

PU-NET DEEP LEARNING ARCHITECTURE FOR GLIOMAS BRAIN TUMOR SEGMENTATION IN MAGNETIC RESONANCE IMAGES

YAMINA AZZI^{✉,1}, ABDELOUAHAB MOUSSAOUI¹ AND MOHAND-TAHAR KECHADI²

¹Department of Computer Science, University of Ferhat Abbes Setif, 19000, Setif, Algeria, ²School of Computer Science, University College Dublin, Belfield, 4, Dublin, Ireland

e-mail: yamina.azzi@univ-setif.dz, abdelouahab.moussaoui@univ-setif.dz, tahar.kechadi@ucd.ie

(Received January 26, 2023; revised July 18, 2023; accepted July 27, 2023)

ABSTRACT

Automatic medical image segmentation is one of the main tasks for many organs and pathology structures delineation. It is also a crucial technique in the posterior clinical examination of brain tumors, like applying radiotherapy or tumor restrictions. Various image segmentation techniques have been proposed and applied to different image types. Recently, it has been shown that the deep learning approach accurately segments images, and its implementation is usually straightforward. In this paper, we proposed a novel approach, called PU-NET, for automatic brain tumor segmentation in multi-modal magnetic resonance images (MRI). We introduced an input processing block to a customized fully convolutional network derived from the U-Net network to handle the multi-modal inputs. We performed experiments over the Brain Tumor Segmentation (BRATS) dataset collected in 2018 and achieved Dice scores of 90.5%, 82.7%, and 80.3% for the whole tumor, tumor core, and enhancing tumor classes, respectively. This study provides promising results compared to the deep learning methods used in this context.

Keywords: Brain tumor, Deep learning, Gliomas, Image segmentation, U-Net.

INTRODUCTION

Gliomas are among the most life-threatening forms of cancer affecting the human brain or spinal cord. They are a common type of primary brain tumors, mainly growing in the glial cells, which adjoin neurons and provide nourishment and protection, unlike metastasis cancers, which grow anywhere in the body and spread to the brain.

Gliomas are typically classified into two categories: low-grade gliomas and high-grade gliomas. Low-grade gliomas are benign tumors characterized by a slow growth rate and regular morphological shapes in the outward appearance. This category can be treated surgically in the case of early diagnosis, and the survival rate may reach up to 10 years. However, it can be recurrent and grow again if regular examinations and safety precautions are not applied. They may spread and become life-threatening. Whereas high-grade gliomas are more aggressive and still considered untreatable because of their fast-growing, complete restriction is impossible because of their irregular contours in the morphology shape, which means tumor cells quickly affect random nearby healthy cells. Some treatment options are often available after surgeries, such as radiotherapy, chemotherapy, or combined approaches. These tumors have a high recurrent probability and a poor survival rate of 2 to

5 years (Bush and Chang, 2016; Menze *et al.*, 2015; Moini and Piran, 2020; Recht and Bernstein, 1995).

Early detection and accurate delineations are crucial for any cancer type, mainly brain tumors, to apply treatments correctly, reduce mortality rates, and ensure good patient quality of life. For this, different imaging modalities are used. These include Magnetic Resonance Imaging (MRI) or Computed tomography (CT) to detect these abnormal masses. MRI is currently the best non-invasive diagnostic tool for gliomas due to its potential in soft tissue imaging based on hydrogen atoms magnetization and its diversity in imaging modalities, which exhibit complementary information about tumor segments. It also provides safe, three-dimensional (3D), high-resolution, detailed anatomical structures without any radiation-related concerns (Brandal G, 2018; Sharma *et al.*, 2010).

In the past, specialists and radiologists performed glioma segmentation manually, which is tedious, time-consuming and greatly depends on human experiences and skills. These are reasons that motivate automatic segmentation. However, automatic segmentation is challenging. Because the anatomy of patients have many variations, the noise due to artefacts that may occur during the acquisition phase may affect image resolution and the low contrast between tissues, which leads to the absence of well-precise bounds between healthy brain tissues and abnormalities (İşin *et al.*, 2016; Withey and Koles, 2007).

Over the years, several automatic medical image segmentation techniques have been proposed and used. However, they are not accurate enough compared to manual segmentation. Therefore, developing more robust approaches to segment images is still a hot topic. Recently, deep learning with convolutional neural networks gained significant popularity as a segmentation technique for general image segmentation applications (Minaee *et al.*, 2021) and is highly used for brain glioma segmentation. The fully convolutional networks enable the localization of pathological structures with high accuracy. They outperform the traditional segmentation techniques.

