







Optimized Brain Tumor Segmentation for Resource Constrained Settings: VGG-Infused U-Net Approach

Mizanu Zelalem Degu¹ , Confidence Raymond^{2,3} , Dong Zhang⁴ ,
Amal Saleh⁵, Uduinna C. Anazodo^{2,3,6,7} , and Gizeaddis Lamesgin Simegn⁸ 

¹ Faculty of Computing and Informatics, Jimma University, Jimma, Ethiopia
mizanu143@gmail.com

² Medical Artificial Intelligence Laboratory (MAI Lab), Lagos, Nigeria

³ Lawson Health Research Institute, London, ON, Canada

⁴ Department of Electrical and Computer Engineering, University of British Columbia,
Vancouver, Canada

⁵ School of Medicine, College of Health Sciences, Addis Ababa University, Addis Ababa,
Ethiopia

⁶ Montreal Neurological Institute, McGill University, Montréal, Canada

⁷ Department of Clinical and Radiation Oncology, University of Cape Town, Cape Town,
South Africa

⁸ School of Biomedical Engineering, Jimma University, Jimma, Ethiopia

Abstract. Semantic segmentation of brain tumors plays a crucial role in assisting treatment by precisely delineating tumor boundaries in brain images. This aids clinicians in formulating surgical plans and targeted therapies, ultimately enhancing patient outcomes. Automatic brain tumor segmentation poses a significant challenge due to its complexity and resource-intensive nature, particularly in low-resource settings where access to high quality data and expertise is limited, hindering effective medical image analysis and treatment planning. In this study, the complex multi-label 3D segmentation problem was simplified into more manageable 2D single-label segmentation tasks and a U-Net architecture was adapted, with the integration of a pre-trained VGG19 model to reduce computational demands while effectively extracting features. Six models were trained in total, three using the conventional U-Net and three employing the VGG-infused U-Net (VIU-Net). The Experiment was performed using a total of 1424 magnetic resonance images. The study shows that the VIU-Net model achieved lower dice losses of 0.188, 0.068, and 0.106 for the segmentation of Necrotic Core, SNFH, and Enhancing Tumor regions respectively, showcasing the effectiveness of our method and its suitability for implementation in low resource settings.

Keywords: Automatic Segmentation · Brain Tumor · Optimized U-Net · Low Resource Setting

1 Introduction

A brain tumor is an anomalous mass of tissue characterized by the uncontrolled and unrestrained proliferation of cells. In this condition, cellular growth and multiplication occur without apparent restraint from the regulatory mechanisms that typically control normal cells. This unregulated growth can lead to the formation of a mass within the brain, potentially impacting its normal functions and necessitating specialized medical attention and intervention [1].

Brain tumors, particularly gliomas, are among the deadliest cancers, with 80% of individuals diagnosed with glioma dying within two years. In contrast, over 90% of those diagnosed with prostate or breast cancer survive for at least five years. In Sub-Saharan Africa, glioma death rates have risen by 25%, sharply contrasting with the 30% declining rates in the Global North [2]. While earlier studies indicated a relatively lower number of cases in developing countries, recent investigations focusing on low- and middle-income countries, particularly in sub-Saharan Africa (SSA), reveal the substantial prevalence of brain tumors in the region [3, 4]. According to recent studies [2] the death rate from glioma in SSA are among the highest in the world and continuing to rise. Delayed presentation, incident of other infectious disease such as HIV, shortage of health facilities, low level of income is among the contributors [5]. In response to this intricate pathology, an accurate diagnosis technique is essential for successful treatment planning, and magnetic resonance imaging (MRI) has become the principal imaging modality for diagnosing brain tumors and their extent.

