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.

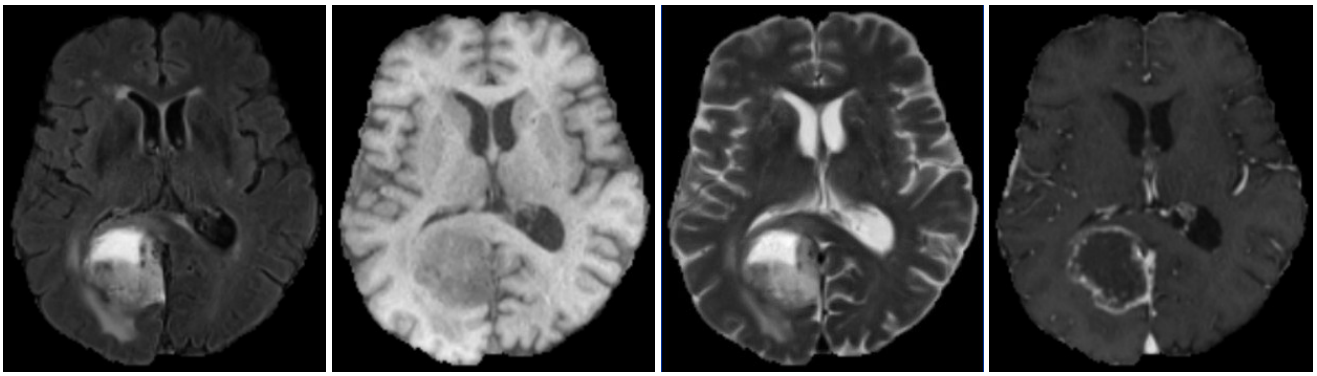
Magnetic Resonance Imaging (MRI) has emerged as the imaging technology of choice for brain tumor diagnosis, and treatment planning [2]. Non-invasive MRI scans produce high-resolution and soft tissue 3D volumes. As depicted in Fig. 1, more than one MRI slices are used to view different tumor regions.

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

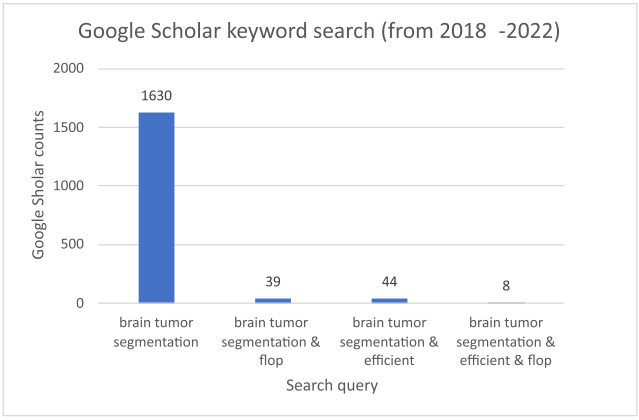


FIGURE 2. Results of Google Scholar searches for a period from 2018 to 2022 for articles with the following search queries: ***brain tumor segmentation*** [in title], ***brain tumor segmentation*** [in title] & ***flop*** [in all fields], ***brain tumor segmentation efficient*** [in title] and ***brain 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 $1 \leq y \leq h, 1 \leq x \leq w, 1 \leq z \leq d, 1 \leq j \leq n$. 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.

