Optimizing Brain Tumor Segmentation: A Comparative Study of 3D U-Net Architectures

Tiago Assis, Beatriz Moreira, Rute Patuleia

MSc in Data Science
Deep Learning 23/24

Ciências ULisboa
Faculdade de Ciências da Universidade de Lisboa

INTRODUCTION

Glioma is the most common brain tumor, with glioblastomas (GBM) being the most aggressive and lethal. Despite advancements in diagnosis and treatment, the average survival rate remains about 16 months. Extensive research to improve diagnosis, characterization, and treatment is paramount to reduce the mortality rate of this disease. However, manual segmentation of gliomas is time-consuming, labor-intensive, and prone to variability. Therefore, automated segmentation using convolutional neural networks (CNN) can improve diagnosis, treatment planning, and patient outcomes [1, 2].

OBJECTIVES

- Implement and evaluate top-ranked approaches in the Brain Tumor Segmentation challenge;
- Explore variations of the original U-Net architecture with added dilated convolutions and deep supervision;
- Ablation studies to understand the contribution of several components to the overall model performances.

DATASET

Brain Tumor Segmentation (BraTS) 2023 challenge dataset [1, 2]:

• Includes 1470 patients (1251 training set, 219 test set);

• Multi-parametric 3D MRI scans with four modalities: native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR);

• Ground truth masks annotated and approved by neuro-radiologists, with the labels for enhancing tumor (ET), peritumoral edematous/invaded tissue (ED), and necrotic tumor core (NCR).

• The sub-regions considered in the challenge evaluation are the "enhancing tumor" (ET), the "tumor core" (TC), and the "whole tumor" (WT). The TC entails both the ET as well as the NCR parts of the tumor. The WT describes the complete extent of the tumor, as it entails all the sub-regions.

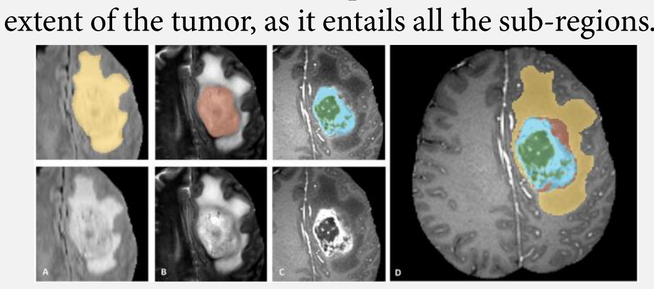


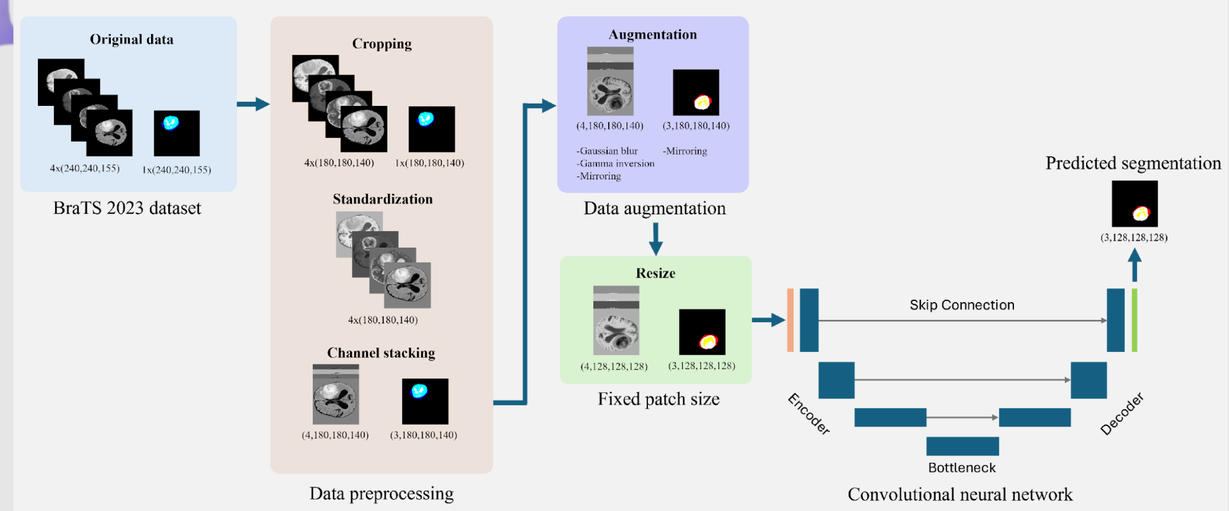
Figure 2. Image patches show from left to right: the whole tumor (A), the tumor core (B), the enhancing tumor structures (C). The segmented sub-regions are combined to generate the final labels of the tumor structures (D). Adapted from [1].