Brain Tumor segmentation from MR images is essential in standard of care, enabling the clinicians to identify tumor location, extent, and subtypes. This not only helps with initial diagnosis but also aids with administering and monitoring treatment progress. Given the critical nature of this task, the precise demarcation of the tumor and its sub-regions is traditionally carried out manually by experienced neuro-radiologists. This manual process is inherently challenging, not only due to its tedious nature but also because it demands the expertise of highly experienced professionals. The complexity arises from multiparametric image contrast, the presence of heterogeneous tumors, and the reliance on labeling that is subject to both inter and intra-rater variability [6–8].

Developing computer-aided decision support systems, particularly those involving automatic semantic segmentation and classification tailored for low resource settings such as Sub-Saharan Africa countries (SSA), is challenging due to the scarcity of human resources, infrastructure, insufficient funding for cancer research and an insufficient number of trained personnel (neuroradiologists, oncologists, and physicists) [3, 9, 10]. To this end, the aim of this study is to develop a brain tumor semantic segmentation model with minimal resource by breaking down the demanding 3D multilabel segmentation problem into more manageable 2D tasks, where tumor features of the image can be effectively extracted. This will enable accessible brain tumor segmentation approach suitable for resource-constrained environments.

2 Related Works

Before the availability of sophisticated computational devices and large dataset, attempts on automatic brain tumor segmentation were limited to handcrafted feature extraction and shallow machine learning algorithms. Since its establishment, the brain tumor segmentation (BraTS) [11] Challenge has served as a significant motivator for the development of state-of-art algorithms for tumor segmentation. The challenge has been providing a community benchmark dataset and environment for adult glioma over the past 11 years [12]. Almost all the proposed models for the challenge used deep learning models as a backbone [13–16]. In recent challenges, a specific deep learning model, U-Net and its modified versions, are being used in addressing the brain tumor segmentation problem. U-Net is a deep learning model based on convolutional neural network specifically designed to better address segmentation problems in biomedical computer vision [16].

In recent studies, an extension or modification of the U-Net architecture has been applied, involving the integration of a 3D U-Net with an additional variational decoder branch. This enhancement is designed to provide extra supervision and regularization to the encoder branch [15]. A two-stage cascaded U-Net was also proposed for segmentation [13]: the first stage was trained to produce coarse segmentation mask and the second stage was trained to refine the output of the first stage. The nnU-Net [17], a self-configuring framework that automatically adapt U-Net to diverse datasets was proposed to streamline the process of medical image segmentation. The winner of BraTS-2021 [6] suggested utilization of large nnU-Net model with group normalization, and axial attention in the expanding path of the network.

Although previous methods on brain tumor segmentation produced promising performance, these models were trained using high quality images captured in resource-rich settings and higher resolution multiparametric (mpMRI) techniques, resulting in unreliable and decreased performance in low-resource settings [18]. In the context low-resource settings, medical image acquisition presents distinct challenges characterized by limited expertise, unoptimized imaging protocols, and the shortage of imaging infrastructure, leading to low-quality images with limited contrast and resolution. Moreover, fine-tuning large models equipped with multiple automated augmentation modules is computationally expensive and nearly impractical in these settings.

3 Methods

3.1 Dataset

The dataset used for training the models was acquired from the BraTS-Africa Challenge [2]. A total of 60 adult glioma cases collected retrospectively from clinical studies in SSA were used in this study. Details of the data collection and image pre-processing are described in the BraTS-Africa Challenge [5]. Briefly, each case consisted of four pre-processed three-dimensional (3D) mpMRI volumes (T1-weighted (T1w), post gadolinium contrast T1 weighted (T1ce), T2-weighted (T2w) and T2-FLAIR (FLAIR)) with one multi-labelled tumor binary mask, comprising of three sub-regions (enhancing tumor, surrounding non-enhancing FLAIR hyperintensity (SNFH), and non-enhancing tumor core) [5].

