HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 2 (2024)

Evaluating the Efficacy of the 3D U-Net Architecture for Glioblastoma Multiforme Tumor Segmentation

Arnav Jadhav[1] and Erick Siavichay[#]

[1]Chantilly High School, USA
[#]Advisor

ABSTRACT

Glioblastoma is the deadliest form of brain cancer which begins as a congregation of cancerous cells within the brain but then progresses into invading and destroying healthy brain tissue [9]. Radiation therapy, the most popular treatment option, is where neuro-oncologists apply intense radiation energy beams directly on the tumor region to kill the cancerous cells. However, for radiation therapy to be effective, the segmented magnetic resonance image (MRI) in which the oncologists base where to apply the radiation must be segmented nearly 100% accurately, or else the energy beams will mistakenly damage healthy brain tissue. Additionally, radiation therapy will be futile if the MRI segmentation is not complete in time for the therapy appointment. Our research problem is evaluating the performance of notable segmentation models for accurate and immediate 3D segmentation of glioblastoma multiforme brain tumors from MRIs while focusing on one particular architecture, the 3D U-Net. Each of the models was given 850 MRIs from the BraTS 2021 dataset, which is an annual competition hosted by the University of Pennsylvania [8]. Our 3D U-Net model achieved a testing accuracy of 0.986) accuracy and took approximately 55 seconds to predict segmentations of the tumor region. The results reveal that the 3D U-Net model is capable of automating glioblastoma tumor segmentation in significantly fewer hours than a human oncologist would take, all while maintaining similar or higher accuracy, where even a minor difference can mean the distinction between life and death.

Introduction

Glioblastoma, a highly aggressive form of brain cancer, poses significant challenges in diagnosis and treatment planning. Magnetic Resonance Imaging (MRI) is pivotal in the assessment and management of this condition, offering detailed insights into the tumor characteristics and progression [5]. The precise segmentation of glioblastoma within magnetic resonance images (MRIs) is a critical yet time-intensive task in the field of neuro-oncology. Currently, neurologists and radiologists often collaborate to manually segment tumor areas in MRI scans, a process requiring extreme precision and considerable effort. This manual segmentation is not only laborious but also prone to variability and potential inaccuracies due to human error and the subjective nature of the interpretation. Glioblastoma, characterized by its heterogeneous appearance and complex spatial relationships in three-dimensional MRI data, further complicates this task. Consequently, there is a significant demand for more efficient and accurate methods to segment these tumors [4]. The 3D U-Net architecture, a machine-learning model specifically designed for volumetric segmentation, has shown promise in medical imaging applications [7]. However, its effectiveness in glioblastoma segmentation, particularly compared to other machine-learning models and manual segmentation methods, has not been fully explored. This experiment aims to investigate the efficacy of the 3D U-Net architecture in the segmentation of glioblastoma from MRI images, comparing its performance with other prevalent machine learning models and traditional manual segmentation

methods used by neurologists and radiologists. The ultimate goal is to enhance the diagnostic and treatment planning process for glioblastoma, potentially leading to better patient outcomes and more efficient use of medical resources in neuro-oncology.

Background

In medical imaging, precise segmentation of glioblastoma multiforme (GBM), a notably aggressive brain cancer, is crucial for diagnosis and treatment planning. This experiment's objective leverages machine learning, specifically convolutional neural networks (CNNs) like the U-Net architecture, known for its encoding-decoding framework that efficiently captures and reconstructs intricate image details. This capability is vital for delineating the complex boundaries of GBM tumors [7]. Prior models such as DeepSeg and DeepSCAN have advanced the use of CNNs for medical image segmentation, using structures and techniques that enhance model performance and generalization. Yet, the particular complexity of GBM calls for further refined segmentation methods [10].

The evaluation of these models often hinges on the Dice Similarity Coefficient, a formula that quantifies the accuracy of segmentation by comparing the predicted and actual tumor areas. Throughout this study, variables and units are consistently defined, such as 'f' for the number of filters in a convolutional layer. The DSC is a statistical metric measuring the similarity between two volumetric geometries in terms of the pixel classification (the model's predicted segmentation mask and the corresponding ground truth image) and is particularly useful in the evaluation of segmentation models where higher values indicate better performance. The mathematical formula for DSC is $2|X \cap Y||X| + |Y|$. The variable X represents the set of the numerical values of our model's predicted segmentation pixels and the variable Y represents the set of the numerical values of the pixels of the ground truth image. The numerator of the formula ($|X \cap Y|$) represents the cardinality of the intersection of the sets X and Y, and the denominator represents the total number of elements of both sets combined (non-unique).

