Deep Learning for Brain Tumor Segmentation and Survival Prediction

Author
Tirivangani Batanai Hendrix Takura Magadza
219098526

Supervisor
Prof. Serestina Viriri

A thesis submitted in fulfillment of the requirement for the degree of
Doctor of Philosophy in Computer Science

School of Mathematics, Statistics and Computer Science
University of KwaZulu-Natal
South Africa

Declaration of Authorship

I, Tirivangani Batanai Handrix Takura MAGADZA, declare that this thesis titled, "Deep learning for Brain Tumor Segmentation" and the work presented in it are my own. I declare that:

1. The research reported in this thesis, except where otherwise indicated or acknowledged, is my original work;

2. This thesis has not been submitted in full or in part for any degree or examination to any other university;

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

   (a) Their words have been re-written but the general information attributed to them has been referenced,

   (b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references sections.

Candidate:
Tirivangani Batanai Handrix Takura MAGADZA

Signed: *TBHT Magadza*

Date: 05/02/2024

As the candidate's supervisor I approve the submission of this thesis for examination.

Supervisor:
Prof. Serestina VIRIRI

Signed: [redacted]

Date: 07/06/2024

List of Publications

I, Tirivangani Batanai Hendrix Takura MAGADZA, declare that the following are publications from this thesis:

1. Magadza T, Viriri S. Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art. Journal of Imaging. 2021; 7(2):19. https://doi.org/10.3390/jimaging7020019.

2. T. Magadza and S. Viriri, "Brain Tumor Segmentation Using Partial Depthwise Separable Convolutions," IEEE Access, vol. 10, pp. 124206-124216, 2022, doi: 10.1109/ACCESS.2022.3223654.

3. T. Magadza and S. Viriri, "Efficient nnU-Net for Brain Tumor Segmentation," IEEE Access, vol. 11, pp. 126386-126397, 2023, doi: 10.1109/ACCESS.2023.3329517.

4. Magadza T, Viriri S. Efficient Survival Prediction for Gliomas using deep learning. Scientific Reports. 2024. (Under review)

Abstract

A brain tumor is an abnormal growth of cells in the brain that multiplies uncontrolled. The death of people due to brain tumors has increased over the past few decades. Early diagnosis of brain tumors is essential in improving treatment possibilities and increasing the survival rate of patients. The life expectancy of patients with glioblastoma multiforme (GBM), the most malignant glioma, using the current standard of care is, on average, 14 months after diagnosis despite aggressive surgery, radiation, and chemotherapies. Despite considerable efforts in brain tumor segmentation research, patient diagnosis remains poor. Accurate segmentation of pathological regions may significantly impact treatment decisions, planning, and outcome monitoring. However, the large spatial and structural variability among brain tumors makes automatic segmentation a challenging problem, leaving brain tumor segmentation an open challenge that warrants further research endeavors.

While several methods automatically segment brain tumors, deep learning methods are becoming widespread in medical imaging due to their resounding performance. However, the boost in performance comes at the cost of high computational complexity. Therefore, to improve the adoption rate of computer-assisted diagnosis in clinical setups, especially in developing countries, there is a need for more computational and memory-efficient models.

In this research, using a few computational resources, we explore various techniques to develop deep learning models accurately for segmenting the different glioma sub-regions, namely the enhancing tumor, the tumor core, and the whole tumor. We quantitatively evaluate the performance of our proposed models against the state-of-the-art methods using magnetic resolution imaging (MRI) datasets provided by the Brain Tumor Segmentation (BraTS) Challenge.

Lastly, we use segmentation labels produced by the segmentation task and MRI multimodal data to extract appropriate imaging/radiomic features to train a deep learning model for overall patient survival prediction.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Prof S. Viriri, for his unwavering mentorship, guidance, and encouragement throughout the journey of completing this thesis. His insightful feedback and constructive criticism have been invaluable in shaping the development of my research.

I am profoundly thankful to my wife, Addlight Mukwazvure, for her constant support and encouragement during the highs and lows of this academic endeavor. Her understanding, patience, and belief in my capabilities have been a driving force in overcoming challenges and achieving milestones.

I extend my heartfelt thanks to my family and friends for their steadfast belief in me. Their encouragement, motivation, and understanding have played a pivotal role in keeping me focused and motivated throughout this academic pursuit.

A special appreciation goes to the Centre of High-Performance Computing for providing reliable computational resources. The access to such resources significantly contributed to the successful execution of complex simulations and analyses essential to this research.

This thesis would not have been possible without the collective support and encouragement from these individuals and institutions. Thank you for being an integral part of this significant academic achievement.

Dedication

This research work is dedicated to Almighty God for His grace, mercy and strength.

Contents

List of Figures

List of Tables

Acronym List

MRI	Magnetic resonance imaging
GBM	Glioblastoma multiforme
WHO	World Health Organization
BraTS	Brain Tumor Segmentation
OS	Overall survival
CNN	Convolutional Neural Networks
FLAIR	Fluid-attenuated inversion recovery
NCR	Necrotic core
ED	Peritumoral edema
ET	Enhancing tumor
NET	Non-enhancing tumor
WT	Whole tumor
TC	Tumor core
GTR	Gross tumor resection
PET	Positron emission tomography
CPU	Central processing unit
GPU	Graphics processing unit
3D	Three-Dimensional
2D	Two-Dimensional
FC	Fully connected
FCN	Fully convolutional network
SDG	Stochastic gradient descent
MSE	Mean squared error
medianSE	Median squared error
stdSE	Standard deviation of the square error
SpearmanR	Spearman rank correlation

Chapter 1

Introduction

1.1 Introduction

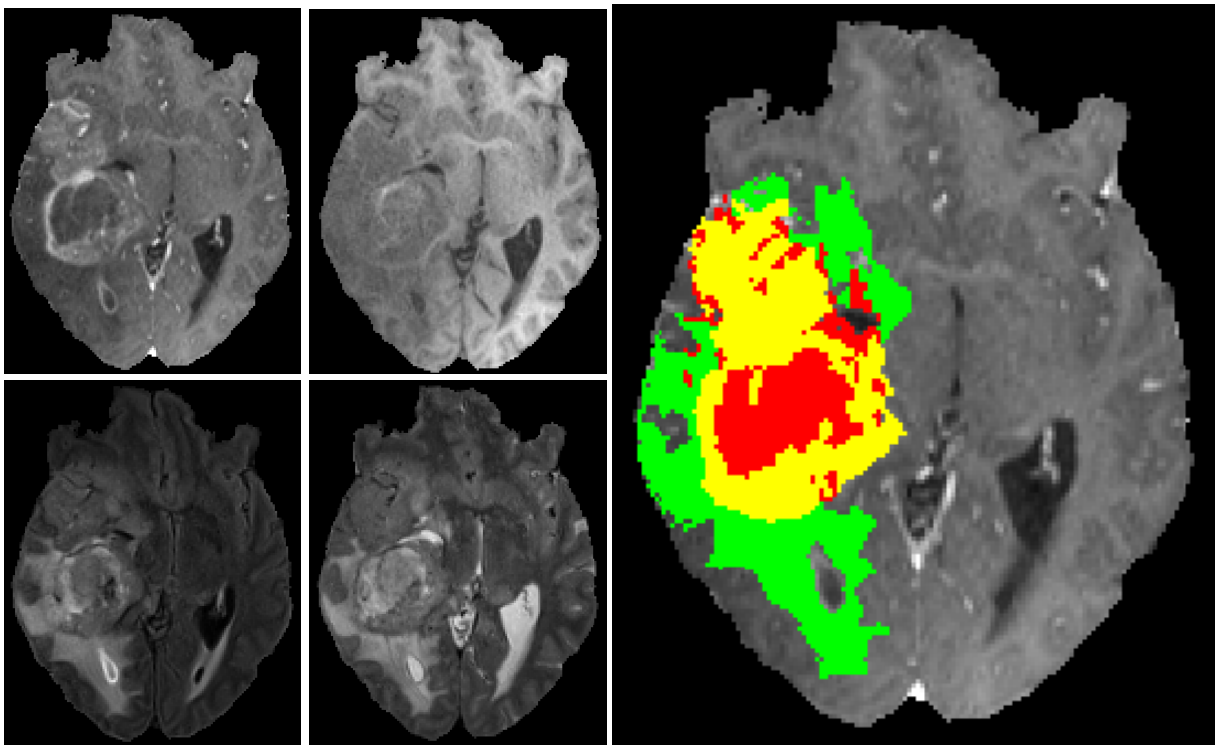
Brain tumors are an abnormal growth of cells in the brain that multiplies in an uncontrolled way. Although their causes are unknown [7], brain tumors have been one of the leading causes of death in the past few decades. Early diagnosis of brain tumors plays an important role in improving treatment possibilities and increasing the survival rate of the patients [19]. There are two main types of tumors: malignant (cancerous) and benign. (noncancerous) tumors. Cancerous tumors can begin in the brain (primary brain tumors) or can start in other parts of your body and spread to your brain (secondary brain tumors). Brain tumors can appear anywhere in the brain with varying sizes and shapes, which makes them challenging to delineate. Gliomas, the most common brain tumors in adults [28], have the highest mortality rate and occurrence among brain tumors. The median survival rate is less than two years, even with aggressive therapy.

Magnetic Resonance Imaging (MRI) is usually the modality of choice used to inform diagnosis and treatment of brain tumors [24] due to its high spatial resolution, soft tissue contrast, and non-invasive characteristics. MRI provides rich information for brain tumor diagnosis and treatment planning. More than one MRI slice is required to view different brain regions, e.g., T1, T2, T1 contrast, and FLAIR images. Figure 1.1 shows sample image data from the BraTS dataset [24, 3, 4].

Generally, a healthy brain comprises three essential components: gray matter, white matter, and cerebrospinal fluid. Brain tumor segmentation aims to identify the location and extent of different tumor regions, including active tumorous tissue, necrotic (dead) tissue, and edema (swelling near tumors). This process involves distinguishing abnormal areas in comparison to normal tissues.

Creating imaging features that facilitate accurate segmentation is a complex task that demands meticulous engineering and specialized expertise. Manual segmentation of brain tumors, though traditionally employed, is characterized by subjectivity, time-intensive procedures, and high costs. Classical machine learning algorithms often utilize handcrafted features for segmentation, introducing certain limitations.

Integrating deep learning techniques in image segmentation offers a promising alternative to overcome the drawbacks associated with classical machine learning. Deep learning's capacity for self-learning of features may address the limitations of handcrafted features, potentially enabling the identification of novel and valuable imaging features for quantitative analysis in brain MRI. This shift towards deep learning can potentially enhance the efficiency and accuracy of brain tumor segmentation processes.



(a) (b)

Figure 1.1: Examples of BraTS multimodal scans and Glioma sub-regions. (a) MRI Slices: T1-weighted (top left), native T1 (top right), FLAIR (bottom left), T2 (bottom right), (b) Segmented regions: edema (green), necrotic (NCR) and the non-enhancing (NET) tumor core (red), enhancing tumor (yellow)

1.2 Motivation

The typical lifespan of individuals diagnosed with glioblastoma multiforme (GBM), recognized as the most malignant glioma with a World Health Organization (WHO) grade IV classification, stands at an average of 14 months following diagnosis, despite undergoing aggressive surgery, radiation, and chemotherapy [31]. Despite extensive research on brain tumor segmentation, the ability to diagnose patients has not improved significantly [24]. The precise delineation of pathological regions is essential for treatment decisions, planning, and outcome monitoring [27]. However, the inherent spatial and structural variations among brain tumors pose a formidable obstacle to automatic segmentation [27], rendering brain tumor segmentation an ongoing challenge that necessitates further research efforts.

1.3 Problem Statement

Diagnosing gliomas typically involves identifying two central pathological regions: the tumor core (enhancing, non-enhancing, and necrotic parts) and the peritumoral edematous area (infiltrating cells). Accurate segmenting these regions and predicting overall patient survival pre- and post-treatment are crucial for informing patient diagnoses. Manual delineation of tumor boundaries is labor-intensive and susceptible to human error and observer bias [35].

However, automatic segmentation of brain tumors poses challenges due to variations in tumor location and size among individuals. In some cases, tumors infiltrate surrounding normal tissues, further complicating delineation [24]. Deep learning techniques, while powerful, demand substantial annotated data, presenting a challenge in the medical domain. Additionally, their computational requirements limit adoption in clinical practices, particularly in developing countries.

Notably, no single algorithm universally outperforms others in simultaneously segmenting all tumor regions [24][4]. Many research efforts solely concentrate on the segmentation task, overlooking the clinical relevance of predicting patient survival. While the performance is still low, classical machine learning techniques dominate existing works on survival prediction, which require robust feature extraction and selection mechanisms.

Due to limited datasets, applying deep learning techniques in survival prediction has been an open challenge. The automatic extraction of numerous features by deep learning methods makes them prone to overfitting, significantly impacting their performance. This research aims to explore the application of deep learning techniques in segmentation and prediction tasks, specifically emphasizing utilizing limited computational resources. The objective is to address the challenges associated with accurate glioma diagnosis and prognosis more resourcefully.

1.4 Aims and Objectives

This research aims to devise and evaluate computationally efficient deep-learning models for automatic brain tumor segmentation and overall survival prediction. The main objectives of this research are to:

1. Conduct a critical analysis of the state-of-the-art automatic brain tumor segmentation techniques.

2. Develop and evaluate computationally efficient deep learning models for automatic brain tumor segmentation.

3. Design and evaluate a lightweight deep learning model for overall survival prediction.

1.5 Methodology

The methodology in this work will be divided into two tasks: segmentation of gliomas and prediction of patient overall survival (OS) from pre-operative scans. The segmentation task will be addressed by using the BraTS Dataset[24] to develop a deep-learning method for segmentation of the different glioma sub-regions, namely the "enhancing tumor," the "tumor core," and the "whole tumor."

Segmentation labels produced by the segmentation task will be combined with the multimodal MRI data provided by the BraTS Dataset to extract appropriate imaging/radiomic features that will be used to train machine learning algorithms for overall patient survival prediction. Predictions will be in the following classes: short survivors (< 10 months), mid-survivors, and long survivors (> 15 months).

1.6 Thesis Contributions

The main contribution is to develop deep learning techniques for accurate diagnosis and prognosis of gliomas in a more resource-efficient manner. The following are the contributions of this thesis to the field of medical image analysis:

1. Chapter 2 presents a comprehensive study of state-of-the-art automatic brain tumor segmentation techniques. An extensive overview of the fundamental building blocks, state-of-the-art techniques, and tools integral to implementing automated brain tumor segmentation algorithms are well presented. Unique challenges and their possible solutions to medical image analysis are also discussed.

2. Chapter 3 addresses the need for efficient yet effective 3D brain tumor segmentation solutions. A computationally efficient 3D brain tumor segmentation network architecture incorporating depthwise separable convolutions to alleviate computational burdens is proposed. Furthermore, the chapter presents several modifications to the nnU-Net framework to reduce computational complexity while maintaining competitive segmentation performance. Again, a computational analysis of recent works for automatic brain tumor segmentation was discussed.

3. Chapter 4 proposes an end-to-end deep learning framework that utilizes adaptive feature selection mechanisms for overall survival prediction. An extensive analysis of several feature fusion strategies is presented.

1.7 Datasets Sources

The datasets used in this research came from the Brain Tumor Segmentation Challenge (BraTS) [3, 4, 24] for the years 2020 and 2021. The datasets are open-source and pub-

licly available. Furthermore, the challenge provides an online evaluation platform to standardize the performance evaluation of proposed techniques.

1.8 Thesis Outline

The rest of this thesis is arranged as follows: Chapter 2 presents an extensive study of the state-of-the-art deep-learning techniques of automatic brain tumor segmentation. Unique challenges and their possible solutions to medical image analysis are also discussed. In Chapter 3, we propose and evaluate computationally efficient deep-learning models for automatic brain tumor segmentation. We also presented a computational study of recent works for automatic brain tumor segmentation. Chapter 4 proposes and evaluates an end-to-end deep learning pipeline for survival prediction. A quantitative analysis of the performance of the proposed method against state-of-the-art was presented. In Chapter 5, we Harmonize the findings derived from the various papers within this thesis. Lastly, in Chapter 6, we concluded this thesis with a discussion of presented methods and recommendations for further research.

Chapter 2

Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art

2.1 Introduction

In recent years, the field of medical image analysis has witnessed significant advancements, with deep learning techniques emerging as powerful tools for automatic segmentation of brain tumors. This chapter provides a comprehensive literature review, offering insights into the state-of-the-art methods employed in the realm of automatic brain tumor segmentation. By elucidating the key building blocks, datasets, tools, and techniques that form the foundation of contemporary research, this review aims to distill the wealth of knowledge in the field. Through a concise exploration of these pivotal aspects, the groundwork is laid for a thorough understanding of the landscape surrounding automatic brain tumor segmentation, setting the stage for the subsequent chapters of this thesis.

Part of this work was published in [1]

[1] Magadza T, Viriri S. Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art. Journal of Imaging. 2021; 7(2):19. https://doi.org/10.3390/jimaging7020019

Journal of *Imaging*

MDPI

Review

Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art

Tirivangani Magadza and Serestina Viriri *

School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban 4000, South Africa; 219098526@stu.ukzn.ac.za

* Correspondence: viriris@ukzn.ac.za

Abstract: Quantitative analysis of the brain tumors provides valuable information for understanding the tumor characteristics and treatment planning better. The accurate segmentation of lesions requires more than one image modalities with varying contrasts. As a result, manual segmentation, which is arguably the most accurate segmentation method, would be impractical for more extensive studies. Deep learning has recently emerged as a solution for quantitative analysis due to its record-shattering performance. However, medical image analysis has its unique challenges. This paper presents a review of state-of-the-art deep learning methods for brain tumor segmentation, clearly highlighting their building blocks and various strategies. We end with a critical discussion of open challenges in medical image analysis.

Keywords: brain tumor segmentation; deep learning; magnetic resonance imaging; survey

Citation: Magadza, T.; Viriri, S. Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art. *J. Imaging* **2021**, *7*, 19. https://doi.org/10.3390/jimaging7020019

Academic Editor: Leonardo Rundo

Received: 23 November 2020

Accepted: 11 January 2021

Published: 29 January 2021

1. Introduction

Brain tumors are an abnormal growth of cells in the brain. Their exact causes are not yet known, but there are factors that can increase the risk of brain tumor, such as exposure to radiation and a family history of brain cancer. There has been an increase in incidences of brain tumors in all ages globally over the past few years [1]. In the United States alone, an estimate of 78,980 new cases of primary malignant and non-malignant tumors were expected to be diagonized in 2018. Despite considerable efforts in brain tumor segmentation research, patient diagnosis remains poor [2]. The most common types of tumors in adults are meningiomas (low grade tumors) and gliomas and glioblastomas (high grade tumors). Low grade tumors are less aggressive and they come with a life expectancy of several years. High grade tumors are much more aggressive and they have a median survival rate of less than two years.

Medical imaging techniques, such as Magnetic Resonance Imaging (MRI), CT scans, Positron emission tomography (PET), among others, play a crucial role in the diagnosis of the tumors. These techniques are used to locate and assess the progression of the tumor before and after treatment. MRI is usually the modality of choice for diagnosis and treatment planning for brain tumors [2] due to its high resolution, soft tissue contrast, and non-invasive characteristics. Surgery is the most common form of treatment for brain tumors, but radiation and chemotherapy can also be used to slow the growth of the tumor [1]. More than one MRI slice is required to view different regions of the brain, e.g., T1, T2, T1 contrast and FLAIR images.

Again, in clinical practice, delineation of the tumor is usually done manually. An experienced radiologist will carefully study the scanned medical images of the patient segmenting all of the affected regions. Apart from being time consuming, manual segmentation is dependent on the radiologist and it is subject to large intra and inter rater variability [3]. Consequently, manual segmentation is limited to qualitative assessment or visual inspection only.

Meanwhile, quantitative assessment of the brain tumors provides valuable information for a better understanding of the tumor characteristics and treatment planning [4]. Quantitative analysis of the affected cells reveals clues about the disease progression, its characteristics, and effects on the particular anatomical structure [5]. This task proved to be difficult, because of large variability in shape, size, and location of lesions. Moreover, more than one image modalities with varying contrast need to be considered for accurate segmentation of lesions [4]. As a result, manual segmentation, which provides arguably the most accurate segmentation results, would be impractical for larger studies. Most research endeavors today now focus on using computer algorithms for the automatic segmentation of tumors with the potential to offer objective, reproducible, and scalable approaches to the quantitative assessment of brain tumors.

These methods categorically fall into traditional machine learning and deep learning methods [6]. The application of statistical learning approaches to low-level brain tumor classification features is common in conventional machine learning methods. They mainly focus on the estimation of tumor boundaries and their localization. Additionally, they heavily depend on preprocessing techniques for contrast enhancement, image sharpening, and edge detection/refining, relying on human expertise for feature engineering. Wadhwa et al. [7] provide a concise overview of methods in this category.

On the other hand, deep learning methods rely on large scale dataset availability for training and require minimum preprocessing steps than traditional methods. Over the past few years, convolutional neural networks (CNNs) have dominated the field of brain tumor segmentation [6]. Alom et al. [8] provide a detailed review of deep learning approaches that span across many application domains.

Preliminary investigations [9,10] saw deep learning as a promising technique for automatic brain tumor segmentation. With deep learning, a hierarchy of increasingly complex features is directly learned from in-domain data [1] bypassing the need of feature engineering as with other automatic segmentation techniques. Accordingly, the focus would be on designing network architectures and fine-turning them for task at hand. Deep learning techniques have been popularized by their ground breaking performance in computer vision tasks. Their success can be attributed to advances in high-tech central processing units (CPU) and graphics processing units (GPUs), the availability of huge datasets, and developments in learning algorithms [11]. However, in the medical field, there is hardly enough training samples to train deep models without suffering from over-fitting. Furthermore, ground truth annotation of three-dimensional (3D) MRI is a time consuming and a specialized task that has to be done by experts (typically neurologists). As such, publicly available image datasets are rare and will often have few subjects [12].

In this survey, we highlight state of the art deep learning techniques, as they apply to MRI brain tumor segmentation. Unique challenges and their possible solutions to medical image analysis are also discussed.

2. Overview of Brain Tumor Segmentation

This section provides a brief introduction to brain tumor segmentation.

2.1. Image Segmentation

A digital image, like an MRI image, can be represented as a two-dimensional function, $f(x, y)$, where x and y are the spatial coordinates and the value of f at any given point (x, y) is the intensity or gray level of the image at that point. Each point in an image represents a picture element, called a pixel. The function f can also be viewed as $M \times N$ matrix , A, where M and N represent the number of rows and columns, respectively. Thus,

$$A = f(x, y) = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots \\ \vdots & \ddots & \\ a_{M,1} & & a_{M,N} \end{bmatrix} \quad (1)$$

In computer vision, image segmentation is the process of partitioning a digital image into multiple disjoint segments, each having certain properties. It is typically used in order to locate objects and their boundaries in images. This is achieved by assigning every pixel. (x, y), in an image A, a label depending on some characteristics or computed property, such as color, texture, or intensity.

The goal of brain tumor segmentation as depicted in Figure 1, is to detect the location, and extension of the tumor regions, namely:

- active tumorous tissue;
- necrotic (dead) tissue; and,
- edema (swelling near the tumor).

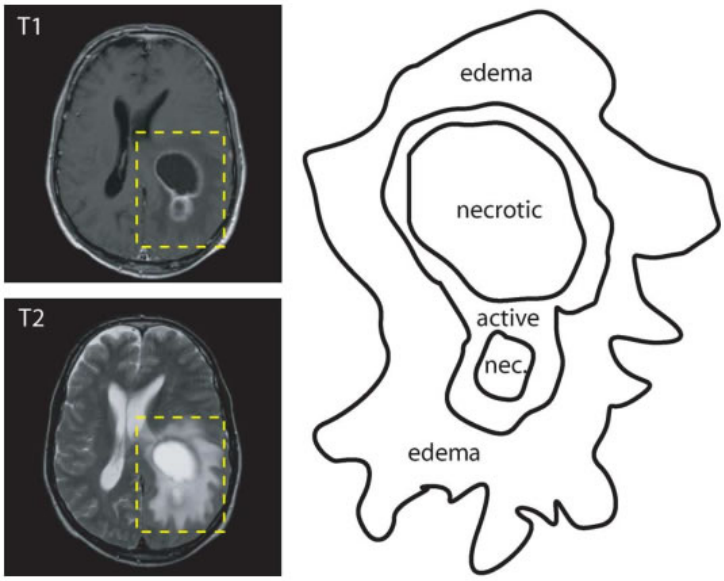


Figure 1. Labeled example of a brain tumor illustrating the importance of the different modalities (adapted from [13]).

This is done by identifying abnormal areas when compared to normal tissues [1]. Some tumors, like glioblastomas, are hard to distinguish from normal tissues, because they infiltrate surrounding tissues causing unclear boundaries. As a solution, more than one image modalities with varying contrasts are often employed. In Figure 1, two MRI modalities (T1 with contrast and T2) were used in order to accurately delineate tumor regions.

2.2. Types of Segmentation

Brain tumor segmentation can be broadly categorised as manual segmentation, semi-automatic segmentation, and fully automatic segmentation, depending on the level of human involvement. Gordillo et al. [14] provide a full description of these methods.

2.2.1. Manual Segmentation

With manual segmentation, a human operator uses specialized tools in order to carefully draw or paint around tumor regions. The accuracy of segmentation results depends heavily on the training and experience of the human operator as well as knowledge of brain anatomy. Apart from being tedious and time consuming, manual segmentation is widely used as a gold standard for semi-automatic and fully automatic segmentation.

2.2.2. Semi-Automatic Segmentation

Semi-automated segmentation combines both computer and human expertise. User interaction is needed for the initialisation of the segmentation process, providing feedback and an evaluation of segmentation results [3]. Although semi-automatic segmentation

methods are less time consuming than manual segmentation, their results are still dependent on the operator.

2.2.3. Fully Automatic Segmentation

In fully automatic brain tumor segmentation, no human interaction is required. Artificial intelligence and prior knowledge are combined in order to solve the segmentation problems [3]. Fully automatic segmentation methods are further divided into discriminating and generative methods. Discriminating methods often rely on supervised learning where relationships between input image and manually annotated data are learnt from a huge dataset. Within this group, classical machine learning algorithms, which rely on hand crafted features, have been extensively used with great success over the past years. However, these methods may not be able to take full advantage of the training data due to the complexity of medical images [15]. More recently, deep learning methods have gained popularity because of their unprecedented performance in computer vision tasks and their ability to learn features directly from data. On the other hand, generative methods use prior knowledge regarding the appearance and distribution of difference tissue types.

3. Deep Learning

Deep learning is a class of machine learning algorithms that uses multiple layers to learn a hierarchy of increasingly complex presentations directly from the raw input. Machine learning models are all about finding appropriate representations for their input data. In this section, we will describe the building blocks, and recent techniques and architectures of deep learning algorithms for brain tumor segmentation that we found in papers surveyed in this work, as summarized in Figure 2.

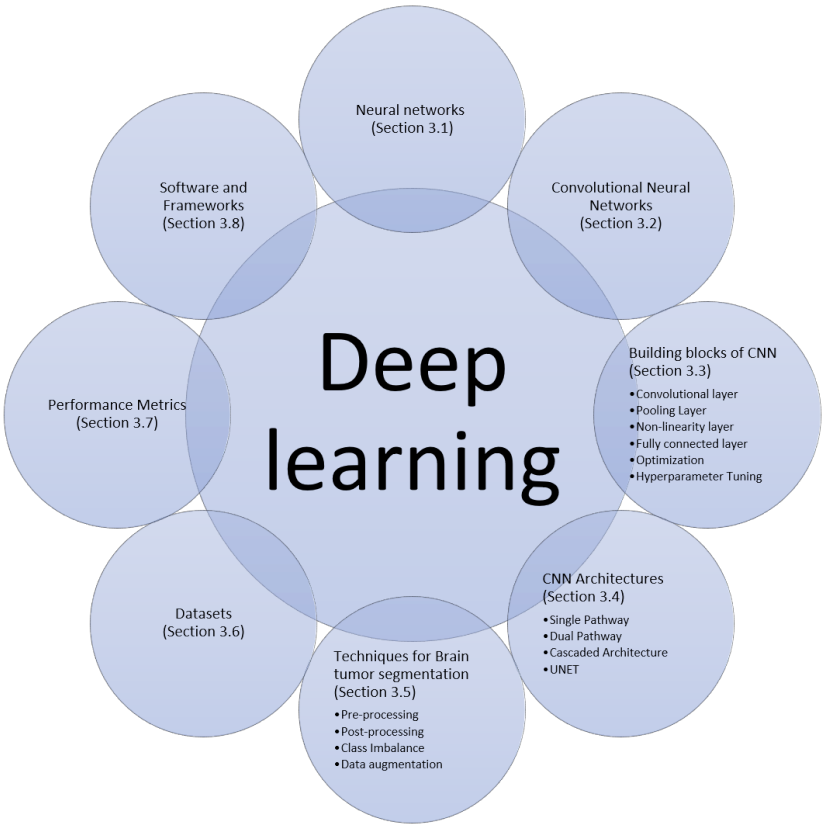


Figure 2. Building blocks, architectures and techniques for deep learning algorithms for brain tumor segmentation.

3.1. Neural Networks

A neural network is a type of a machine learning algorithm that is able to learn useful representations from data [16,17]. The network is formed by connecting processing units, called neutrons, by directed links. Each link is associated with a weight that adjusts as learning proceeds. When the topology of the network forms a directly acyclic graph, the network is referred to as a feed forward neural network (Figure 3). Associated with each neutron is a function $f(x : \theta)$, which maps an input x to an output y and it learns the value of the parameters $\theta = \{w, b\}$, where w is a weight vector and b is a scalar, through a back-propagation algorithm:

$$f(x : \theta) = \sigma(w \cdot x + b) \tag{2}$$

where $\sigma(\cdot)$ is element-wise non-linearity activation function.

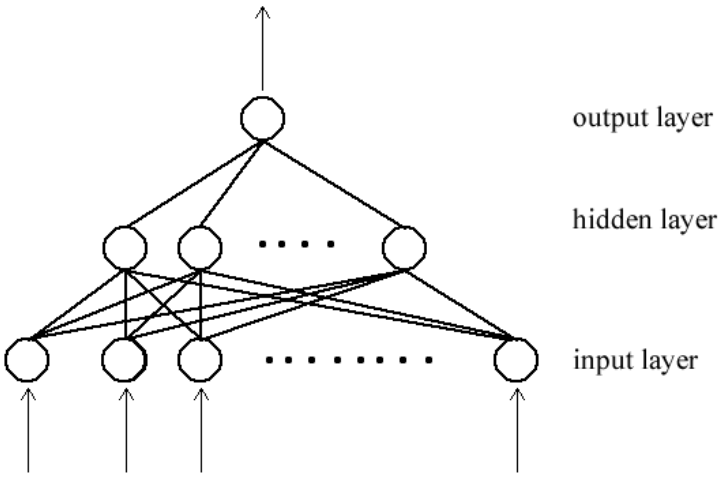


Figure 3. Typical feed-forward neural network composed of three layers. (adapted from [18]).

In a typical neural network, neurons are organized in layers. The input of each neuron in a layer is connected to all or some of the output of neurons in the up-stream layer. Likewise, the output of each neuron is connected to all or some of the input of neurons in the downstream layer. The first layer in the network is the input layer, and the final layer is the output layer. Layers in the middle are referred to as hidden layers. When each neuron in a layer is connected to all of the neurons in the next layer, the network is called fully connected network. A deep neural network is formed when there are many hidden layers, hence the term *deep learning*.

3.2. Convolutional Neural Network (CNN)

A convolutional neural network is a type of a neural network that performs a convolutional operation in some of its layers. The convolutional layer is able to learn local features from the input data. By stacking many convolutional layers one after the other, the network is able to learn a hierarchy of increasingly complex features. A polling layer is usually added in-between successive convolutional layers to summarize important features. This will reduce the number parameters that are passed to downstream layers and, at the same time, introducing translation invariant (able to recognize learned patterns, regardless of their geometric transformations) to the network.

Recently, CNN has become the de factor model for brain tumor segmentation because of its record shattering performance in classical computer vision problems as well as in medical image analysis as compared to other models. CNN models are able to learn spatial hierarchies of features within data, for example, the first convolutional layer will learn small local patterns, like edges, the second layer will learn larger patterns made up of features of the preceding layer and so on. This ability make them a better fit for image

analysis task. Furthermore, units in convolutional layers share weights, thereby reducing the number of parameter to learn and improve the efficiency of the network.

3.3. Building Blocks CNN

3.3.1. Convolutional Layer

This layer consists of a set of learnable filters or kernels (the typical size is usually 3×3 or $3 \times 3 \times 3$, depending whether the input is a two-dimensional (2D) or three-dimensional (3D) image, respectively) that are used to slide over the entire input volume, performing a dot product between entries of the filter and the input at that point. Thus, the convolutional operation first extracts patches from its input in a sliding window fashion, and then applies the same linear transformation to all of these patches. The output of the convolution operation is sometimes referred to as the feature map. The network will learn filters that recognize certain visual patterns present in the input data. When convolutional layers are stacked one after the other, the network is able to learn a hierarchy of increasing complex features, from simple edges to being able to recognize the presence of a face for example.

Over the past few years, there were various attempts meant to improve the performance of deep learning models by replacing the conventional convolutional layer with blocks that increase the network's capacity while using less computational resources. For example, Szegedy et al. [19] introduced the inception block that captured sparse correlation patterns whlie using multi-scale receptive fields. Their network architecture, the GoogleNet, a winner of ILSVRC 2014, had fewer network parameters and required less computational resources than its predecessors AlexNet [20] or VGG [21]. The residual block was another notable improvement [22], which facilitated very deep networks that do not suffer from the vanishing gradient problem. Hu et al. [23] introduced the Squeeze-and-Excitation (SE) block that captured the interdependencies between the network's feature maps.

3.3.2. Pooling Layer

A pooling layer usually follow a convolutional layer or a set of convolutional layers. The goal is to reduce the dimensions of the feature maps, and at the same time, keep important features. A pooling operation is applied to a rectangular neighbourhood in a sliding window fashion. For example, the max pooling is used in order to produce a maximum of a rectangular neighbourhood. Other popular pooling operations include average and weighted average pooling.

3.3.3. Non-Linearity Layer

Typical convolutional layers involves three steps [16]. In the first step, the layer performs convolutional operation on input feature maps to produce a set of linear activations. Second, a non-linear transformation is performed on the output feature maps. Third, a pooling layer is used in order to modify the output further. Non-linear transformations can be obtained by using special class of functions, called activation functions. Non-linearity gives the network the ability to learn nontrivial representations that are sparse. Hence, making the network resilient to slight modifications or noise in the input data as well as improving computational efficiency of the representations.