3.2 Dataset Processing

All data were further processed before model training. Firstly, all image and mask pairs were cropped to remove redundant background voxels limiting noise and computational cost of training the model. To clip voxel values and improve ease of network convergence, we performed the min-max normalization. Then, generated two-dimensional (2D) image slices from the 3D isotropic (1mm^3) image volumes by taking one slice at a time iteratively along the sagittal view where the quality of the image is better relative to trans axial or coronal views. Each iteration of the slicing produced 9,856 number of $128 \times 128 \times 1$ images for each T1CE, T2w, and FLAIR volumes. The T1w volume was not resliced to 2D since the T1CE with the same dimensions, was already included. Individual slices from the three volumes were then added together to produce a 2D merged image. During the slicing process, we intentionally excluded certain 2D frames to ensure a balanced representation of segmented and non-segmented regions. Specifically, we maintained a ratio of at least 5% segmented regions to non-segmented regions. Finally, a total of 1,424 2D images were obtained from slicing the 60 3D volumetric datasets and used in training and evaluating the model. The generation of 2D images from the 3D label data (Mask image) followed a similar process. The multilabel mask images were then split into three single-class masks, facilitating the training of a single-class segmentation model. Figure 1 illustrates the 3D to 2D data decomposition workflow.

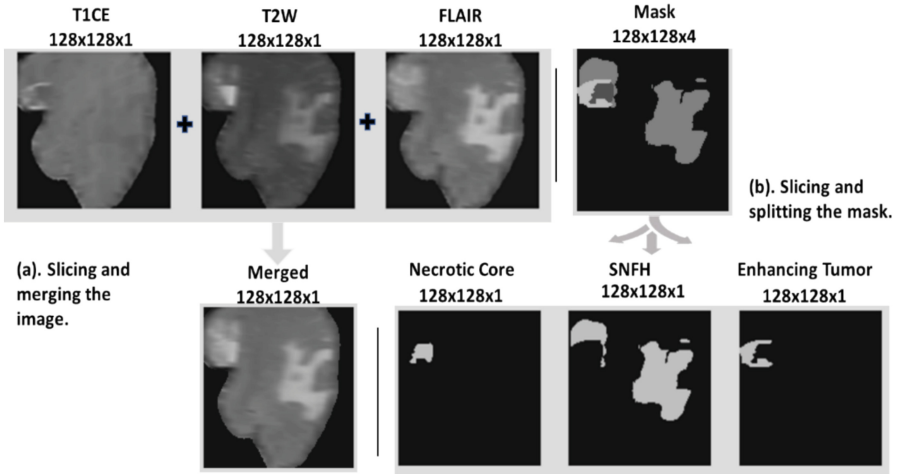


Fig. 1. The 3D to 2D data decomposition workflow illustrated using one case.

3.3 The Segmentation Models

With the objective of designing an efficient brain tumor segmentation model for a low-resource setting scenario, characterized by a small dataset, and limited computational.

resources, the strengths of the U-Net model, which was initially designed for semantic segmentation tasks, was combined with the capabilities of a pretrained image classification model - VGG19, in extracting features. Consequently, our proposed model,

VGG-infused U-Net model (VIU-Net)¹, takes the form of a 2D U-Net with a VGG component serving as the feature extraction enhancement module. Beyond feature extraction another silent novelty and strength of this work is in the deployment of VGG19, which is deployed at a very vital position in the U-Net model (Skip connection) where it helps denoise high dimension features to further improve the performance of the model [19, 20]. The VGG module is employed on the last skip features, which represent the output of the final down-sampling layer. The choice of utilizing VGG over other available pretrained models, such as ResNet [21] and EfficientNet [22], is due to its structural simplicity that allows the extraction of specific parts to be integrated with the U-Net architecture. Following the VGG19 block, a 2D up-sampling layer was added to ensure the output tensor shape aligns with the shape of the first up sampling layer for the concatenation. The architectural detail of the VIU-Net model is summarized in Fig. 2.

