Glioblastoma Tumor Segmentation with Multi-channel 3D U-Net

Nathan Gruenhagen, MD[1] and Manav Bhalla, MD[2]

Medical College of Wisconsin, Milwaukee WI 53213, USA
ngruenhagen@mcw.edu
mbhalla@mcw.edu

Abstract. In this paper, we analyze the use of a multi-channel 3D U-Net convolutional neural network for glioblastoma brain tumor segmentation using multiple MRI sequences including FLAIR, T1, and Contrast Enhanced T1 images. The U-Net architecture is a popularly used method for such tasks, and includes several key features including multi-channel architecture which is well suited to the multiple sequences obtained in MRI studies. In this approach, we apply several preprocessing transformations to multi-sequence 3D MRI data acquired from the 2021 MICCAI Brain Tumor Segmentation Challenge and use this data to train a 3D U-Net Convolutional Neural Network for glioblastoma brain tumor segmentation. Multiclass segmentation is evaluated using the Dice Coefficient metrics for tumor core, enhancing tumor, and whole tumor. Evaluation of multi-sequence vs single-sequence input is also explored. Preliminary 2021 MICCAI BraTS Challenge results are presented.

Keywords: U-Net · Deep Learning · Glioblastoma · Image Segmentation · MRI.

1 Introduction

Gliomas are the most common primary central nervous system tumor, arising from glial progenitor cells including astrocytes and oligodendrocytes. Low grade gliomas include oligodendrogliomas and diffuse astrocytomas, which can be further subdivided based on their genetic classifications relating to IDH mutation and 1p/19q deletion using the current WHO classification scheme. Along a spectrum of severity based on the genetic classification, glioblastomas represent high grade astrocytomas (WHO grade IV) and are also referred to as glioblastoma multiforme (GBM). These tumors can be further subdivided based on the presence of the IDH mutation. Glioblastomas with this mutation tend to occur in younger patients with a relatively better prognosis, and those that do not include the mutation tend to occur in older patients and portend the worst prognosis [9]. Recent studies have also demonstrated an importance in the MGMT promoter methylation genetic sequence, which has been found to be one of the strongest prognostic factors for patients with glioblastoma, and can predict a favorable response to chemotherapy [1].

Paramount to evaluation and treatment of these tumors is an understanding of their manifestation on imaging. Higher grade gliomas (WHO grade III and IV) demonstrate enhancement, with glioblastomas typically demonstrating rapid growth, ring enhancement, and a centrally necrotic core. These tumors demonstrate restricted diffusion owing to their high cellularity, and can cross the midline along white matter tracts. Evaluation of the necrotic core, enhancing tumor, and surrounding T2/FLAIR hyperintense signal becomes central to treatment planning and evaluation of treatment response. Automatic segmentation of these components can aid in treatment planning and evaluation by increasing the efficiency and reproducibility of measurements, and could allow for more exact comparisons to prior studies.

Rapidly developing deep learning technology has made this endeavor more attainable in recent years, with Convolutional Neural Nets (CNNs) representing an important technology relating to image segmentation. This study uses a subtype of CNN known as a U-Net, first described by Ronneberger et al. in 2015 [6], which was later developed into a 3-Dimensional architecture capable of using multichannel three dimensional inputs [7].