In the last few years, brain tumor segmentation has received great interest (Ghaffari *et al.*, 2020), and various challenges have been documented under the BRATS title acronym for Brain Tumor Segmentation. The majority of contributions involved the utilization of deep learning techniques and achieved the top ranks. For example, in BRATS 2017 release, (Kamnitsas *et al.*, 2018) proposed an approach that explores multiple deep network architectures trained separately and Ensemble prediction maps to construct the final segmentation results; this achieved the first place on the challenge. Cascaded neural networks with anisotropic and dilated convolutions in the second place introduced by (Wang *et al.*, 2017), where the cascaded design is used to segment the whole tumor bounding box first, which was used to feed the next model to segment the tumor core, and finally, Enhancing the tumor is segmented with the same process.

In BRATS 2018, (Myronenko, 2019) was the first winner with its proposed 3D encoder-decoder architecture, accompanied by a variational auto-encoder part starting from the encoder endpoint. It aims to reconstruct the original input image to regularize the shared encoder. The second winner was (Isensee *et al.*, 2019), who proposed using 3D baseline U-Net with minor modification and focused on region-based prediction and a co-training process that uses additional data. (McKinley *et al.*, 2019) was the third winner with its preposition, which is a shallow network similar to the famous U-Net architecture that uses a densely connected block of dilated convolutions and introduces a new loss function based on binary cross-entropy loss function to calculate the label's uncertainty. (Wang *et al.*, 2019) propounded a work to segment brain glioma structures. It aims to use test time augmentation by augmenting images with 3D rotation, flipping, scaling, and adding noise to training and test images using different underpinning 3D network structures to perform multi-class segmentation and cascaded networks. (A.Albiol

et al., 2019) preferred to get inspired by the well-known two-dimensional (2D) deep learning models like VGG, inception2, inception3, and dense-like models and extend them to 3D versions. All these models were used to segment brain gliomas separately in addition to a final ensemble result. (Rui Hua *et al.*, 2019) developed a cascaded 3D V-Nets framework to handle the problem. The segmentation was performed in a cascaded way where the whole tumor was segmented first by three ensembled V-Nets. The detected tumor is segmented to the other tumor parts necrosis, edema, and enhancing tumor using two ensembled V-Nets. (Kermi *et al.*, 2019) proposed four channels of 2D U-Net architecture to segment gliomas and used residual blocks instead of plain blocks in the original U-Net. To address the class imbalance problem, Weighted Cross-Entropy (WCE) and generalized Dice loss were used. (Marcinkiewicz *et al.*, 2019) suggested segmenting brain gliomas in 2D two cascaded stages based on a convolutional neural network inspired by U-Net, the first stage applied to detect regions of interest and the second stage used to perform multi-class classification. In the next BRATS challenges 2019, the top-ranked approaches (Jiang *et al.*, 2020; McKinley *et al.*, 2020; Zhao *et al.*, 2020) were also based on deep learning architectures. Many additional proposed methods exist in (Crimi *et al.*, 2018; 2019) of different proposed 2D and 3D-based networks. Outside the BRATS challenge, a multitude of research studies has explored deep learning-based approaches (Akbar *et al.*, 2022; Noori *et al.*, 2019; Zhang *et al.*, 2020; Yogananda *et al.*, 2020), which have exhibited favourable achievements too. These works often employ a 2D and 3D U-Net architecture as a foundation and incorporate customized enhancements.

The significant limitations of these existing works are the increased computational complexity, high memory requirements, and long-running time due to using the 3D format and complex operations in deep learning architectures like dilated convolutions and residual blocks, which entail processing many parameters.

This paper presents a study of applying a simple and efficient deep learning approach to multi-modal MRI images, called PU-NET, to segment images of brain tumors. It is based on a two-dimensional (2D) convolutional neural network using an updated U-Net version and multi-view analysis and fusion. The results are promising compared with the baseline and existing technique. The rest of the paper is organized as follows. Section 2 describes the proposed PU-NET method. Experimental results, along with an evaluation study, are reported in Section 3 and Section

4, respectively. Finally, we conclude and provide future directions in Section 6.

METHODOLOGY

In this work, we propose a deep-learning architecture for brain tumor segmentation based on the encoder-decoder aspect and inspired by the reputed network used to segment biomedical images called U-Net (Ronneberger *et al.*, 2015) with significant changes. This architecture is called PU-NET, which refers to its two main parts: an input processing part and a customized U-Net network part. They are based on a central block known as a plain block. It encompasses two consecutive convolution layers of (3×3) kernel size with stride(1,1), each followed by a Batch normalization layer added to ensure the network data normalization and improve the training convergence and speed, followed by a LeakyRelu layer as an activation function with a slope α set to 0.01 Eq.1.

$$\text{LeakyRelu}(x) = \begin{cases} \alpha x, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases} \quad (1)$$

In contrast with the original proposition, all used convolutions are padded to preserve and mitigate the loss of information at the image borders. This loss is significant as the pooling operation applied later will induce information loss.