3.4 Model Training

In this step, separate 2D segmentation models were trained for each of the three masks (Necrotic Core, SNFH, and Enhancing Tumor). A comparative analysis was conducted between the proposed model and the default U-Net model, resulting in a total of six trained models: three default U-Net models and three VIU-Net models. To maintain consistency for ease of comparison, identical hyperparameters were used for all models. The batch size across all models was set at 64, constrained by memory limitations. Training extended over 200 epochs, incorporating an early stopping technique with a patience value of 3. Gradient optimization employed the Adam optimization technique, initializing with a learning rate of 0.0001. The average Dice Loss and Dice Coefficient values were calculated for the training loss function and for model evaluation, respectively. The calculations are based on the standard computation in segmentation tasks of the similarity between the multi-label segmentation output from the models and the binary tumor masks [23]. The model architecture was implemented using Keras framework with TensorFlow 2.11.0, with CUDA 11.2 on an NVIDIA Tesla T4 GPU.

4 Results

4.1 Results of the Default U-Net Model

The default U-Net model was trained as the baseline for our segmentation problem. Three U-Net models, for each segmentation labels (Necrotic Core, SNFH, and Enhancing Tumor) were trained on 997 of the dataset (70%; training data) and evaluated on the remaining 427 (30%; validation data) dataset.

A dice coefficient of 77.64%, 90.17%, and 83.54 and dice losses of 0.22, 0.098, and 0.16 were obtained for segmenting the Necrotic Core, SNFH, and Enhancing Tumor regions, respectively (Table 1). Figure 3 shows the training and validation curves of the model.

¹ The VIU-Net code can be found at [<https://github.com/CAMERA-MRI/SPARK2023/VIU-Net>].

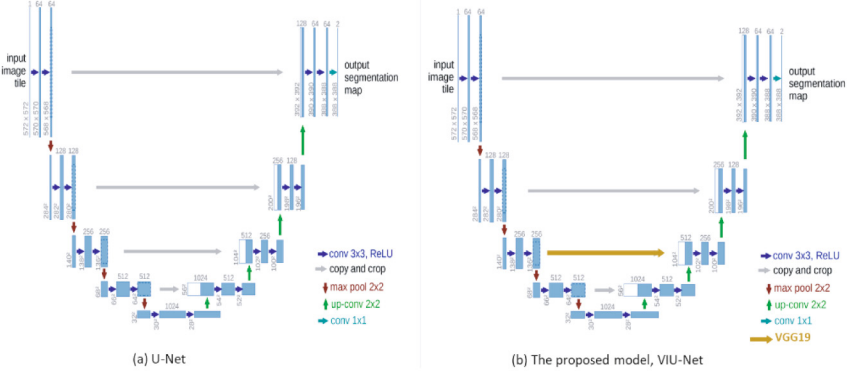


Fig. 2. (a) Architecture of the original U-Net model used with default settings [17] and (b) the architecture of the proposed VGG-Infused U-Net (VIU-Net) model introduced here.

Table 1. Dice Coefficient (Coeff) and Dice loss of the default U-Net and VGG-Infused U-Net (VIU-Net)

Model	Necrotic Core		SNFH		Enhancing Tumor	
	Dice Coeff	Dice Loss	Dice Coeff	Dice Loss	Dice Coeff	Dice Loss
U-Net	77.64%	0.22	90.17%	0.09	83.54%	0.16
VIU-Net	81.31%	0.19	93.16%	0.06	89.39%	0.11