In the past, sigmoid and hyperbolic tangent functions were commonly used for the non-linearity layer. Today, the most popular activation function is the rectified linear unit (ReLU), which is expressed as $f(z) = max(z, 0)$. It was observed in [20,24], where ReLU typically learns faster in network with many layers and does not suffer from vanishing/exploding gradients, as with the sigmoidal activations. However, ReLU presents some potential drawbacks when the network saturates with a constant zero gradient causing the network to converge slowly. As a solution, Maas et al. [25] proposed a Leaky ReLU (LReLU) that allows for small, non-zero gradient to flow when the network is saturated. This function is defined as

$$f(z) = max(z, 0) + \alpha min(0, z) \tag{3}$$

where α is a constant leakiness parameter(typically 0.01). Another common variant of ReLU is Parametric Rectified Linear Unit (PReLU) [26]. This activation function adaptively learns the parameter α in Equation (3), thus improving the accuracy with less computational cost.

3.3.4. Fully Connected Layer

The convolutional layers are used as feature extractors. The features that they produce are then passed to the fully connected (FC) layers for classification. Each unit in the FC layer is connected to all of the units in the previous layer, as shown in Figure 3. The final layer is usually a softmax classifier, which produces a probability vector map over the different classes. All of the features are converted in to a one-dimensional feature vector before being passed to a FC layer. By doing so, spatial information inherent in image data is lost. Another issue with the FC layers is that they have a larger number of parameters as compared to other layers that increase the computational costs and require input images to be of the same size.

As a solution to above problems, Long et al. [27] proposed converting FC layers to 1×1 convolutional layers, thus transforming the the network into a fully convolutional network (FCN). The network takes the input of any arbitrary sizes and outputs a grid of classification maps.

3.3.5. Optimization

The performance of the deep CNN can be improved (or optimized) by training the network on a large dataset. Training involves finding the parameters θ of the model that significantly reduce a cost function $J(\theta)$. Gradient descent is the widely used method for updating network parameters through a back-propagation algorithm. Optimization can be done per single sample, subset, or full set of the training samples. Thus, stochastic, mini-batch, or batch gradient descent, respectively. Today, many optimization algorithms for deep learning use mini-batches and it is now common to just call them stochastic methods [16].

Stochastic gradient descent (SDG) comes with few notable challenges. Choosing an appropriate learning rate can be difficult. A learning rate that is too small leads to very slow convergence (tiny updates to the model parameters) and, at the same time, too large will result in undesired divergence behavior in the loss function. All of the parameter updates are based on the same learning rate, disregarding the fact that some of the features might have higher frequency than other. Another key challenge is that optimization can be trapped in sub-optimal local minima or saddle points, especially for non-convex optimization [28].

Various variants of SDG have been proposed in the literature that address the aforementioned challenges. Memontum-based SDG methods [29] can help in accelerating SDG in relevant direction, dampening undesirable oscillations in local optima. Adagrad [30] addressed the issue of manually turning the learning by adapting the learning rate to the parameters, performing larger updates for infrequent parameters as compared to frequent ones. However, Adagrad suffers from monotonically decreasing learning rate to a point at which the algorithm stops learning. Adadelta [31], RMSprop [32], and Adam [33] addressed the shortcomings of Adagrad by dividing the learning rate by an exponentially decaying average of past gradients.

3.3.6. Loss Function

In machine learning, a loss function is used in order to evaluate how well a specific algorithm models the given data. When the output is far from the true value, loss will be very high and low when the predictions are close to the true values. The primary goal of training a neural network is to minimize the loss (or cost) function of the network as much as possible and, at the same time, ensuring that the network generalizes well with unseen data.

The choice of the cost function depends on the problem area, whether it is a classification or regression problem and the choice of the output unit [16]. The majority of the image classification algorithms use softmax loss, withhs a combination of softmax and CE loss or log-loss [28]. The softmax function produces a probability distribution over a number of given output classes, while the CE loss takes the probability of predictions and penalizes predictions that are confident but wrong. Class imbalance is one major issue in medical image analysis, where one class will have fewer instances than the other. For example, a brain tumor occupies a small portion when compared to healthy tissues. As a result, the classifier will tend to be biased to the majority class. One way of addressing such a problem is to adapt loss functions for class imbalance. Some works [34–36] proposed a loss function that is based on the Dice coefficient. Ronneberger et al. [37] proposed a weighted CE loss, which gives more importance to some pixels in the training data.

3.3.7. Parameter Initialization

Deep learning optimization algorithms are iterative in nature, thus requiring the user to specify initial starting point of the algorithms [16]. The choice of initialization will influence how quickly learning can converge if it can converge at all. Empirical studies have shown that a carefully chosen initialization scheme dramatically improves the rate of convergence [38], while gradient-based optimization starting from random initialization may get stuck near poor solutions [39].

Ref. [38] proposed a normalized initialization scheme (Xavier initialization), which guarantees that weight initialization should not obtain values that are too small or too large, thus reducing saturation and vanishing gradients, thereby improving convergence. This approach was later improved in [26] to perform much better on Relu or PRelu activations and extreme deep models.

3.3.8. Hyperparameter Tuning

Hyperparameters are parameters that are supplied by the user to control the algorithm's behavior before training commences, such as learning rate, batch size, image size, number of epochs, kernel size etc. While the learning algorithms do not adapt these parameters, their choice has varying effects on the resulting model and its performance. The majority of the works studied in this review set their hyperparameters manually or perform a grid search while using the validation set. However, these approaches will become impractical when the number of hyperparameters is large [40] and they rely on human expertise, intuition, or guessing. As a solution to these challenges, automated approaches, like AutoML (http://www.automl.org) and Keras Tuner, (https://keras-team.github.io/keras-tuner/) are beginning to gain much attention.

3.3.9. Regularization

Regularization is a technique for improving the performance of a machine learning algorithm on unseen data. It is a way of reducing over-fitting on training set. Over-fitting occurs when the gap between the training error and test error is too large [16]. When that happens, the model performs well on training data, but poorly on previously unseen data. There are various techniques that can be employed in order to reduce the generalization error, such as reducing the model capacity, which is, reducing the number of learnable parameters in the model; adding L^2 or L^1 weight decay regularization term to the cost function to force the model to only take small weight values; introducing early stopping whenever the model performance stops improving on validation dataset; randomly dropping out (skipping) the output of some units during training [41]. The last approach is one of the most effective and most commonly used technique [17], mainly because it is computationally inexpensive and prevents interdependent learning amongst units. Batch Normalization [42] can also be used as a regularizer by ensuring that the distribution of non-linearity inputs remains more stable as the model trains, thereby improving the training of the model.

Training a machine learning model with more data is the best way to reduce the generalization error. However, in the medical domain, acquiring a training dataset is time-consuming, more expensive, and requires highly trained personnel to annotate ground truth labels. Data augmentation can increase the dataset and reduce over-fitting by flipping, applying small rotations, warping, and using the non-rigid deformation transformation of images. However, great care must be taken when performing transformations of the medical image dataset since the patch's label is determined by the center of pixel [43]. Some recent works used generative models that include variational autoencoders [44] and generative adversarial networks [45] to act as additional regularization that deals with data scarcity.

3.4. Deep CNN Architectures

3.4.1. Single Pathway

A single pathway architecture is a basic network that resembles a feed-forward deep neural network. Data flows from the input layer to the classification layer using a single path. Urban et al. [10] proposed a 3D single path CNN which has fully connected convolutional layer as the classification layer. This gave the network the ability to classify multiple 3D pixel in one go. In [46], each image's modality was fed to a different two-dimensional (2D) CNN. The result of each CNN was then used as features to train a random forest classier. Extracts from XY, XZ, and YZ planes around each center pixel were used as the neighborhood information. Pereira et al. [43] used small kernels in their convolutional layers. As a result, a very deep network, DeepMedic, was obtained, which can learn more feature hierarchies. Their architecture obtained first and second positions in BRATS 2013 and 2015 challenge,respectively.

3.4.2. Dual Pathway

Many segmentation algorithms perform pixel-wise classification, where an input patch is extracted from an MRI image and then predicts the label of the central pixel without considering global neighborhood information. This can be risky because of infiltrating nature of brain tumors, which produces unclear boundaries. Hence, local information cannot be enough to accurately produce good segmentation results. As a solution, other researchers [1,47] introduced neighbourhood information to the mix by using CNN with two data streams (dual pathway) that are combined in order to influence label predictions of each pixel. One of the streams will represent local information, the visual details of the region around the center pixel. The other stream represents the global context, which takes the location of the extracted patch in the brain into account.

3.4.3. Cascaded Architecture

In a cascaded architecture, the output one CNN is concatenated with the other. There many variations with this architecture in the literature, but the most prominent is the input cascade [1,48]. In this architecture the output of one CNN becomes a direct input of another CNN. The Input cascade is used in order to concatenate contextual information to the second CNN as additional image channels. This is an improvement to the dual-path way that performs multi-scale label predictions separately from each other. Another variation of cascaded architecture is the local pathway concatenation [1]. In this architecture, the output of the first CNN is concatenated with the output of the first hidden layer of the second CNN instead of its input.

Hierarchical segmentation [34,49] is another form of a cascaded architecture. In this architecture, the segmentation of brain tumor regions is sequentially done by reducing the multi-class segmentation problem into the multi-stage binary segmentation problem. This architecture takes full advantage of the hierarchical nature of tumor sub-regions and helps in reducing false positives as well as mitigating the inherent class imbalance problem. The first stage of architecture segments the whole tumor from the input MRI modalities, which is then used as a bounding box for the next stage. For the second stage, the output of

the first stage is used as an input to perform either a multi-class intra-tumoral segmentation, as in [49], or perform successive binary segmentation of the remain tumor sub-regions [34]. Wang et al. [34] observed an increase in the training and inference time of a multi-stage binary segmentation as compared to a single multi-class network approach.

3.4.4. UNET

The UNET architecture [37] is an improvement of FCN [27], which resembles an encoder and decoder network designed specifically for biomedical image segmentation. The network consists of a contracting path (encoder) and an expansive path (decoder), which gives it the u-shaped architecture. The contracting path consists of the repeated application of two convolutional layers, followed by a rectified linear unit (ReLU) and max pooling layer. Along the path, the spacial information is reduced, while feature information is increased. The expansive path consists of a series of up-sampling operations combined with high-resolution features from the contracting path through skip connections.

3.5. Techniques for Brain Tumor Segmentation

3.5.1. Pre-Processing

Data preprocessing is a very crucial step of preparing raw input data to be more amenable to neural networks. MRI images contains various artifacts that are caused by the acquisition protocol and the hardware used. These artifacts need to be corrected before the images are fed into the network for better performance. One of the notable artifacts is the presence of smooth intensity variations within the image, which is also known as bias field. Among various techniques for bias field correction, the non-parametric nonuniform normalization (N3) [50] approach has become the technique of choice for bias field correction due to its ease of use and its availability as an open source project [51]. This technique was later improved in [51] and it is also well known as N4ITK. These techniques are limited to a single image. Accordingly, for uniform intensity distribution across patients and acquisitions, the intensity normalization proposed by Nyul et al. [52] can be applied.

Another popular preprocessing technique is to normalize image dataset to have a mean zero and a standard deviation of one. This technique assists in removing the bias from features. Image cropping can also be applied to remove as much background pixels as possible.

3.5.2. Post-Processing

The post-processing step is performed to further refine the segmentation results. It helps in reducing the number of misclassifications or false positives in the segmentation results while using algorithms, like conditional random fields (CRF) [4,34,53], markov random fields (MRF) [54], connected component analysis [1,53,55], and morphological operators [48,56]. CRF and MRF based techniques effectively remove false positives by combining model predictions with low-level image information, like local interations of pixels and edges when making finer adjustments. However, these techniques are computationaly expensive [14]. Connected compents analysis involves finding and extracting connected components and then applying a simple thresholding technique to remove unwanted blobs. Another technique of removing false positive around edges of the segmentation image is to apply morphological operations, erosion, and dilation in succession.

3.5.3. Class Imbalance

The performance of the segmentation task is affected by the class imbalance problem, where there is an unequal distribution of voxel classes in the training dataset. For example, in brain tumor segmentation, healthy voxels constitute 98% of the total voxels [1]. Training the model on this distribution will cause the model to be more biased towards the majority class. Whereas, training with equal distribution results in bias towards tumor classes [57]. Several techniques have been explored in the literature in order to address this problem.

Many works incorporated loss-based methods of addressing the class-imbalance problem. Lin et al. [58] proposed a loss function that addresses the problem by dynamically scaling the loss based on the model's confidence in classifying samples. The scaling factor was reduced when the model's accuracy in classifying classed increases. As a result, the model pays more attention to misclassified samples. In [59], dice loss was used as a means of addressing the problem. Some works [60,61] incorporated a weighted-loss function, where voxels (or pixels) belonging to different classes are assigned weights according to their distribution in the training data. This ensures that each class in the segmentation problem has an equal contribution to the model's loss. Kuzima et al. [62] combined the CE loss with Dice based loss as means of addressing class imbalance problem. Other works explored hard negative mining [63,64] as a solution to the class-imbalance problem. Voxels with largest negative losses and positive voxels are used in order to update the model's weights.

Two-phase training [1,5,57] is also another way of dealing with the class imbalance problem. In the first phase, the network is trained with patches that have equal class distribution and then trained with true class distribution in the second phase. Hussain et al. [57] reported that two-phased training helped in removing most of the false positives.

In [34], Wang et al. pointed out that hierarchical segmentation also assists in addressing the class-imbalance problem.

3.5.4. Data Augmentation

Data augmentation is a technique for reducing the generalization error of a machine learning algorithm. As indicated earlier, one way of effectively increasing the machine learning model's generalization capabilities is to train it on more data. However, acquiring a considerable amount of high-quality training data is nearly impossible in practice, especially for the medical domain. Data augmentation has emerged in order to increase the training data by creating more synthetic data and adding (augment) it to the training set.

Data augmentation can be broadly divided into two categories [65]: the transformation of original data and artificial data generation. With the transformation of original data, new data are generated by applying various transformations on the original data, which include affine transformations (which involves rotation, zooming, cropping, flipping, and translations), elastic transformations (shape variations), and pixel-level transformation (intensity variations). While these transformations assist in mitigating insufficient data challenges, they fundamentally produce very correlated images [66], which results in very little performance improvement [66,67] and sometimes generates anatomically incorrect examples (e.g. using rotation) [65]. However, their use in the literature is widespread, due to the ease of implementation.

On the other hand, artificial data generation [67,68] exploits the Generative adversarial networks (GANs) [69] to generate realistic data that are indistinguishable from the real data and also serves as a effective method for data anonymization [66]. GANs are able to generate a wide variety of realistic samples that can bring invariance and robustness. However, there are scenarios where they can generate samples that are very similar to the real ones, resulting in poor performance [65].

3.6. Datasets

Over the past few years, there have been considerable research interests in automatic brain tumor segmentation. As research output continued to grow, the objective evaluation of different algorithms became a challenge because researchers used private datasets with varying attributes. As a result, benchmarking challenges, such as Multi-modal Brain Tumor Image Segmentation (BRATS), emerged to standardize performance evaluation while using publicly accessible datasets. Table 1 show a summary of the mostly used datasets for brain tumor segmentation.

Since 2012, the BRATS Challenge [2], in conjunction with the International Conference on Medical Image Computing and Computer-Assisted Interventions (MICCAI), has been

the primary bench-marking resource for brain tumor segmentation. It offers the medical research community publicly accessible datasets for training and validation and standardized metrics in order to objectively evaluate model performance against an online evaluation platform. The dataset initially contained as small as 30 clinically acquired scans of glioma patience, and the number has continued to grow over the subsequent years.

Table 1. Summary of commonly used public datasets for brain tumor segmentation.

Name	Total	Training Data	Validation Data	Testing Data
BRATS 2012 [2]	50	35	-	15
BRATS 2013 [2]	60	35	-	25
BRATS 2014 [2]	238	200	-	38
BRATS 2015 [2]	253	200	-	53
BRATS 2016 [2]	391	200	-	191
BRATS 2017 [2]	477	285	46	146
BRATS 2018 [2]	542	285	66	191
BRATS 2019 [2]	653	335	127	191
Decathlon [70]	750	484	-	266

Medical Segmentation Decathlon Challenge offers a relatively large dataset that supports a wide range of segmentation task. The Challenge aims to facilitate research in general-purpose segmentation algorithms that solve various functions without any human intervention. For brain tumor segmentation, the dataset comprises a subset of the 2016 and 2017 BRATS Challenge data.

3.7. Performance Evaluation Metrics

In order to objectively measure the performance of segmentation algorithms, researchers have to group different tumor structures into three mutually inclusive regions:

- the *whole* tumor (includes all tumor structures);
- the *tumor* core (exclusive of edema); and,
- the *active* tumor (only consists of the "enhancing core").

Subsequently, they measure the algorithm's performance on each region against several metrics that include the Dice score, Sensitivity, Specificity, and Hausdorff measure.

3.8. Software and Frameworks

Researchers and engineers have always relied on open-source software frameworks from idea generation to experimentation to production deployments in order to accelerate the deep learning workflow. This section described some of the popular machine learning frameworks that were used in the reviewed papers.

Theano [71] is a free and open-source python framework for the fast computation of large-scale dataflow mathematical expressions compiled and executed naively on both CPUs and GPUs. Moreover, the research community has been utilizing the platform in order to conduct machine learning research. However, it is not a purely a machine learning framework, but rather a compiler for mathematical expressions that are defined in NumPy-like syntax. Several high-level software packages like Pylearn2, Keras, blocks, and Lasagne have been built on top of Theano, leveraging its strengths as an efficient mathematical powerhouse.

Pylearn2 [72] is a free and open-source machine learning library that is built on top of the Theano framework. It started gaining popularity after being used to win a transfer learning challenge and implementing various state of the art computer vision benchmarks. The library focuses on flexibility and extensibility, allowing for researchers to implement arbitrary machine learning models at ease. Unfortunately, the library no longer has an active developer and has, ever since, fallen behind other actively maintained frameworks, like Keras.

Caffe [73] is a C++ deep learning framework that was initially developed for computer vision applications and later spread to other domains like robotics, neuroscience, and astronomy. It offers a complete toolkit for a deep learning pipeline, from training to production deployment. Each processing stage is supplemented with well-documented examples. Moreover, the framework is shipped with implementations of popular deep learning building block and reference models allowing for quick experimentation with state-of-the-art deep learning methods. The definition of models is done in config files, rather than being hard-coded, ensuring the separation of representation from implementation.

Pytorch [74] is yet another fully-fledged open-source deep learning framework. Its design philosophy moved away from the define and execute style, as in many frameworks that create a static computational graph before running the model. While this approach is powerful, it sacrifices usability, the ease of debugging, and flexibility. Instead, Pytorch took an imperative approach by dynamically constructing the computational graph, allowing for the models to be idiomatically defined following the python programming model. The framework also offers a seamless transition from research to production, distributed training, and the seamless execution of models on edge devices.

Tensorflow [75] is an end-to-end distributed deep learning platform for large scale machine learning applications. The platform supports the execution of dataflow graphs across a span of heterogeneous devices, such as mobile devices and large-scale distributed systems, with little or no change. Its design philosophy has been used to simplify model parallelism within a single machine and across thousands of distributed systems. It has a complete toolbox for quick experimentation with state-of-the-art deep learning models, seamless transition from research to heterogeneous deployments, and the visualization and debugging of large-scale models.

Keras [76] is a fast-growing high-level API for deep learning applications. Although it initially supported multiple data-flow graph back-ends, like Theano, it is now deeply woven into the Tensorflow 2 ecosystem. It provides consistent and simple APIs to quickly experiment with new models and leverage Tensorflow in order to export the models to run in browsers and mobile devices. Moreover, it comes bundled with building blocks and pre-trained state-of-the-art models for various machine learning domains. The industry and the research community have adopted the platform, because of its ease of use, user-centric approach, and extensive documentation.

4. Discussion

Deep learning methods to medical image analysis have received tremendous attention over the past few years. This is evident in the considerable increase in the number of published works each year [2]. Deep learning techniques are able to learn a hierarchy of increasingly complex features directly from data, as stated earlier. For example, in brain tumor segmentation, deep learning algorithms can learn to segment MRI images by being trained on a sufficiently large dataset. For this reason, CNN based models have been widely adopted in medical image analysis, following their success in solving many problems in computer vision, speech recognition, and natural language processing. Table 2 shows a summary of deep learning methods that were reviewed in this work. Many techniques differ considerably in terms of architectural design, with recent works following the Unet [37] architecture and ensemble methods as shown in Table 3. Moreover, several techniques have been developed in order to address inherent problems in automated brain MRI analysis.

Table 2. Overview of Deep learning methods for brain tumor segmentation. BN = Batch normalization, GN = Group normalization, outliers = remove top 1%, hist-norms = Histogram normalization,RN = Range normalization, HS = Histogram standardization, slice-norm = Slice-based normalization, PLN = Piece-wise linear normalization, IN = Instant normalization, CE = Cross entropy,BS = Bootstrapping, SS = Sensitivity-specification, NM = Negative Mining, WCE = Weighted cross-entropy, neg-mining = Hard negative mining.

Reference	Input	Preprocessing	Regulization	Loss	Optimizer	Activation
Unet Architecture						
[47]	3D	Z-score				ReLu
[77]	2D		BN	Dice, WCE, BS, SS	Adam	ReLU
[34]	2D	Z-score, hist-norms	dropout	CE	SDG	LReLU
[78]	3D	cropping	BN	Jaccard loss, CE		PReLU
[79]		Z-score, N4ITK,lin-norm				
[80]	2D			Dice	Adam	
[81]	2D	Z-score, HM	BN	CE	Adam	ReLU
[82]	3D	bounding box	dropout	Dice	Adam	
[83]	3D	Z-score, rescaling, outliers	IN, L2	Dice	Adam	LReLU
[84]	2D	slice-norm		CE	Adam	
[85]	3D		BN	Dice	Adam	
[15]	2D	Z-score	BN	CE	Adam	ReLU
[63]	3D	Z-score	GN	CE, neg-mining	SGD	
[36]	2D	bounding-box, cropping, Z-score, intensity-windowing	BN	Dice	Adam	Relu
[86]	2D	N4ITK, Nyúl	BN, spatial-dropout	CE	Adam	ReLU
[60]	2D		BN	CE		ReLU
[87]	2D	Z-score, remove outliers	BN	WCE, Dice	SGD	PReLU
[88]	3D	Z-score	IN, L2	CE, Dice	Adam	LReLU
[5]		N4ITK, remove outliers		WCE	Adam	
[35]	2D	Z-score	BN	Dice	Adam	Relu
[59]	3D	Z-score	BN	Dice	Adam	PReLU
[89]	3D	Z-score	BN, L2	CE, Dice, focal	Adam	ReLU
[90]		Z-score			Adam	RelU
[91]	3D	Z-score	GN, L2, Dropout	Dice	Adam	ReLU
[92]	3D	RN, random axis mirror		CE, Dice	SDG	
[64]	3D	Z-score, N4ITK	BN, L2	CE, NM	Adam	ReLU
Dual-pathay Architecture						
[10]	2D		L1, L2 Dropout		SDG	
[1]	2D	Z-score, N4ITK, outliers	L1, L2, Dropout	log-loss	Maxout	ReLU
[47]	2D	Z-score			Adam	ReLU
[57]	2D	Z-score, N4ITK	BN, Dropout	log-loss	SDG	ReLU
[63]	3D		GN	CE, NM	SDG	
[53]	2D	N4ITK				PReLU
[5]		N4ITK, outliers		WCE	SGD	
[93]	3D	N4ITK, LIN				ReLU
[94]	3D		Dropout	log-loss	SDG	PReLU
[95]	2D	N4ITK	Dropout		SGD	ReLU
[4]	3D	Z-score		log-loss		ReLU
[79]		Z-score, N4ITK, PLN				
[96]	3D	Z-score	BN, L2, Dropout			ReLU

Table 2. *Cont.*

Reference	Input	Preprocessing	Regularization	Loss	Optimizer	Activation
Single-pathway Architecture						
[9]	2D			log-loss	SGD	ReLU
[46]	2D		Dropout	CE	SGD	ReLU
[43]	2D				SGD	ReLU
[64]	3D	Z-score, N4ITK	BN	CE, NM	Adam	ReLU
[97]	2D			CE	Nesterov, RMSProp	ReLu
[98]	2D	Z-score, outliers			Adam, SGD, RMSProp	ReLu
[99]	3D					ReLU
[43]	3d	Z-score, N4ITK, Nyúl	Dropout	CE	Nesterov	LReLU
Ensemble Architecture						
[59]	3D	Z-score	BN	dice	Adam	PReLU
[64]	3D	Z-score, N4ITK	BN	CE, NM	Adam	ReLU
[63]	3D		GN	CE, NM	SDG	
[61]	2D	Z-score, N4ITK, HN,	Dropout	CE	Adam	
[98]	2D	Z-score, outliers			Adam, SGD, RMSProp	ReLu
[44]	3D	Z-score	GN, L2, spatial dropout	Dice	Adam	ReLU
[79]		Z-score, N4ITK, PLN				
Cascaded Architecture						
[34]	2D	HS, Z-score	dropout	CE	SGD	LReLU
[1]	2D	Z-score, N4ITK, remove outliers	Dropout L2, L1	log-loss	Maxout	
[48]	2D				Maxout	RelU
[85]	3D		BN	Dice	Adam	LReLU
[100]	2D	Z-score, BN,outliers	L2, dropout	CE	SGD	ReLU
[34]	2.5D	Z-score	BN	Dice	Adam	PReLU
[59]	3D	Z-score	BN	Dice	Adam	PReLU
[89]	3D	Z-score			Adam	ReLU
[86]	2D	Z-score, N4ITK	BN, spatial dropout	CE	SDG	ReLU
[34]	3D	Z-score	BN	Dice	Adam	PReLU
[86]		N4ITK, Nyúl	BN, dropout	CE	Adam	ReLU
[35]	2D	Z-score	BN	Dice	Adam	ReLU
[91]	3D	Z-score	GN, L2, dropout	Dice	Adam	ReLU

Table 3. A summary of top performing methods on BraTS 2017, 2018, and 2019 validation data as reported by the online evaluation platform. ET—Enhancing tumor, WT—Whole tumor, and TC—Tumor core.

Rank	Reference	Architecture	Dice			Sensitivity			Specificity			Hausdorff 95		
			ET	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC
BraTS 2017														
1	[79]	Ensemble	0.738	0.901	0.797	0.783	0.895	0.762	0.998	0.995	0.998	4.499	4.229	6.562
2	[34]	Cascaded	0.786	0.905	0.838	0.771	0.915	0.822	0.999	0.995	0.998	3.282	3.890	6.479
3	[83]	Unet	0.776	0.903	0.819	0.803	0.902	0.786	0.998	0.996	0.999	3.163	6.767	8.642
3	[101]	SegNet	0.706	0.857	0.716	0.687	0.811	0.660	0.999	0.997	0.999	6.835	5.872	10.925
BraTS 2018														
1	[44]	Ensemble	0.825	0.912	0.870	0.845	0.923	0.864	0.998	0.995	0.998	3.997	4.537	6.761
2	[88]	Unet	0.809	0.913	0.863	0.831	0.919	0.844	0.998	0.995	0.999	2.413	4.268	6.518
3	[102]	Ensemble	0.792	0.901	0.847	0.829	0.911	0.836	0.998	0.994	0.998	3.603	4.063	4.988
3	[103]	Ensemble	0.814	0.909	0.865	0.813	0.914	0.868	0.998	0.995	0.997	2.716	4.172	6.545
BraTS 2019														
1	[91]	two-stage Unet	0.802	0.909	0.865	0.804	0.924	0.862	0.998	0.994	0.997	3.146	4.264	5.439
2	[92]	Unet	0.746	0.904	0.840	0.780	0.901	0.811	0.990	0.987	0.990	27.403	7.485	9.029
3	[104]	Ensemble	0.634	0.790	0.661	0.604	0.727	0.587	0.983	0.980	0.983	47.059	14.256	26.504

Deep learning algorithms require a relatively large amount of training data to generalize well on unseen data. However, this poses many challenge in the medical domain. Firstly, it takes a well trained radiologist a considerable amount of time to annotate even a single MRI volume. Moreover, the work is subject to an intra-rater and inter-rater variability. Therefore, all of the annotations are approved by one to many experienced neuro-radiologists [105], before they can be used in supervised training, which makes the process of creating training and testing datasets not only time consuming, but expensive. Secondly, medical data is protected by data protection laws that restrict the usage and sharing of this kind of data to other parties. Consequently, a lot of time is spent seeking approvals and removing personal identifiable information from medical data. Fortunately, Table 1 shows a consistent increase of training and testing data for the BraTS Challenge. Hopefully, this trend will continue in the coming years. Thus, facilitating training relative deep networks and reducing over-fitting.

Because the lack of large-scale datasets restricts deep learning models' full potential, researchers have adopted data augmentation as an immediate solution to the data challenges that are mentioned above. Other works have recently explored weakly-supervised learning [106–108] as a promising solution to address the need for fully annotated pixel-wise labels. Instead of performing pixel-level annotations, known to be tedious and time-consuming, weakly-supervised annotation uses bounding box or image-level annotations in order to signify the presence or absence of lesions in images. This approach has the benefit of being cheap, contains less labeling noise [107], far larger volumes of data can be generated than pixel-level annotation, and training of deep learning models can leverage both kinds of datasets.

Moreover, deep learning techniques require a huge amount of computational and memory resources [28]. Very deep networks, which are becoming a widespread, have millions of parameters that result in many costly mathematical computations that are restrictive on the kind of computational hardware that can be used by researchers. Furthermore, the use of 3D deep learning models increases the computational and memory requirements by large margins. All of the reviewed literature use deep learning software libraries to provide an infrastructure to define and train deep neural networks in parallel or distributed manner while leveraging multi-core or multi-GPU environments. Currently, researchers are being limited by the amount of GPU memory at their disposal (typically 12 gigabytes). For this reason, batch sizes and model complexities are being limited to what can fit into the available memory.

The performance of brain tumor segmentation algorithms have continued to increase over the past few years due to the availability of more training data and use of more sophisticated CNN architectures and training schemes. However, their robustness is still lagging behind expert performance [105]. Recently, researchers have used the ensemble methods to achieve state-of-the-art performance (see Table 3). Precisely, the ensemble methods fuse the segmentation results of several models to improve the robustness of individual approach, resulting in superior performance as compared to inter-rater agreements [105]. Interestingly, single Unet [37] based models [91] continue to produce exceptional performance, supporting the argument that: "*a well trained UNet is hard to beat*" [88]. The reviewed literature have shown that careful initialization of hyper-parameters, a selection of pre-processing techniques, employing advanced training schemes, as well as dealing with the class imbalance problem will immensely improve the accuracy and robustness of segmentation algorithms.

5. Summary

This paper has discussed several building blocks, state-of-the-art techniques, and tools for implementing automatic brain tumor segmentation algorithms. Despite the tremendous advance in the field, the robustness of deep learning methods are still inferior to expert performance. Some notable architectures, including ensemble methods and UNet based models, have shown great potential for improving the state-of-the-art with careful pre-

processing, weight initialization, advanced training schemes, and techniques in order to address inherent class imbalance problems. The lack of a large-scale medical training dataset is the leading factor in many segmentation algorithms' poor performance.

Author Contributions: Conceptualization, T.M. and S.V.; methodology, T.M. and S.V; formal analysis, S.V.; investigation, T.M.; resources, S.V.; writing original draft preparation, T.M.; writing review and editing, S.V.; supervision, S.V. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data available in publicly accessible repositories.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Havaei, M.; Davy, A.; Warde-Farley, D.; Biard, A.; Courville, A.; Bengio, Y.; Pal, C.; Jodoin, P.M.; Larochelle, H. Brain tumor segmentation with Deep Neural Networks. *Med. Image Anal.* **2017**, *35*, 18–31. [CrossRef] [PubMed]
2. Menze, B.H.; Jakab, A.; Bauer, S.; Kalpathy-Cramer, J.; Farahani, K.; Kirby, J.; Burren, Y.; Porz, N.; Slotboom, J.; Wiest, R.; et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans. Med. Imaging* **2015**, *34*, 1993–2024. [CrossRef] [PubMed]
3. Işın, A.; Direkoğlu, C.; Şah, M. Review of MRI-Based Brain Tumor Image Segmentation Using Deep Learning Methods. *Procedia Comput. Sci.* **2016**, *102*, 317–324. [CrossRef]
4. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.J.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **2017**, *36*, 61–78. [CrossRef] [PubMed]
5. Razzak, M.I.; Imran, M.; Xu, G. Efficient Brain Tumor Segmentation With Multiscale Two-Pathway-Group Conventional Neural Networks. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 1911–1919. [CrossRef]
6. Muhammad, K.; Khan, S.; Ser, J.D.; de Albuquerque, V.H.C. Deep Learning for Multigrade Brain Tumor Classification in Smart Healthcare Systems: A Prospective Survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, 1–16. [CrossRef]
7. Wadhwa, A.; Bhardwaj, A.; Singh Verma, V. A review on brain tumor segmentation of MRI images. *Magn. Reson. Imaging* **2019**, *61*, 247–259. [CrossRef]
8. Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S.; Hasan, M.; Van Essen, B.C.; Awwal, A.A.S.; Asari, V.K. A State-of-the-Art Survey on Deep Learning Theory andArchitectures. *Electronics* **2019**, *8*, 292. [CrossRef]
9. Zikic, D.; Ioannou, Y.; Brown, M.; Criminisi, A. Segmentation of Brain Tumor Tissues with Convolutional Neural Networks. In Proceedings of the BRATS-MICCAI, Boston, MA, USA, 14 September 2014; pp. 36–39.
10. Urban, G.; Bendszus, M.; Hamprecht, F.A.; Kleesiek, J. Multi-Modal Brain Tumor Segmentation Using Deep Convolutional Neural Networks. In Proceedings of the BRATS-MICCAI, Boston, MA, USA, 14 September 2014; pp. 31–35.
11. Shen, D.; Wu, G.; Suk, H.I. Deep learning in medical image analysis. *Annu. Rev. Biomed.* **2017**, *19*, 221–248. [CrossRef]
12. Havaei, M.; Guizard, N.; Larochelle, H.; Jodoin, P.M. Deep Learning Trends for Focal Brain Pathology Segmentation in MRI. In *Machine Learning for Health Informatics*; Holzinger, A., Ed.; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; Volume 9605, pp. 125–148. [CrossRef]
13. Corso, J.J.; Sharon, E.; Dube, S.; El-Saden, S.; Sinha, U.; Yuille, A. Efficient Multilevel Brain Tumor Segmentation With Integrated Bayesian Model Classification. *IEEE Trans. Med. Imaging* **2008**, *27*, 629–640. [CrossRef]
14. Gordillo, N.; Montseny, E.; Sobrevilla, P. State of the Art Survey on MRI Brain Tumor Segmentation. *Magn. Reson. Imaging* **2013**, *31*, 1426–1438. [CrossRef] [PubMed]
15. Chen, L.; Bentley, P.; Mori, K.; Misawa, K.; Fujiwara, M.; Rueckert, D. DRINet for Medical Image Segmentation. *IEEE Trans. Med. Imaging* **2018**, *37*, 2453–2462. [CrossRef] [PubMed]
16. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; Adaptive Computation and Machine Learning; The MIT Press: Cambridge, MA, USA, 2016.
17. Chollet, F. *Deep Learning with Python*; Manning Publications Co.: Shelter Island, NY, USA, 2018.
18. Svozil, D.; Kvasnicka, V.; Pospichal, J. Introduction to Multi-Layer Feed-Forward Neural Networks. *Chemom. Intell. Lab. Syst.* **1997**, *39*, 43–62. [CrossRef]
19. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.

20. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. In *Proceedings of the Advances in Neural Information Processing Systems*; Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2012; Volume 25, pp. 1097–1105.
21. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2015**, arXiv:1409.1556.
22. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385.
23. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-Excitation Networks. *arXiv* **2019**, arXiv:1709.01507.
24. Glorot, X.; Bordes, A.; Bengio, Y. Deep Sparse Rectifier Neural Networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings, Ft. Lauderdale, FL, USA, 11–13 April 2011; pp. 315–323.
25. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier Nonlinearities Improve Neural Network Acoustic Models. In Proceedings of the ICML Workshop on Deep Learning for Audio, Speech and Language Processing, Atlanta, GA, USA, 16 June 2013.
26. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *arXiv* **2015**, arXiv:1502.01852.
27. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *arXiv* **2015**, arXiv:1411.4038.
28. Bernal, J.; Kushibar, K.; Asfaw, D.S.; Valverde, S.; Oliver, A.; Martí, R.; Lladó, X. Deep Convolutional Neural Networks for Brain Image Analysis on Magnetic Resonance Imaging: A Review. *Artif. Intell. Med.* **2019**, *95*, 64–81. [CrossRef]
29. Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the Importance of Initialization and Momentum in Deep Learning. In *Proceedings of the 30th International Conference on Machine Learning*; Dasgupta, S., McAllester, D., Eds.; PMLR: Atlanta, GA, USA, 2013; Volume 28, pp. 1139–1147.
30. Duchi, J.; Hazan, E.; Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* **2011**, *12*, 2121–2159.
31. Zeiler, M.D. ADADELTA: An Adaptive Learning Rate Method. *arXiv* **2012**, arXiv:1212.5701.
32. Tieleman, T.; Hinton, G. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA Neural Netw. Mach. Learn.* **2012**, *4*, 26–31.
33. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
34. Wang, G.; Li, W.; Ourselin, S.; Vercauteren, T. Automatic Brain Tumor Segmentation Based on Cascaded Convolutional Neural Networks With Uncertainty Estimation. *Front. Comput. Neurosci.* **2019**, *13*, 56. [CrossRef] [PubMed]
35. Li, H.; Li, A.; Wang, M. A novel end-to-end brain tumor segmentation method using improved fully convolutional networks. *Comput. Biol. Med.* **2019**, *108*, 150–160. [CrossRef]
36. Cahall, D.E.; Rasool, G.; Bouaynaya, N.C.; Fathallah-Shaykh, H.M. Inception Modules Enhance Brain Tumor Segmentation. *Front. Comput. Neurosci.* **2019**, *13*, 44. [CrossRef]
37. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv* **2015**, arXiv:1505.04597.
38. Glorot, X.; Bengio, Y. Understanding the Difficulty of Training Deep Feedforward Neural Networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010; Volume 9, pp. 249–256.
39. Bengio, Y.; Lamblin, P.; Popovici, D.; Larochelle, H. Greedy Layer-Wise Training of Deep Networks. In *Advances in Neural Information Processing Systems 19*; Schölkopf, B., Platt, J.C., Hoffman, T., Eds.; MIT Press: Cambridge, MA, USA, 2007; pp. 153–160.
40. Claesen, M.; De Moor, B. Hyperparameter Search in Machine Learning. *arXiv* **2015**, arXiv:1502.02127.
41. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
42. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv* **2015**, arXiv:1502.03167.
43. Pereira, S.; Pinto, A.; Alves, V.; Silva, C.A. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Trans. Med. Imaging* **2016**, *35*, 1240–1251. [CrossRef] [PubMed]
44. Myronenko, A. 3D MRI Brain Tumor Segmentation Using Autoencoder Regularization. *arXiv* **2018**, arXiv:1810.11654.
45. Rezaei, M.; Harmuth, K.; Gierke, W.; Kellermeier, T.; Fischer, M.; Yang, H.; Meinel, C. Conditional Adversarial Network for Semantic Segmentation of Brain Tumor. *arXiv* **2017**, arXiv:1708.05227.
46. Rao, V.; Sarabi, M.S.; Jaiswal, A. Brain tumor segmentation with deep learning. In Proceedings of the MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), 2015; pp. 56–59. Available online: https://www.researchgate.net/profile/Mona_Sharifi2/publication/309456897_Brain_tumor_segmentation_with_deep_learning/links/5b444445458515f71cb8a65d/Brain-tumor-segmentation-with-deep-learning.pdf (accessed on 1 June 2020).
47. Casamitjana, A.; Puch, S.; Aduriz, A.; Sayrol, E.; Vilaplana, V. 3D Convolutional Networks for Brain Tumor Segmentation. In Proceedings of the MICCAI Challenge on Multimodal Brain Tumor Image Segmentation (BRATS), 2016; pp. 65–68. Available online: https://imatge.upc.edu/web/sites/default/files/pub/cCasamitjana16.pdf (accessed on 1 June 2020).
48. Hussain, S.; Anwar, S.M.; Majid, M. Brain Tumor Segmentation Using Cascaded Deep Convolutional Neural Network. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, Korea, 11–15 July 2017; pp. 1998–2001. [CrossRef]

49. Pereira, S.; Oliveira, A.; Alves, V.; Silva, C.A. On hierarchical brain tumor segmentation in MRI using fully convolutional neural networks: A preliminary study. In Proceedings of the 2017 IEEE 5th Portuguese Meeting on Bioengineering (ENBENG), Coimbra, Portugal, 16–18 February 2017; pp. 1–4. [CrossRef]
50. Sled, J.; Zijdenbos, A.; Evans, A. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* **1998**, *17*, 87–97. [CrossRef]
51. Tustison, N.J.; Avants, B.B.; Cook, P.A.; Zheng, Y.; Egan, A.; Yushkevich, P.A.; Gee, J.C. N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imaging* **2010**, *29*, 1310–1320. [CrossRef]
52. Nyul, L.; Udupa, J.; Zhang, X. New variants of a method of MRI scale standardization. *IEEE Trans. Med. Imaging* **2000**, *19*, 143–150. [CrossRef]
53. Zhao, X.; Wu, Y.; Song, G.; Li, Z.; Zhang, Y.; Fan, Y. A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. *Med. Image Anal.* **2018**, *43*, 98–111. [CrossRef]
54. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *arXiv* **2016**, arXiv:1606.04797.
55. Vaidhya, K.; Thirunavukkarasu, S.; Alex, V.; Krishnamurthi, G. Multi-Modal Brain Tumor Segmentation Using Stacked Denoising Autoencoders. In *Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Menze, B., Maier, O., Reyes, M., Handels, H., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 181–194. [CrossRef]
56. Pereira, S.; Pinto, A.; Alves, V.; Silva, C.A. Deep Convolutional Neural Networks for the Segmentation of Gliomas in Multi-Sequence MRI. In *Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Menze, B., Maier, O., Reyes, M., Handels, H., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 131–143. [CrossRef]
57. Hussain, S.; Anwar, S.M.; Majid, M. Segmentation of glioma tumors in brain using deep convolutional neural network. *Neurocomputing* **2018**, *282*, 248–261. [CrossRef]
58. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. *arXiv* **2018**, arXiv:1708.02002.
59. Sun, L.; Zhang, S.; Chen, H.; Luo, L. Brain Tumor Segmentation and Survival Prediction Using Multimodal MRI Scans with Deep Learning. *Front. Neurosci.* **2019**, *13*, 810. [CrossRef] [PubMed]
60. Mlynarski, P.; Delingette, H.; Criminisi, A.; Ayache, N. Deep learning with mixed supervision for brain tumor segmentation. *J. Med. Imaging* **2019**, *6*, 034002. [CrossRef] [PubMed]
61. Iqbal, S.; Ghani Khan, M.U.; Saba, T.; Mehmood, Z.; Javaid, N.; Rehman, A.; Abbasi, R. Deep learning model integrating features and novel classifiers fusion for brain tumor segmentation. *Microsc. Res. Tech.* **2019**, *82*, 1302–1315. [CrossRef] [PubMed]
62. Kuzina, A.; Egorov, E.; Burnaev, E. Bayesian Generative Models for Knowledge Transfer in MRI Semantic Segmentation Problems. *Front. Neurosci.* **2019**, *13*, 844. [CrossRef]
63. Kao, P.Y.; Ngo, T.; Zhang, A.; Chen, J.W.; Manjunath, B.S., Brain Tumor Segmentation and Tractographic Feature Extraction from Structural MR Images for Overall Survival Prediction. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; Volume 11384, pp. 128–141. [CrossRef]
64. Kao, P.Y.; Shailja, F.; Jiang, J.; Zhang, A.; Khan, A.; Chen, J.W.; Manjunath, B.S. Improving Patch-Based Convolutional Neural Networks for MRI Brain Tumor Segmentation by Leveraging Location Information. *Front. Neurosci.* **2020**, *13*. [CrossRef]
65. Nalepa, J.; Marcinkiewicz, M.; Kawulok, M. Data Augmentation for Brain-Tumor Segmentation: A Review. *Front. Comput. Neurosci.* **2019**, *13*, 83. [CrossRef]
66. Shin, H.C.; Tenenholtz, N.A.; Rogers, J.K.; Schwarz, C.G.; Senjem, M.L.; Gunter, J.L.; Andriole, K.; Michalski, M. Medical Image Synthesis for Data Augmentation and Anonymization Using Generative Adversarial Networks. *arXiv* **2018**, arXiv:1807.10225.
67. Han, C.; Rundo, L.; Araki, R.; Nagano, Y.; Furukawa, Y.; Mauri, G.; Nakayama, H.; Hayashi, H. Combining Noise-to-Image and Image-to-Image GANs: Brain MR Image Augmentation for Tumor Detection. *IEEE Access* **2019**, *7*, 156966–156977. [CrossRef]
68. Han, C.; Murao, K.; Noguchi, T.; Kawata, Y.; Uchiyama, F.; Rundo, L.; Nakayama, H.; Satoh, S. Learning More with Less: Conditional PGGAN-Based Data Augmentation for Brain Metastases Detection Using Highly-Rough Annotation on MR Images. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing China, 3–7 November 2019; pp. 119–127. [CrossRef]
69. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *arXiv* **2014**, arXiv:1406.2661.
70. Simpson, A.L.; Antonelli, M.; Bakas, S.; Bilello, M.; Farahani, K.; van Ginneken, B.; Kopp-Schneider, A.; Landman, B.A.; Litjens, G.; Menze, B.; et al. A Large Annotated Medical Image Dataset for the Development and Evaluation of Segmentation Algorithms. *arXiv* **2019**, arXiv:1902.09063.
71. Team, T.T.D.; Al-Rfou, R.; Alain, G.; Almahairi, A.; Angermueller, C.; Bahdanau, D.; Bastien, F.; Bayer, J.; Belikov, A.; Belopolsky, A.; et al. Theano: A Python Framework for Fast Computation of Mathematical Expressions. *arXiv* **2016**, arXiv:1605.02688.
72. Goodfellow, I.J.; Warde-Farley, D.; Lamblin, P.; Dumoulin, V.; Mirza, M.; Pascanu, R.; Bergstra, J.; Bastien, F.; Bengio, Y. Pylearn2: A Machine Learning Research Library. *arXiv* **2013**, arXiv:1308.4214.
73. Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; Darrell, T. Caffe: Convolutional Architecture for Fast Feature Embedding. *arXiv* **2014**, arXiv:1408.5093.

74. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *arXiv* **2019**, arXiv:1912.01703.
75. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. Tensorflow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv* **2016**, arXiv:1603.04467.
76. Chollet, F. Keras: The Python Deep Learning API. 2020. Available online: https://keras.io/ (accessed on 1 June 2020).
77. Zhang, J.; Shen, X.; Zhuo, T.; Zhou, H. Brain tumor segmentation based on refined fully convolutional neural networks with a hierarchical dice loss. *arXiv* **2017**, arXiv:1712.09093.
78. Kayalibay, B.; Jensen, G.; Smagt, P.V.D. CNN-based segmentation of medical imaging data. *arXiv* **2017**, arXiv:1701.03056.
79. Kamnitsas, K.; Bai, W.; Ferrante, E.; McDonagh, S.; Sinclair, M.; Pawlowski, N.; Rajchl, M.; Lee, M.; Kainz, B.; Rueckert, D.; et al. Ensembles of Multiple Models and Architectures for Robust Brain Tumour Segmentation. *arXiv* **2017**, arXiv:1711.01468.
80. Dong, H.; Yang, G.; Liu, F.; Mo, Y.; Guo, Y. Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. In *Proceedings of the Medical Image Understanding and Analysis*; Valdés Hernández, M., González-Castro, V., Eds.; Communications in Computer and Information Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2017; pp. 506–517. [CrossRef]
81. Alex, V.; Safwan, M.; Krishnamurthi, G. Automatic Segmentation and Overall Survival Prediction in Gliomas Using Fully Convolutional Neural Network and Texture Analysis. *arXiv* **2017**, arXiv:1712.02066.
82. Erden, B.; Gamboa, N.; Wood, S. *3D Convolutional Neural Network for Brain Tumor Segmentation*; Technical Report; Computer Science, Stanford University: Stanford, CA, USA, 2017.
83. Isensee, F.; Kickingereder, P.; Wick, W.; Bendszus, M.; Maier-Hein, K.H. Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge. *arXiv* **2018**, arXiv:1802.10508.
84. Meng, Z.; Fan, Z.; Zhao, Z.; Su, F. ENS-Unet: End-to-End Noise Suppression U-Net for Brain Tumor Segmentation. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 5886–5889. [CrossRef]
85. Liu, J.; Chen, F.; Pan, C.; Zhu, M.; Zhang, X.; Zhang, L.; Liao, H. A Cascaded Deep Convolutional Neural Network for Joint Segmentation and Genotype Prediction of Brainstem Gliomas. *IEEE Trans. Bio-Med. Eng.* **2018**, *65*, 1943–1952. [CrossRef] [PubMed]
86. Pereira, S.; Pinto, A.; Amorim, J.; Ribeiro, A.; Alves, V.; Silva, C.A. Adaptive feature recombination and recalibration for semantic segmentation with Fully Convolutional Networks. *IEEE Trans. Med. Imaging* **2019**. [CrossRef] [PubMed]
87. Kermi, A.; Mahmoudi, I.; Khadir, M.T. Deep Convolutional Neural Networks Using U-Net for Automatic Brain Tumor Segmentation in Multimodal MRI Volumes. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 37–48. [CrossRef]
88. Isensee, F.; Kickingereder, P.; Wick, W.; Bendszus, M.; Maier-Hein, K.H. No New-Net. *arXiv* **2019**, arXiv:1809.10483.
89. Wang, L.; Wang, S.; Chen, R.; Qu, X.; Chen, Y.; Huang, S.; Liu, C. Nested Dilation Networks for Brain Tumor Segmentation Based on Magnetic Resonance Imaging. *Front. Neurosci.* **2019**, *13*, 285. [CrossRef]
90. Ribalta Lorenzo, P.; Nalepa, J.; Bobek-Billewicz, B.; Wawrzyniak, P.; Mrukwa, G.; Kawulok, M.; Ulrych, P.; Hayball, M.P. Segmenting brain tumors from FLAIR MRI using fully convolutional neural networks. *Comput. Methods Programs Biomed.* **2019**, *176*, 135–148. [CrossRef]
91. Jiang, Z.; Ding, C.; Liu, M.; Tao, D. Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task. In *Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Bakas, S., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 231–241. [CrossRef]
92. Zhao, Y.X.; Zhang, Y.M.; Liu, C.L. Bag of Tricks for 3D MRI Brain Tumor Segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Lecture Notes in Computer Science; Crimi, A., Bakas, S., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 210–220. [CrossRef]
93. Zhuge, Y.; Krauze, A.V.; Ning, H.; Cheng, J.Y.; Arora, B.C.; Camphausen, K.; Miller, R.W. Brain tumor segmentation using holistically nested neural networks in MRI images. *Med. Phys.* **2017**, *44*, 5234–5243. [CrossRef]
94. Liu, Y.; Stojadinovic, S.; Hrycushko, B.; Wardak, Z.; Lau, S.; Lu, W.; Yan, Y.; Jiang, S.B.; Zhen, X.; Timmerman, R.; et al. A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery. *PLoS ONE* **2017**, *12*, e0185844. [CrossRef]
95. Li, Z.; Wang, Y.; Yu, J.; Guo, Y.; Cao, W. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci. Rep.* **2017**, *7*, 5467. [CrossRef]
96. Kamnitsas, K.; Chen, L.; Ledig, C.; Rueckert, D.; Glocker, B. Multi-Scale 3D Convolutional Neural Networks for Lesion Segmentation in Brain MRI. *Ischemic Stroke Lesion Segm.* **2015**, *13*, 46.
97. Hoseini, F.; Shahbahrami, A.; Bayat, P. AdaptAhead Optimization Algorithm for Learning Deep CNN Applied to MRI Segmentation. *J. Digit. Imaging* **2019**, *32*, 105–115. [CrossRef]
98. Naceur, M.B.; Saouli, R.; Akil, M.; Kachouri, R. Fully Automatic Brain Tumor Segmentation using End-To-End Incremental Deep Neural Networks in MRI images. *Comput. Methods Programs Biomed.* **2018**, *166*, 39–49. [CrossRef] [PubMed]
99. Yi, D.; Zhou, M.; Chen, Z.; Gevaert, O. 3-D convolutional neural networks for glioblastoma segmentation. *arXiv* **2016**, arXiv:1611.04534.

100. Cui, S.; Mao, L.; Jiang, J.; Liu, C.; Xiong, S. Automatic Semantic Segmentation of Brain Gliomas from MRI Images Using a Deep Cascaded Neural Network. *J. Healthc. Eng.* **2018**, *2018*, 4940593. [CrossRef] [PubMed]
101. Yang, T.; Ou, Y.; Huang, T. Automatic Segmentation of Brain Tumor from MR Images Using SegNet: Selection of Training Data Sets. In Proceedings of the 6th MICCAI BraTS Challenge, Quebec City, QC, Canada, 14 September 2017; pp. 309–312.
102. McKinley, R.; Meier, R.; Wiest, R. Ensembles of Densely-Connected CNNs with Label-Uncertainty for Brain Tumor Segmentation. In *Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 456–465. [CrossRef]
103. Zhou, C.; Chen, S.; Ding, C.; Tao, D. Learning Contextual and Attentive Information for Brain Tumor Segmentation. In *Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 497–507. [CrossRef]
104. McKinley, R.; Rebsamen, M.; Meier, R.; Wiest, R. Triplanar Ensemble of 3D-to-2D CNNs with Label-Uncertainty for Brain Tumor Segmentation. In *Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Bakas, S., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 379–387. [CrossRef]
105. Bakas, S.; Reyes, M.; Jakab, A.; Bauer, S.; Rempfler, M.; Crimi, A.; Shinohara, R.T.; Berger, C.; Rozycki, M.; Prastawa, M.; et al. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. *arXiv* **2019**, arXiv:1811.02629.
106. Ji, Z.; Shen, Y.; Ma, C.; Gao, M. Scribble-Based Hierarchical Weakly Supervised Learning for Brain Tumor Segmentation. *arXiv* **2019**, arXiv:1911.02014.
107. Pavlov, S.; Artemov, A.; Sharaev, M.; Bernstein, A.; Burnaev, E. Weakly Supervised Fine Tuning Approach for Brain Tumor Segmentation Problem. *arXiv* **2019**, arXiv:1911.01738.
108. Wu, K.; Du, B.; Luo, M.; Wen, H.; Shen, Y.; Feng, J. Weakly Supervised Brain Lesion Segmentation via Attentional Representation Learning. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2019*; Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.T., Khan, A., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 211–219. [CrossRef]

2.2 Conclusion

In conclusion, this chapter has provided a comprehensive overview of the fundamental building blocks, state-of-the-art techniques, and tools for implementing automatic brain tumor segmentation algorithms. Despite the remarkable strides made in the field, it is evident that the robustness of current deep-learning methods falls short of expert performance. Noteworthy architectures such as ensemble methods and UNet-based models exhibit promising potential, especially with meticulous preprocessing, refined weight initialization, and advanced training schemes to mitigate inherent class imbalance challenges. Significantly, the pervasive issue of limited large-scale medical training datasets remains a critical factor contributing to the suboptimal performance of numerous segmentation algorithms. As we delve further into this thesis, exploring and integrating innovative methodologies will be essential to bridge the existing gaps and propel the field toward more accurate and reliable automatic brain tumor segmentation solutions.

Chapter 3

Automatic Brain Tumor Segmentation

3.1 Introduction

Brain tumors, particularly gliomas, represent a formidable medical challenge, being the most common and aggressive form with median survival rates of less than two years for the highest grade. Accurate segmentation of these tumors is crucial for effective treatment planning and diagnosis. Yet, it remains a complex task due to their diverse locations, varying shapes, and the diffusing nature of aggressive tumors. Traditional segmentation methods relying on pixel intensity values face limitations in capturing the intricate boundaries of such lesions. In response to these challenges, deep learning methods have emerged as powerful tools for automatic brain tumor segmentation, offering superior performance despite the associated computational complexities.

This chapter addresses the need for efficient yet effective 3D brain tumor segmentation solutions. While deep learning methods, especially encoder-decoder networks like U-Nets, have shown remarkable performance, the computational demands raise concerns for practical applications. In this context, we propose a novel network architecture that leverages depthwise separable convolutions to enhance computational efficiency without compromising segmentation accuracy. The study also extends the U-Net model within the nnU-Net framework, introducing bottleneck units and a shuffle attention mechanism to mitigate the impact of loss segmentation accuracy due to the reduction of parameter count.

3.2 Brain Tumor Segmentation Using Partial Depthwise Separable Convolutions

3.2.1 Introduction

In fully automatic brain tumor segmentation, recent research efforts encounter significant constraints imposed by the available computation budget. The limitations imposed by GPU memory capacity have led to a necessary compromise, with batch sizes and model complexities restricted to what can be accommodated within these computational bounds. The utilization of 3D MRI volumes, recognized for their empirical superiority over 2D counterparts, introduces additional challenges due to large patch sizes, rendering the

training of these models difficult and, in some cases, impractical.

There is a pressing need for computational and memory-efficient models in response to these constraints, focusing on enhancing the feasibility of computer-assisted diagnosis, particularly in resource-constrained environments and developing countries. This section introduces an efficient network architecture for 3D brain tumor segmentation. Our proposed model incorporates depthwise separable convolutions to alleviate computational burdens, fostering improved efficiency without compromising segmentation performance. We provide a comparative analysis with state-of-art methods.

Part of this work was published in[1].

[1]T. Magadza and S. Viriri, "Brain Tumor Segmentation Using Partial Depthwise Separable Convolutions," in IEEE Access, vol. 10, pp. 124206-124216, 2022, doi: 10.1109/ACCESS.2022.3223654.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 October 2022, accepted 10 November 2022, date of publication 21 November 2022,
date of current version 30 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3223654

RESEARCH ARTICLE

Brain Tumor Segmentation Using Partial Depthwise Separable Convolutions

TIRIVANGANI MAGADZA AND SERESTINA VIRIRI, (Senior Member, IEEE)
School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban 4041, South Africa
Corresponding author: Serestina Viriri (viriris@ukzn.ac.za)

ABSTRACT Gliomas are the most common and aggressive form of all brain tumors, with medial survival rates of less than two years for the highest grade. While accurate and reproducible segmentation of brain tumors is paramount for an effective treatment plan and diagnosis, automatic brain tumor segmentation is challenging because the lesion can appear anywhere in the brain with varying shapes and sizes from one patient to another. Moreover, segmentation is only done by analyzing pixel intensity values of surrounding tissues, and the diffusing nature of aggressive brain tumors makes it even more challenging to delineate tumor boundaries. Nevertheless, deep learning methods have superior performance in automatic brain tumor segmentation. However, their boost in performance comes at the cost of high computational complexity. This paper proposes efficient network architecture for 3D brain tumor segmentation, partially utilizing depthwise separable convolutions to reduce computational costs. The experimental results on the BraTS 2020 dataset show that our methods could achieve comparable results with the state-of-the-art methods with minimum computational complexity. Furthermore, we provide a critical analysis of the current efficient model designs. The code for this project is available at https://github.com/tmagadza/partialDepthwiseNet.

INDEX TERMS Brain tumor segmentation, deep learning, depth-wise separable convolution, magnetic resonance imaging, 3D U-Net.

I. INTRODUCTION

Gliomas are adults' most common primary tumors. Although their exact causes are still a mystery [1], risk factors include exposure to ionizing radiation and a family history of tumors. These tumors can appear anywhere in the brain with varying shapes and sizes, making them difficult to segment. The World Health Organization (WHO) has classified the tumors into four grades, from grade I to grade IV, depending on growth and aggressiveness. Low-grade gliomas (LGG), which constitute grades I and II, are less aggressive and have survival rates of several years. While high-grade gliomas (HGG) (grade III and IV) are much more aggressive and have median survival rates of less than two years even after treatment.

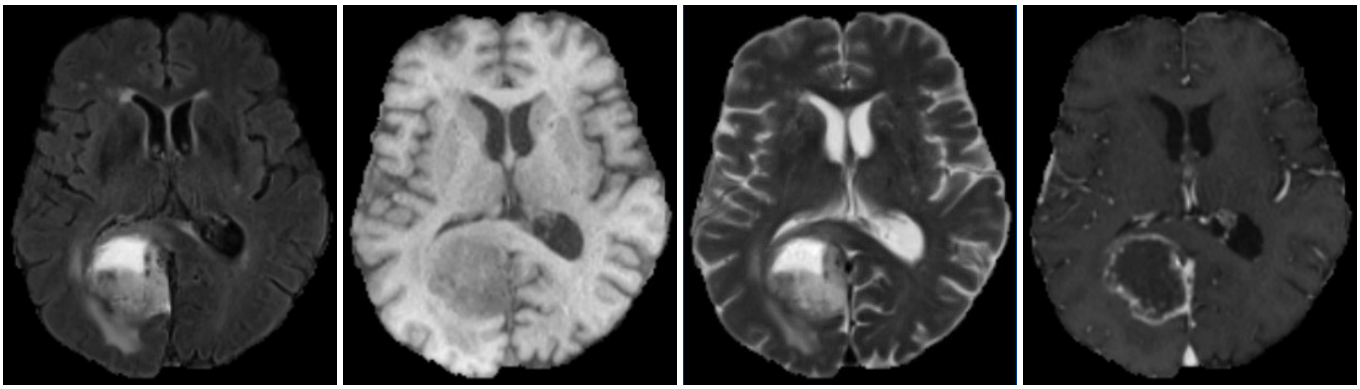
Magnetic Resonance Imaging (MRI) has emerged as the imaging technology of choice for brain tumor diagnosis, and treatment planning [2]. Non-invasive MRI scans produce high-resolution and soft tissue 3D volumes. As depicted in Fig. 1, more than one MRI slices are used to view different tumor regions.

In clinical practice, highly trained radiologists do brain tumor segmentation manually. Although manual segmentation arguably produces the most accurate segmentation results, it suffers from intra, and inter-rater variability [2], [3]. Moreover, it is tedious and time-consuming, and results depend on the radiologist's experience and knowledge. To this end, manual segmentation is mainly used for visual inspection and is a gold standard for semi-automatic and fully automatic segmentation.

Meanwhile, automatic segmentation methods require little to no human involvement. They have the benefits of being objective, reproducible, and well-suited for quantitative assessment of brain tumors. They have shown great potential in improving diagnosis and treatment planning.

Recently deep learning methods, particularly the Convolutional Neural Networks (CNNs) [4], [5], [6], [7], are being used to automatically analyze brain scans (usually

The associate editor coordinating the review of this manuscript and approving it for publication was Wai-keung Fung.



(a) FLAIR (b) T1 (c) T2 (d) T1ce

FIGURE 1. Examples of different MRI imaging modalities.

MRI scans) due to their record-shattering performance. They require no feature engineering: they automatically learn features directly from data. However, these methods have high memory and computation complexity. Furthermore, they require a huge amount of training data for better performance, which is a challenge in medical imaging.

Currently, research efforts in fully automatic brain tumor segmentation are limited to the available computation budget [8]. Batch sizes and model complexities are now being limited to what can fit into the available GPU memory. The use of 3D MRI volumes with large patch sizes in CNN models, which were empirically shown to outperform 2D counterparts, makes it even more difficult, if not impossible, to train these models.

Therefore, to improve the adoption rate of computer-assisted diagnosis in clinical setups, especially in developing countries, there is a need for more computational and memory-efficient models. Luckily, there has been an increase of research efforts to optimize the current state-of-the-art deep learning models in computer vision task [9], [10], [11], [12], [13], [14].

The contributions of this research work are:

1) We proposed efficient network architecture for 3D brain tumor segmentation, partially utilizing depthwise separable convolutions to reduce computational costs.
2) We quantitatively analyze the computational complexity of the proposed method and compare the segmentation performance with the state-of-the-art.
3) We provide critical analysis of the latest methods that employ efficient model design.

The rest of the paper is organized as follows: Section II reviews related work in efficient networks. Section III describes the proposed architecture for an optimized 3D brain tumor segmentation. Section IV presents the experimental results discussed in Section V. Lastly; Section VI provides concluding remarks.

II. LITERATURE REVIEW AND RELATED WORKS

Brain tumor segmentation is the process of classifying every pixel in a medical image as a normal or tumorous pixel. The process is done before and after treatment to determine the disease's progression and evaluate the effectiveness of the chosen treatment strategy [2]. It is very challenging to accurately segment brain tumor for several reasons: (1) segmentation is only achieved by the analysis of intensity variations between surrounding tissues [2], (2) brain tumors comes in various shapes and sizes from one patient to another, (2) aggressive brain tumors often diffuse into surrounding normal tissues making it even more difficult to delineate tumor boundaries. Fig 1 clearly shows that a single imaging modality is insufficient to delineate tumor boundaries accurately. When done manually, brain tumor segmentation is tedious and suffers from intra, and inter-rater variability [3]. Accurate and reproducible segmentation of brain tumors is critical for effective treatment planning, diagnosis, and monitoring of disease progression. In recent years, computer-assisted diagnosis has become mainstream in assisting medical practitioners in interpreting medical images [8], [15], [16]. While there are several methods for the automatic segmentation of brain tumors, deep learning methods are becoming widespread in the medical imaging domain [17] due to their resounding performance. However, the boost in performance comes at the cost of high computational complexity, as we shall see later.

Among the deep learning family, U-Net architecture [18] has emerged as the architecture of choice, primarily for the semantic segmentation of medical images. The architecture is composed of downsampling and upsampling paths. The downsampling path, which resembles a typical convolutional network, is used for feature extraction. At the same time, the upsampling path is used to recover the spatial resolution lost during feature extraction. The network heavily depends on data augmentation for better generalization. Since its inception in 2015, the architecture has inspired many research efforts in medical imaging. In [19], the U-Net network was improved to take 3D volumes as input to fully exploit the volumetric data inherent in medical images. However, volumetric segmentation substantially increases the computation requirements. Kamnistas et al. [20] proposed

TABLE 1. A summary of recent works for automatic brain tumor segmentation with computational analysis.

Reference	Year	Type	Batch size	Epochs	Params (M)	FLOPS (G)	GPU (G)	Input	Techniques
Chen et al. [29]	2019	3D	12	500	3.88	27	44	$128x128x128$	Channel grouping, Multi-fiber, Dilated Convolution
Cheng et al. [33]	2020	3D	1			227	12	$144x160x128$	Multitask learning (Multi-branch Decoder)
Peng et al. [30]	2020	3D	1	100	1.24	121	11	$128x128x128$	Depthwise separable convolution
Nguyen et al. [34]	2020	3D	4	500	1.38	15	16	$128x128x128$	Dilated multi-fiber
Wang et al. [35]	2021	3D	16	8000	15.14	208	192	$128x128x128$	Transformer, self-attention mechanism
Zhou et al. [36]	2021	3D	2	500	17.3	371	8	$128x128x128$	ShuffleNetV2
Luo et al. [37]	2021	3D	10	800	0.29	24	24	$128x128x128$	Hierarchical decoupled convolutions
Jia et al. [38]	2021	3D	4	450	26.07	621	44	$128x128x128$	Combination of single cascaded models
Liu et al. [39]	2021	3D		500	5.23	36	11	$128x128x128$	Learnable Group convolution and deep supervision
Li et al. [40]	2021	3D	8	500	0.71	10	44	$128x128x128$	Multi-branch sharing network
Xiao et al. [41]	2021	3D	9	900	0.35	31	33	$128x128x128$	Multi-view fusion convolution
Fang et al. [42]	2021	2D	16	50	72.8	38	16	$160x160$	Self-supervised
Tong [43]	2022	2D		75	0.62	146	16	$200x168$	Multipath feature extraction
Sun and Wang [44]	2022	2D		50	0.2	29	12	$168x200$	depthwise convolution
Li et al. [40]	2022	3D	8	500	4.77	151	96	$128x128x128$	Supervised Attention Module
Yang et al. [45]	2022	3D	2	400	5	5	16	$128x128x128$	modality disentanglement
Raza et al. [46]	2022	3D	4	100	30.47	374	16	$128x128x128$	Modified UNet
Jia et al. [47]	2022	3D	1	250	17.91	450	96	$128x128x128$	Cascaded multi-scale fusion, attention mechanism
Cai et al. [48]	2022	3D	2		27.7	902	40	$128x128x128$	hierarchical fully connected module
Zhang et al. [49]	2022	3D	1	1000	106	748	64	$128x128x128$	incomplete multimodal learning, Intra-modal Transformer
Hu et al. [50]	2022	3D	6	1000	28.35	76	22	$64x64x64$	Ensemble learning
Subhan Akbar et al. [51]	2022	2.5D	8	900	0.168	5	11	$16x196x196$	Attention mechanisms, dilated convolution
Liang et al. [52]	2022	3D	4	400	66.7	9	22	$128x128x128$	U-Shaped Transformer network

an ensemble of multiple heterogeneous models (including the U-Net-based models) for robust semantic segmentation. Despite winning the BraTS 2017 challenge, their model is highly inefficient as each model has to be trained separately. In [21], Wang et al. exploited the hierarchical nature of brain tumor structures by proposing a cascade of U-Net models. Isensee et al. [22] incorporated context and localization modules for better segmentation performance. Myronenko [23], the winner of the BraTS 2018 challenge, used an autoencoder to regularize a shared decoder in the U-Net variant. His model suffered from high computational complexity due to the large patch size ($160x192x128$), standard convolutional operations, and additional overhead due to the use of an autoencoder. Isensee et al. [7] clearly showed that a U-Net architecture with minor alterations can achieve superior performance. However, large patch sizes ($128x128x128$) and standard convolutional operations will result in high computational and memory requirements. Jiang et al. [6] proposed a cascaded U-Net that took advantage of the hierarchical nature of brain tumor substructure. Despite winning the BraTS 2019 Challenge, their model is still computationally expensive. Zhao et al. [4] exploited various heuristics in data processing, model designing, and optimization to improve segmentation performance. Their work came second in the BraTS 2019 Challenge. Isensee et al. [24], the winner of BraTS 2020 challenge, used the nnU-Net framework [25] with BraTS specific modifications in post-processing, region-based training, and data argumentation demonstrating the competitiveness of the U-Net model. The models that follow the encoder-decoder-like structure, as in the U-Net have achieved state-of-the-art performance. However, most of the works focused mainly on improving the segmentation performance and the expense of the computational complexity. In this work, we introduced yet another U-Net model that follows on the works by Myronenko [23] and Ellis and Aizenberg [26] for a more efficient volumetric segmentation.