We also proposed using the LeakyRelu activation instead of the Relu function in the original proposition to prevent the dying Relu (Lu *et al.*, 2019; Mastromichalakis, 2020) problem that may happen where most of the Relu Eq. 2 neurons only output zero because of negative inputs during the learning process, which may cause the inactivity of a significant part of the network neurons and negatively affect the results.

$$\text{Relu}(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases} \quad (2)$$

The input processing part handles the multi-modal MRI scans quickly and efficiently. It consists of four input layers, each for a specific modality 2D slice (T1ce, T1, T2, Flair), each followed by a plain block to extract relevant features from each modality separately as a first step. Then, a concatenation layer is used to merge the four outputs and transfer the output to the U-Net part as the primary input.

The U-Net network part keeps the same divisions as the original proposal (Ronneberger *et al.*, 2015).

It comprises three parts: An encoder path(contracting path), a decoder path(an expansive path), and a simple bridge between them. The encoder path is similar to the convolution neural network (CNN) feed-forward pass. It contains three levels against four in the original U-Net to alleviate the network's training parameters and reduce memory consumption, complexity, and running time. Each level includes a plain convolution block, except for the first one, where the concatenation output replaces it, followed by a max-pooling layer with stride two and a dropout layer that regularizes the network to avoid over-fitting. This path is used for extracting relevant features and capturing the context of the input images to enable the segmentation task.

The expansive path is symmetric to the contracting path in the number of levels, each comprising a stack of layers. It starts with a padded convolution transpose of (3×3) kernel size with stride(2,2) used to recover some loosed information in previous convolution layers from the feature map. It is also followed by a concatenation layer that merges the output of the convolution plain block in the contracting side with the corresponding output of the convolution transpose to get more information about the spatial resolution after a convolution plain block is applied, followed by a dropout layer. The two paths are related by a bridge, which is a plain block.

Finally, a (1×1) convolution with sigmoid activation is applied to generate the output probabilities for the segmentation map. Figure 1 exposes the detailed PU-NET architecture layers.

EXPERIMENTAL RESULTS

DATASET PRE-PROCESSING

In this study, we propose to use the brain tumor segmentation (BRATS) dataset 2018 version (Bakas *et al.*, 2017; 2018; Menze *et al.*, 2015), which includes two main files named Training and Validation data. The training data contains MRI images of patients diagnosed with high-grade and low-grade gliomas (Glioblastoma). Each patient file encompasses four co-registered MRI modalities: native pre-contrast (T1), post-contrast T1-weighted (T1ce), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR) in addition to the manual segmentation file where all pixels were segmented into four classes with four rates (0, 1, 2, 4) summarized in Table 1. The data was acquired with different clinical protocols and scanners at 19 institutions. All BRATS multi-modal scans are available as Nifty files (.nii.gz), and they are