2 Methods

2.1 Datset and Preprocessing

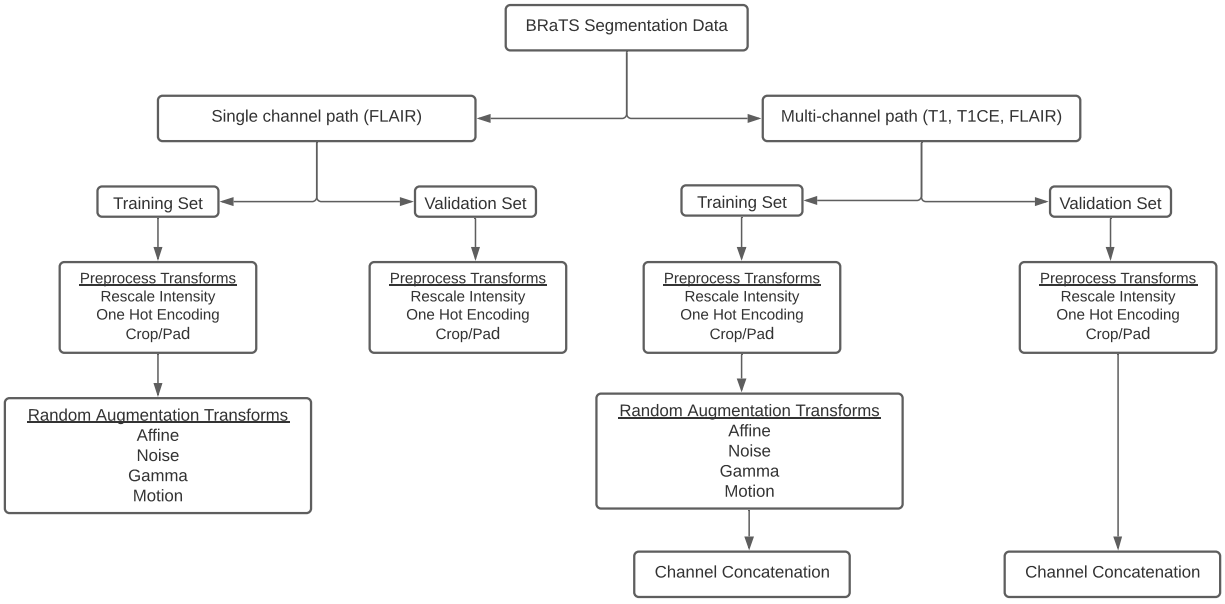

Fig 1. Overview of data pipeline including preprocessing and augmentation. The pipeline is based on TORCHIO documentation [8].

The data used in this study is acquired from the 2021 MICCAI BrATS Challenge data set which includes T1, T2, T1 Contrast Enhanced, and FLAIR MRI sequences with expert annotations of tumor core, enhancing tumor, and FLAIR signal [1-5]. For this paper, 960 subjects were used as the training set and 258 were used as the validation set. Using the TORCHIO framework, several standardized preprocessing transforms were applied to the data set including rescaling and one hot encoding.

Data augmentation was performed on the training set data, which allows for decreased overfitting. Augmentations were applied randomly to the data utilizing the TORCHIO framework as detailed in figure 1. Gamma compression and expansion was performed to augment contrast of random images [8]. To account for motion artifacts encountered in MRI, motion artifacts were simulated on random images. Affine transformations were also applied to random images, which is a type of image transformation that preserves colinearity and ratios of distances [8]. Concatenation of channels was performed after application of the transformations.

2.2 U-Net Architecture

Originally described in 2015 by Ronneberger et. Al, the U-Net architecture includes both a contraction path where the input is down-sampled followed by an opposite, symmetric expanding path where the image is up-sampled, with intervening skip connections. In 2016, a three dimensional U-net is described by Ronneberger et. Al, which is capable of learning from multichannel, 3 dimensional image data.

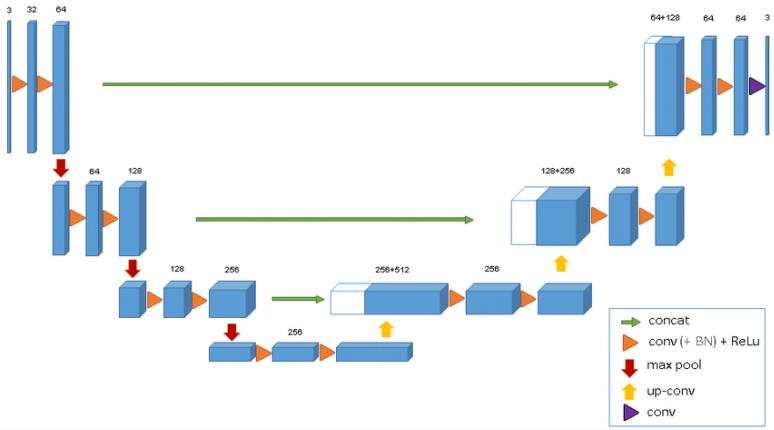

Fig 2. 3D U-Net architecture demonstrating downsampling and upsampling phases based on the paper by O. Ronneberger, et al. [7]