To learn recent trends in efficient model design for brain tumor segmentation, we performed a Google Scholar search for recent works with *efficient* in their title or mentioned *FLOP* in their body for a period from 2018 to 2022. Fig 2. depicts the results of the search. The figure clearly shows that of 1630 works for brain tumor segmentation, only 39 (2%) reported on the computational complexity of their methods. Surprisingly, of 44 works with *efficient* in their title report, only 8 (18%) reported on the computational efficiency of their models. These results indicate that the majority of works emphasize more on improving segmentation performance while sacrificing computational costs.

In Table 1, we summarized the works that provided an analysis of the computational complexity of their methods which is measured by the number of parameters, floating-point operations per second (FLOPS), and the GPU memory requirements for a given model. From the table, most works for the period use 3D patches with input size cropped from $240x240x155$ to $128x128x128$ pixels to fit on the GPU memory. The batch size depends on the available GPU memory. Since a large patch size consumes much of the memory, the researcher has to make the trade-off between increasing the batch size and reducing the input patch size, which in turn hurts the segmentation performance [23]. Another way is to maintain the large patch size and increase the number of GPUs. In reality, most researchers have a very

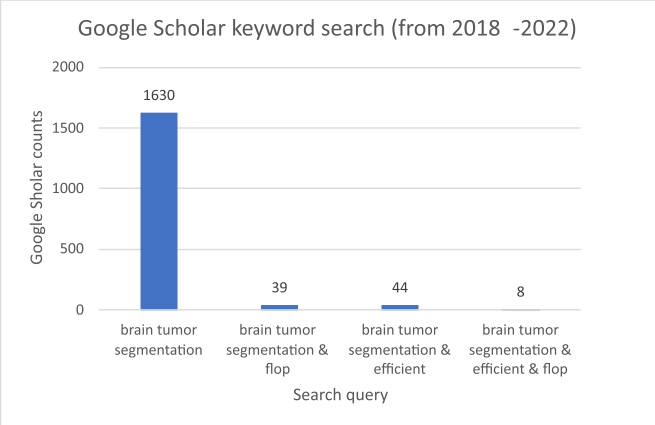


FIGURE 2. Results of Google Scholar searches for a period from 2018 to 2022 for articles with the following search queries: ***brain tumor segmentation*** [in title], ***brain tumor segmentation*** [in title] & ***flop*** [in all fields], ***brain tumor segmentation efficient*** [in title] and ***brain tumor segmentation efficient*** [in title] & ***flop***.

tight computational budget. We have observed that several works [27], [28], [29], [30] exploited channel grouping to minimize the interaction between the feature maps when performing convolutional operations, thereby the reducing the number of parameters and FLOPs.

Our work is inspired by depthwise separable convolutions introduced by Sifre and Mallat in [31] and subsequently used to improve the efficiency and reduce the model size of 2D convolutional networks in [10] and [9]. Furthermore, we extensively use residual connections introduced by He et al. [32] to improve the flow of gradients in deep networks.

III. METHODS AND TECHNIQUES

A. STANDARD CONVOLUTION

Consider the input feature maps $\mathbf{I} \in \mathbb{R}^{h\times w\times d\times c}$, where h, w, d, and c are the height, width, depth and number of channels of the input feature maps respectively, and the convolutional kernel $\mathbf{K} \in \mathbb{R}^{k\times k\times k\times c\times n}$, where k is the size of the convolutional kernel and n is the number of output channels. The operation of a standard convolutional layer $\mathbf{O} \in \mathbb{R}^{h\times w\times d\times n} = \mathbf{K} * \mathbf{I}$ is given by:

$$\mathbf{O}(y,x,z,j) = \sum_{i=1}^{c}\sum_{u,v,w=1}^{k} \mathbf{K}(u,v,w,i,j). \\ \mathbf{I}(y+u-1, x+v-1, z+w-1, i). \quad (1)$$

where $1 \le y \le h, 1 \le x \le w, 1 \le z \le d, 1 \le j \le n$. The computational complexity of a convolutional layer in terms of the number of multiplications is

$$nck^3hwd. \quad (2)$$

The complexity of the standard convolution is cubic, with the kernel size limiting the kernel size of most CNN in medical image analysis to $3 \times 3 \times 3$.

B. DEPTHWISE SEPARABLE CONVOLUTION

The depthwise separable convolution splits the standard convolutional operation into depthwise and pointwise convolutions. First, it independently applies a spatial convolution to each input channel. It then performs a 1×1 convolution to combine the results. A standard convolution performs these operations in a single pass. Factorization of the convolutional operation has the benefit of improving efficiency and reducing the model size.

Depthwise convolution with one filter per input channel can be expressed as

$$\mathbf{O}_D(y,x,z,c) = \sum_{u,v,w=1}^{k} \mathbf{K}_D(u,v,w,c). \\ \mathbf{I}(y+u-1, x+v-1, z+w-1, c). \quad (3)$$

where $\mathbf{K}_D \in \mathbb{R}^{k\times k\times k\times c}$ is the depthwise convolutional kernel where the c_{th} filter in $\mathbf{K}_D$ is applied to the c_{th} channel in $\mathbf{I}$ to produce the c_{th} of the output feature map $\mathbf{O}_D \in \mathbb{R}^{h\times w\times d\times c}$. The computational cost of the depthwise convolution is:

$$ck^3hwd. \quad (4)$$

whereas a pointwise convolution can be expressed as:

$$\hat{\mathbf{O}}(y,x,z,n) = \sum_{i=1}^{c} \mathbf{K}_P(i,n)\mathbf{O}_D(y,x,z,i). \quad (5)$$

where $\mathbf{K}_P \in \mathbb{R}^{1\times 1\times 1\times c\times n}$ is the pointwise convolutional kernel. The computational complexity of this operation is, therefore:

$$nchwd. \quad (6)$$

The combination of depthwise convolution and pointwise (1×1) convolution is called the depthwise separable convolution. The computational complexity of the depthwise separable convolution is

$$ck^3hwd + nchwd. \quad (7)$$

C. MODEL ARCHITECTURE

Our work follows a 3D U-Net [19] structure as shown in Fig. 3. The network is made up of five layers, with two ResNet-like [32] style convolutional blocks in both the encoding and decoding path. The encoding path takes in a random four-channel 3D MRI patch with a receptive field of $128 \times 128 \times 128$. Each layer along the encoding path reduces the spatial resolution by half using stride convolution and doubles the number of the channels starting with a base width of 32 channels. As in [26], each residual block consists of two consecutive convolutional blocks performing group normalization, followed by rectified linear unit activation, and a $3 \times 3 \times 3$ convolution (see fig 5a).

Along the decoding path, each layer reduces the number of feature maps by half before upscaling the spacial resolution using trilinear interpolation and concatenates the result with gated high-resolution feature maps from the encoding path.

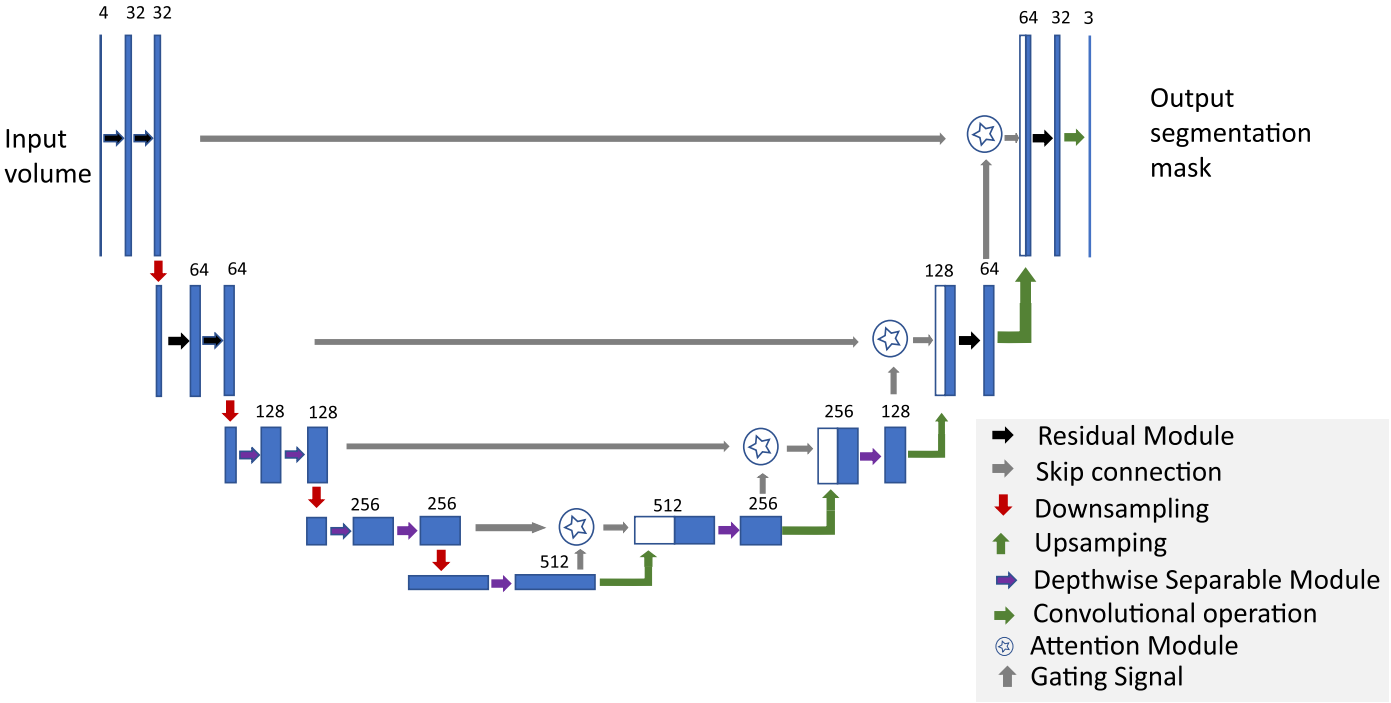


FIGURE 3. Schematic visualization of the network architecture. Input is a four-channel 3D MRI crop, followed by double residual modules with a base of 32 filters. The bottom three layers replaced all the convolutional blocks in residual modules with depthwise separable convolutions. Skip connections are rescaled by learned weights from the attention module. The network's output has three channel segmentation maps (with the same spatial resolution as the input) followed by the sigmoid activation function. (adapted from [18]).

In the last layer, the network uses a $1 \times 1 \times 1$ convolution to reduce the number of feature maps to three, followed by a sigmoid activation function.

To improve the computational efficiency of the network, one can replace all the standard $3 \times 3 \times 3$ convolutions in residual modules with depthwise separable convolutions. However, empirical studies reviewed that the group convolutions in PyTorch[1] deep learning framework, which models 3D depthwise separable convolutions, tend to use more GPU memory than standard convolutions. Therefore, to allow our network to fit available GPU memory, we only replaced the bottom three layers of the network with depthwise separable convolutions. Fig 5b depicts the structure of the depthwise separable module.

D. ATTENTION MECHANISM

In deep learning, the attention mechanism forces the network to focus more on certain input parts while suppressing the rest. We adopted the spatial attention [53] on skip connections to enhance salient feature responses and suppress noisy ones before concatenating with feature response from the decoding path. The module combines feature responses from the skip connections and the decoding path to learning gating weights and then applies them to the skip connections feature responses. See Fig. 4 for the structure and operations performed by the spatial attention module.

[1]https://pytorch.org/

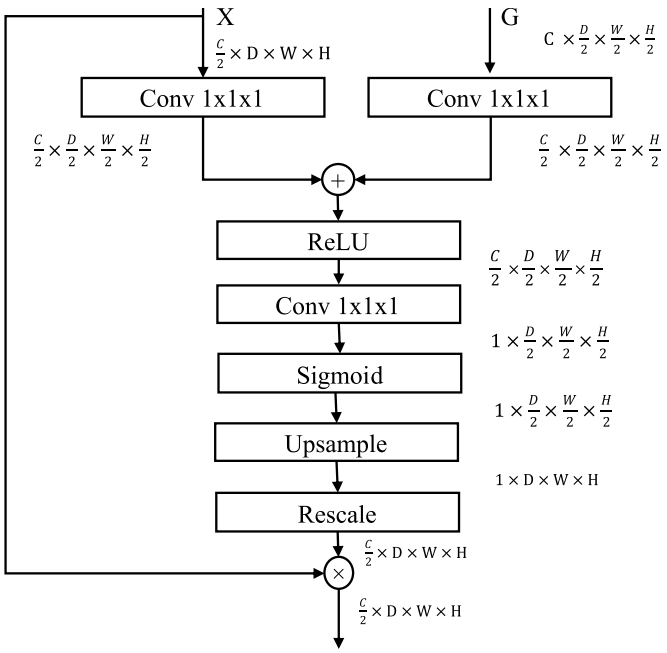


FIGURE 4. Illustration of the spatial attention module with two inputs *X* (skip connection) and *G* (gating signal from coarse scale). The module outputs weighted feature responses from the skip connections.

E. LOSS

We use the multi-class soft dice loss:

$$\mathbf{L}_{dice} = 1 - \frac{2}{c} \sum \frac{\sum y_{true} * y_{pred}}{\sum y_{true}^2 + \sum y_{pred}^2 + \epsilon} \tag{8}$$

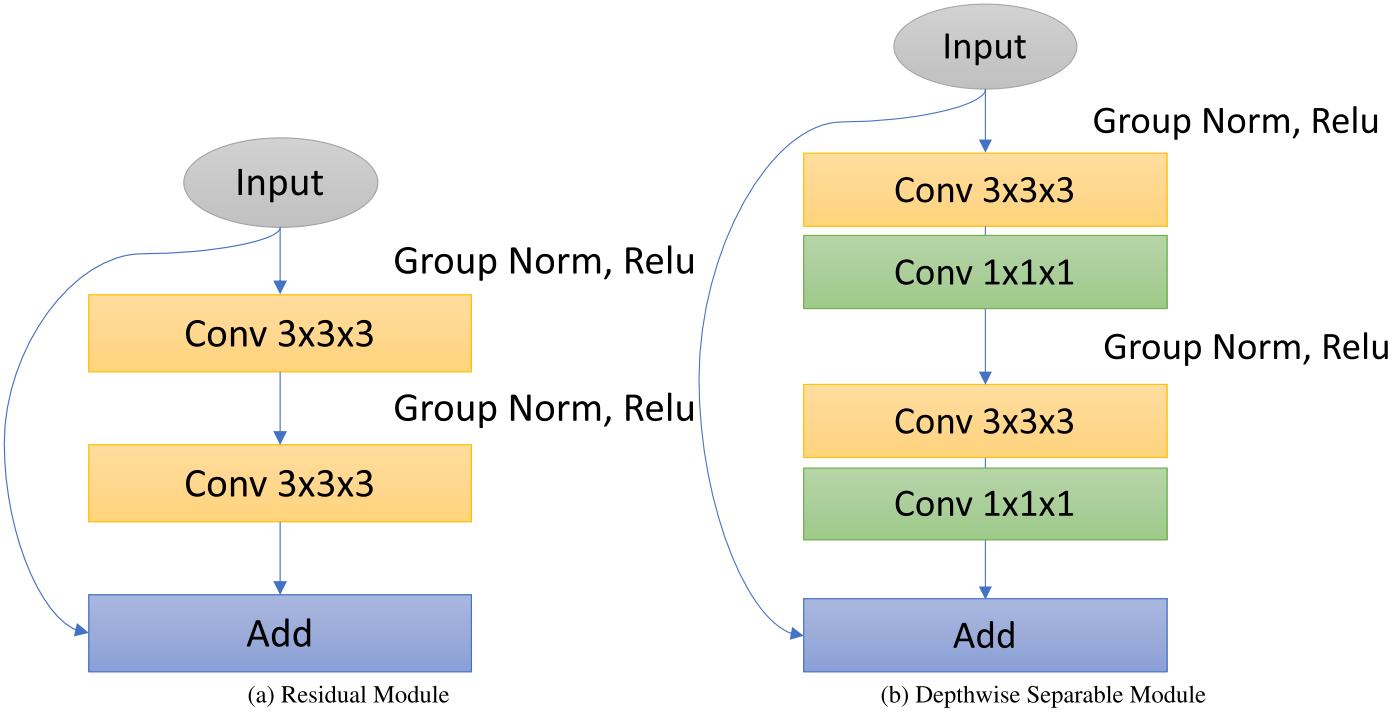


FIGURE 5. Comparison of different convolutional blocks incorporated in the network.

where $\mathbf{L}_{dice} \in \mathbb{R}$ is the mean loss across c classes, $y_{true} \in \mathbb{R}^{c\times n\times h\times w\times d}$ is the ground truth, $y_{true} \in \mathbb{R}^{c\times n\times h\times w\times d}$ is the predicted segmentation maps, and ϵ is a small value to prevent division by zero.

F. DATA AUGMENTATION

Data augmentation is an effective technique to increase the training dataset, thereby improving model generalization ability. In this paper, we apply data augmentations techniques that are relatively easy to implement and have low computational complexity. Specifically, we adopted the data augmentation scheme of Ellis and Aizenberg [26]. Random Gaussian noise and blurring were applied to input images with a 50% probability per training iteration. Input images were independently randomly scaled on each axis, with a standard deviation of 0.1 and a 50% probability per training iteration. Moreover, images were randomly flipped and translated independently of each direction.

IV. EXPERIMENTS AND RESULTS

A. DATA AND IMPLEMENTATION DETAILS

We used BraTS 2020 [2], [54], [55] dataset with 369 training and 125 validation subjects. Each training subject contains native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR) volumes, along with manually labeled tumor segmentation maps. The validation set contains all the multimodal scans except the ground truth annotations, as in the training set. We evaluated the performance of our model on the validation set through submissions of segmentation maps to the BraTS challenge online portal.[2] All the scans in both the training and validation sets were co-registered to the same anatomical template, interpolated to the same resolution ($1mm^3$), and skull-stripped.

Our network was implemented in Pytorch[3] using an open source deep learning framework[4] [26]. We used the Adam optimizer with an initial learning rate of $\alpha = 1e-4$, which was decreased by a factor of 0.5 every time the validation loss plateaued for 20 epochs and a weight decay of $1e-3$. The batch size was 2. We trained our network on an NVIDIA Tesla V100 16GB GPU. The code for this project is available at https://github.com/tmagadza/partialDepthwiseNet.

TABLE 2. Comparative performance of the baseline and proposed model. Average computational requirements for each model trained for 100 epochs.

Model	Size (MB)	Training Time (h)	Pred Time (s)	Params (M)	FLOPs (G)
3DU-Net [26]	268	8.6	1.4	23	828
Proposed (with att.)	26	**6.8**	**1.3**	6.9	617
Proposed (without att.)	**25**	**6.8**	**1.3**	**6.7**	**616**

Best values are shown in bold. att. - attention mechanism.

B. SIZE AND SPEED

In Table 2, we compare the size and speed of the baseline model and the proposed method. We used the network

[2]https://ipp.cbica.upenn.edu/
[3]www.pytorch.com
[4]https://github.com/ellisdg/3DUnetCNN

TABLE 3. Ablation analysis of our proposed network on the BraTS 2020 Validation set in terms of Dice Similarity Coefficient. All models were trained for 100 epochs. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core. BL - baseline model, DS - Depthwise separable module, AT - attention module, WD - Weight decay, Ensemble10 - An ensemble of 10 models. †: 5 fold cross-validation. *: Input size of 96 × 128 × 80.

	Dice Similarity Coefficient			
Model	ET	WT	TC	Mean
BL	0.7492	0.8934	0.8190	0.8205
BL+DS	0.7589	0.8958	0.8160	0.8236
BL†+DS	0.7692	0.9020	0.8190	0.8301
BL+DS+AT	0.7486	0.8975	0.8084	0.8181
BL*+DS+AT	0.7622	0.8958	0.8199	0.8260
BL+DS+AT+WD	0.7744	0.8976	0.8238	0.8319
Ensemble10	**0.7745**	**0.9042**	**0.8286**	**0.8357**

Best values are shown in bold, and underlined are second best.

TABLE 4. Ablation analysis of our proposed network on the BraTS 2020 Validation set measured by Hausdorff distance (95%). All models were trained for 100 epochs. BL - baseline model, DS - Depthwise separable module, AT - attention module, WD - weight Decay, Ensemble10 - Ensemble of 10 models. †: 5-fold cross-validation. *: Input size of 96 × 128 × 80.

	Hausdorff distance (95%)			
Model	ET	WT	TC	Mean
BL	22.49	6.22	13.55	14.09
BL+DS	27.95	6.09	10.78	14.94
BL†+DS	27.79	5.90	9.95	14.55
BL+DS+AT	33.32	6.44	7.54	15.76
BL*+DS+AT	**22.33**	**5.51**	7.19	**11.68**
BL+DS+AT+WD	29.82	6.78	7.36	14.65
Ensemble10	25.07	5.61	**7.10**	12.60

Best values are shown in bold.

TABLE 5. Mean performance metrics on BraTS 2020 Validation dataset of our proposed method as compared to the state-of-the-art methods in terms of dice similarity score. We trained our model for 400 epochs. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core. Ensemble of 10 models.

	Dice			
Method	ET	WT	TC	Mean
Isensee et al. [56]	**0.7989**	0.9124	0.8506	**0.8540**
Jia et al. [38]	0.7875	**0.9129**	**0.8546**	0.8517
Y. Yuan [57]	0.7927	0.9108	0.8529	0.8521
Wang et al. [58]	0.7873	0.9009	0.8173	0.8352
Ensemble10 (ours)	0.7745	0.9042	0.8286	0.8357

Best values are shown in bold

architecture proposed by Ellis and Aizenberg [26] as the baseline model. All the models were trained for 100 epochs. Our model outperforms the baseline model in all metrics. The proposed model substantially decreases the model size and parameter count by roughly 90% and 70%, respectively. Moreover, it needed lesser time to complete 100 epochs of training. Removal of the attention mechanism barely reduces the computational complexity of the proposed method.

C. ABLATION STUDY

We performed an ablation analysis to determine the performance contribution of each component of the proposed network. We trained each model for 100 epochs on the BraTS 2020 validation set while maintaining all other network parameters constant. To improve segmentation performance on the enhancing tumor, we replaced all enhancing tumor voxels with necrosis if the total number of predicted voxels were less than a threshold of 300 voxels. We refer to the stripped-down version of our proposed model as a baseline. To maintain consistency with other previous works, we only report on metrics computed by the online evaluation platform (https://ipp.cbica.upenn.edu/).

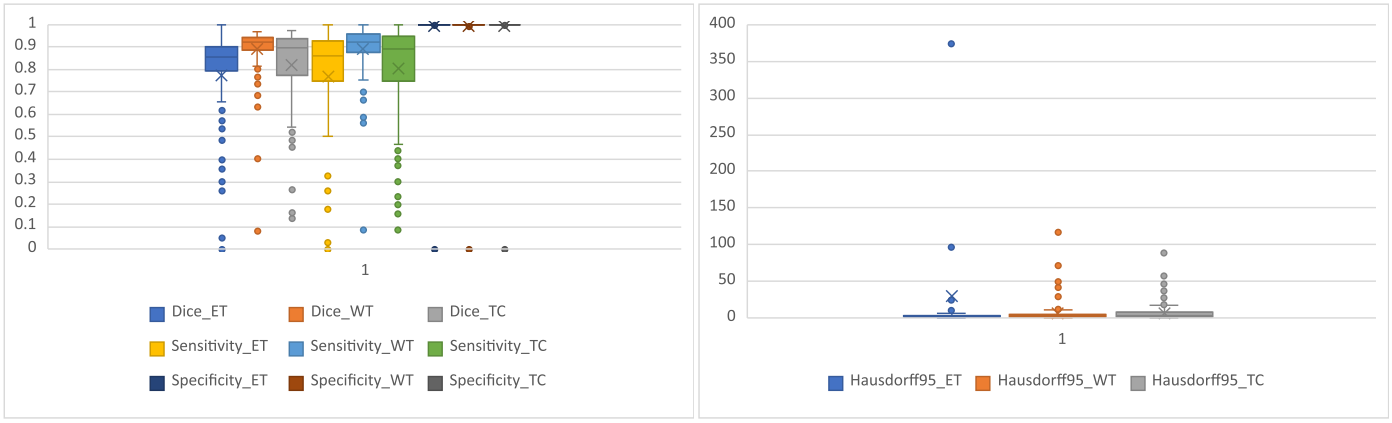
Table 3 shows the Dice Similarity Coefficient results on the BraTS 2020 validation set. The performance of the baseline in all regions was quite strong. Adding depthwise separable modules improved the dice scores marginally for the enhancing and whole tumor regions. We observed more gain when we trained the model with 5-fold cross-validation. Adding the attention mechanism decreased dice scores for the enhancing tumor and tumor core regions. However, by reducing the receptive field to $96 \times 128 \times 80$, we observed an interesting boost in dice scores for the enhancing tumor. Applying $L2$ weight regularization to the proposed model resulted in good segmentation performance in all tumor regions. Moreover, there was an increase in performance by creating an ensemble of 10 models (5 single models + 5 models resulting from 5-fold cross-validation) aggregated by hierarchical majority vote.

Table 4 reports the performance of the proposed network as measured by the Hausdorff distance (95%) metric. Interestingly, our proposed model trained with small input patches outperformed all models, including the ensemble of 10 models in all tumor regions. We observed a reduced Hausdorff distance in tumor core regions due to attention mechanism and weight regularization. The ensemble of the model did not yield many expected benefits save for the tumor core regions only.

Fig. 6a and 6b show the box plots results of the proposed methods on BraTS 2020 validation dataset. It can be seen in Fig. 6a that the predictions of our method in all metrics are left-skewed, indicating that the predictions are concentrated in higher areas. The model shows a very high ability to predict background voxel very well. The plots also show very low fluctuations in the whole tumor predictions, indicating segmentation of whole tumor regions is fairly easy. On the other hand, our model exhibit comparatively high variability in the sensitivity of the enhancing tumor. Fig 6b shows that our model has very low variability in terms of the Hausdorff distance (95%) metric.

D. COMPARISON WITH THE STATE-OF-THE-ART

Table 5 reports on the dice similarity score performance of our models trained for 100 epochs against previous methods using the BraTS 2020 dataset. The online evaluation platform computed all metrics. No single model outperformed all methods in all metrics. Our model ensemble performed better than the method proposed by Wang et al. [5] overall and in both the whole tumor and the tumor core regions.



(a) Dice Similarity Score

(b) Hausdorff distance (95%)

FIGURE 6. Box plots for the BraTS 2020 Validation results of our proposed method.

TABLE 6. Mean performance metrics on BraTS 2020 Validation dataset of our proposed method as compared to the state-of-the-art methods in terms of Hausdorff distance (95%). We trained our model for 400 epochs. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core. BL - baseline model, DS - Depthwise separable module, AT - attention module, Ensemble10 - Ensemble of 10 models. *: Input size of 96 × 128 × 80.

Method	Hausdorff95			
	ET	WT	TC	Mean
Isensee et al. [56]	23.50	**3.69**	7.82	11.67
Jia et al. [38]	26.58	4.18	**4.97**	11.91
Y. Yuan [57]	18.20	4.10	5.99	**9.43**
Wang et al. [58]	**17.95**	4.96	9.77	10.89
BL*+DS+AT (ours)	22.33	5.51	7.19	11.68
Ensemble10 (ours)	25.07	5.61	7.10	12.60

Best values are shown in bold.

Table 6 gives an aggregate summary of the performance of our methods in terms of 95% Hausdorff distance (mm) against previous methods. Again, no single method outperformed all methods in all regions. An ensemble of 11 models by Yuan [57] achieved the best performance overall. Our single model trained with small input patches performed well on this metric again. Specifically, It outperformed the ensemble of 25 models by Isensee et al. [56] in both the enhancing tumor and tumor core regions. It also performed well in the tumor core region as compared to the method by Wang et al. [58].

V. DISCUSSIONS

Accurate and reproducible segmentation of brain tumors is paramount for an effective treatment plan and diagnosis. Deep learning methods have shown promising results as compared to the inter-rater agreement. While several state-of-the-art automatic brain tumor segmentation exists in the literature, most focus on improving segmentation results at the cost of high computational complexity. Some works tried to incorporate techniques known to enhance network efficiency, like residual learning [32] in their design. We believe more emphasis should place on efficient model design as well. A competitive and lightweight model will result in cost savings in the long run. For example, the HPC Cluster[5] we use to train the model poses a 12h limit for each job. Moreover, every user falls under a Principal Investigator who applies for CPU-h resource allocation for their research programme. Thus, one would prefer the best accuracy under a limited computational budget. Table 2 clearly shows that our method needs less time to train and requires just 26MB of disk space. Often the best-performing models are an ensemble of multiple models, which will result in more bandwidth utilization if the trained weights are to be moved to another location. For example, the nnU-Net model[6] used by Isensee et al. [56] to win the BraTS 2020 Challenge, comprises 25 models, which amount to 2 Gig in compressed form. In real-life situations where the model is trained is not usually where it will be deployed. For these reasons, we have proposed an efficient network incorporating the depthwise separable modules to reduce the model size and the parameter number while improving training and inference speed. Specifically, we replaced the convolution blocks of the bottom three layers of the U-Net structure with depthwise separable convolutions. We evaluated the performance of our network on the BraTS 2020 dataset. Results show that our model significantly reduced the model size and the number of parameters by more significant margins than the baseline model (as shown in Table 1).

As for the segmentation results, our model performed poorly in dice scores for the enhancing tumor. This is a common problem [30] that may be caused by an intratumoral class imbalance since LGG images do not have an enhancing region. One way of addressing the issue is to replace the enhancing tumor with necrosis if the prediction of enhancing tumor class is less than a certain threshold [6]. In Table 4, we observed substantial improvement in the Hausdorff distance (95%) score in all tumor regions when we trained our proposed model with small patch sizes. Moreover, qualitative inspection of randomly selected predictions on the training

[5]https://www.chpc.ac.za/

[6]https://zenodo.org/record/4003545#.Y1emJHZBzcc

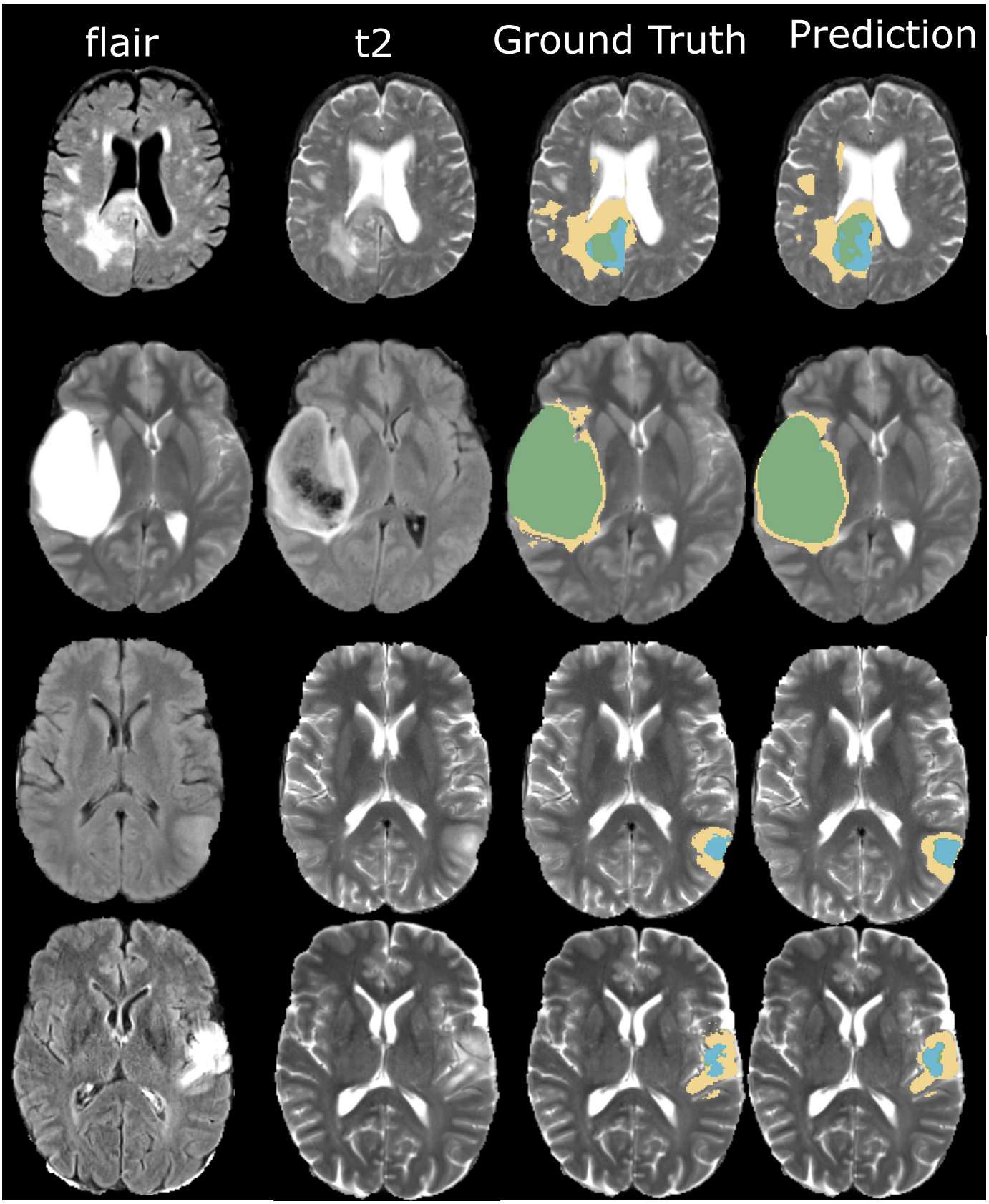


FIGURE 7. Qualitative inspections of two randomly selected predictions on the training set. Edema is shown in yellow, necrosis in green, and enhancing tumor in blue.

set (see Fig. 7) reviews that our model sometimes gives highly accurate segmentation and, on the other, performs poorly. The use of model ensemble [2] is known to mitigate the problem.

VI. CONCLUSION

This paper proposes an efficient model for brain tumor segmentation using partial Depthwise Separable Convolutions. Our proposed network partially replaced some convolutional blocks in a standard U-Net structure with depthwise separable blocks. The experimental results on the BraTS 2020 dataset show that our methods could achieve comparable results with the state-of-the-art methods with minimum computational complexity. Additionally, we have provided an extensive computational analysis of current methods. In the future, we will explore the fusing of multiple resolutions to capture long-range dependencies to improve segmentation performance.

REFERENCES

[1] D. Cahill and S. Turcan, "Origin of gliomas," *Seminars Neurol.*, vol. 38, no. 1, pp. 5–10, 2018.