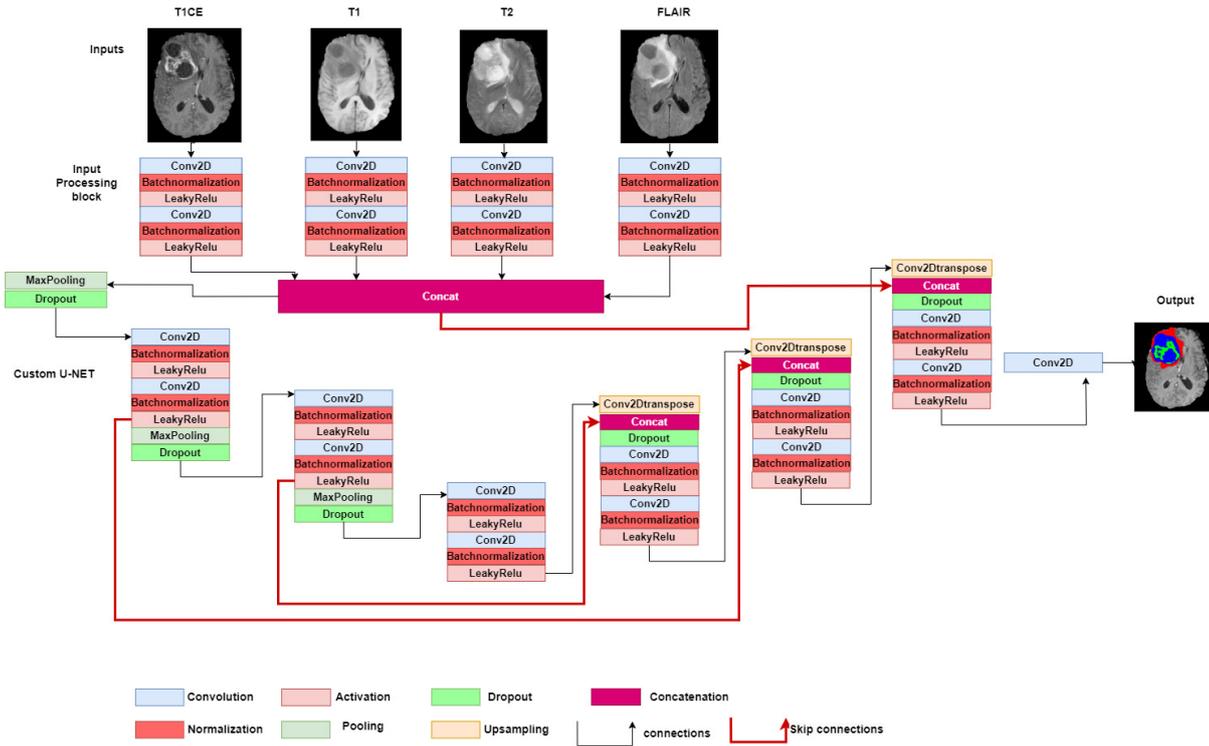


Fig. 1: The PU-NET model architecture used to segment brain tumor structures

pre-processed: co-registered to the same anatomical template interpolated to the exact resolution ($1mm^3$) and skull-stripped. Table 1 below summarizes the dataset information:

Table 1: The MICCAI_BRATS_2018_Data_Training information

MICCAI_BRATS_2018_Data_Training		
Gliomas grade	(HGG)	(LGG)
Number of patients	210	75
Modalities	T1, T1CE, T2, FLAIR	
Dimension	(240,240,155)	
Format	Nifty (Nii.gz)	
Ground truth	Available 0 =background and healthy tissue	
Labels	1= Necrotic /non-enhancing (NCR/NET). 2= Peritumoral edema (ED) 4 =Enhancing tumor (ET).	

Normalizing the data is an essential step in all machine learning tasks to avoid dominating some features by others and gain the correct information from all relevant features. In our case, we propose to use z-score normalization (Pal and Sudeep, 2016; Reinhold *et al.*, 2019) to get a normal distribution of all voxels in each image modality by computing the mean

and standard deviation only for the brain region (non-zero part). (Weninger *et al.*, 2019) and update each voxel value by a new value computed by Equation 3.

$$Z = \frac{v - \mu}{\sigma} \quad (3)$$

Where v is the voxel value and μ, σ are the voxels' mean and standard deviation, respectively.

IMPLEMENTATION DETAILS

Our deep learning proposed approach was implemented using python3 over the Tensorflow Keras library and executed on the UB2-HPC (University of BATNA 2) GPU node, which contains 4 GPUs configured with CUDA 10.0 and CUDNN 5.6.7. In our case, the official BRATS training data (MICCAI_BRATS_2018_Data_Training) was randomly split into Train set 80% and Valid set 20% with 42 as a random state to build the model, which is later tested over the official validation set (MICCAI_BRATS_2018_Validation_Data), provided by the organization of 66 patients without Ground truth, which is used in our case as a test set.

Table 2: The PU-NET Model used datasets

Datasets	# of patients	Data source
Train set	228	BRATS_Training_Data (80%)
Valid set	57	BRATS_Training_Data (20%)
Test set	66	BRATS_Validation_Data

To obtain results that are more significant in a very efficient way, we proposed resolving the problem as a binary segmentation problem rather than multi-class segmentation, where we turned the original label classes to the proposed classes in the challenge, we trained the model for each class separately, and the results were fused. The new classes are exhibited in Table 3.

Table 3: *The segmentation Classes*

New Class	Labels
Whole Tumor (WT)	(NCR/NET 1 +edema 2 +Enhancing Tumor 4)
Tumor Core (TC)	(NCR/NET 1 +Enhancing Tumor 4)
Enhancing Tumor (ET)	Enhancing Tumor 4 label

We also propose to use this training process over the three MRI views (Axial, Coronal, and Sagittal) separately and ensemble the results over models to ensure the collection of the 3D contextual information. All image modalities and their corresponding ground truth were cropped from (240,240,155) size to (128,128,128) with the following indices [56:184, 56:184, 12:140] to reduce memory consumption. Table 4 shows our network parameters:

Table 4: *PU-NET network parameters*

Parameters	Values
Views	Axial/Coronal/Sagittal
Training Samples	29184
Validation Samples	7296
Initial filters N°	16
Batch size	32
Epochs	200
Dropout rate	0.4
Early stopping	Active (patience epochs=10)
Model checkpoint	Active
Optimizer	Adam (default)

Where: Initial filters N° represents the number of convolution kernels, which doubles for each encoder plain block and halved for each decoder plain block and transposed layer. Batch size is the number of training samples used in one iteration. Epochs are the number of times the algorithm trains over all samples. The dropout rate is the proportion of randomly selected nodes set to zero with early stopping used to stop the learning process if the validation dice score didn't improve after ten patience epochs and the Model checkpoint active to save the best model with the best validation dice score, all used to prevent over-fitting. Finally, the Adam optimizer with default parameters was used to fine-tune the network weights.