CONCLUSION

This project demonstrated the effectiveness of 3D U-Net architectures in accurately segmenting glioma tumor sub-regions, including whole tumor, tumor core, and enhancing tumor. The addition of dilated convolutions in a U-Net architecture had a good performance in capturing complex details within tumor regions, by reducing the loss of spatial information caused by down-sampling.

Challenges remain, particularly in delineating the ET region, which exhibits high variability. Incorporating data augmentation and deep supervision techniques enhanced model robustness and performance despite the small dataset.

METHODOLOGY



Data preprocessing:

• MRI scans are cropped to maximize brain regions, standardized, stacked in 4 channels, and down-sampled to a fixed size for input.

• Ground truth is transformed into a 3-channel one-hot-encoded segmentation for a more focused classification of the tumor sub-regions.

Data augmentation:

• Random transformations are applied to virtually increase the training set's diversity and enhance model generalizability.

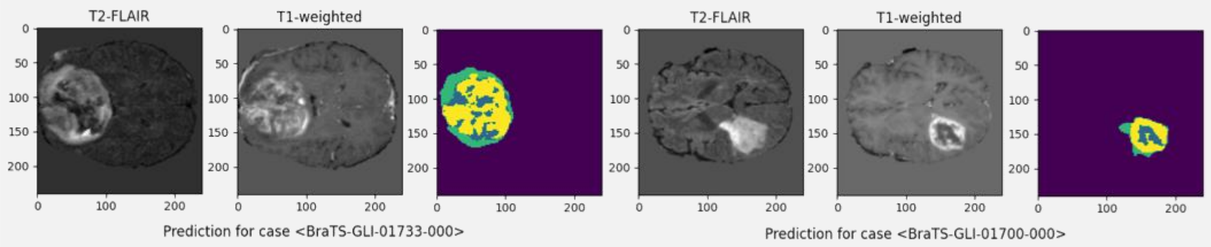
Network architectures:

• 3D U-Net-like networks are implemented and trained, including a U-Net with dilated convolutions (Dilated U-Net) [3] in the bottleneck and the baseline nnU-Net [4] architecture.

Evaluation:

• The Dice Loss and Binary Cross-Entropy Loss are used to evaluate the training procedure, with the Dice Coefficient and Hausdorff Distance (HD) as main performance metrics.

RESULTS



Model	Mean Dice	Mean HD	Sens	Spec
Dilated U-Net	0.7448	16.98	**0.8283**	0.9987
Dilated U-Net – S	0.6988	23.31	0.7625	0.9983
Dilated U-Net + SGD	0.7504	17.22	0.7677	**0.9992**
Dilated U-Net + 2L	**0.7536**	**15.80**	0.7703	0.9991
Dilated U-Net + WD + 2L	0.7468	17.69	0.7781	0.9990
Dilated U-Net + P + 1L	0.7037	24.62	0.7762	0.9985
nnU-Net	0.7314	20.11	0.7667	0.9989
nnU-Net – S	0.7392	18.39	0.7547	0.9991
nnU-Net + CA	0.7480	15.76	0.7644	0.9990
nnU-Net + A	0.7488	17.60	**0.8068**	0.9989
nnUNet + A + WD	**0.7513**	**15.21**	0.7473	**0.9993**
nnU-Net + A + CA + 1L	0.7077	22.58	0.7790	0.9984

Sens – Sensitivity
Spec - Specificity

S – Deep Supervision
A – Adam
SGD – Stochastic Gradient Descent
WD – Weight Decay
CA – Cosine Annealing
P – Polynomial
1L – 1 Loss Function
2L – 2 Loss Functions

- Dilated U-Net yields slightly better overall results.
- Deep Supervision is a necessary performance gain for Dilated U-Net, and the use of SGD and 2 Loss Functions are improvements to the default implementation.
- Adopting only 1 Loss Function in nnU-Net results in a substantial loss in performance, while using Adam and Weight Decay instead of the default SGD achieves superior results.

REFERENCES

[1] Menze B. H., et al. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BraTS).
[2] Bakas S., et al. (2017). Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features.
[3] Henry T., et al. (2020). Top 10 BraTS 2020 challenge solution: Brain tumor segmentation with self-ensembled, deeply-supervised 3D-Unet like neural networks.
[4] Isensee F., et al. (2020). nnU-Net for Brain Tumor Segmentation.

U LISBOA | UNIVERSIDADE DE LISBOA
Ciências ULisboa

Acknowledgements. This poster session is sponsored by:
Quidgest
INNOVATION MAKERS Fueling Financial Business Transformation
SENSE FINITY
NTT DATA