[2] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, and L. Lanczi, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6975210/

[3] A. Işın, C. Direkoğlu, and M. Şah, "Review of MRI-based brain tumor image segmentation using deep learning methods," *Proc. Comput. Sci.*, vol. 102, pp. 317–324, Jan. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S187705091632587X

[4] Y.-X. Zhao, Y.-M. Zhang, and C.-L. Liu, "Bag of tricks for 3D MRI brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2020, pp. 210–220.

[5] F. Wang, R. Jiang, L. Zheng, C. Meng, and B. Biswal, "3D U-Net based brain tumor segmentation and survival days prediction," 2019, *arXiv:1909.12901*.

[6] Z. Jiang, C. Ding, M. Liu, and D. Tao, "Two-stage cascaded U-Net: 1st place solution to BraTS challenge 2019 segmentation task," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2020, pp. 231–241.

[7] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "No new-net," 2018, *arXiv:1809.10483*.

[8] T. Magadza and S. Viriri, "Deep learning for brain tumor segmentation: A survey of state-of-the-art," *J. Imag.*, vol. 7, no. 2, p. 19, Jan. 2021. [Online]. Available: https://www.mdpi.com/2313-433X/7/2/19

[9] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[10] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807. [Online]. Available: http://ieeexplore.ieee.org/document/8099678/

[11] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," 2016, *arXiv:1602.07360*.

[12] M. Wang, B. Liu, and H. Foroosh, "Factorized convolutional neural networks," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 545–553. [Online]. Available: http://ieeexplore.ieee.org/document/8265281/

[13] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation networks," 2017, *arXiv:1709.01507*.

[14] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," 2017, *arXiv:1707.01083*.

[15] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, no. 1, pp. 221–248, Mar. 2017.

[16] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[17] B. Sahiner, A. Pezeshk, L. M. Hadjiiski, X. Wang, K. Drukker, K. H. Cha, R. M. Summers, and M. L. Giger, "Deep learning in medical imaging and radiation therapy," *Med. Phys.*, vol. 46, no. 1, pp. e1–e36, Jan. 2019.

[18] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," 2015, *arXiv:1505.04597*.

[19] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," 2016, *arXiv:1606.06650*.

[20] K. Kamnitsas, W. Bai, E. Ferrante, S. McDonagh, M. Sinclair, N. Pawlowski, M. Rajchl, M. Lee, B. Kainz, D. Rueckert, and B. Glocker, "Ensembles of multiple models and architectures for robust brain Tumour segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi, S. Bakas, H. Kuijf, B. Menze, and M. Reyes, Eds. Cham, Switzerland: Springer, 2018, pp. 450–462.

[21] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," 2017, *arXiv:1709.00382*.

[22] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain tumor segmentation and radiomics survival prediction: Contribution to the BRATS 2017 challenge," 2018, *arXiv:1802.10508*.

[23] A. Myronenko, "3D MRI brain tumor segmentation using autoencoder regularization," 2018, *arXiv:1810.11654*.

[24] F. Isensee, P. F. Jager, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnU-Net for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 118–132.

[25] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Dec. 2020. [Online]. Available: http://www.nature.com/articles/s41592-020-01008-z

[26] D. G. Ellis and M. R. Aizenberg, "Trialing U-Net training modifications for segmenting gliomas using open source deep learning framework," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 40–49.

[27] C. Zhao, Z. Zhao, Q. Zeng, and Y. Feng, "MVP U-Net: Multi-view pointwise U-Net for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 93–103.

[28] X. Zhou, X. Li, K. Hu, Y. Zhang, Z. Chen, and X. Gao, "ERV-Net: An efficient 3D residual neural network for brain tumor segmentation," *Expert Syst. Appl.*, vol. 170, May 2021, Art. no. 114566. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417421000075

[29] C. Chen, X. Liu, M. Ding, J. Zheng, and J. Li, "3D dilated multi-fiber network for real-time brain tumor segmentation in MRI," 2019, *arXiv:1904.03355*.

[30] S. Peng, W. Chen, J. Sun, and B. Liu, "Multi-scale 3D U-nets: An approach to automatic segmentation of brain tumor," *Int. J. Imag. Syst. Technol.*, vol. 30, no. 1, pp. 5–17, Mar. 2020, doi: 10.1002/ima.22368.

[31] L. SIfre and S. Mallat, "Rigid-motion scattering for texture classification," 2014, *arXiv:1403.1687*.

[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, *arXiv:1512.03385*.

[33] G. Cheng, J. Cheng, M. Luo, L. He, Y. Tian, and R. Wang, "Effective and efficient multitask learning for brain tumor segmentation," *J. Real-Time Image Process.*, vol. 17, no. 6, pp. 1951–1960, Dec. 2020, doi: 10.1007/s11554-020-00961-4.

[34] T. H. Nguyen, C. H. Le, D. V. Sang, T. Yao, W. Li, and Z. Wang, "Efficient brain tumor segmentation with dilated multi-fiber network and weighted bi-directional feature pyramid network," in *Proc. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov. 2020, pp. 1–7.

[35] W. Wang, C. Chen, M. Ding, H. Yu, S. Zha, and J. Li, "TransBTS: Multimodal brain tumor segmentation using transformer," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021* (Lecture Notes in Computer Science), M. de Bruijne, P. C. Cattin, S. Cotin, N. Padoy, S. Speidel, Y. Zheng, and C. Essert, Eds. Springer, 2021, pp. 109–119.

[36] X. Zhou, X. Li, K. Hu, Y. Zhang, Z. Chen, and X. Gao, "ERV-Net: An efficient 3D residual neural network for brain tumor segmentation," *Expert Syst. Appl.*, vol. 170, May 2021, Art. no. 114566. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0957417421000075
[37] Z. Luo, Z. Jia, Z. Yuan, and J. Peng, "HDC-Net: Hierarchical decoupled convolution network for brain tumor segmentation," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 3, pp. 737–745, Mar. 2021.
[38] H. Jia, W. Cai, H. Huang, and Y. Xia, "H$$Ĵ2$$NF-Net for brain tumor segmentation using multimodal MR imaging: 2nd place solution to BraTS challenge 2020 segmentation task," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 58–68.
[39] H. Liu, Q. Li, and I.-C. Wang, "A deep-learning model with learnable group convolution and deep supervision for brain tumor segmentation," *Math. Problems Eng.*, vol. 2021, pp. 1–11, Feb. 2021. [Online]. Available: https://www.hindawi.com/journals/mpe/2021/6661083/
[40] J. Li, H. Yu, C. Chen, M. Ding, and S. Zha, "Category guided attention network for brain tumor segmentation in MRI," *Phys. Med. Biol.*, vol. 67, no. 8, Apr. 2022, Art. no. 085014, doi: 10.1088/1361-6560/ac628a.
[41] Z. Xiao, K. He, J. Liu, and W. Zhang, "Multi-view hierarchical split network for brain tumor segmentation," *Biomed. Signal Process. Control*, vol. 69, Aug. 2021, Art. no. 102897. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1746809421004948
[42] F. Fang, Y. Yao, T. Zhou, G. Xie, and J. Lu, "Self-supervised multi-modal hybrid fusion network for brain tumor segmentation," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 11, pp. 5310–5320, Nov. 2022.
[43] J. Tong, "A performance-consistent and computation-efficient CNN system for high-quality brain tumor segmentation," M.S. thesis, Concordia Univ., Montreal, QC, Canada, 2022.
[44] Y. Sun and C. Wang, "A computation-efficient CNN system for high-quality brain tumor segmentation," *Biomed. Signal Process. Control*, vol. 74, Apr. 2022, Art. no. 103475. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1746809421010727
[45] Q. Yang, X. Guo, Z. Chen, P. Y. M. Woo, and Y. Yuan, "D2-Net: Dual disentanglement network for brain tumor segmentation with missing modalities," *IEEE Trans. Med. Imag.*, vol. 41, no. 10, pp. 2953–2964, Oct. 2022.
[46] R. Raza, U. I. Bajwa, Y. Mehmood, M. W. Anwar, and M. H. Jamal, "DResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 103861. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1746809422003809
[47] H. Jia, C. Bai, W. Cai, H. Huang, and Y. Xia, "HNF-Netv2 for brain tumor segmentation using multi-modal MR imaging," 2022, *arXiv:2202.05268*.
[48] J. Cai, Z. He, Z. Zheng, Q. Xu, C. Hu, and M. Huo, "Learning global dependencies based on hierarchical full connection for brain tumor segmentation," *Comput. Methods Programs Biomed.*, vol. 221, Jun. 2022, Art. no. 106925. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260722003078
[49] Y. Zhang, N. He, J. Yang, Y. Li, D. Wei, Y. Huang, Y. Zhang, Z. He, and Y. Zheng, "MmFormer: Multimodal medical transformer for incomplete multimodal learning of brain tumor segmentation," 2022, *arXiv:2206.02425*.
[50] J. Hu, X. Gu, and X. Gu, "Mutual ensemble learning for brain tumor segmentation," *Neurocomputing*, vol. 504, pp. 68–81, Sep. 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0925231222007871
[51] A. S. Akbar, C. Fatichah, and N. Suciati, "SDA-UNET2.5D: Shallow dilated with attention UNet2.5D for brain tumor segmentation," *Int. J. Intell. Eng. Syst.*, vol. 15, pp. 135–149, Jan. 2022.
[52] J. Liang, C. Yang, M. Zeng, and X. Wang, "TransConver: Transformer and convolution parallel network for developing automatic brain tumor segmentation in MRI images," *Quant. Imag. Med. Surg.*, vol. 12, no. 4, pp. 2397–2415, Apr. 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8923874/
[53] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*.
[54] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data*, vol. 4, no. 1, Dec. 2017, Art. no. 170117. [Online]. Available: http://www.nature.com/articles/sdata2017117
[55] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, and A. Crimi, "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge," 2018, *arXiv:1811.02629*.
[56] F. Isensee, P. F. Jäger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "Automated design of deep learning methods for biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021.
[57] Y. Yuan, "Automatic brain tumor segmentation with scale attention network," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Springer, 2021, pp. 285–294.
[58] Y. Wang, Y. Zhang, F. Hou, Y. Liu, J. Tian, C. Zhong, Y. Zhang, and Z. He, "Modality-pairing learning for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 230–240.

TIRIVANGANI MAGADZA received the B.Tech. degree in computer science from the Harare Institute of Technology, Zimbabwe, in 2012, and the M.Tech. degree in computer science from Jawaharlal Nehru Technological University, Hyderabad, India, in 2016. He is currently pursuing the Ph.D. degree in computer science with the School of Mathematics, Statistics, and Computer Science, University of KwaZulu-Natal, Durban, South Africa. His research interests include medical image analysis, computer vision, high-performance computing, wireless sensor networks, and natural language processing.

SERESTINA VIRIRI (Senior Member, IEEE) received the B.Sc. degree in mathematics and computer science and the M.Sc. and Ph.D. degrees in computer science. He is currently a Full Professor of computer science with the School of Mathematics, Statistics, and Computer Science, University of KwaZulu-Natal, South Africa. He has been in academia, since 1998. His main research interests include artificial intelligence, computer vision, image processing, machine learning, medical image analysis, pattern recognition, and other image processing related fields, such as biometrics, medical imaging, and nuclear medicine. He has published extensively in several artificial intelligence and computer vision-related accredited journals and international and national conference proceedings. He is a reviewer for several machine learning and computer vision-related journals. He has also served on program committees for numerous international and national conferences. He is a Rated Researcher by the National Research Foundation (NRF) of South Africa.

3.2.2 Conclusion

In conclusion, this section introduces a pragmatic and efficient model for brain tumor segmentation by incorporating depthwise separable convolutions. Our approach involves selectively replacing conventional convolution blocks within a standard U-Net structure with depthwise separable blocks, addressing the imperative need for computational and memory efficiency. The experimental validation on the challenging BraTS 2020 dataset underscores the efficacy of our proposed method, demonstrating comparable segmentation results with state-of-the-art approaches while minimizing computational complexity.

Moreover, we presented a comprehensive computational analysis of existing methods, offering valuable insights into the trade-offs between segmentation performance and computational demands. In the future, we will explore multi-resolution fusion techniques to capture long-range dependencies. By pushing the boundaries of segmentation methodologies, we anticipate further improvements in performance, fostering advancements in the utilization of computer-assisted diagnosis, particularly in environments where computational resources are constrained.

3.3 Efficient nnU-Net for Brain Tumor Segmentation

3.3.1 Introduction

The nnU-Unet is one of the leading open-source frameworks for medical image analysis. It has been used to win several medical image analysis challenges. However, despite its success, the framework utilized conventional convolution operations, which are computationally expensive. This section presented several modifications to the nnU-UNet framework for computational efficiency. Furthermore, we present extensive experimental results to study the efficacy of our techniques.

This work has been published in[2].

[2]T. Magadza and S. Viriri, "Efficient nnU-Net for Brain Tumor Segmentation," in IEEE Access, vol. 11, pp. 126386-126397, 2023, doi: 10.1109/ACCESS.2023.3329517.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 September 2023, accepted 23 October 2023, date of publication 6 November 2023,
date of current version 15 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3329517

RESEARCH ARTICLE

Efficient nnU-Net for Brain Tumor Segmentation

TIRIVANGANI MAGADZA AND SERESTINA VIRIRI, (Senior Member, IEEE)
School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban 4000, South Africa
Corresponding author: Serestina Viriri (viriris@ukzn.ac.za)

ABSTRACT Brain tumors are one of the leading causes of death in adults. They come in various shapes and sizes from one patient to another. Sometimes, they infiltrate surrounding normal tissues, making it challenging to delineate tumor boundaries. Despite extensive research, the prognosis is still low. Accurate and timely brain tumor segmentation is critical for treatment planning and disease progression monitoring. Automatic segmentation of brain tumors using deep learning methods has produced high-quality and reproducible segmentation results. Specifically, the encoder-decoder networks, like the U-Nets, have dominated the previous BraTS Challenges because of their superior performance. Due to the importance of high-quality segmentation, most state-of-the-art models focus more on pushing the boundaries of the current methods at the expense of computational complexity. The computational budget for practical applications is minimal, requiring technological solutions that balance accuracy and available computational resources. In this study, we extended the U-Net model in the nnU-Net by replacing the basic 3D convolution blocks with bottleneck units utilizing depthwise-separable convolutions. Furthermore, we introduced the shuffle attention mechanism in the skip connections to compensate for the slight loss in segmentation accuracy due to a reduction in the number of parameters. On the brain tumor dataset BraTS 2020, our network achieves dice scores of 79.2%, 91.2%, and 84.8% for enhancing tumor (ET), whole tumor (WT), and tumor core (TC), respectively, with only 2.51M parameters and 55.26G FLOPS. Extensive experimental results of the BraTS 2020 dataset reviewed that the proposed modifications achieved competitive performance at a lower computational cost. The code for this project is available at https://github.com/tmagadza/EfficientNNUNET.git.

INDEX TERMS Brain tumor segmentation, depthwise-separable convolutions, group convolution, shuffle attention, U-Net.

I. INTRODUCTION

A brain tumor is the abnormal growth of cells in any part of the brain. Their exact causes are not yet known [1]. However, the risk factors include a family history of brain tumors, metastases, and exposure to ionizing radiation. There are about 120 types of tumors, with gliomas being the most common and one of the leading causes of death among adults [2]. The World Health Organization broadly classifies gliomas into low-grade (Grade I and II) and high-grade (Grade III and IV) tumors. The low-grade tumors are less aggressive, with a life expectancy that spans many years. On the other hand, high-grade tumors are much more aggressive, with a median survival rate of fewer than two years, and require immediate treatment [3].

Timely, accurate, and reproducible segmentation of brain tumors is critical for diagnosis, treatment planning, and monitoring of disease progression. In clinical practice, segmentation is done manually by a high-trained radiologist. This process is tedious and time-consuming and suffers from intra and inter-rater variability [3], [4]. Consequently, manual segmentation is only used for qualitative assessment or visual inspection.

Meanwhile, in recent years, automatic brain tumor segmentation has been slowly becoming a viable solution to manual segmentation. It requires minimal human involvement if not none at all. However, it also presented its unique challenges. Brain tumors come in different shapes, sizes, and locations from one patient to another, limiting the use of prior

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino.

knowledge of the shape and location of anatomic tissues. The most aggressive tumors often diffuse into surrounding tissues, making delineating tumor boundaries difficult. Furthermore, segmentation only depends on comparing pixel intensities between normal brain parts and lesions. Despite these challenges, automatic brain tumor segmentation is still a promising solution for quantitatively assessing brain tumors.

More recently, deep learning methods for automatic brain tumor segmentation have attracted much attention among the research community owing to their success in various computer vision applications. Applying deep learning techniques to medical image analysis requires expertise in choosing the appropriate network for the task at hand and making numerous decisions regarding hyper-parameters, preprocessing and post-processing techniques, training schemes, data augmentation, etc. [5]. A slight mistake in the configuration of these methods will lead to a significant drop in performance. For example, methods based on U-Net [6] like structure have been dominating the BraTS challenge [7]. Still, the performance of these methods varies significantly, signifying the importance of expected knowledge for the task at hand [5].

In 2020, Isensee et al. [8] proposed an open-source self-configuring deep learning framework for biomedical image segmentation, which they dubbed nnU-Net.[1] Their framework automates the entire segmentation pipeline, including configuring any medical dataset, preprocessing, network architecture, training, and post-processing without human input. nnU-Net has set a new state of the art in various semantic segmentation challenges [8]. In the context of Brain Tumor segmentation, Isensee et al. [7] investigated the suitability of nnU-Net for brain tumor segmentation while applying BraTS-specific modification, and their method came first in BraTS 2020 Challange. Again, in BraTS 2021 Challenge, Luu and Park [9] proposed several modifications to the nnU-Net, including using a larger network, swapping batch normalization with group normalization, and adopting axial attention in the decoder. Their method also came first.

Despite several benefits that nnU-Net brings to medical image segmentation, it needs more computational costs. At its core, nnU-Net is an instance of basic U-Net architecture. It makes use of standard convolution, which is computationally expensive. By using 3D convolutions, which have been shown to perform better than 2D counterparts, the number of parameters increases substantially, making it practically impossible to train the model reasonably for a given computational budget.

This work investigated the effects of reducing the nnU-Net framework's computational complexity on the model's segmentation performance on brain tumor segmentation tasks.

[1]https://github.com/MIC-DKFZ/nnUNet

Our main contributions can be summarized as follows:

1) We propose swapping all standard convolutions with depthwise separate convolutions to reduce the number of network parameters and improve the efficiency of the network.
2) We introduce bottleneck units to reduce the number of parameters further.
3) We adopt the 3D shuffle attention mechanism in skip connections to improve the segmentation performance of the network. Moreover, we introduced residual connections to avoid network degradation.
4) We extensively evaluate the proposed modifications using BraTS 2020 dataset.

The rest of the paper is organized as follows: Section II reviews related work. Section III describes the dataset used and the proposed modifications to the nnU-Net framework. Section IV presents the experimental results, which are discussed in Section V. Lastly, Section VI provides concluding remarks.

II. RELATED WORK

A. U-NET LIKE ARCHITECTURE

Since the introduction U-Net [6] in 2015, the encoder-decoder-like structure became the de-facto standard for biomedical segmentation. The U-Net architecture uses an encoder pathway to extract rich semantic and global information by successively reducing the spatial resolution by half and doubling the number of feature maps. The decoder gradually doubles the spatial resolution to recover the spatial resolution while reducing the feature maps by half. Skip connections combine the encoder's finer features and the decoder's course features. Dong et al. [10] proposed a 2D U-Net that was optimized using soft dice loss to mitigate the unbalanced nature of the BraTS 2015 dataset. Their methods applied extensive data augmentation techniques to improve segmentation performance. Myronenko [11], the winner of BraTS 2018, proposed an encoder-decoder network with an asymmetrically larger encoder to extract more deep features. Their method uses a variational autoencoder branch to regularize the shared encoder. The author observed that increasing the width of the network improved performance. Their approach is computationally expensive due to standard convolutions and large input patch sizes. Insensee et al. [12] developed a U-Net Like 3D architecture, which was trained using large patch size, dice loss, and extensive data augmentation. Deep supervision was used to improve gradient propagation to lower layers further. Li et al. [13] proposed an up-skip connection between the encoder and decoder to improve the information flow. Their network incorporated an inception module and used cascading training strategy to segment tumor regions sequentially. Zhao, Y. et al. [14] investigated the usefulness of various schemes in data processing, model designing, and optimization as applied to general DCNN design and training for the 3d brain tumor segmentation. Their method won second place in the BraTS 19 Challenge.

B. REGION-BASED TRAINING

Wang et al. [15] developed a method that exploits brain tumors' hierarchical nature by segmenting partially overlapping regions one after the other in a cascading fashion. Their method uses anisotropic convolution to balance between accuracy and model complexity. Multi-scale feature fusion was exploited for robust segmentation. The shallow layers learn to represent local and low-level features while deep layers learn to represent more global and high-level features. Their method was not end-to-end. Each network was trained separately, increasing the time required for both training and testing [16]. Wang et al. [17] extended their previous work [15] to incorporate uncertainty estimation gathered from test time augmentation. The paper showed that uncertainty estimation could identify false positives and improve segmentation performance. Unfortunately, their method required a longer time to train. In [18], Zhoul et al. adopted the multi-task learning approach instead of training three networks separately, combining the three tasks in a single model. The paper adopted the curriculum learning scheme, gradually introducing each task as the learning proceeded.

C. LIGHT-WEIGHT NETWORKS

Chen et al. [16] used anisotropic convolutions to split the standard 3D convolution into three parallel branches, each extracting features from different orthogonal views. The use of separable convolution has the benefit of reducing the number of parameters. Their model replaces all the standard convolution operations in the U-Net structure with separable convolutions. Chen et al. [19] exploited group convolution to reduce model complexity. Each group is split into two three branches using weighted 3D dilated convolution for multi-scale learning. A multiplexer unit facilitates information sharing between each group or fiber. Zhou et al. [20] utilized the shufflenetV2 units in the encoder to reduce the number of parameters, while in the decoder, residual units are used to address network degradation. Luo, Z et al. [21] proposed hierarchical decoupled convolution to reduce the number of parameters in an encoder-decoder structure. Peng et al. [22] proposed a U-Net variant that utilizes weighted dilated convolutions to learn multi-scale features. The authors used group convolutions to reduce the number of parameters in the network. Furthermore, the authors used dense residual blocks to improve segmentation performance. In [23], the authors used a 3D inverted residual module to reduce the computational complexity of 3D models. Their methods achieved competitive results on BraTS 2018 while using few computational resources. Zhang et al. [24] exploited shuffle units and depthwise separable convolutions to reduce the number of network parameters and operations.

D. ATTENTION MECHANISM

Noori et al. [25] proposed a 2d encoder-decoder networks structure that utilizes residual units to improve network training and apply channel attention after concatenating low-level and high-level features. The authors argue that it is improper to concatenate features from low-level and high-level features without weighing them. Empirical results demonstrate the effectiveness of channel attention in improving segmentation performance. Zhang et al. [26] proposed a 2d encoder-decoder network structure that incorporates residual units and attention gates in the skip connection. Experiment results showed the effectiveness of attention gates in improving network performance. Cao et al. [27] proposed a UNet-like network structure that utilizes 3D Shuffle Attention in the encoder and skip connections. The authors adopted an optimized shuffle unit as a basic building block. The authors did not report on the complexity analysis of their method.

III. MATERIALS AND METHODS

A. DATA

We used the BraTS 2020 dataset [3], [28], [29] that contains 369 training and 125 validation subjects for training and validation of our model. As illustrated in Figure 1, all subjects have native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR) volumes that were acquired using varying clinical protocols and scanners from nineteen (19) institutions. The training set was also comprised of manually annotated ground truth by one to four raters applying the same annotation protocol, and experienced radiologists approved their annotations. In contrast, the ground truth for the validation set was not made public. Instead, the researchers can use the online evaluation platform[2] to evaluate models. All scans were co-registered to the same anatomical template, interpolated to the same resolution (1mm3), skull-stripped, and had an original image size of $240 \times 240 \times 155$. We also used the BraTS 2021 dataset [3], [29], [30] which includes 1251 training cases and 219 validaton cases. The structure and format of BraTS 2021 is consistent with the BraTS 2020 dataset.

B. NNU-NET BASELINE

Our baseline model was an instantiation of nnU-Net [7], a winning model for BraTS 2020. The model follows an encoder-decoder structure with skip connections linking the two pathways, as presented in Figure 2. As a 3D U-Net [31], the model takes in a large input patch of $128 \times 128 \times 128$, with four 3D MRI image modalities concatenated in the channel dimension. The network comprises five (5) resolution levels. In the encoding pathway, each level reduces the spatial resolution by half using strided convolution and doubles the feature maps starting from base feature maps of 32 up to a maximum of 320. Two consecutive convolution blocks were applied in each layer, each performing $3 \times 3 \times 3$ convolution followed by instance normalization [32] and then Leaky Relu non-linearity. In the

[2]https://ipp.cbica.upenn.edu/

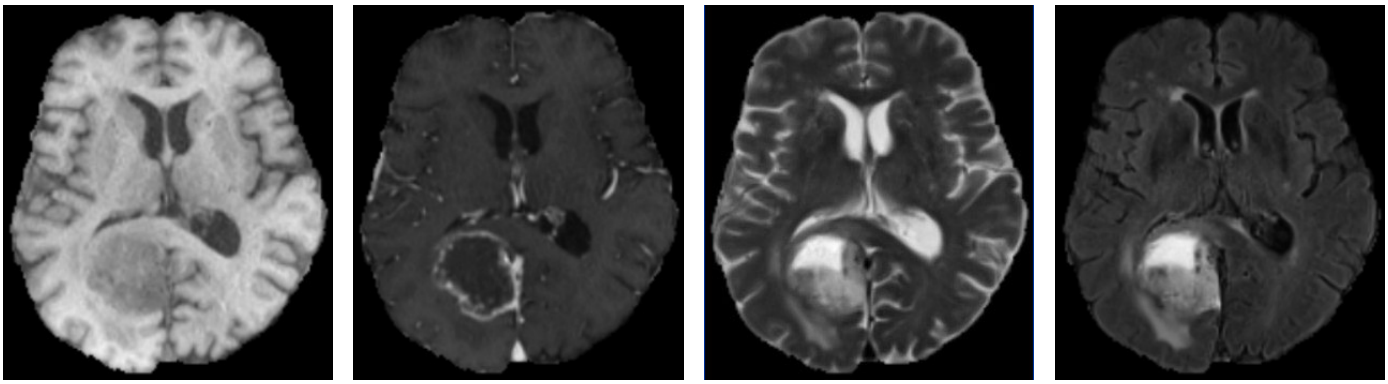


FIGURE 1. Examples of different MRI imaging modalities. From left to right: T1, T1ce, T2, and FLAIR.

decoding path, each layer gradually reduces the number of feature maps by half while doubling the spatial resolution with transpose convolutions. Convolution blocks in the decoding path follow the same structure as the encoding path. $1 \times 1 \times 1$ convolution followed by sigmoid non-linearity is performed after the last layer to reduce the number of feature maps to 3. Deep supervision was also used to improve network training in all layers along the decoding path except the two lowest resolutions. To improve the segmentation performance, we directly optimize the three partially overlapping regions: whole tumor, tumor core, and enhancing tumor, instead of providing labels that include: edema, non-enhancing tumor, and necrosis and enhancing tumor. Aggressive data augmentation techniques were applied on the fly using the batchgenerators framework.[3] Specifically, we applied rotation, scaling, elastic deformation, additive brightness augmentation, and gamma augmentation as described in [7]. The loss function was a summation of dice and binary cross-entropy losses, which has been shown to improve segmentation performance [33].

C. NNU-NET MODIFICATIONS

1) REDUCED COMPUTATIONAL COMPLEXITY

A standard convolution operation is computationally expensive since it simultaneously performs spatial and channel-wise correlation in one go. An excessive amount of computation is required when using 3D MRI volumes with large patch sizes, which were shown to perform well as compared to 2D counterparts, making it difficult to train the resulting models. To reduce the number of parameters as well as computational complexity, we replaced all the standard convolution operations with depthwise separable convolutions, which apply $3 \times 3 \times 3$ convolution on each channel separately followed by $1 \times 1 \times 1$ convolution to project the output channels from previous operation to another channel space as illustrated in Figure 3(b). A depthwise separable convolution can be generalized as a group convolution with a group size equal to the number of input channels. We adopted the bottleneck unit as our basic building block with depthwise separable convolution in the middle, as shown in Figure 3(c). The module introduced an addition hyper-parameter, reduction ratio r, to reduce the number of input channels for the middle layer. We have fixed the value of r to 4.

[3]https://github.com/MIC-DKFZ/batchgenerators

2) SHUFFLE ATTENTION MECHANISM

The use of depthwise separable convolutions will significantly reduce network parameters, which may slightly reduce the segmentation accuracy. To compensate for the loss in performance, we introduced the shuffle attention (SA) [34] mechanism, which simultaneously applies spatial and channel attention. The attention will help the network focus more on all salient features of the task. The network can learn to capture the pixel-level correlations and channel dependency by combining spatial and channel attention. Numerous studies [35], [36], [37], [38] have shown that attention mechanisms can considerably enhance network performance.

Given an input feature map $I \in \mathbb{R}^{C\times H\times W\times D}$, where H, W, D, and C are the height, width, depth, and number of channels of the input feature map, respectively, SA first divides I into G groups along the channel dimension, i.e., $I = [I_1, \cdots, I_G], I_k \in \mathbb{R}^{C/G\times H\times W\times D}$. Then each sub group I_k is further split into two branches, denoted by $I_{k1}, I_{k2} \in \mathbb{R}^{C/2G\times H\times W\times D}$. As shown in Figure 4, the first branch is used to generate the channel attention map by applying global average pooling (GAP), which generates channel-wise statistics $s \in \mathbb{R}^{C/2G\times 1\times 1\times 1}$, to the input feature map, which can be calculated by shrinking I_{k1} through spatial dimension $H \times W \times D$:

$$s = \mathcal{F}_{gp}(I_{k1}) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} \sum_{t=1}^{D} I_{k1}(i, j, t) \qquad (1)$$

Furthermore, a linear transformation $\mathcal{F}_c(\cdot)$ is perform, followed by a simple gating mechanism with sigmoid activation σ to produce the channel attention:

$$I'_{k1} = \sigma(\mathcal{F}_c(s)) \cdot I_{k1} = \sigma(W_1 s + b_1) \cdot I_{k1} \qquad (2)$$

where $W_1 \in \mathbb{R}^{C/2G\times 1\times 1\times 1}$ and $b_1 \in \mathbb{R}^{C/2G\times 1\times 1\times 1}$ are parameters used to scale and shift s.

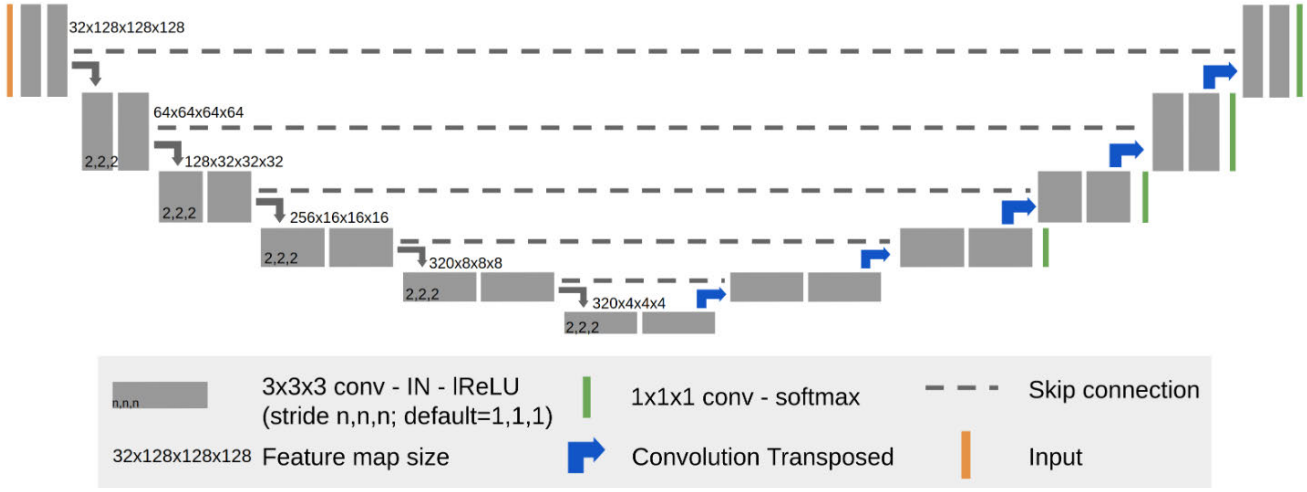


FIGURE 2. Baseline model as generated by the nnU-Net framework. (adapted from [7]).

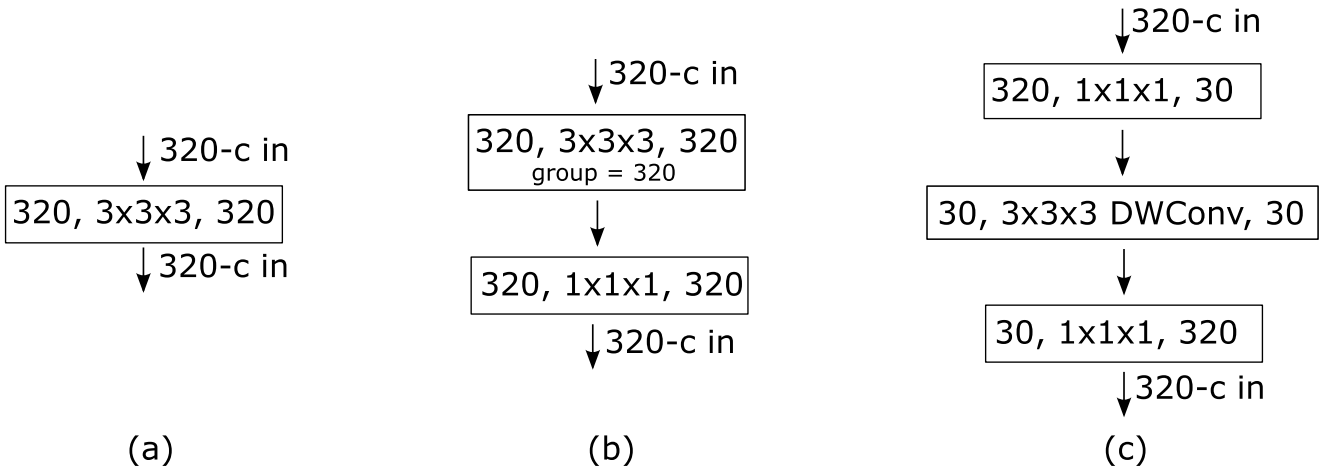


FIGURE 3. Basic building blocks. (a) Standard convolution block. (b) DWConv: Depthwise separable convolution block. (c) Bottlenet unit with depthwise separable convolution block in the middle.

The second branch generates the spatial attention by firstly obtaining spatial-wise statistics through Group Norm (GN) over I_{k2} followed by a linear transformation $\mathcal{F}_c(\cdot)$. The final output of spatial attention is given by:

$$I'_{k2} = \sigma(W_2 \cdot GN(I_{k2}) + b_2) \cdot I_{k2} \tag{3}$$

where W_2 and b_2 are parameters with shape $\mathbb{R}^{C/2G\times1\times1\times1}$.