EVALUATION METRICS

All evaluation metrics used are those proposed by the challenge organization (Menze *et al.*, 2015). The evaluation metrics used are Dice Score, Sensitivity, Specificity, and Hausdorff 95 distance; those metrics are highly recommended for medical image segmentation (Taha and Hanbury, 2015; Huttenlocher *et al.*, 1993). Most of the time, the healthy tissue pixels are larger than the tumor's ones, which means that the data suffer from the class imbalance distribution problem, which prevents the right learning and leads the model to learn only about the frequent class label where the image segmentation problem focuses on the infrequent class (Small and Ventura, 2017). All these metrics are well-adapted to address the class imbalance problem in such cases. They only consider the segmentation class and not the background class (Jadon, 2020),

EXPERIMENTAL RESULTS

Table 5 exhibits the results of our PU-NET approach on the training, validation, and test data respectively (Table 2) with the proposed evaluation metrics in section 3.3. The results are also reported over the MRI views separately in addition to the validation and Test ensemble (Ens) results realized by multi-view label fusion using majority voting, proving the efficiency of the multi-view exploration and fusion in improving the result. The ensembled 5-fold validation result is also reported. It is important to note that all Test results were collected from the online submission system CBICA server <https://ipp.cbica.upenn.edu/> which is an Image Processing Portal available for authorized users to access the Center for Biomedical Image Computing and Analytics computing cluster and imaging analytics.

The following Fig. 2 exhibits some examples from the test set Table 2 segmented by our PU-NET. Colours in the figure refer to labels in table 1: red: label 2, Green: label 1, blue: label 4.

PU-NET AGAINST BASELINE TECHNIQUES

In this section, our primary goal is to evaluate the effectiveness of our 2D PU-NET model by specifically examining the impact of the input processing block. We compare baseline techniques that propose different approaches for processing multi-modal MRI data.

The first baseline technique involves stacking each corresponding slice from each modality (T1ce, T1, T2, and Flair) as RGBA color images. These stacked images are provided directly to the U-Net part of our

Table 5: PU-NET training, validation, and Test Results.

Metrics	Classes	Dice			HD95(mm)			Sensitivity			Specificity		
		WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET
Training	Axial	0.927	0.934	0.877	3.524	2.388	1.913	0.922	0.935	0.885	0.953	0.979	0.991
	Coronal	0.924	0.918	0.864	3.579	2.448	1.954	0.933	0.921	0.895	0.951	0.979	0.990
	Sagittal	0.937	0.910	0.880	3.636	2.564	1.958	0.933	0.923	0.893	0.952	0.979	0.991
Validation	Axial	0.917	0.806	0.859	3.388	2.323	1.756	0.908	0.792	0.836	0.950	0.981	0.993
	Coronal	0.918	0.835	0.861	3.423	2.339	1.792	0.921	0.830	0.861	0.949	0.981	0.992
	Sagittal	0.916	0.812	0.857	3.496	2.433	1.800	0.910	0.805	0.837	0.950	0.981	0.992
	Ensemble	0.925	0.839	0.868	3.546	2.360	1.771	0.920	0.820	0.849	0.950	0.981	0.992
Test	Axial	0.890	0.791	0.781	15.912	12.288	6.231	0.907	0.807	0.810	0.993	0.997	0.997
	Coronal	0.889	0.787	0.782	12.134	18.774	11.002	0.923	0.811	0.821	0.991	0.996	0.997
	Sagittal	0.894	0.792	0.776	11.174	13.732	3.486	0.917	0.829	0.803	0.992	0.996	0.997
	ensemble	0.904	0.815	0.803	5.054	7.732	3.019	0.924	0.817	0.820	0.993	0.997	0.997
	5-fold	0.905	0.827	0.803	4.638	8.679	3.573	0.921	0.817	0.815	0.993	0.997	0.998

proposed model as input without utilizing the input processing block.

The second baseline technique involves providing the multi-modal MRI data in the 3D format to the PU-NET architecture. This means the data is preserved in its volumetric structure instead of processed as 2D slices. Additionally, the architecture operations are modified to accommodate the 3D format.

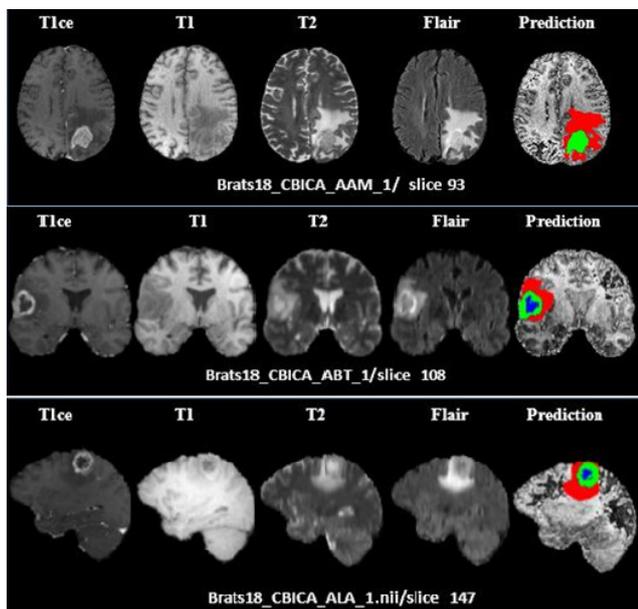


Fig. 2: Segmentation result of the brain tumor structures by the Proposed PU-NET on three different patients.