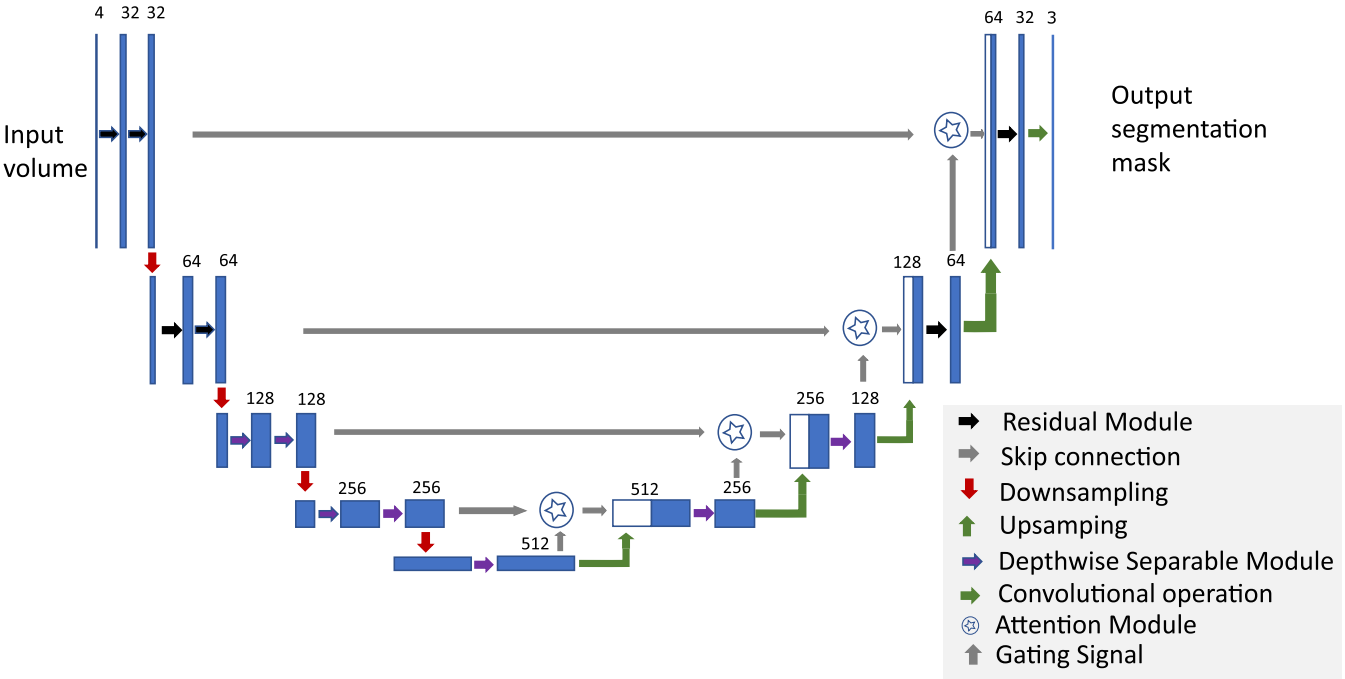


FIGURE 3. Schematic visualization of the network architecture. Input is a four-channel 3D MRI crop, followed by double residual modules with a base of 32 filters. The bottom three layers replaced all the convolutional blocks in residual modules with depthwise separable convolutions. Skip connections 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/

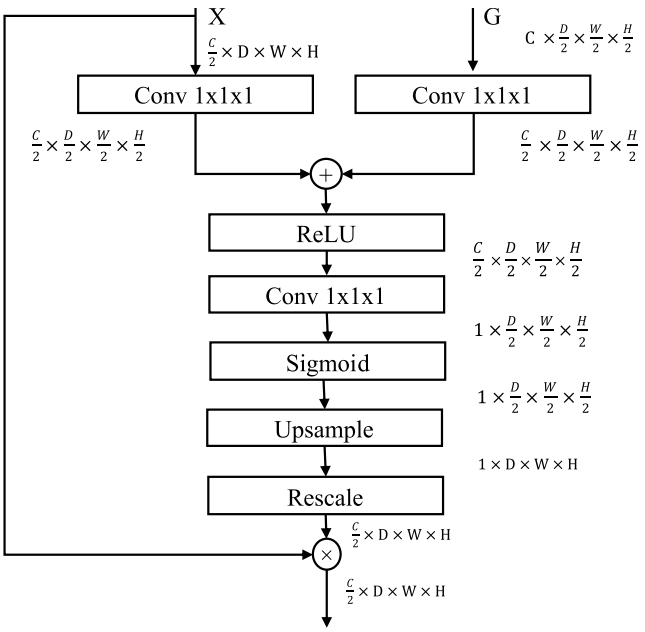


FIGURE 4. Illustration of the spatial attention module with two inputs *X* (skip connection) and *G* (gating signal from coarse scale). The module outputs weighted feature responses from the skip connections.