Diagrams of the U-Net architecture supplement the text, illustrating the data flow and crucial network components, streamlining the complex science into tangible concepts. This research aims to push the boundaries of GBM segmentation, contributing to the field of computational neuroscience and the application of AI in healthcare [11].

Dataset

We used a dataset from the annual competition Brain Tumor Segmentation Challenge (BraTS). This competition is created and hosted by the University of Pennsylvania's Perelman School of Medicine [8]. The dataset consists of 1250 unique MRIs with five different contrasts applied to each MRI: FLAIR, T1, T1-CE, T2, and the corresponding segmentation mask. Each of these contrasts enhances different features of the brain; for example, the T1 contrast enhances the fats within the brain, while the T2 contrast enhances the waters and fluids [8].

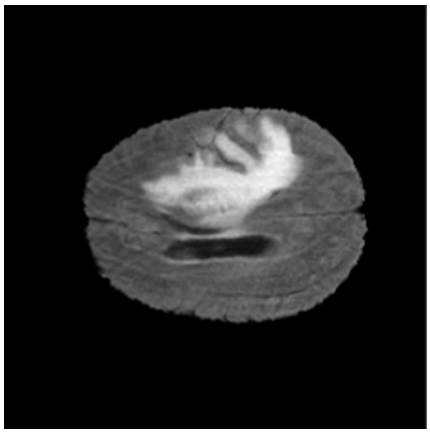

Figure 1. One of the MRIs from the BraTS 2021 Dataset. The general tumor region (large white/light grey area on the top of the brain) is clearly visible.

We split the dataset into three different sets: training, validation, and testing. The training set consisted of 850 MRIs (68% of total data), the validation set had 250 MRIs (20% of total data), and the testing set had 150 MRIs (12% of total data).

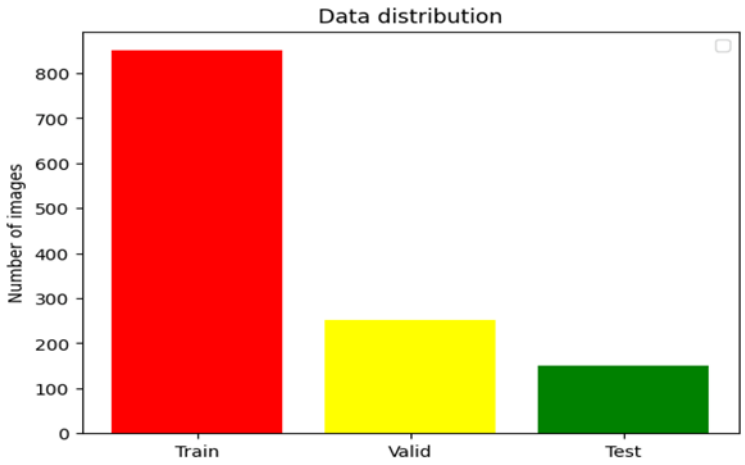

Figure 2. Bar Graph Representing the Distribution of Training, Validation, and Testing Images

Methodology

To address the challenge of accurately segmenting glioblastoma multiforme, we developed a segmentation model based on the U-Net architecture. The U-Net model, particularly in its 3D form, is tailored for the segmentation of volumetric data, such as MRI scans. This model is renowned for its efficiency and accuracy in medical image analysis, making it an ideal choice for our research problem.

Model Development and Evaluation

The 3D U-Net model extends the original U-Net design to analyze spatial hierarchies and patterns in three dimensions, making it adept at capturing context and spatial continuity essential for accurately segmenting

tumors, organs, and other structures in biomedical images. This model comprises an encoder (contraction) path and a decoder (expansion) path, forming a "U" shape. The encoder uses consecutive 3D convolutions with ReLU activation functions and batch normalization to extract features and reduce spatial dimensions, capturing the input images' context. ReLU introduces non-linearity, allowing complex pattern learning, while batch normalization stabilizes learning by normalizing the activations, leading to faster convergence and reduced initialization sensitivity.