The results obtained in Table 6 provide compelling evidence that our PU-NET model performs exceptionally well when compared to stacking slices as RGBA color images. These results further validate the efficiency and effectiveness of our input processing block.

However, it is worth noting that the 3D version of our model exhibits some drawbacks. Specifically, it requires a significantly longer training time, exceeding 20 hours under GPU for each class label. In contrast, our 2D PU-NET model takes a maximum of 1 hour for execution. Furthermore, the results obtained from the 3D version diverge significantly from those of our 2D PU-NET model, particularly in the TC and ET classes.

PU-NET AGAINST EXISTING METHODS

In this section, table 7 summarizes the results of the state-of-the-art techniques that refer to the currently existing approach or method that has achieved high performance in the field to compare them against our PU-NET results to analyze the strengths and weaknesses of both methods and determine if our approach outperforms state-of-the-art or provides any notable improvements.

DISCUSSION

From Table 7, our proposed method PU-NET outperforms the 2D approaches in terms of both Dice and Hausdorff95 metrics. While it is not away from the top-ranked approaches in the BRATS 2018 challenge and this is due to the fine parameters tuning; it is important to consider, too, that (Myronenko, 2019), (Isensee *et al.*, 2019), and (McKinley *et al.*, 2019) utilized data augmentation techniques and additional data in the co-training process. In contrast, our work did not incorporate any additional or augmented data. Furthermore, when computing the statistical p-values between the dice scores, we found values of 0.61, 0.72 and 0.66, respectively, suggesting that the results are not highly significant compared to our findings.

Regarding (Wang *et al.*, 2019), our 2D approach outperformed all of their proposed 3D architectures with data augmentation, except for

Table 6: PU-NET versus baseline technique

Metrics	Dice			HD95(mm)			Sensitivity			Specificity		
	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET
Classes												
Axial	0.891	0.786	0.761	8.210	14.113	5.584	0.909	0.800	0.798	0.992	0.996	0.998
Coronal	0.892	0.793	0.761	13.670	18.720	12.491	0.918	0.795	0.816	0.992	0.997	0.997
Sagittal	0.890	0.789	0.724	11.597	18.722	10.457	0.918	0.800	0.815	0.992	0.996	0.997
RGBA(Ens)	0.903	0.801	0.772	5.442	6.996	3.967	0.921	0.766	0.781	0.993	0.998	0.998
3D PU-NET	0.884	0.710	0.734	7.898	28.930	29.695	0.911	0.780	0.750	0.980	0.980	0.982
PU-NET	0.905	0.827	0.803	4.638	8.679	3.573	0.921	0.817	0.815	0.993	0.997	0.998

Table 7: PU-NET and state-of-the-art techniques comparison results

Method	type	DICE			HD95(mm)		
		WT	TC	ET	WT	TC	ET
(Myronenko, 2019)	3D	0.910	0.866	0.823	4.51	6.85	3.92
(Isensee <i>et al.</i> , 2019)	3D	0.912	0.863	0.808	4.27	6.52	2.41
(McKinley <i>et al.</i> , 2019)	3D	0.903	0.847	0.796	3.55	4.17	4.93
	3D U-Net+ TTA	0.873	0.783	0.754	5.90	8.03	4.53
(Wang <i>et al.</i> , 2019)	3D WNet+TTA	0.895	0.730	0.770	4.92	11.13	4.44
	3D Cas+TTA	0.902	0.858	0.797	6.18	6.37	3.13
	3D VGG	0.872	0.760	0.751	-	-	-
	3D inception2	0.877	0.773	0.753	-	-	-
(A.Albiol <i>et al.</i> , 2019)	3D inception3	0.873	0.776	0.781	-	-	-
	3D Densely	0.874	0.755	0.729	-	-	-
	Ens	0.881	0.777	0.773	-	-	-
(Rui Hua <i>et al.</i> , 2019)	3D	0.904	0.836	0.776	5.17	6.27	3.51
(Akbar <i>et al.</i> , 2022)	3D	0.895	0.797	0.777	9.13	8.67	3.90
(Yogananda <i>et al.</i> , 2020)	3D	0.900	0.820	0.800	6.0	7.5	4.4
(Kermi <i>et al.</i> , 2019)	2D	0.868	0.805	0.783	8.12	9.84	3.72
(Marcinkiewicz <i>et al.</i> , 2019)	2D	0.898	0.811	0.751	-	-	-
(Noori <i>et al.</i> , 2019)	2D	0.895	0.823	0.813	4.05	6.34	2.93
(Zhang <i>et al.</i> , 2020)	2D	0.872	0.808	0.772	5.62	8.36	3.57
PU-NET	2D	0.905	0.827	0.803	4.63	8.67	3.57