Each level of the U-Net utilizes a convolutional layer which a applies a filter kernel to a subregion of the image, multiplying and summing up the values of the operation to generate an output image [6]. This is performed by subregions across the image prior to each ReLU activation function. Downsampling is achieved using a max-pooling method by applying a maximum filter to subregions. This reduces the computational cost by reducing the size of the image and thus the number of parameters that a model must learn, and also helps to avoid overfitting. The network used in this study utilizes a 3D U-Net architecture implemented via the MONAI Python framework, which also utilizes residual units and parametric ReLU activation functions [10].

The model was trained and implemented using the Pytorch deep learning framework built on the Python programming language, with TORCHIO and MONAI frameworks utilized for data preprocessing and model training. Two separate paths of model training were explored- single channel training using FLAIR images, and multi-channel training using T1, T1 Contrast Enhanced, and FLAIR images. For multi-channel training, a random batch data loader was created using the concatenated T1, T1 Contrast Enhanced, and FLAIR images available for each subject. T2 images were elected to be omitted from the study secondary to hardware constraints and relative redundancy of data relative to FLAIR imaging. After concatenating the images into 3 channel, 3 dimensional inputs, the training data set was used to train the model.

Both single channel and multi-channel training pathways were trained on an NVIDIA GTX 1060 6GB GPU and an AMD Ryzen 5 1500X 3.50 GHz Quad-Core Processor. All models were trained using a batch size of 1 and a learning rate of 0.01. An adaptive moment estimation (ADAM) optimizer algorithm was used for back propagation, as it is computationally more efficient than classical stochastic gradient descent. For the loss function, Dice Loss was used which is commonly implemented to evaluate computer vision segmentation tasks. Each pathway utilized 20 epochs for model training.

3 Experimental Results

3.1 Local Test Segmentation Results

Below are examples showing a comparison of expert labeled annotation images and model predicted segmentation images for both the single channel and multi-channel training pathways on validation set images.

Figure 3 demonstrates validation set results utilizing the single channel training pathway. Mismatch between expert annotations and U-Net labels are apparent along the superfical aspect of the tumor adjacent to the vertex. Figure 4 demonstrates the superior validation set results of the multichannel pathway.
Predictions were generated on the validation data set for both training pathways. The single channel training pathway achieved an average Dice score of 0.50, while the multi-channel pathway achieved an average Dice score of 0.71.

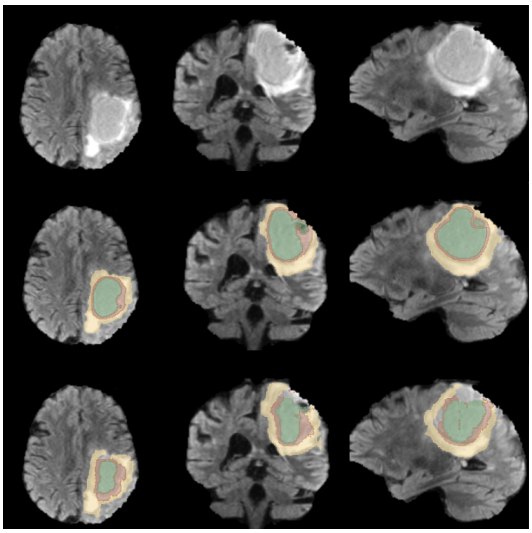

Fig 3. Single channel training pathway validation set example. Expert annotations and U-Net predictions are displayed in the middle and third rows, respectively.

