaggle. BraTS 2020 Dataset - Training and Validation, 2020. Accessed: 2024-09-20.

[3] Sidratul Montaha, Sami Azam, A.K.M. Rakibul Haque Rafid, Md Hasan, and Asif Karim. Brain tumor segmentation from 3d mri scans using u-net. *SN Computer Science*, 4, 05 2023.

[4] Smarta Sangui, Tamim Iqbal, Piyush Chandra Chandra, Swarup Kr Ghosh, and Anupam Ghosh. 3d mri segmentation using u-net architecture for the detection of brain tumor. *Procedia Computer Science*, 218:542–553, 2023. International Conference on Machine Learning and Data Engineering.

[5] Sagar Sharma, Simone Sharma, and Anidhya Athaiya. Activation functions in neural networks. *Towards Data Sci*, 6(12):310–316, 2017.

[6] Shradha Verma, Anuradha Chug, Amit Singh, and Dinesh Singh. Pds-mcnet: a hybrid framework using mobilenetv2 with silu6 activation function and capsule networks for disease severity estimation in plants. *Neural Computing and Applications*, 35:1–24, 06 2023.

[7] paperswithcode. Celu (continuously differentiable exponential linear unit). `https://paperswithcode.com/method/celu`, 2023. Accessed: 2025-01-15.

[8] S T Padmapriya, T Chandrakumar, and T Kalaiselvi. Improving the prediction accuracy of mri brain tumor detection and segmentation. *International Journal of Computing and Digital Systems*, 15(1):1–10, 2024.

[9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[10] Nahian Siddique, Sidike Paheding, Colin P. Elkin, and Vijay Devabhaktuni. U-net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*, 9:82031–82057, 2021.

[11] Atika Akter, Nazeela Nosheen, Sabbir Ahmed, Mariom Hossain, Mohammad Abu Yousuf, Mohammad Ali Abdullah Almoyad, Khondokar Fida Hasan, and Mohammad Ali Moni. Robust clinical applicable cnn and u-net based algorithm for mri classification and segmentation for brain tumor. *Expert Systems with Applications*, 238:122347, 2024.

[12] Ilyasse Aboussaleh, Jamal Riffi, Khalid El Fazazy, Mohamed Adnane Mahraz, and Hamid Tairi. Efficient u-net architecture with multiple encoders and attention mechanism decoders for brain tumor segmentation. *Diagnostics*, 13(5), 2023.

[13] Abdul Haseeb Nizamani, Zhigang Chen, Ahsan Ahmed Nizamani, and Uzair Aslam Bhatti. Advance brain tumor segmentation using feature fusion methods with deep u-net model with cnn for mri data. *Journal of King Saud University - Computer and Information Sciences*, 35(9):101793, 2023.

[14] Jianxin Zhang, Xiaogang Lv, Hengbo Zhang, and Bin Liu. Aresu-net: Attention residual u-net for brain tumor segmentation. *Symmetry*, 12(5), 2020.

[15] Jakhongir Nodirov, Akmalbek Bobomirzaevich Abdusalomov, and Taeg Keun Whangbo. Attention 3d u-net with multiple skip connections for segmentation of brain tumor images. *Sensors*, 22(17), 2022.

[16] Abdul Haseeb Nizamani, Zhigang Chen, Ahsan Ahmed Nizamani, and Uzair Aslam Bhatti. Advance brain tumor segmentation using feature fusion methods with deep u-net model with cnn for mri data. *Journal of King Saud University - Computer and Information Sciences*, 35(9):101793, 2023.

[17] Nagwa M. Aboelenein, Piao Songhao, Anis Koubaa, Alam Noor, and Ahmed Afifi. Httu-net: Hybrid two track u-net for automatic brain tumor segmentation. *IEEE Access*, 8:101406–101415, 2020.

[18] Reza Azad, Nika Khosravi, and Dorit Merhof. Smu-net: Style matching u-net for brain tumor segmentation with missing modalities. In Ender Konukoglu, Bjoern Menze, Archana Venkataraman, Christian Baumgartner, Qi Dou, and Shadi Albarqouni, editors, *Proceedings of The 5th International Conference on Medical Imaging with Deep Learning*, volume 172 of *Proceedings of Machine Learning Research*, pages 48–62. PMLR, 06–08 Jul 2022.

[19] Shuyue Li, Jia Liu, and Zhen Song. Brain tumor segmentation based on region of interest-aided localization and segmentation u-net. *International Journal of Machine Learning and Cybernetics*, 13:2435–2445, September 2022.

[20] Mobeen Ur Rehman, SeungBin Cho, Jee Hong Kim, and Kil To Chong. Bu-net: Brain tumor segmentation using modified u-net architecture. *Electronics*, 9(12), 2020.

[21] Muhammad Usman Saeed, Ghulam Ali, Wang Bin, Sultan H. Almotiri, Mohammed A. AlGhamdi, Arfan Ali Nagra, Khalid Masood, and Riaz ul Amin. Rmu-net: A novel residual mobile u-net model for brain tumor segmentation from mr images. *Electronics*, 10(16), 2021.

[22] M. Jorge Cardoso, Wenqi Li, Richard Brown, Nic Ma, Eric Kerfoot, Yiheng Wang, Benjamin Murray, Andriy Myronenko, Can Zhao, Dong Yang, Vishwesh Nath, Yufan He, Ziyue Xu, Ali Hatamizadeh, Wentao Zhu, Yun Liu, Mingxin Zheng, Yucheng Tang, Isaac Yang, Michael Zephyr, Behrooz Hashemian, Sachidanand Alle, Mohammad Zalbagi Darestani, Charlie Budd, Marc Modat, Tom Vercauteren, Guotai Wang, Yiwen Li, Yipeng Hu, Yunguan Fu, Benjamin Gorman, Hans Johnson, Brad Genereaux, Barbaros S. Erdal, Vikash Gupta, Andres Diaz-Pinto, Andre Dourson, Lena Maier-Hein, Paul F. Jaeger, Michael Baumgartner, Jayashree Kalpathy-Cramer, Mona Flores, Justin Kirby, Lee A.D. Cooper, Holger R. Roth, Daguang Xu, David Bericat, Ralf Floca, S. Kevin Zhou, Haris Shuaib, Keyvan Farahani, Klaus H. Maier-Hein, Stephen Aylward, Prerna Dogra, Sebastien Ourselin, and Andrew Feng. MONAI: An open-source framework for deep learning in healthcare, November 2022.