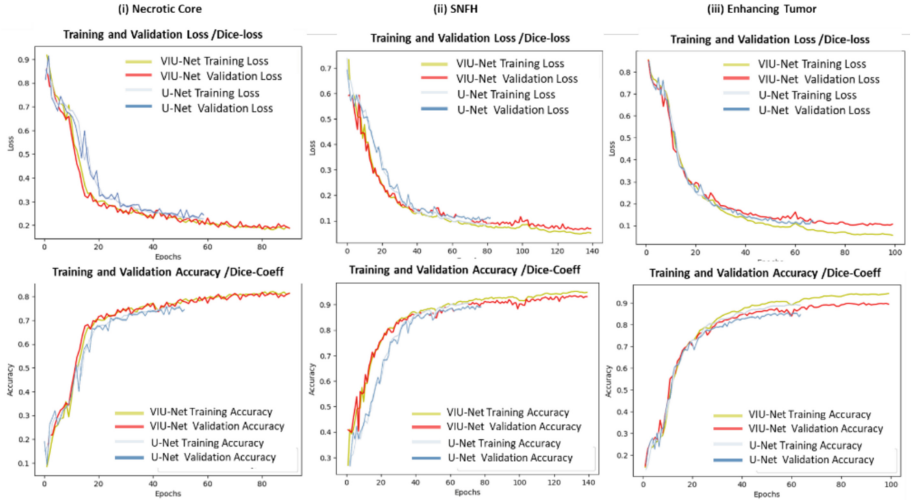


Fig. 3. Training history plot of the U-Net and VIU-Net model.

4.2 Results of the Proposed Model (VGG-Infused U-Net (VIU-Net))

The three VIU-Net models for each segmentation labels were trained and validated on the same dataset at the default U-Net models. The dice coefficients of 81.31%, 93.16%, and 89.39% and dice losses of 0.188, 0.068, and 0.106 were obtained for Necrotic Core, SNFH, and Enhancing Tumor regions, respectively (Table 1). Figure 3 shows the training and validation curves of the proposed model. As shown in Table 1, the proposed VIU-Net model in general, produced higher Dice coefficient and lower Dice losses compared to the default U-Net model. An illustration on a single case (Fig. 4), shows better representation of the predicted tumor sub-region masks using the VIU-Net model compared to the default U-Net model.

Table 2. Comparison of the performance the proposed approach with other reported models for two-dimensional (2D) Brain Tumor segmentation.

Study	Dataset	Data size	Method	Dice Coefficient		
				ET	NC	SNFH
Tseng et al. [24]	BraTs 2015	274	Fused based attention	0.69	0.68	0.85
Li et al. ([25])	BraTs 2017	285	Pair + Fusion	0.75	0.71	0.88
Wang et al. ([26])	BraTS 2018	285	Feature fusion	0.89	-	-
Liu et al. [27]	BraTS 2020	285	Attention based feature fusion	0.76	0.8	0.88
The proposed Model, VIU-Net	BraTS- Africa 2023	60	VGG infused U-Net	0.89	0.81	0.93

Data size is the number of 2D images in the training data used by each study. Enhancing Tumor (ET), Necrotic Core (NC), surrounding non-enhancing FLAIR hyperintensity (SNFH).

5 Discussion

Image segmentation plays a pivotal role in understanding the extent and nature of brain tumors for further treatment decisions. In low-resource settings, where data and computational resources are constrained, the challenge for accurate brain tumor segmentation increases. This study explores the importance of incorporating pre-trained models into the architecture of segmentation models, with a specific emphasis on their effectiveness in improving the performance of models for accurate brain tumor segmentation in environments with limited resources.

The inherent ability of pre-trained models to capture features in images can enhance the performance of underlying segmentation models. A VGG19 pretrained model was integrated with a U-Net architecture for this purpose. The selection of VGG19 was motivated by its simplicity, enabling the extraction and seamless integration of internal components into the U-Net. This approach is particularly advantageous compared to