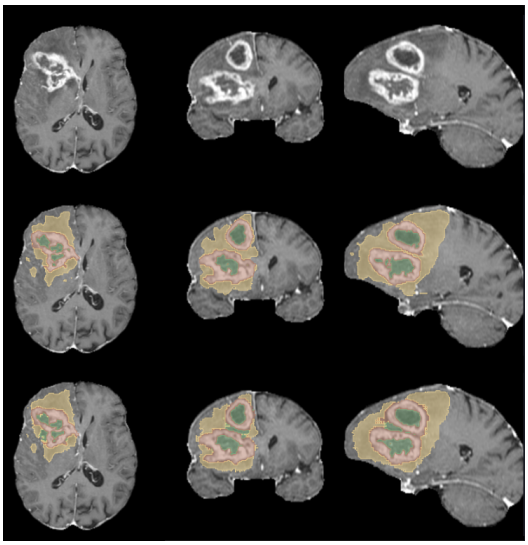

Fig 4. Multi-channel training pathway validation set example. Expert annotations and U-Net predictions are displayed in the middle and third rows, respectively.

3.2 MICCAI BraTS 2021 Validation Set Results

The 2021 MICCAI BraTS Challenge validation images were segmented by each training pathway model. The single channel training pathway achieved average

Dice scores of 0.26 for enhancing tumor, 0.45 for tumor core, and 0.81 for whole tumor label metrics. The multi-channel training pathway achieved average Dice scores of 0.63 for enhancing tumor, 0.72 for tumor core, and 0.84 for whole tumor label metrics. The multi-channel pathway demonstrates superior Dice scores across all categories. Detailed scores for each training pathway are shown below.

Qualitative results are shown below from a single case. Figure 5 demonstrates prediction segmentations generated by the U-Net superimposed on the post contrast images, which illustrates the accuracy of the tumor core and enhancing tumor segmentations. Figure 6 demonstrates predictions superimposed on FLAIR images, allowing a visualization of the accuracy of segmentation of the component of infiltrating tumor. Figure 7 shows the segmentations superimposed on T1 precontrast images, where tumor core is most apparent.

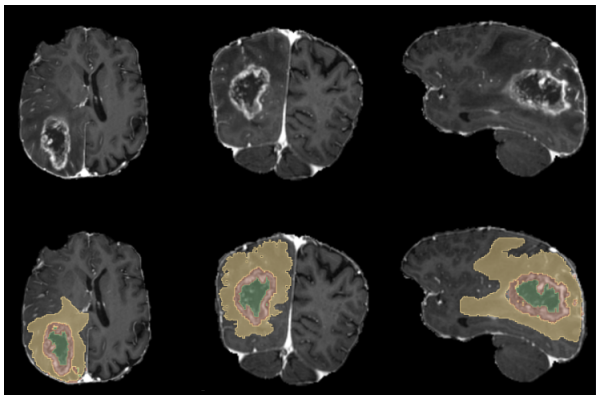

Fig 5. Multi-channel training pathway submission t1ce example.

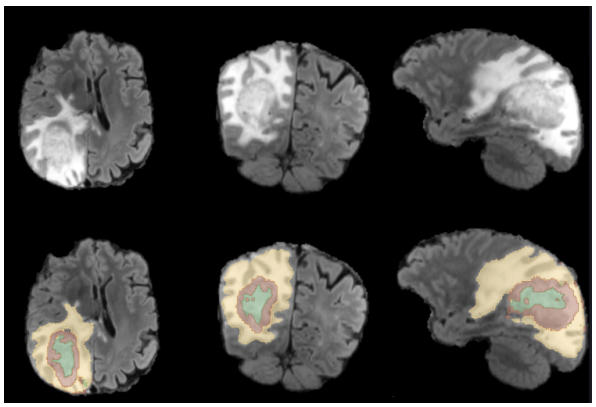

Fig 6. Multi-channel training pathway submission FLAIR example.