Then, a concatenation operation is applied to the two branches to make the number of channels as the same as the number of input, i.e., $I'_k = [I'_{k1}, I'_{k2}] \in \mathbb{R}^{C/2G\times H\times W\times D}$. All the sub-groups are then aggregated, followed by the "channel shuffle" operation to enable information communication between different sub-groups. The final output of the SA module is the same size as I.

D. TRAINING

Our model was implemented in Pytorch[4] using opensource framework for biomedical segmentation[5] [8]. Each network takes an input patch of 128 × 128 × 128, with four 3D MRI image modalities concatenated in the channel dimension. We normalize each input channel independently by subtracting the mean and dividing it by the standard deviation. Data augmentation, which comprised random rotation and scaling, elastic deformation, additive brightness augmentation, and gamma scaling, was applied on the fly. The loss function was a summation of batched dice and cross-entropy loss. We optimize all the networks with stochastic gradient descent with an initial learning rate of 0.01 and Nesterov momentum of 0.99. The learning rate was decayed with a polynomial schedule:

$$lr = 0.01 \times (1 - \frac{epoch}{400})^{0.9} \tag{4}$$

Each network was trained for a total of 400 epochs, with each epoch defined as 250 iterations, on an NVIDIA Tesla V100 16GB GPU. During inference, we post-processed each subject by replacing the enhancing tumor with the tumor core when the predicted volume was less than some threshold. The configuration of our models was as follows:

[4]https://pytorch.org/
[5]https://github.com/MIC-DKFZ/nnUNet

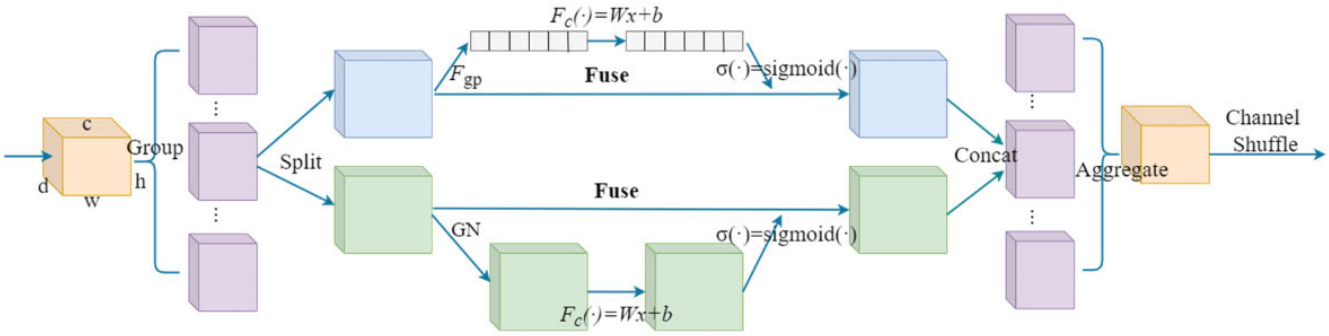


FIGURE 4. The 3D Shuffle Attention Module. The input feature map is first divided into sub-groups along the channel dimension. Then, each sub-group is further divided into two branches, the channel, and spatial attention branches. A concatenation operation is used to join features from the two branches. Afterward, all sub-groups are aggregated, followed by a channel shuffle operation to enable information communication between different sub-groups. (adapted from [34]).

- **BL**: baseline nnUnet-Net without modifications (see Section III-B).
- **BL + DS**: replaced all standard convolution with depthwise separable convolutions
- **BL + BU**: baseline with bottleneck unit as a basic building block
- **BL** $+DS+$ **BU**: baseline with bottleneck unit with depthwise separable convolution in the middle as described in Section III-C1
- **BL** $+DS+BU+$ **SA**: baseline with depthwise separable convolutions, bottleneck unit, shuffle attention in both the encoder and skip connections.
- **BL + R**: baseline with residua connection
- **BL** $+DS+$ **R / BL** $+DS+$ **R***: baseline with depthwise separable convolutions and residual connection. * indicates that the ReLu is applied after the addition of residual units.
- **BL** $+DS+$ **R*+ AA**: baseline with depthwise separable convolutions, residual connection, and shuffle attention skip connections.

IV. RESULTS

A. PERFORMANCE COMPARISON OF THE PROPOSED METHOD

Due to a 12-hour limitation on CHPC,[6] we trained all model configurations for a maximum of 400 epochs. Each configuration was trained with all 369 training cases and evaluated with the 125 validation cases. The validation results of each configuration, as computed by the online evaluation platform, are presented in Table 1. The results show that the baseline configuration (BL) performance for both the dice score and Hausdauf distance is relatively high. Introducing the bottleneck unit (BU) to the baseline showed a decrease in dice score for enhancing tumor and tumor core by 1.9% and 0.4%, respectively. From the results, we can also see a slight increase in Hausdorff distance in enhancing tumor, whole tumor, and tumor core by 9.80 mm, 0.12 mm, and 0.74 mm, respectively.

On the other hand, replacing all the convolution blocks in the baseline with the depthwise separable convolutions (DS) produced similar if not better, results. For Example, the dice score in the tumor core improved by 0.5% while remaining the same for the whole tumor and marginally decreased by 0.4% in enhancing the tumor. As for the Haursdoff distance, the results in Table 1 show an increase of 5.80 mm in enhancing tumor and an improvement of 2.14 mm in the tumor core. The BL + DS + BU model achieved slightly less performance as compared to the BL + DS model. At the same time, the BL + DS + BU + SA model shows a slight improvement in performance in dice score and the Haursdoff distance for enhancing tumor compared to the BL + DS + BU model.

Residual units can help reduce degradation in deep networks like U-Net structure [39]. From Table 1, we did not observe any benefits of residual connections with ReLu before addition (BL + R and BL + DS + R) except for Hausdourff distance, where we observed a decrease of 2.80 mm in the whole tumor and an increase of 3.00 mm in the tumor core. Interestingly, applying ReLu after the addition further decreases the performance. Introducing Shuffle Attention to the skip connection of BL + DS + R* substantially improved performance in the Hausdorff distance of the tumor core.

B. MODEL COMPLEXITY

Table 1 also reports on the complexity of the different model configurations in terms of floating-point operations (FLOPS) and a number of parameters (Params) as computed by the THOP[7] python library. The table shows that the BL + DS model balanced model complexity and segmentation performance well. Specifically, it achieved 82% and 90% reduction in floating-point operations and several parameters, respectively, without affecting the segmentation performance. A combination of depthwise separable convolutions and bottleneck units (BL + DS + BU) further reduced the model complexity at a slight reduction in segmentation performance. The results show that the Shuffle

[6]https://www.chpc.ac.za/

[7]https://github.com/Lyken17/pytorch-OpCounter

TABLE 1. Performance comparison on the BraTS 2020 validation set (125 cases). Metrics are computed by the online evaluation platform. See Section III-D for decoding the abbreviations. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core.

Model	Dice			HD95				
	ET	WT	TC	ET	WT	TC	FLOPS	Params.
BL	**0.796**	**0.912**	0.843	23.515	**4.337**	8.340	308.711G	25.708M
BL + BU	0.777	0.911	0.839	32.313	4.453	9.081	65.608G	3.804M
BL + DS	0.792	**0.912**	**0.848**	29.312	4.410	6.200	55.263G	2.513M
BL + DS + BU	0.782	0.910	0.847	26.676	4.758	8.881	**52.333G**	**2.510M**
BL + DS + BU + SA	0.788	0.905	0.837	26.525	5.626	9.101	52.501G	2.510M
BL + R	**0.796**	0.910	0.844	**23.490**	4.591	5.923	328.842G	31.088M
BL + DS + R	0.792	0.910	0.841	26.421	5.203	9.192	75.394G	7.894M
BL + DS + R*	0.784	0.909	0.841	32.331	4.580	9.368	75.394G	7.894M
BL + DS + R* + AA	0.788	0.908	0.845	29.612	5.108	**5.919**	75.461G	7.894M

Best values are shown in bold, and second best are underlined.

Attention barely increases the computation cost. Because of strided convolutions for downsampling and upsampling, introducing residual units resulted in a slight increase in floating point operation due to the $1 \times 1 \times 1$ convolution to match the dimensions in both branches before addition.

C. PERFORMANCE COMPARISON WITH THE STATE-OF-THE-ART

1) PERFORMANCE COMPARISON WITH THE STATE-OF-THE-ART METHODS WITHOUT MODEL ENSEMBLE

For a fair comparison, Table 2 list the results without a model ensemble of the top performances in the BraTS 2020 validation dataset except for the result for Isensee et al. [7], since they did not present the results for a single model. Yuan [40] and Wang et al. [41] are the top participants of the BraTS 2020 challenge, and only results without model ensemble are listed here. We also included single model results of our previous work [42], Raza et al. [43], and Daza et al. [44]. From the results, it is evident that our proposed method achieves superior performance with minimum computation complexity.

Y. Yuan [40] won third place in the BraTS 2020 challenge by aggregating the output feature maps from all the encoding layers with high-level feature maps of each decoding layer using skip connections. Yuan's method achieves superior performance against state-the-art in Hausdorff distance for enhancing tour. However, our lightweight method outperforms Yuan's method in the other metrics.

Wang et al. [41] won second place in the BraTS 2020 challenge. Their methods utilize two interconnected pathways, which take a pair of modalities each. From the results, it is clear that our method demonstrated superior performance in all metrics.

In our previous work [42], we partially utilized depthwise separable convolutions in both the encoder and the decoder. Although our previous work shows competitive performance, it has many floating point operations. In contrast, our proposed work is superior in all metrics.

On the other hand, Raza et al. [43] adopted residual units in the encoding pathway resulting in superior performance in dice score for the enhancing tumor, the worst performance in dice score for the whole tumor, and a comparable performance in the remaining metrics. The computational complexity of their method is relatively high. Similarly, Daza et al. [44] proposed a lightweight method with superior performance in the Hausdorff distance for enhancing tumors. Our approach remains superior in other metrics.

2) PERFORMANCE COMPARISON WITH THE STATE-OF-THE-ART METHODS WITH MODEL ENSEMBLE

Table 3 reports on the aggregate performance of the state-of-the-art methods with the model ensemble on the BraTS 2020 validation dataset. Isensee et al. [7] won the first price, followed by Jia et al. [45] and Wang et al. [41] for the second price, and then Yuan [40] for the third place in the BraTS 2020 challenge. We have also included model ensemble results for Wang et al. [46]. From the table, it is clear that the state-of-the-art methods achieved the best performance at the cost of computational complexity.

Isensee et al. [7] applied nnU-Net [8] with BraTS specific modifications and extensive data augmentation to the brain tumor segmentation problem. Their winning method, which is an ensemble of 25 models, uses basic U-Net structures, with each model trained for 1000 epochs. Results show that Isensee et al.'s method is superior in dice for enhancing tumor and Hausdorff distance for the whole tumor as compared to the state-of-the-art methods. Additionally, their 5 model ensemble also shows similar performance. However, our method achieves the same results as Isensee et al.'s method in dice for the whole tumor and a slight improvement in Hausdorff distance for the tumor core and comparable performance for the other metrics while using significantly fewer computational resources.

Jia et al. [45] proposed a two-stage cascaded model that maintains high-resolution feature representation and uses a Non-local attention mechanism to aggregate contextual information from all layers. Again, their best method was an ensemble of 27 models, each trained for 450 epochs. Despite high performance in dice for the whole tumor and Hausdorff distance for the tumor core, their method is

TABLE 2. Mean performance metrics on BraTS 2020 Validation dataset as compared to the state-of-the-art without model ensemble. See Section III-D for decoding the abbreviations. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core.

Model	Dice			HD95					
	ET	WT	TC	ET	WT	TC	Epochs	FLOPS	Params.
Y. Yuan [40]	0.785	0.904	0.842	**20.35**	5.49	8.34	300	-	16.50M
Wang et al. [41]	0.785	0.907	0.837	32.25	4.39	8.34	1000	-	-
Magadza et al. [42]	0.774	0.898	0.824	29.82	6.78	7.36	100	616.00G	6.90M
Raza et al. [43]	**0.800**	0.866	0.836	29.82	6.78	7.36	100	374.04G	30.47M
Daza et al. [44]	0.794	0.897	0.845	29.82	**3.59**	6.47	-	**49.82G**	4.02M
BL + DS (ours)	0.792	**0.912**	**0.848**	29.31	4.41	6.20	400	55.26G	**2.51M**
BL + DS + R (ours)	0.792	0.910	0.841	26.42	5.20	9.19	400	75.39G	7.89M
BL + DS + R* + AA (ours)	0.788	0.908	0.845	29.61	5.11	**5.92**	400	75.46G	7.89M

Best values are shown in bold, and second best are underlined.

still computationally expensive. Compared to our method, as shown in Table 3, our best single model achieves comparable results at low computation costs.

Wang et al. [41] and Yuan [40] ensemble 9 models and 11 models to secure second and third places receptively. Their ensemble improved performance in all metrics except for Wang et al. [41], which marginally increased Hausdorff distance for the enhancing tumor and tumor core. In comparison, our model achieved competitive results.

Lastly, Wang et al. [46] introduced Transformer to the encoder-decoder structure for brain tumor segmentation to model long-range dependencies. Their 5 model ensemble achieved superior performance in the Hausdorff distance for the enhancing tumor with relatively huge computational costs. In contrast, our lightweight method demonstrated superior performance in other metrics at significantly low computation costs.

3) PERFORMANCE COMPARISON ON BRATS 2021 DATASET.

Table 4 compares our best-performing model with the state-of-the-art models on the BraTS 2021 dataset 5-fold cross-validation results. By the time of writing this paper, the online evaluation platform for the BraTS 2021 dataset[8] was no longer available. The table shows that our model performed well on the enhancing tumor region and performed slightly poorly on both the whole tumor and tumor core. On the other hand, our model uses very light resources as compared to the state-of-the-art.

D. QUALITATIVE ANALYSIS

In Figure 5, we show qualitative overview of the segmentation performance of BL + DS model on the validation set. To avoid cherry picking [7], we systematically selected cases by first computing an average over the three validation regions and then picked the best, worst, median, and 75th and 25th percentile. The clearly show that the segmentation quality of our model is quite high overall. However, in the worst scenarios, it completely fail to segment small enhancing tumor lesion.

[8]https://www.synapse.org/#!Synapse:syn25829067/wiki/610863

V. DISCUSSION

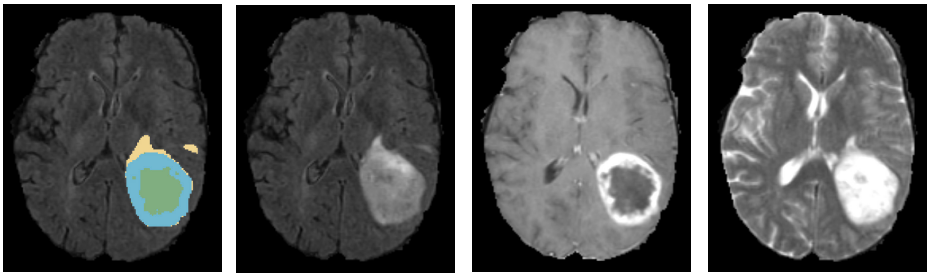
Automatic brain tumor segmentation is paramount for the timely, reproducible, and accurate delineation of tumor sub-structures. Although deep learning methods have demonstrated superior performance than traditional methods in the past few years, automatic brain tumor segmentation is still an open challenge. Brain liaisons appear in different shapes, sizes, and locations from one patient to another, rendering prior knowledge useless. Moreover, deep learning methods require massive training datasets and computational resources [20]. One would need a model with competitive performance for a limited computational budget for practical application.

Unfortunately, as shown in Table 3, most state-of-the-art methods focus more on improving segmentation performance at the cost of high computation resources. These methods are usually an ensemble of multiple models. For Example, Isensee et al.'s method [7], which won the first prize in the BraTS 2020 challenge, is an ensemble of 25 models. Each model needed to be trained separately for 1000 epochs before their results could be aggregated. Furthermore, their method is computationally expensive when applied to 3D MRI scans due to standard convolutions.

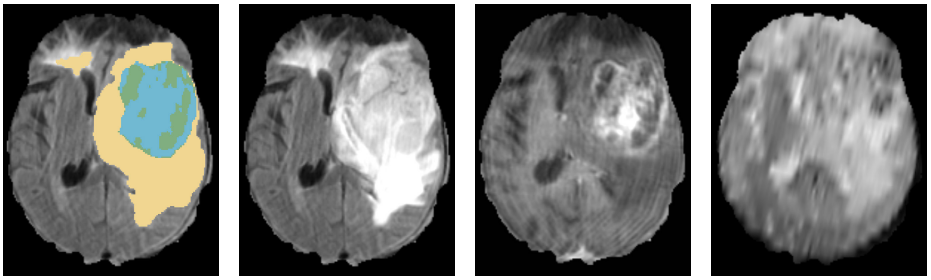
Similarly, Jia et al. [45] ensemble 27 models to win second place in the same challenge. These models chew a significant amount of computation resources to train them. With an increase in the training dataset set, as in BraTS 2021, computation resources are needed even more. The computational requirements may be prohibitive for clinical applications or out of reach for many researchers resulting in poor adoption rates.

Motivated by the above observations, we extended Isensee et al.'s work [7] by introducing depthwise separable convolutions to reduce the computational costs significantly. We also experimented with bottleneck units to further reduce the number of parameters at the expense of a slight loss in segmentation performance. As shown in Table 1, our model configuration with depthwise separable convolutions demonstrated a good balance between computation cost and segmentation performance compared to other configurations. The results are consistent with other previous studies [50], [51], [52]. Although residual units [53] may assist in

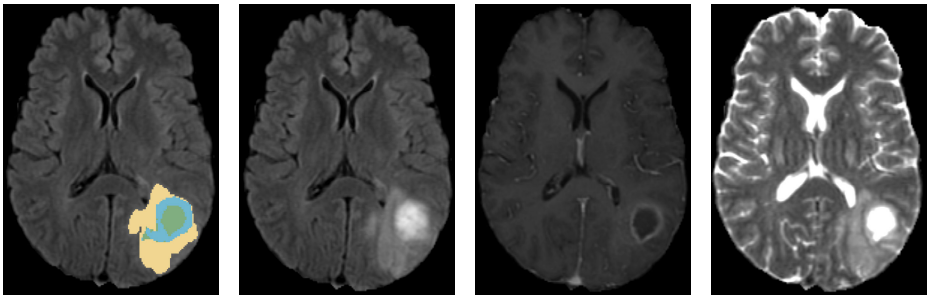
Best: BraTS20_Validation_040, whole: 0.97, core: 0.98, enh: 0.96



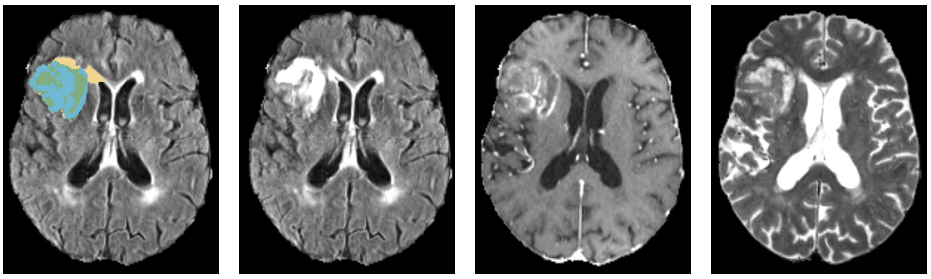
75th percentile: BraTS20_Validation_105, whole: 0.94, core: 0.92, enh: 0.91



Median: BraTS20_Validation_052, whole: 0.95, core:0.91, enh: 0.82



25th percentile: BraTS20_Validation_052, whole: 0.89, core:0.94, enh: 0.66



Worst: BraTS20_Validation_090, whole: 0.91, core:0.25, enh: 0.00

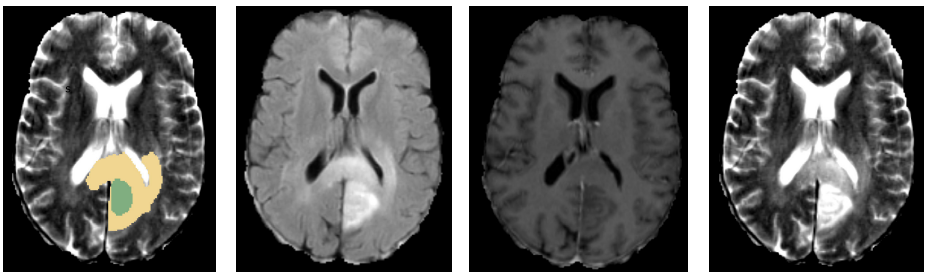


FIGURE 5. Qualitative validation set result. Selection criteria for cases were based on best,worst, median, and 75th and 25th percentile. From left to right:FLAIR image with overlay of generated segmentation, FLAIR image, T1ce image, and T2 image. Edema is shown in yellow, necrosis in green and enhancing tumor in blue.

TABLE 3. Mean performance metrics on BraTS 2020 Validation dataset as compared to the state-of-the-art with the model ensemble. See Section III-D for decoding the abbreviations. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core. † - Ensemble of models trained with 5-fold cross-validation.

Model	Dice			HD95					
	ET	WT	TC	ET	WT	TC	Epochs	FLOPS	Params.
Isensee et al. [7]	**0.799**	0.912	0.851	23.50	**3.69**	7.82	1000	539.56G	31.20M
Isensee et al. [7] †	**0.799**	0.912	**0.857**	26.41	3.73	5.64	1000	539.56G	31.20M
Jia et al. [45]	0.788	**0.913**	0.855	26.58	4.18	**4.97**	450	621.09G	26.07M
Jia et al. [45]†	0.784	**0.913**	0.835	26.50	4.18	5.52	450	621.09G	26.07M
Y. Yuan [40]	0.793	0.911	0.853	18.20	4.10	5.99	300	-	16.50M
Wang et al. [41]	0.787	0.908	0.856	35.01	4.71	5.70	1000	-	-
Wang et al. [46]†	0.787	0.901	0.817	**17.95**	4.97	9.77	8000	333.00G	32.99M
BL + DS (ours)	0.792	0.912	0.848	29.31	4.41	6.20	400	**55.26G**	**2.51M**
BL + DS + R (ours)	0.792	0.910	0.841	26.42	5.20	9.19	400	75.39G	7.89M
BL + DS + R* + AA (ours)	0.788	0.908	0.845	29.61	5.11	5.92	400	75.46G	7.89M

Best values are shown in bold, and second best are underlined.

TABLE 4. Cross-validation results on BraTS 2021 dataset. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core.

Model	Dice						
	ET	WT	TC	Avg.	Epochs	FLOPS	Params.
Jia et al. [47]	0.877	0.934	0.910	0.907	250	436.59G	16.85M
Siddiquee et al. [48]	**0.888**	0.935	0.921	0.915	300	-	-
Luu et al. [9]	0.882	**0.938**	**0.924**	0.915	1000	-	-
Liang et al. [49]	0.883	0.908	0.904	**0.926**	600	68.60G	20.40M
BL + DS (ours)	0.885	0.923	0.913	0.907	200	**55.25G**	**2.51M**

Best values are shown in bold, and second best are underlined.

mitigating network degradation [20], we did not observe any meaningful benefits. He et al. [54] observed that applying 1×1 convolution in residual skip connection will result in poor performance, especially when the number of residual units is high. In the future, we will experiment with residual units in the decoding path as in [20].

The benefits of attention mechanism have been studied extensively in natural language processing [55], computer vision [34], [56], [57], [58] as well in medical image segmentation [23], [27], [59]. In this work, we adopted a lightweight Shuffle Attention mechanism [34] to squeeze in extra segmentation performance without introducing noticeable computational costs. Table 1 shows that the attention mechanism in the skip connections significantly improved the Hausdorff distance for the tumor core.

It is evident in Table 2 and Table 3 that our proposed method is both competitive and efficient regarding computation resources. Table 2 shows that our single model nearly outperforms state-of-the-art methods without model ensemble in all metrics. However, as shown in Table 3, the model ensemble is essential to garner extra segmentation performance. Nevertheless, our single model achieved comparable performance using significantly few computational resources. Table 4, our model performed slight poor for both the whole tumor and tumor core regions. However, performance can be boosted by increasing the number of training iterations. Visual inspection in Figure 5 demonstrated that the segmentation quality of our method is high overall. Sometimes, our method fails to segment small regions of enhancing tumors. In the future, we will experiment with an ensemble of lightweight models to improve segmentation performance.

VI. CONCLUSION

This paper proposed some modifications to the nnU-Net framework to reduce computational complexity while maintaining competitive segmentation performance. Specifically, we replaced all convolution blocks with depthwise separable convolutions. We adopted the bottleneck units to minimize the trainable network parameters further. We applied the Shuffle Attention mechanism to the skip connections to improve performance without introducing additional computational costs. Moreover, we utilized residual units to prevent network degradation. Experimental results on BraTS 2020 validate the effectiveness of the proposed method. Our method achieves competitive results while consuming significantly few computational resources.

ACKNOWLEDGMENT

The authors would like to thank CHPC for the provision of state-of-the-art computing infrastructure.

REFERENCES

[1] S. Turcan and D. Cahill, "Origin of gliomas," *Seminars Neurol.*, vol. 38, no. 1, pp. 5–10, Feb. 2018.

[2] *Brain Tumors and Brain Cancer*. Accessed: Mar. 27, 2023. [Online]. Available: https://www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor

[3] B. H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.

[4] A. Işın, C. Direkoğlu, and M. Şah, "Review of MRI-based brain tumor image segmentation using deep learning methods," *Proc. Comput. Sci.*, vol. 102, pp. 317–324, Dec. 2016.
[5] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.
[6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," 2015, *arXiv:1505.04597*.
[7] F. Isensee, P. F. Jäger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnU-Net for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 118–132.
[8] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021.
[9] H. M. Luu and S.-H. Park, "Extending nn-UNet for brain tumor segmentation," 2021, *arXiv:2112.04653*.
[10] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks," in *Medical Image Understanding and Analysis* (Communications in Computer and Information Science), M. V. Hernández and V. González-Castro, Eds. Cham, Switzerland: Springer, 2017, pp. 506–517.
[11] A. Myronenko, "3D MRI brain tumor segmentation using autoencoder regularization," 2018, *arXiv:1810.11654*.
[12] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain tumor segmentation and radiomics survival prediction: Contribution to the BRATS 2017 challenge," 2018, *arXiv:1802.10508*.
[13] H. Li, A. Li, and M. Wang, "A novel end-to-end brain tumor segmentation method using improved fully convolutional networks," *Comput. Biol. Med.*, vol. 108, pp. 150–160, May 2019.
[14] Y.-X. Zhao, Y.-M. Zhang, and C.-L. Liu, "Bag of tricks for 3D MRI brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2020, pp. 210–220.
[15] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," 2017, *arXiv:1709.00382*.
[16] W. Chen, B. Liu, S. Peng, J. Sun, and X. Qiao, "S3D-UNet: Separable 3D U-Net for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11384, A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, and T. van Walsum, Eds. Cham, Switzerland: Springer, 2019, pp. 358–368.
[17] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation based on cascaded convolutional neural networks with uncertainty estimation," *Frontiers Comput. Neurosci.*, vol. 13, p. 56, Aug. 2019.
[18] C. Zhou, C. Ding, X. Wang, Z. Lu, and D. Tao, "One-pass multi-task networks with cross-task guided attention for brain tumor segmentation," *IEEE Trans. Image Process.*, vol. 29, pp. 4516–4529, 2020.
[19] C. Chen, X. Liu, M. Ding, J. Zheng, and J. Li, "3D dilated multi-fiber network for real-time brain tumor segmentation in MRI," 2019, *arXiv:1904.03355*.
[20] X. Zhou, X. Li, K. Hu, Y. Zhang, Z. Chen, and X. Gao, "ERV-Net: An efficient 3D residual neural network for brain tumor segmentation," *Expert Syst. Appl.*, vol. 170, May 2021, Art. no. 114566.
[21] Z. Luo, Z. Jia, Z. Yuan, and J. Peng, "HDC-Net: Hierarchical decoupled convolution network for brain tumor segmentation," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 3, pp. 737–745, Mar. 2021.
[22] Y. Peng and J. Sun, "The multimodal MRI brain tumor segmentation based on AD-Net," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104336.
[23] Y. Liu, X. Du, D.-H. Wang, and S. Zhu, "A lightweight brain tumor segmentation network based on 3D inverted residual modules," in *Proc. 11th Int. Conf. Comput. Pattern Recognit.* New York, NY, USA: Association for Computing Machinery, May 2023, pp. 149–155.
[24] R. Zhang, S. Jia, M. J. Adamu, W. Nie, Q. Li, and T. Wu, "HMNet: Hierarchical multi-scale brain tumor segmentation network," *J. Clin. Med.*, vol. 12, no. 2, p. 538, Jan. 2023.
[25] M. Noori, A. Bahri, and K. Mohammadi, "Attention-guided version of 2D UNet for automatic brain tumor segmentation," in *Proc. 9th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, Oct. 2019, pp. 269–275.
[26] J. Zhang, Z. Jiang, J. Dong, Y. Hou, and B. Liu, "Attention gate ResU-Net for automatic MRI brain tumor segmentation," *IEEE Access*, vol. 8, pp. 58533–58545, 2020.
[27] Y. Cao, W. Zhou, M. Zang, D. An, Y. Feng, and B. Yu, "MBANet: A 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104296.
[28] S. Bakas et al., "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge," 2018, *arXiv:1811.02629*.
[29] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data*, vol. 4, no. 1, Sep. 2017, Art. no. 170117.
[30] U. Baid et al., "The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification," 2021, *arXiv:2107.02314*.
[31] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," 2016, *arXiv:1606.06650*.
[32] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," 2016, *arXiv:1607.08022*.
[33] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "No new-Net," 2018, *arXiv:1809.10483*.
[34] Q.-L. Zhang and Y.-B. Yang, "SA-Net: Shuffle attention for deep convolutional neural networks," 2021, *arXiv:2102.00240*.
[35] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," 2017, *arXiv:1704.06904*.
[36] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in *Computer Vision—ECCV*, vol. 11211, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham, Switzerland: Springer, 2018, pp. 3–19.
[37] Y. Yuan, "Evaluating scale attention network for automatic brain tumor segmentation with large multi-parametric MRI database," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12963, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2022, pp. 42–53.
[38] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning where to look for the pancreas," in *Proc. 1st Conf. Med. Imag. Deep Learn.*, May 2018, pp. 1–10.
[39] L. H. Shehab, O. M. Fahmy, S. M. Gasser, and M. S. El-Mahallawy, "An efficient brain tumor image segmentation based on deep residual networks (ResNets)," *J. King Saud Univ.-Eng. Sci.*, vol. 33, no. 6, pp. 404–412, Sep. 2021.
[40] Y. Yuan, "Automatic brain tumor segmentation with scale attention network," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 285–294.
[41] Y. Wang, Y. Zhang, F. Hou, Y. Liu, J. Tian, C. Zhong, Y. Zhang, and Z. He, "Modality-pairing learning for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 230–240.
[42] T. Magadza and S. Viriri, "Brain tumor segmentation using partial depthwise separable convolutions," *IEEE Access*, vol. 10, pp. 124206–124216, 2022.
[43] R. Raza, U. I. Bajwa, Y. Mehmood, M. W. Anwar, and M. H. Jamal, "DResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 103861.
[44] L. Daza, C. Gómez, and P. Arbeláez, "Cerberus: A multi-headed network for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 342–351.

[45] H. Jia, W. Cai, H. Huang, and Y. Xia, "H^2NF-Net for brain tumor segmentation using multimodal MR imaging: 2nd place solution to BraTS challenge 2020 segmentation task," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Lecture Notes in Computer Science), A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 58–68.
[46] W. Wang, C. Chen, M. Ding, H. Yu, S. Zha, and J. Li, "TransBTS: Multimodal brain tumor segmentation using transformer," in *Medical Image Computing and Computer Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), M. de Bruijne, P. C. Cattin, S. Cotin, N. Padoy, S. Speidel, Y. Zheng, and C. Essert, Eds. Cham, Switzerland: Springer, 2021, pp. 109–119.
[47] H. Jia, C. Bai, W. Cai, H. Huang, and Y. Xia, "HNF-Netv2 for brain tumor segmentation using multi-modal MR imaging," 2022, *arXiv:2202.05268*.
[48] M. M. R. Siddiquee and A. Myronenko, "Redundancy reduction in semantic segmentation of 3D brain tumor MRIs," 2021, *arXiv:2111.00742*.
[49] J. Liang, C. Yang, and L. Zeng, "3D PSwinBTS: An efficient transformer-based Unet using 3D parallel shifted windows for brain tumor segmentation," *Digit. Signal Process.*, vol. 131, Nov. 2022, Art. no. 103784.
[50] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.
[51] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "ShuffleNet V2: Practical guidelines for efficient CNN architecture design," in *Computer Vision—ECCV*, vol. 11218, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham, Switzerland: Springer, 2018, pp. 122–138.
[52] D. Zhang, Y. Song, D. Liu, C. Zhang, Y. Wu, H. Wang, F. Zhang, Y. Xia, L. J. O'Donnell, and W. Cai, "Efficient 3D depthwise and separable convolutions with dilation for brain tumor segmentation," in *AI 2019: Advances in Artificial Intelligence* (Lecture Notes in Computer Science), J. Liu and J. Bailey, Eds. Cham, Switzerland: Springer, 2019, pp. 563–573.
[53] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, *arXiv:1512.03385*.
[54] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," 2016, *arXiv:1603.05027*.
[55] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30. Red Hook, NY, USA: Curran Associates, 2017.
[56] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, and H. Lu, "Dual attention network for scene segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3141–3149.
[57] J. Ho, N. Kalchbrenner, D. Weissenborn, and T. Salimans, "Axial attention in multidimensional transformers," 2019, *arXiv:1912.12180*.
[58] L. Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu, and T.-S. Chua, "SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6298–6306.
[59] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*.

TIRIVANGANI MAGADZA received the B.Tech. degree in computer science from the Harare Institute of Technology, Zimbabwe, in 2012, and the M.Tech. degree in computer science from Jawaharlal Nehru Technological University, Hyderabad, India, in 2016. He is currently pursuing the Ph.D. degree in computer science with the School of Mathematics, Statistics, and Computer Science, University of KwaZulu-Natal, Durban, South Africa. His research interests include medical image analysis, computer vision, high-performance computing, wireless sensor networks, and natural language processing.

SERESTINA VIRIRI (Senior Member, IEEE) received the B.Sc. degree in mathematics and computer science and the M.Sc. and Ph.D. degrees in computer science. He is currently a Full Professor of computer science with the School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, South Africa. He is also a rated Researcher by the National Research Foundation (NRF) of South Africa. Since 1998, he has been with the academia. He has published extensively in several artificial intelligence, and computer vision-related accredited journals and international and national conference proceedings. His main research interests include artificial intelligence, computer vision, image processing, machine learning, medical image analysis, pattern recognition, and other image processing-related fields, such as biometrics, medical imaging, and nuclear medicine. He is a reviewer of several machine learning and computer vision-related journals. He has also served on program committees for numerous international and national conferences.