the cascaded architecture, which exhibited slightly better performance than our PU-NET in TC class segmentation, similar to the findings of (Rui Hua *et al.*, 2019). However, it is important to note that the p-values of 0.87 and 0.90 suggest these differences are not statistically significant. Moreover, our PU-NET approach demonstrated greater efficiency than the extended architectures proposed by (A.Albiol *et al.*, 2019), which are among the popular deep learning architectures used in the context of image segmentation, both (Akbar *et al.*, 2022; Yogananda *et al.*, 2020) works which are out of the BRATS 2018 challenge proposed 3D-based architectures with complex U-Net designs. Still, our results surpass them and support our the idea that sometimes complex architectures and using 3D format can increase the computational complexity without necessarily being efficient, so it is justifiable to shift the research focus towards 2D simple architectures.

In our perspective, the strength of our PU-NET

model lies in incorporating the input processing block as an initial separate feature extractor for each MRI modality. It serves as a robust mechanism for integrating the multi-modal complementary information and improving the tumor segmentation accuracy. In addition, the simplicity of our U-Net architecture part and the use of 2D slices address major limitations in deep learning-based approaches, specifically regarding time and memory requirements, which are advantageous when working with large-size medical datasets. As with any approach, there are certain limitations associated with our proposed method; the input processing block in our approach is specifically designed to handle datasets with multi-modal images. While this benefits such datasets, it may not be suitable or optimal for datasets with single-modality images. Similar to other deep learning architectures, the interpretability of our approach is limited. Deep learning models, including our PU-NET, often function as complex black boxes, challenging

understanding of the underlying decision-making process. Interpretability is an ongoing area of research in deep learning, and further efforts are needed to enhance the transparency and explainability of models like ours. Indeed, while our approach may have limitations, it is important to recognize that its performance remains competitive. The results obtained with our 2D PU-NET architecture demonstrate promise and suggest potential avenues for further improvement in future research.

CONCLUSION

This study investigated a novel approach inspired by the reputed U-Net architecture to tackle the brain tumor segmentation problem from multi-modal MRI images called PU-NET. We introduced an input processing block to an updated U-Net that deals with multi-input 2D images collected from the different modalities. We suggested exploring this task in the three MRI views coronal, sagittal, and axial and aggregating the final predictions to generate the final segmentation and benefit from the 3D contextual information. The results were later compared against the RGBA and the 3D versions and verified against the existing works where our method achieved interesting results, and using data augmentation and additional training data seemed to be the next verified steps with our method for more results enhancement.

REFERENCES

- A. Albiol A, Antonio A, Francisco A (2019). Extending 2D Deep Learning Architectures to 3D Image Segmentation Problems. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc, vol. 1. Springer International Publishing.
- Akbar AS, Fatichah C, Suciati N (2022). Single level unet3d with multipath residual attention block for brain tumor segmentation. Journal of King Saud University Computer and Information Sciences 34:3247–58.
- Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, Freymann JB, Farahani K, DC (2017). Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. Scientific Data 4:170117.
- Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, Shinohara RT, Berger C, Ha SM, Rozycki M, et al. (2018). Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. arXiv .
- Brandal G PW (2018). Conventional and advanced magnetic resonance imaging in patients with high-grade glioma. Q J Nucl Med Mol Imaging 62:239.
- Bush NAO, Chang S (2016). Treatment strategies for low-grade glioma in adults. Journal of Oncology Practice 12:1235–41.
- Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, eds. (2019). Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, vol. 11384 of *Lect Notes Comput Sc*. Cham: Springer International Publishing.
- Crimi A, Bakas S, Kuijf H, Menze B, Reyes M, eds. (2018). Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, vol. 10670 of *Lect Notes Comput Sc*. Cham: Springer International Publishing.
- Ghaffari M, Sowmya A, Oliver R (2020). Automated brain tumor segmentation using multimodal brain scans: A survey based on models submitted to the brats 2012–2018 challenges. IEEE Rev Biomed Eng 13:156–68.
- Huttenlocher D, Klanderman G, Rucklidge W (1993). Comparing images using the Hausdorff distance. IEEE T Pattern Anal 15:850–63.
- Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein KH (2019). No new-net. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc, vol. 11384.
- Işin A, Direkoğlu C, Şah M (2016). Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. Procedia Comput Sci 102:317–24.
- Jadon S (2020). A survey of loss functions for semantic segmentation. 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology CIBCB .
- Jiang Z, Ding C, Liu M, Tao D (2020). Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lect Notes Comput Sc, vol. 11992. Springer International Publishing.
- Kamnitsas K, Bai W, Ferrante E, McDonagh S, Sinclair M, Pawlowski N, Rajchl M, Lee M, Kainz B, Rueckert D, Glocker B (2018). Ensembles of Multiple Models and Architectures for Robust Brain Tumour Segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic