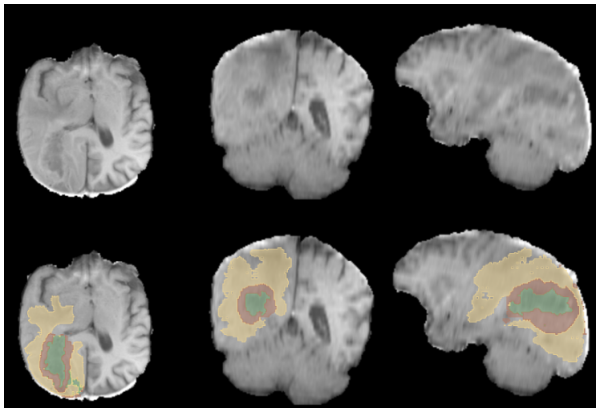

Fig 7. Multi-channel training pathway submission t1 example.

	Dice_ET	Dice_TC	Dice_WT
Single Channel	0.26	0.45	0.81
Multi-Channel	0.63	0.72	0.84

Table 1. Dice scores of the single channel and mutlti-channel training pathways with respect to the different label categories.

	Hausdorff95_ET	Hausdorff95_TC	Hausdorff95_WT
Single Channel	43.14	19.07	9.85
Multi-Channel	39.79	16.19	10.35

Table 2. Hausdorff distance metrics of the single channel and mutlti-channel training pathways.

	Sensitivity_ET	Sensitivity_TC	Sensitivity_WT
Single Channel	0.27	0.39	0.77
Multi-Channel	0.76	0.80	0.83

Table 3. Sensitivity metrics of the single channel and mutlti-channel training pathways.

	Specificity_ET	Specificity_TC	Specificity_WT
Single Channel	0.9989	0.9994	0.9994
Multi-Channel	0.9990	0.9991	0.9991

Table 4. Specificity metrics of the single channel and mutlti-channel training pathways.

4 Discussion and Conclusion

In this paper we explored the use of three dimensional multi-channel U-Net architecture for segmentation of glioblastoma components using the 2021 BRaTS Challenge data utilizing multiple Python deep learning frameworks for multi-sequence MRI image data preprocessing and neural network training. Hyperparameters including learning rate, batch size, and activation functions were kept constant across the single channel and multi-channel training pathways to allow for direct comparison of performance.

The single channel training pathway was strongest in segmentation of the whole tumor dice metric and weakest in the enhancing tumor metric. Thus, this training pathway and data structure can be well suited for whole tumor segmentation tasks for clinical scenarios in which only this metric is required. While not an ideal outcome for multi-sequence multi-component glioblastoma segmentation, this process could be used in many other clinical scenarios. It should be noted that implementation of this pathway in practice gives the advantage of decreased resource requirements including a significantly decreased amount of time required for expert annotation, decreased hardware requirements, and decreased time for model training given the overall decreased amount of parameters to be utilized for training.

Superior performance was seen across all categories within the multi-channel training pathway. State of the art performance can be achieved in whole tumor, enhancing tumor, and core tumor segmentation allowing for precise evaluation of glioblastoma on multi-sequence MRI. Increased efficiency and reproducibility of these segmentation metrics has many advantages in both the realm of Radiation Oncology and Diagnostic Radiology. Deep learning powered tumor mapping can greatly increase the efficiency and precision of radiation treatment planning. Precise automatic mapping can also allow rapid evaluation of treatment response when used to compare to prior images, and would allow for invaluable optimization of workflows in Diagnostic Radiology.

References

1. U.Baid, et al., The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification, arXiv:2107.02314, 2021.
2. B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015) DOI: 10.1109/TMI.2014.2377694
3. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data, 4:170117 (2017) DOI: 10.1038/sdata.2017.117
4. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, et al., "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-GBM collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.KLXWJJ1Q

5. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, et al., "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-LGG collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.GJQ7R0EF
6. Ronneberger, O., Fischer, P., and Brox, T., "U-Net: Convolutional Networks for Biomedical Image Segmentation", arXiv:1505.04597, 2015.
7. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O., "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation", arXiv:1606.06650, 2016.
8. Perez-Garcia, F. Sparks, R. Ourselin, S. TorchIO: a Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. Technical report, 2020.
9. Osborn, A., Hedlund, G., Salzman, K.: Osborn's Brain. 2nd edn. Elsevier, Salt Lake City, Utah (2017). pp. 509-549.
10. Monai: Medical Open Network for AI, https://monai.io/