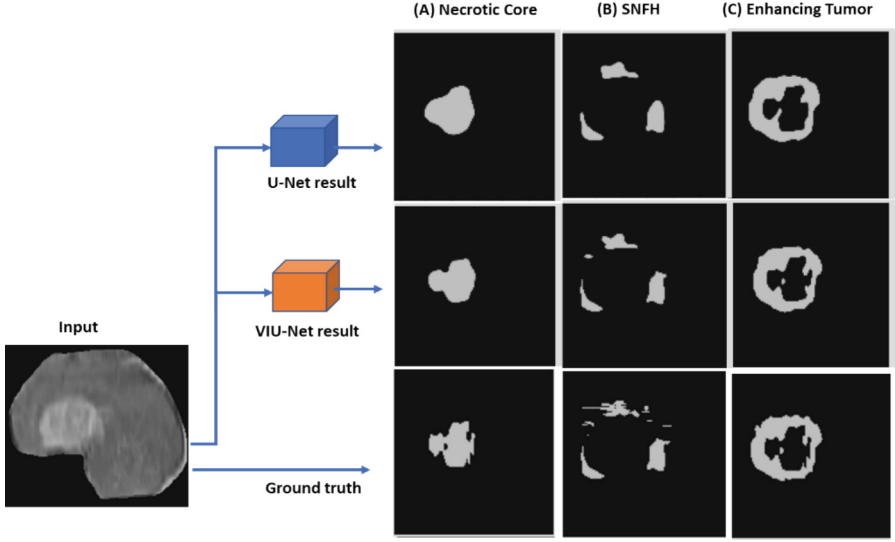


Fig. 4. Comparison the U-Net and the VIU-Net with respect to the ground truth on one slice in a single case.

more complex models like ResNet [21] or EfficientNet [22], where interconnections between layers can pose challenges for selective integration. The data pre-training processing step of decomposing the 3D images into 2D and the transformation of multilabel segment masks into its single-label component masks, enabled focused feature extraction of the tumor sub-regions. Consistent hyperparameters were maintained across all models to facilitate a fair and meaningful comparisons.

While the U-Net model exhibited commendable performance, the proposed model outperformed the baseline model in producing more detailed segmented regions for all three sub-regions, at greater accuracy. This demonstrates how insertion of the pretrained image classification model, VGG19, helps the U-Net to capture complex features and return better accuracy in limited number of datasets, although at nearly twice the training speed (number of epochs). The VGG infused U-Net also produced comparable results to others reported methods were 2D brain tumor segmentation was implemented with much larger datasets (Table 2), even outperforming some methods, suggesting that the proposed model may have promising use for automated brain tumor segmentation in lower resource settings.

We acknowledge that the study's focus on low-resource settings may fully address challenges in these settings, particularly relating to data quality or computational resources. However, these challenges are unmet needs, affecting up to 80% of the world's population where medical imaging capacity is still severely limited and for which further research to explore techniques that can effectively mitigate these constraints is necessary. In addition, the proposed model's generalizability to diverse datasets and its robustness under varying conditions or tumor types requires further investigation.

6 Conclusion

This study aimed to address the challenges of brain tumor segmentation, specifically in low-resourced settings, by proposing a U-Net architecture enhanced with a VGG19 pretrained model. The integration of a VGG19 pretrained model with a 2D U-Net architecture presents a viable solution for improving brain tumor segmentation in resource-constrained environments. The decomposition of higher dimensional data to smaller subsets made segmentation of complex data more feasible in constrained settings. In general, the proposed VGG infused U-Net outperformed the baseline U-Net as well as other previously reported methods developed on much larger dataset. These results contribute valuable insights into the development of efficient segmentation models for medical imaging that have potential real-world applications across many clinical settings.

Acknowledgement. . The author would like to thank the Lacuna Fund for Health and Equity (PI: Udunna Anazodo, 0508-S-001), the National Science and Engineering Research Council of Canada (NSERC) Discovery Launch Supplement (PI: Udunna Anazo-do, DGEER-2022-00136), and the Digital Research Alliance of Canada (the Alliance) for their support. The McGill University Doctoral Internship Program and McMedHacks, Linshan Liu, Lab Manager at the MiND Lab, Montreal Neurological Institute, McGill University, are also thanked for their invaluable contributions. Finally, sincere appreciation is expressed to the SPARK Academy 2023 Instructors and assistants for their dedication and hard work in providing us with valuable training.