- Brain Injuries. BrainLes 2017, Lect Notes Comput Sc. Springer International Publishing.
- Kermi A, Mahmoudi I, Khadir MT (2019). 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. BrainLes 2018. Lect Notes Comput Sc, vol. 2. Springer International Publishing.
- Lu L, Shin Y, Su Y, Karniadakis GE (2019). Dying relu and initialization: Theory and numerical examples. ArXiv abs/1903.06733.
- Marcinkiewicz M, Nalepa J, Lorenzo PR, Dudzik W, Mrukwa G (2019). Segmenting Brain Tumors from MRI Using Cascaded Multi-modal U-Nets. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc, vol. 2. Springer, Cham.
- Mastromichalakis S (2020). Alrelu: A different approach on leaky relu activation function to improve neural networks performance. ArXiv abs/2012.07564.
- McKinley R, Meier R, Wiest R (2019). Ensembles of Densely-Connected CNNs with Label-Uncertainty for Brain Tumor Segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018, Lect Notes Comput Sc, vol. 2. Springer, Cham.
- McKinley R, Rebsamen M, Meier R, Wiest R (2020). Triplanar ensemble of 3d-to-2d cnns with label-uncertainty for brain tumor segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019, Lect Notes Comput Sc, vol. 11992. Springer International Publishing, 379–87.
- Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE T Med Imaging 34:1993–2024.
- Minaee S, Boykov YY, Porikli F, Plaza AJ, Kehtarnavaz N, Terzopoulos D (2021). Image Segmentation Using Deep Learning: A Survey. IEEE T Med Imaging 8828:1–.
- Moini J, Piran P (2020). Chapter 1 - histophysiology. In: Functional and Clinical Neuroanatomy. Academic Press, 1–49.
- Myronenko A (2019). 3D MRI Brain Tumor Segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc. Springer, Cham.
- Noori M, Bahri A, Mohammadi K (2019). Attention-guided version of 2d unet for automatic brain tumor segmentation. In: 2019 9th international conference on computer and knowledge engineering (ICCKE). IEEE.
- Pal KK, Sudeep KS (2016). Preprocessing for image classification by convolutional neural networks. In: 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). IEEE.
- Recht LD, Bernstein M (1995). Low-Grade Gliomas. Neurologic Clinics 13:847–59.
- Reinhold JC, Dewey BE, Carass A, Prince JL (2019). Evaluating the impact of intensity normalization on mr image synthesis. Proc SPIE Int Soc Opt Eng 10949.
- Ronneberger O, Fischer P, Brox T (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. IEEE Access 9:16591–603.
- Rui Hua, Huo Q, Gao Y, Sun Y, Shi F (2019). Multimodal Brain Tumor Segmentation Using Cascaded V-Nets. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc, vol. 11384. Springer, Cham.
- Sharma N, Ray A, Shukla K, Sharma S, Pradhan S, Srivastva A, Aggarwal L (2010). Automated medical image segmentation techniques. J Med Phys 35:3.
- Small H, Ventura J (2017). Handling Unbalanced Data in Deep Image Segmentation.
- Taha AA, Hanbury A (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med Imaging 15:29.
- Wang G, Li W, Ourselin S, Vercauteren T (2017). Automatic Brain Tumor Segmentation using Cascaded Anisotropic Convolutional Neural Networks. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2017. Lect Notes Comput Sc.
- Wang G, Li W, Ourselin S, Vercauteren T (2019). Automatic brain tumor segmentation using convolutional neural networks with test-time augmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc, vol. 11384.
- Weninger L, Rippel O, Koppers S, Merhof D (2019). Segmentation of Brain Tumors and Patient Survival Prediction: Methods for the BraTS 2018 Challenge. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lect Notes Comput Sc, vol. 11384. Cham: Springer International Publishing.

- Withey D, Koles Z (2007). Medical Image Segmentation: Methods and Software. In: Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, vol. 112. IEEE.
- Yogananda CGB, Shah BR, Vejdani-Jahromi M, Nalawade SS, Murugesan GK, Yu FF, Pinho MC, Wagner BC, Emblem KE, Bjørnerud, Fei B MA, JA M (2020). A fully automated deep learning network for brain tumor segmentation. *Tomography* 6:186–93.
- Zhang J, Jiang Z, Dong J, Hou Y, Liu B (2020). Attention gate resu-net for automatic mri brain tumor segmentation. *IEEE Access* 8:58533–45.
- Zhao YX, Zhang YM, Liu CL (2020). Bag of Tricks for 3D MRI Brain Tumor Segmentation. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019, Lect Notes Comput Sc*, vol. 11992. Springer International Publishing, 210–20.