E. LOSS

We use the multi-class soft dice loss:

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

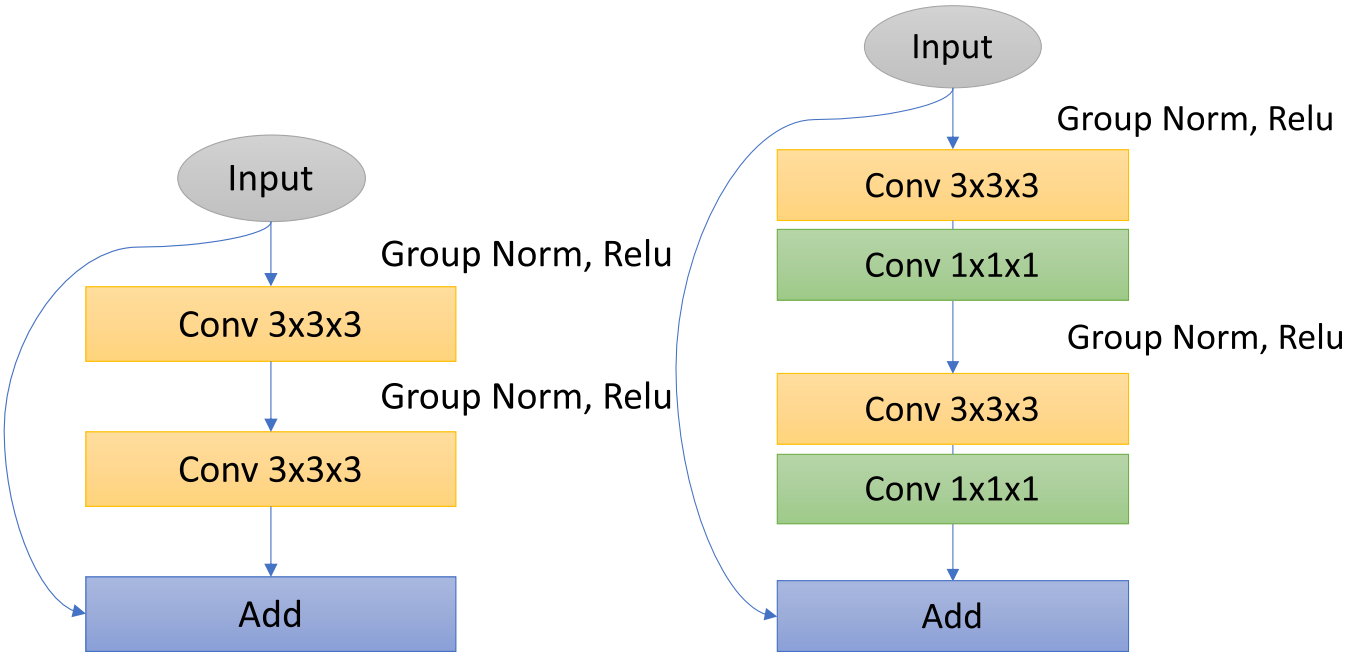


FIGURE 5. Comparison of different convolutional blocks incorporated in the network.

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/).

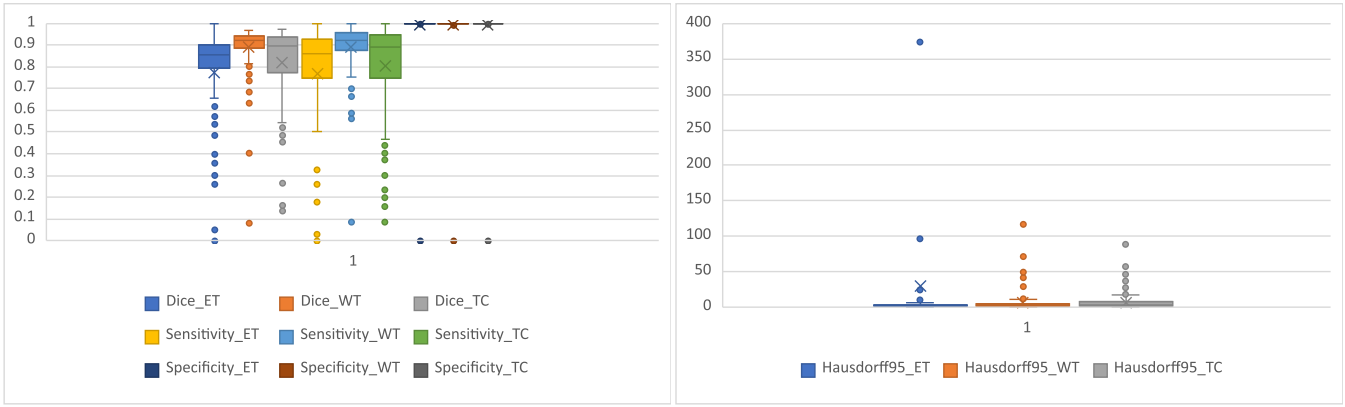
Table 3 shows the Dice Similarity Coefficient results on the BraTS 2020 validation set. The performance of the baseline in all regions was quite strong. Adding depthwise separable modules improved the dice scores marginally for the enhancing and whole tumor regions. We observed more gain when we trained 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