References

1. AANS. Brain Tumors (2023). <https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Brain-Tumors>. Accessed 21 Dec 2023
2. Adewole, M., et al.: The Brain Tumor Segmentation (BraTS) Challenge 2023: Glioma Segmentation in Sub-Saharan Africa Patient Population (BraTS-Africa). PubMed (2023)
3. Kanmounye, U.S., et al.: Adult brain tumors in Sub-Saharan Africa: a scoping review. *Neuro Oncol.* **24**(10), 1799–1806 (2022)
4. Mbi Feh, M.K.N., Lyon, K.A., Brahmarroutu, A.V., Tadipatri, R., Fonkem, E.: The need for a central brain tumor registry in Africa: a review of central nervous system tumors in Africa from 1960 to 2017 (2021)
5. Oprea, C., et al.: Brain opportunistic infections and tumors in people living with HIV—still a challenge in efficient antiretroviral therapy era. *J. NeuroVirol.* **29**(3), 297–307 (2023)
6. Luu, H.M., Park, S.H.: Extending nn-UNet for brain tumor segmentation (2021)
7. Rios Piedra, E.A., Taira, R.K., El-Saden, S., Ellingson, B.M., Bui, A.A.T., Hsu, W.: Assessing variability in brain tumor segmentation to improve volumetric accuracy and characterization of change. In: IEEE-EMBS International Conference on Biomedical and Health Informatics (2016)
8. Visser, M., et al.: Inter-rater agreement in glioma segmentations on longitudinal MRI. *NeuroImage: Clin.* **22**, 101727 (2019)
9. Njei, B., et al.: Artificial intelligence for healthcare in Africa: a scientometric analysis. *Heal. Technol.* **13**, 947–955 (2023)
10. Omotoso, O., et al.: Addressing cancer care inequities in sub-Saharan Africa: current challenges and proposed solutions. *Int. J. Equity Health* **22**(1), 189 (2023)
11. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* **34**(10), 1993–2004 (2014)

12. Kazerooni, A.F., et al.: The brain tumor segmentation (BraTS) challenge 2023: Focus on Pediatrics (CBTN-CONNECT-DIPGR-ASNR-MICCAI BraTS-PEDs). PubMed (2023)
13. Jiang, Z., Ding, C., Liu, M., Tao, D.: Two-Stage cascaded U-Net: 1st place solution to BraTS challenge 2019 Segmentation Task. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham (2020)
14. Kamnitsas, K., et al.: Ensembles of multiple models and architectures for robust brain tumour segmentation. In: International MICCAI Brainlesion Workshop. Springer, Cham (2018)
15. Myronenko, A.: 3D MRI Brain Tumor Segmentation Using Autoencoder Regularization, in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham (2019)
16. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer, Cham (2015)
17. Isensee, F., et al.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods (2021)
18. Zhang, D., Confidence, R., Anazodo, U.: Stroke lesion segmentation from low-quality and few-shot MRIs via similarity-weighted self-ensembling framework. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2022. Springer, Cham (2022)
19. Trung, N.T., et al.: Low-dose CT image denoising using deep convolutional neural networks with extended receptive fields. SIViP **16**(7), 1963–1971 (2022)
20. Bodavarapu, P., Srinivas, P.V.V.S.: Facial expression recognition for low resolution images using convolutional neural networks and denoising techniques. Indian J. Sci. Technol. **14**, 971–983 (2021)
21. He, K., et al.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
22. Tan, M., Le, Q.V.: EfficientNet: rethinking model scaling for convolutional neural networks. ArXiv, 2019. abs/1905.11946
23. Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M.: Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Lecture Notes in Computer Science, pp. 240–248. Springer, Cham (2017)
24. Zhou, T., Canu, S., Ruan, S.: Fusion based on attention mechanism and context constraint for multi-modal brain tumor segmentation. Comput. Med. Imaging Graph. **86**, 101811 (2020)
25. Li, Y., Shen, L.: Deep learning based multimodal brain tumor diagnosis. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3. Springer (2018)
26. Wang, H., et al.: Global and local multi-scale feature fusion enhancement for brain tumor segmentation and pancreas segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham (2020)
27. Liu, C., et al.: Brain tumor segmentation network using attention-based fusion and spatial relationship constraint. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham (2021)