3.3.2 Conclusion

In this section, we presented an efficient nnU-Net model for automatic brain tumor segmentation. The networking incorporated depthwise separable convolutions to reduce network parameters and a 3D shuffle attention mechanism, which sequentially applies spatial and channel attention to boost the segmentation performance. Extensive quantitative analysis was performed to study the significance of our modifications. Experimental results on BraTS 2020 confirmed the effectiveness of our methods compared to the state-of-the-art ones.

3.4 Conclusion

This chapter introduces an efficient model for brain tumor segmentation by integrating depthwise separable convolutions into a standard U-Net structure, significantly enhancing computational and memory efficiency. By selectively replacing conventional convolution blocks, the approach achieves segmentation results comparable to state-of-the-art methods on the BraTS 2020 dataset while reducing computational complexity. Additionally, an optimized nnU-Net model is presented, incorporating depthwise separable convolutions and a 3D shuffle attention mechanism to improve performance. Extensive quantitative analysis confirms the effectiveness of these modifications, demonstrating their superior performance compared to current leading methods.

Chapter 4

Survival Prediction

4.1 Introduction

Despite extensive research on brain tumors, the prognosis remains very poor. Therefore, an accurate estimate of patients' overall survival (OS) time becomes critical for pain and symptom management and preparation for death [30]. Automatic and reproducible survival prediction is very challenging due to the need for large datasets and robust feature extraction and selection mechanisms. This chapter presents our deep learning model for survival prediction, which utilizes adaptive feature selection mechanisms in an end-to-end fashion. We presented a quantitative analysis of our techniques.

Part of the work was submitted for publication in the Scientific Reports Journal.

Efficient Survival Prediction for Gliomas using Deep Learning

Tirivangan Magadza[1] **and Serestina Viriri**[1,*]

[1]School of Mathematics, Statistics and Computer Science; University of KwaZulu-Natal; Durban, South Africa.
[*]corresponding: viriris@ukzn.ac.za

ABSTRACT

Gliomas, especially in older individuals, pose a diagnostic challenge, necessitating accurate survival predictions for effective treatment planning. Traditional machine learning methods for survival prediction in gliomas, though widely used, require complex feature engineering and are not end-to-end. Although deep learning has shown success in brain tumor segmentation, its application to survival prediction is hindered by limited datasets and the curse of dimensionality. To overcome these challenges, we propose an efficient deep-learning method for glioma survival prediction, incorporating a squeeze and excitation block for adaptive feature selection. Our approach explores end-to-end adaptive feature selection and fusion. Extensive evaluation on the BraTS 2020 dataset demonstrates the effectiveness of our method, surpassing state-of-the-art techniques in classification accuracy.

INTRODUCTION

Gliomas, originating from glial cells, are malignant brain tumors primarily affecting the older generation[1]. Despite their prevalence, the exact causes remain unknown[2], and diagnosing them poses significant challenges due to various factors. These tumors exhibit diverse shapes and sizes[3], necessitating distinct diagnostic and treatment strategies tailored to each patient. In some cases, they infiltrate surrounding normal tissues, making it difficult to delineate their boundaries.

Patients diagnosed with high-grade tumors, characterized by increased aggressiveness, face a median survival rate of fewer than two years[3]. Survival prediction plays a critical role in formulating appropriate treatment plans. It aids in determining the aggressiveness of treatment and guides decisions regarding surgery, radiation therapy, and chemotherapy. Additionally, survival prediction is also a crucial factor in palliative care considerations[4].

Consequently, automatic, reproducible, high-quality overall survival prediction techniques are in demand. Most existing techniques are based on conventional machine learning techniques[5]. These techniques work well with limited datasets[6] and require less computation power. Unfortunately, they require robust feature engineering, which is time complicated and time-consuming.

Meanwhile, deep learning techniques have received significant attention in computer vision and natural language applications. In brain tumor segmentation, the first step in survival prediction, all the state-of-the-art methods are based on deep learning techniques, particularly convolution neural networks (CNN). Applying deep learning techniques in overall survival prediction, except in extracting deep features, is still low[7] due to very few samples in publicly available datasets. They extract vast amounts of features, resulting in the curse of dimensionality[6]. Some of these features may be redundant or inappropriate for the task. Often, they overfit the training dataset quickly[7], resulting in poor performance. Moreover, these techniques require more computational power and memory.

Despite the shortcomings of deep learning techniques for survival prediction, some other works[6,8–12] demonstrated the possibility of applying these techniques for survival prediction. In literature, attention mechanisms have been exploited, which forces deep learning methods to pay more attention to salient features while suppressing irrelevant ones[13].

Motivated by these observations, this paper proposes an efficient deep-learning method for survival prediction that leverages adaptive feature selection in an end-to-end fashion. The main contributions of this work can be summarized as follows:

- The incorporation of a squeeze and excitation block for adaptive feature selection is proposed.
- The efficacy of end-to-end adaptive feature selection and feature fusion for glioma survival prediction was investigated.
- An extensive evaluation of the proposed techniques using the BraTS 2020 dataset was conducted.

BACKGROUND WORK

Survival Prediction

Accurate prediction of overall survival for glioma patients is crucial for an effective treatment plan. Since 2017, the BraTS Challenge[3,14,15] introduced survival prediction tasks and segmentation tasks to promote the development and benchmarking of ML algorithms for survival prediction that incorporate radiomic features. The tasks require researchers to predict overall survival time in days for patients with gross tumor resection (GTR), and the evaluation of the techniques is based on the classification of patients into short-survivors (<10 months), mid-survivors (between 10 and 15 months), and long-survival (> 15 months). Predictors for the task should only come from imaging in addition to the age information provided by the dataset.

Machine Learning Techniques

Classical machine learning techniques dominate most works for survival prediction. They can work with a limited dataset[6]. These methods follow a pipeline that first extracts features from either MRI imaging modalities, segmentation maps, or both. This process can generate many features, leading to the curse of dimensionality, thereby requiring robust feature selection mechanisms[5]. Selected features are combined with other clinical information and fed into a regression or classification model for training.

In literature, several machine learning models are used. These includes the random forests[5,16–25], extra-trees[26,27], support vector machines[5,7,28–30], ordinary least squares[4,31], linear regression[5,32], Cox Proportional hazard(CoxPH)[33,34] and Extreme gradient boosting (XGBoost)[5]. Ensemble methods like random forests are widely used because they combine several decision trees for better generalization[21]. Although the above techniques can work with limited datasets, they are not end-to-end. They require robust feature engineering and selection mechanisms, which can be complicated and time-consuming[5].

Neural Network and Deep learning Techniques

On the other hand, very few works exploited neural networks[35–37] and convolution neural networks (CNN)[6,8–12]. By nature, neural network-based techniques with many layers are more robust. However, they generate many parameters requiring vast training data[6] and computational power. These methods quickly overfit datasets with very few samples, rendering them unfit. For survival prediction, neural network-based techniques are limited to a few layers and are trained for a few epochs.

On the brighter side, deep learning techniques (based on CNNs) can automatically extract increasingly complex features. Instead of wasting time on feature engineering, which may require domain expertise, researchers will focus on developing methods for the task. Additionally, deep learning techniques make it easy to develop an end-to-end pipeline. In[38], an end-to-end pipeline that simultaneously segments and predicts overall survival was proposed.

Attention Mechanism

Since deep learning techniques extract massive amounts of features, selecting salient features for the task becomes paramount. Previous studies exploited the Attention mechanism that forces the network to focus more on certain input parts. Channel attention[39] forces pay more attention to specific channels of the input feature maps. Whereas, with spatial attention[13], the network focuses more on a region of interest of the input feature maps. Dual attention[40] mechanism combines both the channel and spatial attention mechanism.

METHODS AND TECHNIQUES

Materials and Methods

Data

The proposed model was trained and validated using the BraTS 2020 Dataset[3,14,15], which comprises preoperative MRI images of glioma patients. It was observed in Table 1 that the dataset had been divided into training and validation sets. The training set, consisting of 369 subjects with manually annotated ground truth, was illustrated in Figure 1. Each patient's MRI images included four modalities, specifically the native (T1), post-contrast T1-weighted (T1Gd), T2 weighted (T2), and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR) scans acquired using varying clinical protocols and scanners from nineteen (19) institutions. Additionally, survival information for 236 patients, including age, resection status, and survival in days, was present in the training set.

In contrast, the validation set consists of 125 subjects with the same MRI volumes as the training set. The ground truth was not made public. Instead, the researchers can use the online evaluation platform[1] to evaluate models. As with the training set, the validation set also contains survival information of 29 subjects stating their demographics. Survival information was withheld.

[1] https://ipp.cbica.upenn.edu/

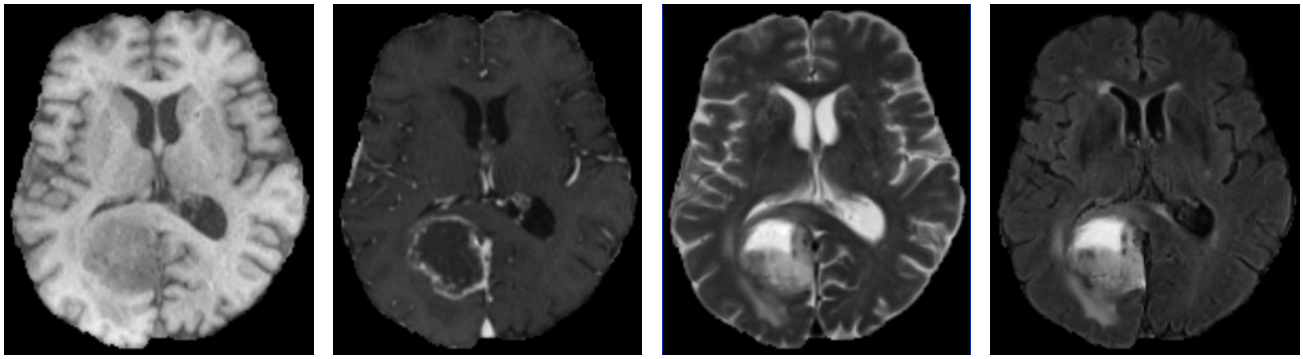


Figure 1. Examples of different MRI imaging modalities. From left to right: T1, T1ce, T2, and FLAIR. (adapted from[41])

Cohort	Subjects	Volume	Survival information
Training	369	$240 \times 240 \times 155$	236 (Age, Resection status, Survival in days)
Validation	125	$240 \times 240 \times 155$	29 (Age, Resection status)

Table 1. Summary of BraTS 2020 Training and Validation dataset.

Preprocessing

All scans in the BraTS 2020 dataset came co-registered to the same anatomical template, interpolated to the exact resolution ($1mm^3$), and scull-stripped. We further normalize the MRI scan by subtracting the mean with the standard deviation. Before training, we concatenate all the MRI modalities and ground truth along the channel axis and then extract a 75^{th} slice along the third dimension to give us a $240 \times 240 \times 5$ input image. We combined segmentation maps from pre-trained models with MRI modalities for the validation dataset. We preprocessed all the subjects and saved them on disk to speed up the training. Furthermore, we scaled the age information by dividing by 1767. After prediction we multiply the out by the same number.

Brain tumor segmentation

Since this paper concentrated on proposing a survival prediction model, we utilized a pre-trained model from our previous work[41] to extract segmentation maps, which we then combined with 4 MRI scans input for survival prediction.

Survival Prediction

Our survival model follows a 2D CNN[12] called the squeeze and excitation model for survival prediction (SE-SPNet). Figure 2 summarizes our pipeline for survival prediction using a 2D CNN feature extraction. We extracted 32 imaging features from $240 \times 240 \times 5$ MRI input by first applying 2×2 convolution operation, batch normalization, and Relu non-linearity followed by 4×4 max pooling operation four times. For three times, we apply 2×2 convolution operation, batch normalization, and Relu non-linearity with 6×6, 5×5, and 6×6 average pooling operations in between, respectively. Extracted features are then scaled using squeeze and excitation block[39] before being flattened and rescaled by multiplying them with age information. We then applied batch normalization, three fully connected layers, and sigmoid non-linearity. The output of the network is the overall survival time in days. The survival prediction algorithm is presented in Algorithm 1.

Adaptive Feature Selection

Convolution Neural Networks are excellent for the automatic extraction of a large number of imaging features[24]. However, only some features may be relevant to the task at hand. Therefore, a robust and efficient feature selection mechanism is paramount for survival prediction[24,29]. In literature, several feature selection strategies were employed: univariate Cox regression model[17], extra tree classifier based[24], principal component analysis[42] and correlation matrix[23]. However, these techniques are more than end-to-end. In this work, we applied automatic feature selection using the squeeze and excitation block (SELayer)[39]. The layer automatically learns how to suppress irrelevant features by giving more prominence to salient features for the task at hand.

Given an input feature map $\mathbf{I} \in \mathbb{R}^{H \times W \times C}$, where H, W, and C are the height, width, and number of channels of the input feature map, respectively, the SELayer first applies the global average pooling $\mathbf{F}_{sq}$ to generate channel-wise statistics $\mathbf{z}_c \in \mathbb{R}^C$:

$$\mathbf{z}_c = \mathbf{F}_{sq}(\mathbf{i_c}) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} i_c(i,j) \tag{1}$$

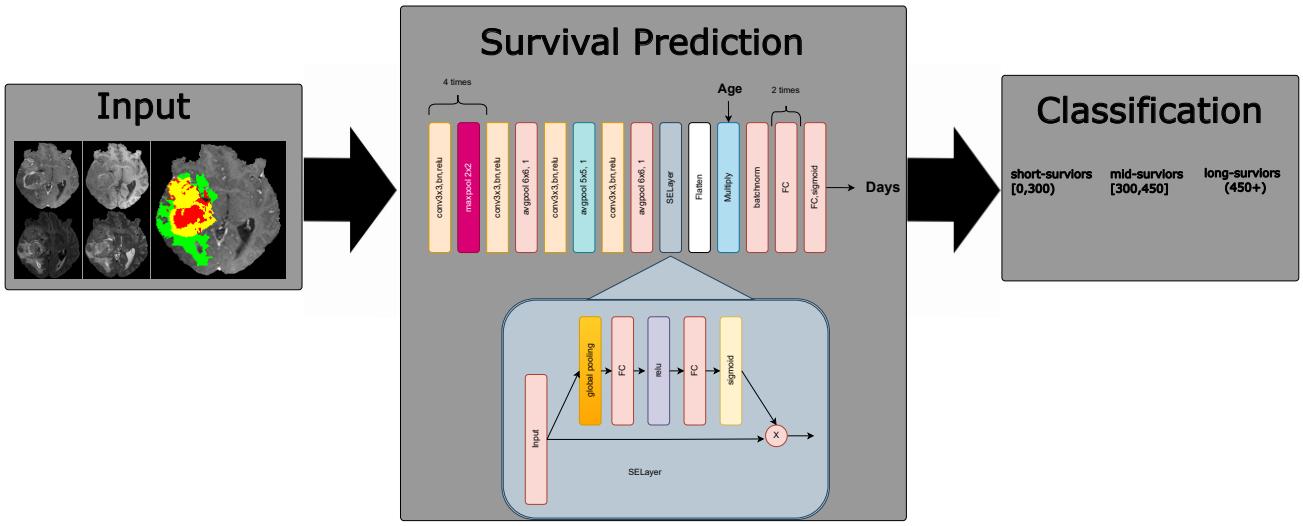


Figure 2. Pipeline for survival prediction.

Algorithm 1 Survival Prediction

Input: Image tensor with dimensions $240 \times 240 \times 5$.
Output: Survival prediction in days.
Apply Convolutional layer, Batch Normalization, and ReLU activation: Output shape $240 \times 240 \times 16$.
Apply MaxPooling: Output shape $120 \times 120 \times 16$.
Apply Convolutional layer, Batch Normalization, and ReLU activation: Output shape $120 \times 120 \times 24$.
Apply MaxPooling: Output shape $60 \times 60 \times 24$.
Apply Convolutional layer, Batch Normalization, and ReLU activation: Output shape $60 \times 60 \times 32$.
Apply MaxPooling: Output shape $30 \times 30 \times 32$.
Apply Convolutional layer, Batch Normalization, and ReLU activation: Output shape $30 \times 30 \times 32$.
Apply MaxPooling: Output shape $15 \times 15 \times 32$.
Apply Convolutional layer, Batch Normalization, and ReLU activation: Output shape $15 \times 15 \times 32$.
Apply Average Pooling: Output shape $10 \times 10 \times 32$.
Apply Convolutional layer, Batch Normalization, and ReLU activation: Output shape $10 \times 10 \times 32$.
Apply Average Pooling: Output shape $6 \times 6 \times 32$.
Apply Convolutional layer, Batch Normalization, and ReLU activation: Output shape $6 \times 6 \times 32$.
Apply Average Pooling: Output shape $1 \times 1 \times 32$.
Apply Squeeze and Excitation (SE) Layer: Output shape $1 \times 1 \times 32$.
Flatten the tensor: Output shape 32.
Multiply selected features with age information: Output shape 32.
Apply Batch Normalization: Output shape 32.
Apply Fully Connected (Linear) layer: Output shape 32.
Apply Fully Connected (Linear) layer: Output shape 16.
Apply Fully Connected (Linear) layer: Output shape 1.
Apply Sigmoid activation: Output shape 1.

where $\mathbf{i}_c \in \mathbf{I} = [\mathbf{i}_1, \mathbf{i}_2, ..., \mathbf{i}_c]$. The output of the squeeze operation is then passed into a sigmoid gating mechanism σ parameterized by a bottleneck with two fully connected (FC) layers around the Relu non-linearity δ. Thus,

$$\mathbf{s} = \sigma(\mathbf{W}_2\delta(\mathbf{W}_1\mathbf{z})) \tag{2}$$

where $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r}\times C}$ and $\mathbf{W}_2 \in \mathbb{R}^{C\times\frac{C}{r}}$. The reduction ratio r is a hyperparameter that controls the capacity and computational cost of the SELayer. We set r to 16. The activation $\mathbf{s}$ is then used to rescale $\mathbf{I}$ to give output feature map $\mathbf{O} \in \mathbb{R}^{H\times W\times C}$:

$$\mathbf{o_c} = s_c \times \mathbf{i}_c \tag{3}$$

where $\mathbf{O} = [\mathbf{o}_1, \mathbf{o}_2, ..., \mathbf{o}_c]$, the result of channel-wise multiplication between the scalar s_c and the feature map $\mathbf{i_c} \in \mathbb{R}^{h\times W}$.

Feature fusion

Previous studies[4,6] demonstrated that a patient's age is an excellent predictor of overall survival. We experimented with various feature fusion strategies like addition, concatenation, and multiplication. Our experimental results demonstrated that rescaling imaging features by multiplying them with age information gives a good performance.

Loss

We used the mean square error as our objective function:

$$\ell = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{4}$$

where n is the number of samples, Y_i is the ground truth and $\hat{Y}_i$ is the predicted value.

Training

The SE-SPNet model was implemented in Pytorch and trained on an NVidia GeForce GTX 1660 Ti. The network took an input patch of 240 × 240, containing four 3D MRI image modalities and a segmentation map concatenated in the channel dimension. All networks were optimized using stochastic gradient descent with an initial learning rate of 0.001 and a Nesterov momentum of 0.9. The training spanned 300 epochs with batch sizes of 64, involving 236 patients with clinical data from the BraTS 2020 training dataset.

EXPERIMENTAL RESULTS AND DISCUSSIONS

Model Complexity

Table 2 compares the number of floating-point operations (FLOPS) and parameter count of 2D CNN and SE-SPNet using the THOP python library. The results show that both models have similar computational requirements despite introducing SELayer as discussed in Section .

Model	Flops (K)	Params (K)
2D CNN[12]	133.440	50.217
SE-SPNet (proposed)	133.440	50.345

Table 2. A comparative computational requirements for the baseline and proposed model.

Comparison with State-of-the-art Techniques

Table 3 compares the proposed model with state-of-the-art techniques on the BraTS 2020 validation dataset. The list excludes McKinley et al.[31] and Marti Asenjo et al.[43], who were tied for the first position in the BraTS 2020 Challenge because they only reported on the BraTS 2020 test dataset which was not available at the time of writing this paper. Bomminemi et al.[32] and Ali et al.[19] were among the top participants in the survival prediction task in the BraTS 2020 Challenge. All metrics were computed by the online evaluation platform[2]. While all the models show relatively low performance in all metrics, our proposed model demonstrated better classification accuracy and SpearmanR correlation results.

[2]https://ipp.cbica.upenn.edu/

Agravat R. and Raval M[18] combined the statistical and radiomic features and age information to train a random forest regression model. Compared to our method, their method performed better in terms of MSE, medianSE and stdSE.

Miron et al.[27] utilized Extra Trees[26] trained with radiomic features for overall survival prediction. Their method demonstrated superior performance in terms of medianSE.

Bommineni et al.[32] won second place in the BraTS 2020 Challenge. Their method used non-imaging features to train a linear regression model. From the results, their methods performed well in terms of MSE, medianSE, and stdSE but could have been better in accuracy than ours.

Ali et al.[19], ranked third in the BraTS 2020 Challenge, fused radiomic and image-based features to train Random Forest (RF) regressor. The results showed that their methods demonstrated better results in MSE, medianSE, and stdSE.

On the other hand, Kaur et al. proposed a 2D CNN that utilized age as the only predictor for survival prediction. From the results, our model shows better accuracy, medianSE, and SpearmanR performance.

Model	Accuracy	MSE	medianSE	stdSE	SpearmanR
Agravat et al.[18]	0.517	116,083.48	43,974.09	168,176.16	0.217
Miron et al.[27]	0.517	197,777.08	**36,397.01**	148,209.10	0.249
Bommineni et al.[32]	0.379	**93,859.54**	67,348.26	**102,092.41**	-
Ali et al.[19]	0.483	105,079.40	37,004.93	146,376.00	0.134
Kaur et al.[12]	0.517	136,783.42	106,608.60	139,210.80	0.299
SE-SPNet (proposed)	**0.586**	160,518.06	92,752.96	200,511.70	**0.336**

Table 3. Comparison with State-of-the-art Techniques on BraTS 2020 Validation set (29 cases) .

Ablation Studies

Effects of feature fusion strategies

To study the efficacy of the proposed feature fusion strategy described in Section , several experiments were conducted. Table 4 shows that multiplying deep features with age produces better performance in terms of accuracy and SpearmanR. However, adding age produces better results in terms of MSE and MedianSE. On the other hand, concatenating features produced better accuracy than addition.

Fusion Strategy	Accuracy	MSE	medianSE	stdSE	SpearmanR
Multiply	**0.59**	160,518.06	92,752.96	200,511.70	**0.34**
Concatenate	0.45	147,925.90	37,708.67	248,356.06	0.27
Add	0.41	**83,131.46**	**25,356.69**	109,818.22	0.30

Table 4. Effects of feature fusion strategies on BraTS 2020 Validation set (29 cases) .

Effects of Adaptive feature selection (AFS)

This paper also examined the effects of adding the squeeze and excitation block (SElayer)[39] to the 2D CNN[12] for adaptive feature selection as described in Section . Table 5 clearly shows that adding the SELayer gives better performance in terms of accuracy, stdSE, and SpearmanR. In contrast, the model without the SElayer performed better in terms of MSE and MedianSE.

Model	Accuracy	MSE	medianSE	stdSE	SpearmanR
Without ADF	0.41	**157,441.74**	**77,405.55**	253,655.87	0.29
With ADF	**0.59**	160,518.06	92,752.96	**200,511.70**	**0.34**

Table 5. Effects of Adaptive feature selections on BraTS 2020 Validation set (29 cases) .

Effects of the optimizer

The effects of SGD and ADAM optimizers on the BraTS 2020 Validation set and training loss were studied, and the results are reported in Table 6 and Figure 3, respectively. While the SGD optimizer produced a model with superior performance in terms of accuracy and SpearmanR, it was noted that the model's training took time to converge. In contrast, better performance in

terms of MSE, medianSE, and stdSE was observed with the ADAM optimizer. However, it was observed that the network quickly overfits just after a few epochs.

Table 6. Effects of training optimizer .

Fusion Strategy	Accuracy	MSE	medianSE	stdSE	SpearmanR
SGD	**0.59**	160,518.06	92,752.96	200,511.70	**0.34**
Adam	0.35	**119,984.57**	**58,543.50**	**125,475.92**	-0.03

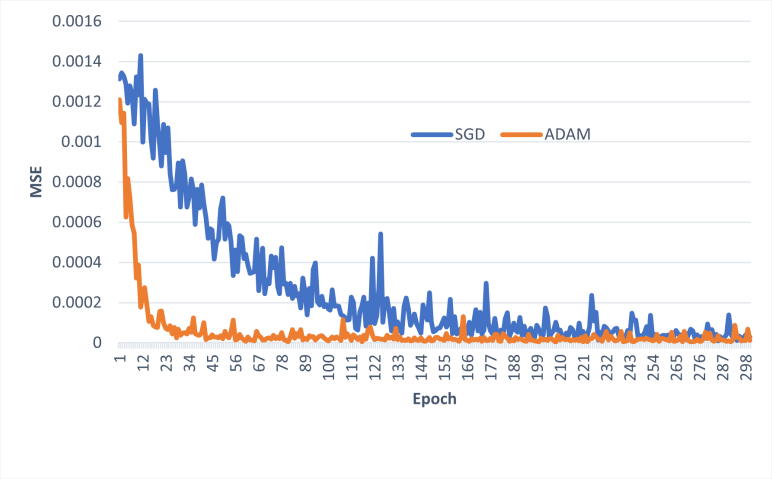


Figure 3. Effects of the optimiser on the training loss.

Discussion

Accurate and reproducible overall survival prediction is critical for treatment planning and palliative medicine[4]. Survival prediction can help identify high-risk patients who can receive a more aggressive treatment regime to improve their quality of life[44]. In this work, we proposed a 2D CNN for overall survival prediction. Since deep learning features extract a huge amount of imaging features, robust ways of selecting relevant for survival prediction are needed. To this end, we propose incorporating squeeze and excitation blocks to adaptively select salient features for the task while suppressing less important ones.

Table 4 provides the performance analysis of our proposed techniques against the state-of-the-art methods. Although our method outperforms the state-of-the-art methods in classification accuracy, the accuracy of 59% is still low. We can attribute this low performance to several reasons. First, the BraTS 2020 dataset only contains age and overall survival days. No other information regarding sexual, any underlying conditions, and treatment received[25]. Second, Imaging features are not enough to fully characterize the heterogineity of gliomas tumors[25,45]. Third, the BraTS 2020 dataset only contains a few training samples, resulting in models overfitting quickly, especially deep learning models[5,6]. In Figure 3, we see a model trained using the ADAM optimizer overfits just after a few epochs. Last, OS times accuracy depends on the quality of tumor segmentation.

Prior studies[6] have observed that there is a high correction between age and survival prediction. We experiment with various ways of combining age information and imaging features. As shown in Table 5, multiplying age information with deep features yielded better classification accuracy. This strategy has the effect of giving more prominence to salient features from the adaptive feature selection module.

Since the primary goal of overall survival prediction in BraTS challenges is to classify patients into short-survivors (<10 months), mid-survivors (between 10 and 15 months), and long-survival (> 15 months), in the future, we would like to directly optimize the classification accuracy by incorporating cross-entropy loss.

CONCLUSIONS

This work presented an efficient deep-learning approach for predicting the overall survival of glioma patients. By incorporating a squeeze and excitation block, the proposed method enables adaptive feature selection, enhancing the model's ability to capture relevant information for survival prediction. Moreover, our approach adopts an end-to-end framework, eliminating the need for complex feature engineering, which is common in traditional machine learning methods for survival prediction. A notable aspect of our methodology involves integrating age information by multiplying re-weighted deep features, contributing to improved predictive performance. The comprehensive evaluation conducted on the BraTS 2020 dataset underscores the efficacy of our proposed method, showcasing its superiority over state-of-the-art techniques in terms of classification accuracy.

References

1. Holland, E. C. Progenitor cells and glioma formation. *Curr. Opin. Neurol.* **14**, 683 (2001).
2. Cahill, D. & Turcan, S. Origin of Gliomas. *Semin. Neurol.* **38**, 5–10, DOI: 10.1055/s-0037-1620238 (2018).
3. Menze, B. H. *et al.* The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS), DOI: 10.1109/TMI.2014.2377694 (2015).
4. Kofler, F. *et al.* A Baseline for Predicting Glioblastoma Patient Survival Time with Classical Statistical Models and Primitive Features Ignoring Image Information. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11992, 254–261, DOI: 10.1007/978-3-030-46640-4_24 (Springer International Publishing, Cham, 2020).
5. Behrad, F. & Saniee Abadeh, M. Evolutionary convolutional neural network for efficient brain tumor segmentation and overall survival prediction. *Expert. Syst. with Appl.* **213**, 118996, DOI: 10.1016/j.eswa.2022.118996 (2023).
6. González, S. R., Zemmoura, I. & Tauber, C. 3D Brain Tumor Segmentation and Survival Prediction Using Ensembles of Convolutional Neural Networks. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, 241–254, DOI: 10.1007/978-3-030-72087-2_21 (Springer International Publishing, Cham, 2021).
7. Suter, Y. *et al.* Deep Learning versus Classical Regression for Brain Tumor Patient Survival Prediction (2018). 1811.04907.
8. Li, Y. & Shen, L. Deep Learning Based Multimodal Brain Tumor Diagnosis. In Crimi, A., Bakas, S., Kuijf, H., Menze, B. & Reyes, M. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, 149–158, DOI: 10.1007/978-3-319-75238-9_13 (Springer International Publishing, Cham, 2018).
9. Zhao, G., Jiang, B., Zhang, J. & Xia, Y. Segmentation then Prediction: A Multi-task Solution to Brain Tumor Segmentation and Survival Prediction. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, 492–502, DOI: 10.1007/978-3-030-72084-1_44 (Springer International Publishing, Cham, 2021).
10. Hermoza, R., Maicas, G., Nascimento, J. C. & Carneiro, G. Censor-Aware Semi-supervised Learning for Survival Time Prediction from Medical Images. In Wang, L., Dou, Q., Fletcher, P. T., Speidel, S. & Li, S. (eds.) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, Lecture Notes in Computer Science, 213–222, DOI: 10.1007/978-3-031-16449-1_21 (Springer Nature Switzerland, Cham, 2022).
11. Trinh, D.-L., Kim, S.-H., Yang, H.-J. & Lee, G.-S. The Efficacy of Shape Radiomics and Deep Features for Glioblastoma Survival Prediction by Deep Learning. *Electronics* **11**, 1038, DOI: 10.3390/electronics11071038 (2022).
12. Kaur, G., Rana, P. S. & Arora, V. Automated Neural Network-based Survival Prediction of Glioblastoma Patients Using Pre-operative MRI and Clinical Data. *IETE J. Res.* **0**, 1–17, DOI: 10.1080/03772063.2023.2217142 (2023).
13. Oktay, O. *et al.* Attention U-Net: Learning Where to Look for the Pancreas (2018). 1804.03999.
14. Bakas, S. *et al.* Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge (2019). 1811.02629.
15. Bakas, S. *et al.* Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci. Data* **4**, 170117, DOI: 10.1038/sdata.2017.117 (2017).
16. Isensee, F., Kickingereder, P., Wick, W., Bendszus, M. & Maier-Hein, K. H. Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge. *arXiv:1802.10508 [cs]* (2018). 1802.10508.
17. Shboul, Z., Vidyaratne, L., Alam, M., Reza, S. M. S. & Iftekharuddin, K. M. Glioblastoma and Survival Prediction. *Brainlesion: Glioma, Multiple Sclerosis, Stroke Trauma. Brain Inj. BrainLes (Workshop)* **10670**, 358–368, DOI: 10.1007/978-3-319-75238-9_31 (2018).

18. Agravat, R. R. & Raval, M. S. 3D Semantic Segmentation of Brain Tumor for Overall Survival Prediction. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, 215–227, DOI: 10.1007/978-3-030-72087-2_19 (Springer International Publishing, Cham, 2021).

19. Ali, M. J., Akram, M. T., Saleem, H., Raza, B. & Shahid, A. R. Glioma Segmentation Using Ensemble of 2D/3D U-Nets and Survival Prediction Using Multiple Features Fusion. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, 189–199, DOI: 10.1007/978-3-030-72087-2_17 (Springer International Publishing, Cham, 2021).

20. Anand, V. K. *et al.* Brain Tumor Segmentation and Survival Prediction Using Automatic Hard Mining in 3D CNN Architecture. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, 310–319, DOI: 10.1007/978-3-030-72087-2_27 (Springer International Publishing, Cham, 2021).

21. Huang, H. *et al.* Overall Survival Prediction for Gliomas Using a Novel Compound Approach. *Front. Oncol.* **11**, 724191, DOI: 10.3389/fonc.2021.724191 (2021).

22. Parmar, B. & Parikh, M. Brain Tumor Segmentation and Survival Prediction Using Patch Based Modified 3D U-Net. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, 398–409, DOI: 10.1007/978-3-030-72087-2_35 (Springer International Publishing, Cham, 2021).

23. Rafi, A., Madni, T. M., Janjua, U. I., Ali, M. J. & Abid, M. N. Multi-level dilated convolutional neural network for brain tumour segmentation and multi-view-based radiomics for overall survival prediction. *Int. J. Imaging Syst. Technol.* **31**, 1519–1535, DOI: 10.1002/ima.22549 (2021).

24. Fiaz, K. *et al.* Brain tumor segmentation and multiview multiscale-based radiomic model for patient's overall survival prediction. *Int. J. Imaging Syst. Technol.* **32**, 982–999, DOI: 10.1002/ima.22678 (2022).

25. Glory Precious, J., Keren Evangeline, I. & Kirubha, S. P. A. Brain tumour segmentation and survival prognostication using 3D radiomics features and machine learning algorithms. *Comput. Methods Biomech. Biomed. Eng. Imaging & Vis.* **11**, 1803–1817, DOI: 10.1080/21681163.2023.2189487 (2023).

26. Geurts, P., Ernst, D. & Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **63**, 3–42, DOI: 10.1007/s10994-006-6226-1 (2006).

27. Miron, R., Albert, R. & Breaban, M. A Two-Stage Atrous Convolution Neural Network for Brain Tumor Segmentation and Survival Prediction. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, 290–299, DOI: 10.1007/978-3-030-72087-2_25 (Springer International Publishing, Cham, 2021).

28. Bakas, S. *et al.* Overall survival prediction in glioblastoma patients using structural magnetic resonance imaging (MRI): Advanced radiomic features may compensate for lack of advanced MRI modalities. *J. Med. Imaging* **7**, 031505, DOI: 10.1117/1.JMI.7.3.031505 (2020).

29. Rathore, F. A. *et al.* Survival Prediction of Glioma Patients from Integrated Radiology and Pathology Images Using Machine Learning Ensemble Regression Methods. *Appl. Sci.* **12**, 10357, DOI: 10.3390/app122010357 (2022).

30. Chen, H. *et al.* A subregion-based survival prediction framework for GBM via multi-sequence MRI space optimization and clustering-based feature bundling and construction. *Phys. Medicine & Biol.* **68**, 125005, DOI: 10.1088/1361-6560/acd6d2 (2023).