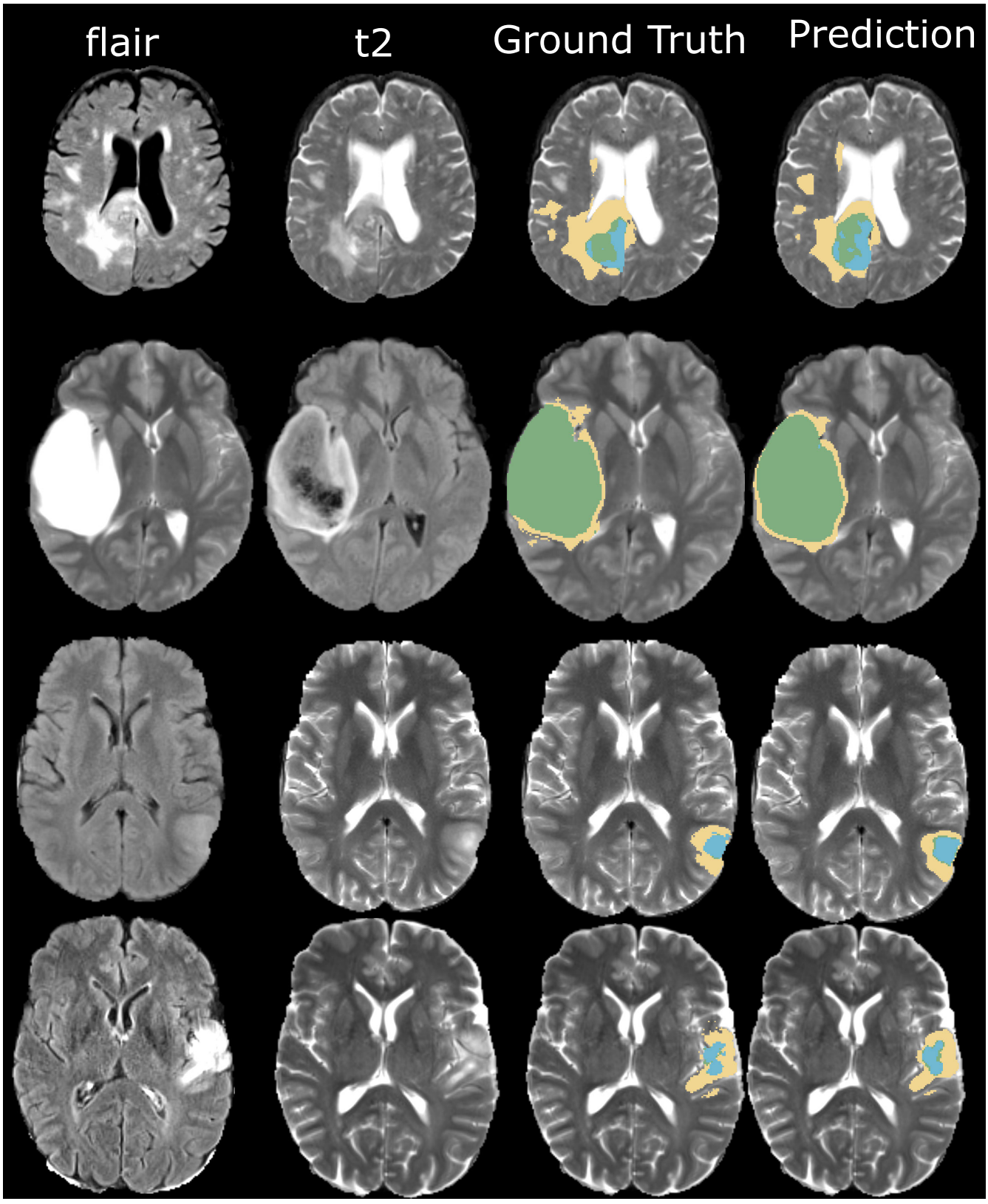


FIGURE 7. Qualitative inspections of two randomly selected predictions on the training set. Edema is shown in yellow, necrosis in green, and enhancing tumor in blue.

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$$Ĵ$$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.