31. McKinley, R. *et al.* Uncertainty-Driven Refinement of Tumor-Core Segmentation Using 3D-to-2D Networks with Label Uncertainty. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12658, 401–411, DOI: 10.1007/978-3-030-72084-1_36 (Springer International Publishing, Cham, 2021).

32. Bommineni, V. L. PieceNet: A Redundant UNet Ensemble. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, 331–341, DOI: 10.1007/978-3-030-72087-2_29 (Springer International Publishing, Cham, 2021).

33. Hussain, S. *et al.* ETISTP: An Enhanced Model for Brain Tumor Identification and Survival Time Prediction. *Diagnostics* **13**, 1456, DOI: 10.3390/diagnostics13081456 (2023).

34. Yan, T. *et al.* Survival prediction for patients with glioblastoma multiforme using a Cox proportional hazards denoising autoencoder network. *Front. Comput. Neurosci.* **16** (2023).

35. Jungo, A. *et al.* Towards Uncertainty-Assisted Brain Tumor Segmentation and Survival Prediction. In Crimi, A., Bakas, S., Kuijf, H., Menze, B. & Reyes, M. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 10670, 474–485, DOI: 10.1007/978-3-319-75238-9_40 (Springer International Publishing, Cham, 2018).

36. Chato, L., Kachroo, P. & Latifi, S. An Automatic Overall Survival Time Prediction System for Glioma Brain Tumor Patients Based on Volumetric and Shape Features. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, 352–365, DOI: 10.1007/978-3-030-72087-2_31 (Springer International Publishing, Cham, 2021).

37. Priya, S. *et al.* Survival prediction in glioblastoma on post-contrast magnetic resonance imaging using filtration based first-order texture analysis: Comparison of multiple machine learning models. *The Neuroradiol. J.* **34**, 355–362, DOI: 10.1177/1971400921990766 (2021).

38. Wang, Q. *et al.* MPSurv: End-to-End Multi-model Pseudo-Label Model for Brain Tumor Survival Prediction with Population Information Integration. In Qin, W. *et al.* (eds.) *Computational Mathematics Modeling in Cancer Analysis*, vol. 14243, 120–130, DOI: 10.1007/978-3-031-45087-7_13 (Springer Nature Switzerland, Cham, 2023).

39. Hu, J., Shen, L., Albanie, S., Sun, G. & Wu, E. Squeeze-and-Excitation Networks. *arXiv:1709.01507 [cs]* (2019). 1709.01507.

40. Fu, J. *et al.* Dual Attention Network for Scene Segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3146–3154 (2019).

41. Magadza, T. & Viriri, S. Efficient nnU-Net for Brain Tumor Segmentation. *IEEE Access* **11**, 126386–126397, DOI: 10.1109/ACCESS.2023.3329517 (2023).

42. Das, S., Bose, S., Nayak, G. K., Satapathy, S. C. & Saxena, S. Brain tumor segmentation and overall survival period prediction in glioblastoma multiforme using radiomic features. *Concurr. Comput. Pract. Exp.* **34**, e6501, DOI: 10.1002/cpe.6501 (2022).

43. Marti Asenjo, J. & Martinez-Larraz Solís, A. MRI Brain Tumor Segmentation Using a 2D-3D U-Net Ensemble. In Crimi, A. & Bakas, S. (eds.) *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, 354–366, DOI: 10.1007/978-3-030-72084-1_32 (Springer International Publishing, Cham, 2021).

44. Liang, J. *et al.* Prognostic factors of patients with Gliomas – an analysis on 335 patients with Glioblastoma and other forms of Gliomas. *BMC Cancer* **20**, 35, DOI: 10.1186/s12885-019-6511-6 (2020).

45. Hermoza, R., Maicas, G., Nascimento, J. C. & Carneiro, G. Post-Hoc Overall Survival Time Prediction From Brain MRI. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 1476–1480, DOI: 10.1109/ISBI48211.2021.9433877 (2021).

Author contributions

The following statements should be used "Conceptualization, T.M. and S.V.; methodology, T.M.; software, T.M.; validation, T.M. and S.V.; formal analysis, T.M.;investigation, T.M.; resources, S.V.; data curation, T.M.; writing—original draft preparation, T.M.; writing—review and editing, S.V.; visualization, T.M.; supervision,S.V.; project administration, T.M.; funding acquisition, S.V.

Author contributions

The authors declare no competing interests.

Data availability

The dataset used in this research work is benchmarked and publicly available.

Additional information

Correspondence and requests for materials should be addressed to S.V

4.2 Conclusion

This chapter presented an end-to-end deep learning model for overall survival prediction. The model incorporates the squeeze-and-excitation module for adaptive feature extraction. We experimented with several feature fusion strategies and provided an extensive analysis of our proposed techniques. Experimental results on BraTS 2020 datasets demonstrate the efficacy of our proposed methods.

Chapter 5

Results and Discussion

5.1 Introduction

This chapter delves into a comprehensive analysis of the outcomes spanning this thesis and harmonizes the findings derived from the various papers within this thesis to illustrate the strategies employed to tackle challenges associated with the automatic segmentation of brain tumors and survival prediction. It begins by discussing experimental results for the brain tumor segmentation task in Section 5.2 and Section 5.3 followed by survival prediction in Section 5.4.

5.2 Brain Tumor Segmentation Using Partial Depthwise Separable Convolutions

This section presents the results of brain tumor segmentation experiments described in Section 3.2.

The network follows a 3D U-Net [36] architecture with five layers in total. Each layer contains two ResNet-like [14] style convolution blocks in both the encoding and decoding paths. To reduce the computational complexity of the network, all the bottom three layers are systematically replaced with depthwise separable convolutions. A spatial attention mechanism [26] is incorporated in the skip connection to compensate for the slight performance decrease caused by replacing conventional convolution operations with depthwise separable convolutions.

The network was trained on BraTS 2020 [3, 4, 24] datasets with 369 training and 125 validation samples. Low-cost data augmentation techniques, as proposed by Ellis et al. [11], were performed. To validate the model's performance, segmentation results were submitted to an online evaluation platform[1], which benchmarked the proposed segmentation maps in terms of the Dice score and the Hausdorff distance.

Figure 5.1 shows a qualitative inspection of two randomly selected predictions on the training set. Our model sometimes performs exceptionally well and, on the other hand, poorly. Ensemble methods [24] are known to improve the robustness of segmentation performance.

Table 5.1 shows the quantitative segmentation performance of our proposed model as compared to the state-of-the-art methods on BraTS 2020 validation datasets as computed

[1] https://ipp.cbica.upenn.edu/

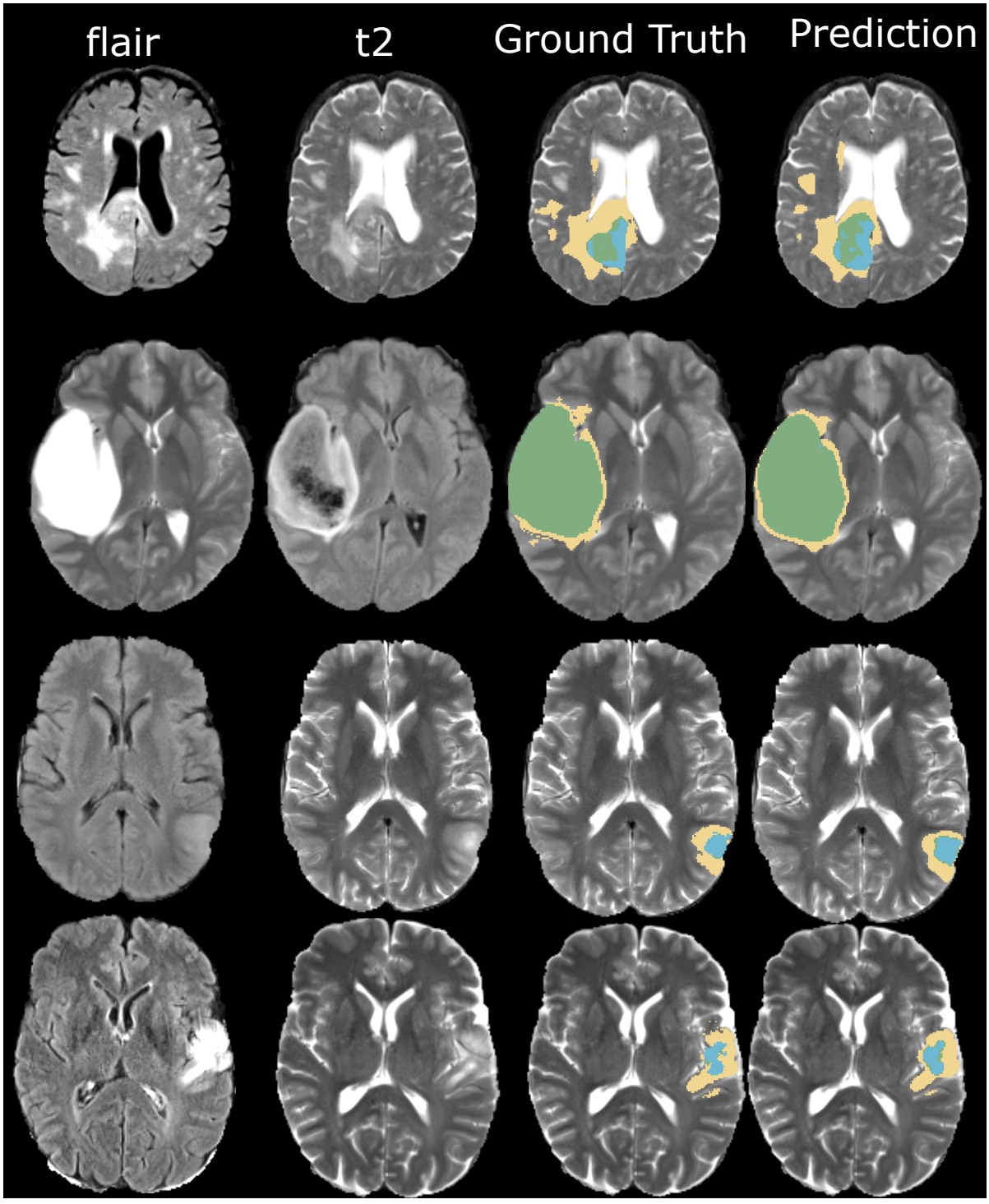


Figure 5.1: Qualitative inspections of two randomly selected predictions on the training set. Edema is shown in yellow, necrosis in green, and enhancing tumor in blue.

Table 5.1: Mean performance metrics on the BraTS 2020 Validation dataset of the proposed method were compared to the state-of-the-art methods in terms of the dice similarity score. The model was trained for 400 epochs. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core. Ensemble10 - Ensemble of 10 models.

	Dice			
Method	ET	WT	TC	Mean
Isensee et al. [18]	**0.7989**	0.9124	0.8506	**0.8540**
Jia et al. [20]	0.7875	**0.9129**	**0.8546**	0.8517
Y. Yuan [34]	0.7927	0.9108	0.8529	0.8521
Wang et al. [33]	0.7873	0.9009	0.8173	0.8352
Ensemble10 (proposed)	0.7745	0.9042	0.8286	0.8357

Best values are shown in bold

Table 5.2: Mean performance metrics on BraTS 2020 Validation dataset of our proposed method compared to the state-of-the-art methods in terms of Hausdorff distance (95%). ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core. Ensemble10 - Ensemble of 10 models. *: Input size of $96 \times 128 \times 80$.

	Hausdorff95			
Method	ET	WT	TC	Mean
Isensee et al. [18]	23.50	**3.69**	7.82	11.67
Jia et al. [20]	26.58	4.18	**4.97**	11.91
Y. Yuan [34]	18.20	4.10	5.99	**9.43**
Wang et al. [33]	**17.95**	4.96	9.77	10.89
Single model* (proposed)	22.33	5.51	7.19	11.68
Ensemble10 (proposed)	25.07	5.61	7.10	12.60

Best values are shown in bold.

by the online evaluation platform. From the results, no single method outperformed all methods in all metrics. Our ensemble of ten light-weight models demonstrated a competitive performance regarding dice score.

Isensee et al. [18] used an ensemble of 25 models, which amount to 2 Gig in the compressed form of disk space, to win the BraTS 2020 challenge. Compared to our proposed methods, which only require 25 MB of disk space with each of the ten models trained for a few epochs, their method produced a mean performance gain of 1.83%

Table 5.2 shows the mean performance evaluation of our proposed models against state-of-the-art techniques in terms of Hausdorff distance (95%). No single model outranking other methods in all metrics. Y. Yuan [34] achieved the best mean performance overall. Interestingly, our single model, trained with an input patch size $96 \times 128 \times 80$, performed better than an ensemble of 25 models by Isensee et al. [18] in both the enhancing tumor and tumor core regions. The inherent class imbalance makes it difficult to detect small liaison by using large import patch size [8].

5.3 Efficient nnU-Net for Brain Tumor Segmentation

nnU-Net [16] is one of the leading open-source deep learning frameworks for medical image analysis. It has been used to win various challenges, including the BraTS 2020 [17] and 2021 [23] brain tumor segmentation task. Nevertheless, as discussed in Section 3.3 and Section 5.2, despite its state-of-the-art performance, it is computationally expensive. The framework uses standard convolution operations, which are computationally expensive because they perform spatial and channel-wise correlation in one go [9].

Motivated by the above problem, several modifications to the nnU-Net were proposed to lower its computational complexity while retaining its state-of-the-art segmentation performance. Section 3.3 described in detail the proposed changes and experiments conducted to study the efficacy of these techniques. The modifications are summarized as follows:

- **Model 1**: all the convolution operations in nnU-Net were replaced with depthwise separable convolutions.
- **Model 2**: residual connections were added to Model 1.
- **Model 3**: the residual connection in Model 2 was modified by applying the ReLU operation after adding the residual units. Furthermore, a shuffle attention mechanism was introduced in the skip connections.

All models were trained for 400 epochs. The online BraTS 2020 evaluation platform was used to validate the segmentation performance of the modifications in terms of dice score and Hausdorff distance. Additionally, the THOP[2] Python library was used to calculate model complexity in terms of floating-point operations (FLOPS) and the number of parameters (Params).

Table 5.3 and Table 5.4 report the computational analysis of the proposed modifications against the state-of-the-art methods with and without model ensemble, respectively. The models demonstrated superiority in all metrics, with the proposed changes effectively reducing computational complexity.

In Table 5.5 and Table 5.6, the segmentation performance of the models is reported against the state-of-the-art methods with and without model ensemble, respectively. As shown in Table 5.5, the approach produced comparable results to the method proposed by Isensee et al. in terms of Dice coefficient for the entire tumor. Additionally, a slight enhancement in the Hausdorff distance for the tumor core was observed, with similar performance across other metrics. Importantly, the method achieved these outcomes while utilizing considerably fewer computational resources. Table 5.6 shows that the proposed method achieved superior performance with minimal computational complexity.

5.4 Efficient Survival Prediction for Gliomas using deep learning

Survival prediction plays a crucial role in shaping effective treatment plans and guiding decisions on the intensity of interventions such as surgery, radiation therapy, and

[2]https://github.com/Lyken17/pytorch-OpCounter

Table 5.3: Computational analysis of the proposed models against the state-of-the-art methods with the model ensemble. † - Ensemble of models trained with 5-fold cross-validation.

Model	Epochs	FLOPS	Params.
Isensee et al. [17]	1000	539.56G	31.20M
Isensee et al. [17] †	1000	539.56G	31.20M
Jia et al. [20]	450	621.09G	26.07M
Jia et al. [20]†	450	621.09G	26.07M
Y. Yuan [34]	300	-	16.50M
Wang et al. [33]	1000	-	-
Wang et al. [32]†	8000	333.00G	32.99M
Model 1 (proposed)	400	**55.26G**	**2.51M**
Model 2 (proposed)	400	<u>75.39G</u>	<u>7.89M</u>
Model 3 (proposed)	400	75.46G	<u>7.89M</u>

Best values are shown in bold, and second best are underlined.

Table 5.4: Computational analysis of the proposed models against the state-of-the-art methods without model ensemble.

Model	Epochs	FLOPS	Params.
Y. Yuan [34]	300	-	16.50M
Wang et al. [33]	1000	-	-
Raza et al. [29]	100	374.04G	30.47M
Daza et al. [10]	-	**49.82G**	<u>4.02M</u>
Model 1 (proposed)	400	**55.26G**	**2.51M**
Model 2 (proposed)	400	<u>75.39G</u>	<u>7.89M</u>
Model 3 (proposed)	400	75.46G	<u>7.89M</u>

Best values are shown in bold, and second best are underlined.

Table 5.5: Segmentation performance on BraTS 2020 Validation dataset compared to the state-of-the-art with the model ensemble. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core. † - Ensemble of models trained with 5-fold cross-validation.

Model	**Dice**			**HD95**		
	ET	WT	TC	ET	WT	TC
Isensee et al. [17]	**0.799**	<u>0.912</u>	0.851	23.50	**3.69**	7.82
Isensee et al. [17] †	**0.799**	<u>0.912</u>	**0.857**	26.41	<u>3.73</u>	5.64
Jia et al. [20]	0.788	**0.913**	<u>0.855</u>	26.58	4.18	**4.97**
Jia et al. [20]†	0.784	**0.913**	0.835	26.50	4.18	<u>5.52</u>
Y. Yuan [34]	<u>0.793</u>	0.911	0.853	<u>18.20</u>	4.10	5.99
Wang et al. [33]	0.787	0.908	0.856	35.01	4.71	5.70
Wang et al. [32]†	0.787	0.901	0.817	**17.95**	4.97	9.77
Model 1 (proposed)	0.792	<u>0.912</u>	0.848	29.31	4.41	6.20
Model 2 (proposed)	0.792	0.910	0.841	26.42	5.20	9.19
Model 3 (proposed)	0.788	0.908	0.845	29.61	5.11	5.92

Best values are shown in bold, and second best are underlined.

Table 5.6: Segmentation performance on BraTS 2020 Validation dataset compared to the state-of-the-art without model ensemble. ET - Enhancing tumor, WT - Whole tumor, TC - Tumor core.

Model	**Dice**			**HD95**		
	ET	WT	TC	ET	WT	TC
Y. Yuan [34]	0.785	0.904	0.842	**20.35**	5.49	8.34
Wang et al. [33]	0.785	0.907	0.837	32.25	<u>4.39</u>	8.34
Raza et al. [29]	**0.800**	0.866	0.836	29.82	6.78	7.36
Daza et al. [10]	<u>0.794</u>	0.897	0.845	29.82	**3.59**	6.47
Model 1 (proposed)	0.792	**0.912**	**0.848**	29.31	4.41	<u>6.20</u>
Model 2 (proposed)	0.792	<u>0.910</u>	0.841	<u>26.42</u>	5.20	9.19
Model 3 (proposed)	0.788	0.908	<u>0.845</u>	29.61	5.11	**5.92**

Best values are shown in bold, and second best are underlined.

chemotherapy. Additionally, survival prediction holds significant importance in the realm of palliative care [22].

Despite the success of deep learning techniques in achieving state-of-the-art segmentation performance, their application in survival prediction remains limited. This is primarily attributed to their high demand for training samples, a challenging aspect in the medical domain. Furthermore, the inherent nature of deep learning models often leads to overfitting issues due to generating numerous features and parameters.

In contrast, traditional machine-learning techniques have been prevalent in survival prediction. However, their reliance on intricate and robust feature extraction and selection processes for competitive performance poses a drawback. These techniques typically involve separate phases for feature extraction, selection, and model training.

An efficient deep-learning model for survival prediction is presented in response to these challenges. The approach incorporated the squeeze-and-excitation block for adaptive feature selection in an end-to-end fashion by leveraging the segmentation maps obtained from the segmentation task detailed in Chapter 3.

As discussed in Chapter 4, various feature fusion strategies were experimented with, and the efficacy of the adaptive feature selection on the BraTS 2020 dataset was extensively studied. Survival predictions in days were submitted to the online evaluation platform, which benchmarked the methods in terms of classification accuracy, mean square error (MSE), median square error (medianSE), standard deviation of the square error (stdSE), and Spearman rank correlation (SpearmanR).

The main objective for the survival prediction, as given by the BraTS Challenge, is to accurately classify patients into short-survivors (< 10 months), mid-survivors (between 10 and 15 months), and long-survivors (> 15 months).

Table 5.7 presents a quantitative analysis of the performance of our proposed method against state-of-the-art methods. Our method achieves superior classification accuracy and SpearmanR correlation performance but performs poorly in terms of MSE.

Despite surpassing the state-of-the-art methods in classification accuracy, our method still exhibits a relatively low accuracy of 59%. Several factors contribute to this suboptimal performance. Firstly, the BraTS 2020 dataset is limited to age and overall survival days, lacking essential information on sex, underlying conditions, and treatment details [12]. Secondly, relying solely on imaging features proves insufficient for fully characteriz-

Table 5.7: Comparison with State-of-the-art Techniques on BraTS 2020 Validation.

Model	Accuracy	MSE	medianSE	stdSE	SpearmanR
Agravat et al. [1]	0.517	116,083.48	43,974.09	168,176.16	0.217
Miron et al. [25]	0.517	197,777.08	**36,397.01**	148,209.10	0.249
Bommineni et al. [6]	0.379	**93,859.54**	67,348.26	**102,092.41**	-
Ali et al. [2]	0.483	105,079.40	37,004.93	146,376.00	0.134
Kaur et al. [21]	0.517	136,783.42	106,608.60	139,210.80	0.299
Proposed	**0.586**	160,518.06	92,752.96	200,511.70	**0.336**

Best values are shown in bold.

ing the heterogeneity of glioma tumors [12, 15]. Lastly, the scarcity of training samples in the BraTS 2020 dataset, particularly for deep learning models, leads to rapid overfitting [5, 13].

5.5 Conclusion

This chapter presented experimental results on automatic brain tumor segmentation and survival prediction. A detailed discussion of challenges in the domain and strategies to address them was presented. Rigorous evaluation against state-of-the-art methods was performed using the BraTS 2020 dataset.

Our brain tumor segmentation methods demonstrated competitive performance, achieving accurate results with minimal computational resources. This accomplishment highlights the efficiency of our segmentation approach in accurately delineating tumor boundaries.

Simultaneously, our survival prediction model exhibited superior performance, especially regarding classification accuracy. This success underscores the robustness and effectiveness of our proposed model, suggesting its potential for valuable insights into patient prognosis.

In summary, our experimental results demonstrate the effectiveness of our approaches in tackling challenges related to automatic brain tumor segmentation and survival prediction. Our methods compete favorably with state-of-the-art techniques and showcase efficiency in computational resource utilization and superior predictive capabilities, emphasizing their potential significance in clinical applications.

Chapter 6

Conclusion and Future Work

6.1 Introduction

In this concluding chapter, we reflect on our research's key findings and contributions, addressing the challenges in automatic brain tumor segmentation and survival prediction. We explore the significance of our work in the broader context of medical imaging and prognosis.

6.2 Overview

Our research aimed to develop computationally efficient deep-learning models for automatic brain tumor segmentation and overall survival prediction. We tackled the limitations of existing methods by introducing novel approaches that balance performance and computational complexity.

Our proposed methods demonstrated competitive accuracy in brain tumor segmentation while operating with minimal computational resources. The success in accurately delineating tumor boundaries, particularly in the challenging context of glioma sub-regions, underscores the efficiency of our segmentation approach.

Simultaneously, our survival prediction model exhibited superior performance, particularly in classification accuracy. This achievement not only adds a valuable tool for prognosis but also emphasizes the robustness and efficacy of our proposed model.

6.3 Contribution to Knowledge

Our research makes several notable contributions to the field of medical imaging and prognosis:

- **Efficient Segmentation Models**: We introduced novel network architectures that leverage depthwise separable convolutions and innovative extensions to the U-Net model, balancing accuracy with computational efficiency in brain tumor segmentation.

- **Survival Prediction Model**: The end-to-end deep learning model for overall survival prediction, incorporating the squeeze-and-excitation module, contributes to the growing work to improve patient outcome predictions.

- **Rigorous Evaluation**: Our thorough evaluation against state-of-the-art methods using the BraTS 2020 dataset establishes our proposed techniques' credibility and potential applicability in clinical setups.

.

6.4 Future work

As we conclude this research, several avenues for future exploration emerge:

- **Ensemble Models for Segmentation**: Extensive experimentation with an ensemble of lightweight models could enhance segmentation performance, providing a robust and reliable solution for diverse brain tumor cases.
- **Multiresolution Approaches**: Exploring multiple resolutions to capture long-range dependencies may improve segmentation accuracy, especially in cases where structural variability among tumors poses a challenge.
- **Incorporating Radiomic Features**: Integrating radiomic features into the survival prediction model can provide a more comprehensive understanding of tumor characteristics, potentially improving the accuracy of prognostic assessments.

6.5 Final Thoughts

In conclusion, our research has made significant strides in the development of efficient and accurate models for automatic brain tumor segmentation and survival prediction. The findings presented in this thesis contribute to the ongoing efforts to advance the field of medical imaging and prognosis. As we look to the future, continued exploration and innovation are essential to address the evolving challenges and further enhance the capabilities of computer-assisted diagnosis in clinical settings.

List of References

[1] Rupal R. Agravat and Mehul S. Raval. 3D Semantic Segmentation of Brain Tumor for Overall Survival Prediction. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 12659, pages 215–227. Springer International Publishing, 2021.

[2] Muhammad Junaid Ali, Muhammad Tahir Akram, Hira Saleem, Basit Raza, and Ahmad Raza Shahid. Glioma Segmentation Using Ensemble of 2D/3D U-Nets and Survival Prediction Using Multiple Features Fusion. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 12659, pages 189–199. Springer International Publishing, 2021.

[3] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S. Kirby, John B. Freymann, Keyvan Farahani, and Christos Davatzikos. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Scientific Data*, 4(1):170117, 2017.

[4] Spyridon Bakas, Mauricio Reyes, Andras Jakab, Stefan Bauer, Markus Rempfler, Alessandro Crimi, Russell Takeshi Shinohara, Christoph Berger, Sung Min Ha, Martin Rozycki, Marcel Prastawa, Esther Alberts, Jana Lipkova, John Freymann, Justin Kirby, Michel Bilello, Hassan Fathallah-Shaykh, Roland Wiest, Jan Kirschke, Benedikt Wiestler, Rivka Colen, Aikaterini Kotrotsou, Pamela Lamontagne, Daniel Marcus, Mikhail Milchenko, Arash Nazeri, Marc-Andre Weber, Abhishek Mahajan, Ujjwal Baid, Elizabeth Gerstner, Dongjin Kwon, Gagan Acharya, Manu Agarwal, Mahbubul Alam, Alberto Albiol, Antonio Albiol, Francisco J. Albiol, Varghese Alex, Nigel Allinson, Pedro H. A. Amorim, Abhijit Amrutkar, Ganesh Anand, et al. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. *arXiv:1811.02629[cs]*, 2019.

[5] Fatemeh Behrad and Mohammad Saniee Abadeh. Evolutionary convolutional neural network for efficient brain tumor segmentation and overall survival prediction. *Expert Systems with Applications*, 213:118996, 2023.

[6] Vikas L. Bommineni. PieceNet: A Redundant UNet Ensemble. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, pages 331–341. Springer International Publishing, 2021.

[7] Daniel Cahill and Sevin Turcan. Origin of Gliomas. In *Seminars in Neurology*, volume 38, pages 5–10, 2018.

[8] Fang Chen, Chenqiang Gao, Fangcen Liu, Yue Zhao, Yuxi Zhou, Deyu Meng, and Wangmeng Zuo. Local patch network with global attention for infrared small target detection. *IEEE Transactions on Aerospace and Electronic Systems*, 58(5):3979–3991, 2022.

[9] Francois Chollet. Xception: Deep Learning with Depthwise Separable Convolutions. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1800–1807. IEEE, 2017.

[10] Laura Daza, Catalina Gómez, and Pablo Arbeláez. Cerberus: A Multi-headed Network for Brain Tumor Segmentation. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 12659, pages 342–351. Springer International Publishing, 2021.

[11] David G. Ellis and Michele R. Aizenberg. Trialing U-Net Training Modifications for Segmenting Gliomas Using Open Source Deep Learning Framework. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 12659, pages 40–49. Springer International Publishing, 2021.

[12] J. Glory Precious, I. Keren Evangeline, and S. P. Angeline Kirubha. Brain tumour segmentation and survival prognostication using 3D radiomics features and machine learning algorithms. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 11(5):1803–1817, 2023.

[13] S. Rosas González, I. Zemmoura, and C. Tauber. 3D Brain Tumor Segmentation and Survival Prediction Using Ensembles of Convolutional Neural Networks. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 12659, pages 241–254. Springer International Publishing, 2021.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *arXiv:1512.03385 [cs]*, 2015.

[15] Renato Hermoza, Gabriel Maicas, Jacinto C. Nascimento, and Gustavo Carneiro. Post-Hoc Overall Survival Time Prediction From Brain MRI. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 1476–1480, 2021-04.

[16] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2):203–211, 2021.

[17] Fabian Isensee, Paul F. Jäger, Peter M. Full, Philipp Vollmuth, and Klaus H. Maier-Hein. nnU-Net for Brain Tumor Segmentation. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, pages 118–132. Springer International Publishing, 2021.

[18] Fabian Isensee, Paul F. Jäger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. Automated Design of Deep Learning Methods for Biomedical Image Segmentation. *Nature Methods*, 18(2):203–211, 2021.

[19] Ali Işın, Cem Direkoğlu, and Melike Şah. Review of MRI-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, 102:317–324, 2016.

[20] Haozhe Jia, Weidong Cai, Heng Huang, and Yong Xia. H$$2̂$$NF-Net for Brain Tumor Segmentation Using Multimodal MR Imaging: 2nd Place Solution to BraTS Challenge 2020 Segmentation Task. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, pages 58–68. Springer International Publishing, 2021.

[21] Gurinderjeet Kaur, Prashant Singh Rana, and Vinay Arora. Automated Neural Network-based Survival Prediction of Glioblastoma Patients Using Pre-operative MRI and Clinical Data. *IETE Journal of Research*, pages 1–17, 2023.

[22] Florian Kofler, Johannes C. Paetzold, Ivan Ezhov, Suprosanna Shit, Daniel Krahulec, Jan S. Kirschke, Claus Zimmer, Benedikt Wiestler, and Bjoern H. Menze. A Baseline for Predicting Glioblastoma Patient Survival Time with Classical Statistical Models and Primitive Features Ignoring Image Information. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 11992, pages 254–261. Springer International Publishing, 2020.

[23] Huan Minh Luu and Sung-Hong Park. Extending nn-UNet for Brain Tumor Segmentation. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, pages 173–186. Springer International Publishing, 2022.

[24] Bjoern H. Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, Levente Lanczi, Elizabeth Gerstner, Marc-André Weber, Tal Arbel, Brian B. Avants, Nicholas Ayache, Patricia Buendia, D. Louis Collins, Nicolas Cordier, Jason J. Corso, Antonio Criminisi, Tilak Das, Hervé Delingette, Çağatay Demiralp, Christopher R. Durst, Michel Dojat, Senan Doyle, Joana Festa, Florence Forbes, Ezequiel Geremia, Ben Glocker, Polina Golland, Xiaotao Guo, Andac Hamamci, Khan M. Iftekharuddin, Raj Jena, Nigel M. John, Ender Konukoglu, Danial Lashkari, José António Mariz, Raphael Meier, Sérgio Pereira, Doina Precup, Stephen J. Price, Tammy Riklin Raviv, Syed M. S. Reza, Michael Ryan, Duygu Sarikaya, Lawrence Schwartz, Hoo-Chang Shin, Jamie Shotton, Carlos A. Silva, Nuno Sousa, Nagesh K. Subbanna, Gabor Szekely, Thomas J. Taylor, Owen M. Thomas, Nicholas J. Tustison, Gozde Unal, Flor Vasseur, Max Wintermark, Dong Hye Ye, Liang Zhao, Binsheng Zhao, Darko Zikic, Marcel Prastawa, Mauricio Reyes, and Koen Van Leemput. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, 2015.

[25] Radu Miron, Ramona Albert, and Mihaela Breaban. A Two-Stage Atrous Convolution Neural Network for Brain Tumor Segmentation and Survival Prediction. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, volume 12659, pages 290–299. Springer International Publishing, 2021.

[26] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y. Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv:1804.03999*, 2018.

[27] Sergio Pereira, Adriano Pinto, Victor Alves, and Carlos A. Silva. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE transactions on medical imaging*, 35(5):1240–1251, 2016.

[28] Alasdair Philips, Denis L. Henshaw, Graham Lamburn, and Michael J. O'Carroll. Brain Tumours: Rise in Glioblastoma Multiforme Incidence in England 1995–2015 Suggests an Adverse Environmental or Lifestyle Factor. *Journal of Environmental and Public Health*, 2018:1–10, 2018.

[29] Rehan Raza, Usama Ijaz Bajwa, Yasar Mehmood, Muhammad Waqas Anwar, and M. Hassan Jamal. dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI. *Biomedical Signal Processing and Control*, 79:103861, 2023.

[30] Karen E. Steinhauser, Elizabeth C. Clipp, Maya McNeilly, Nicholas A. Christakis, Lauren M. McIntyre, and James A. Tulsky. In Search of a Good Death: Observations of Patients, Families, and Providers. *Annals of Internal Medicine*, 132(10):825, 2000.

[31] Erwin G. Van Meir, Costas G. Hadjipanayis, Andrew D. Norden, Hui-Kuo Shu, Patrick Y. Wen, and Jeffrey J. Olson. Exciting New Advances in Neuro-Oncology. *CA: a cancer journal for clinicians*, 60(3):166–193, 2010.

[32] Wenxuan Wang, Chen Chen, Meng Ding, Hong Yu, Sen Zha, and Jiangyun Li. TransBTS: Multimodal Brain Tumor Segmentation Using Transformer. In Marleen Bruijne, Philippe C. Cattin, Stéphane Cotin, Nicolas Padoy, Stefanie Speidel, Yefeng Zheng, and Caroline Essert, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Lecture Notes in Computer Science, pages 109–119. Springer International Publishing, 2021.

[33] Yixin Wang, Yao Zhang, Feng Hou, Yang Liu, Jiang Tian, Cheng Zhong, Yang Zhang, and Zhiqiang He. Modality-Pairing Learning for Brain Tumor Segmentation. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, pages 230–240. Springer International Publishing, 2021.

[34] Yading Yuan. Automatic Brain Tumor Segmentation with Scale Attention Network. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, pages 285–294. Springer International Publishing, 2021.

[35] Ke Zeng, Spyridon Bakas, Aristeidis Sotiras, Hamed Akbari, Martin Rozycki, Saima Rathore, Sarthak Pati, and Christos Davatzikos. Segmentation of Gliomas in Pre-operative and Post-operative Multimodal Magnetic Resonance Imaging Volumes Based on a Hybrid Generative-Discriminative Framework. In Alessandro Crimi, Bjoern Menze, Oskar Maier, Mauricio Reyes, Stefan Winzeck, and Heinz Handels, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 184–194. Springer International Publishing, 2016.

[36] Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3D U-net: Learning dense volumetric segmentation from sparse annotation. In Sebastien Ourselin, Leo Joskowicz, Mert R. Sabuncu, Gozde Unal, and William Wells, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pages 424–432. Springer International Publishing, 2016.