The decoder path upsamples the feature maps to restore the input volume's original spatial dimensions, using transposed convolutions or up-convolutions. It also incorporates skip connections from the encoder path, merging high-resolution features with upsampled ones to preserve spatial information lost during downsampling [6]. This ensures detailed segmentation and accurate localization. Dropout is integrated into deeper layers as a regularization technique to prevent overfitting. By randomly setting a fraction of input units to zero, dropout forces the model to learn more generalized features, enhancing its robustness.

Overall, the 3D U-Net architecture's combination of 3D convolutions, ReLU, batch normalization, skip connections, and dropout makes it highly effective for 3D data segmentation. Its ability to process volumetric data thoroughly, maintain spatial information, and ensure model generalization has made it widely used in medical image analysis.

For our model specifically, we empirically determined the optimal hyperparameters to yield the most accurate results as possible. Ultimately, we found that eight epochs (an epoch is each time the entire training data gets passed through to the model) was the right number of epochs to ensure maximum accuracy. Other hyperparameters that we determined to be the most viable were doubling the number of filters after each block in the encoding layer (starting with 8 in the first convolutional layer then finishing with 64 before the bottleneck layer), setting the dropout rate (a metric that determines the probability of removing neurons during training) equal to 0.3 to prevent overfitting, setting n_channels equal to 2 to indicate that the input images have two channels of information, and implementing callbacks (which are useful for monitoring the internal state and statistics of the model during training) from Keras, which is an open-source library from TensorFlow [3].

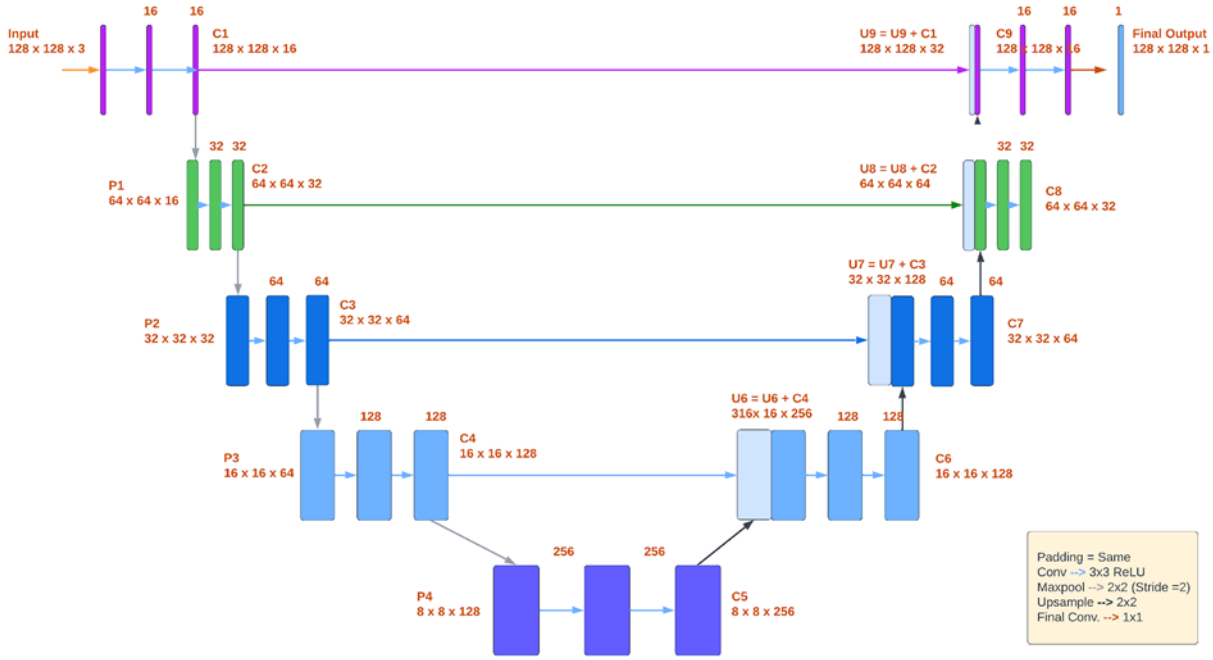

Figure 3. A visual representation of the 3D U-Net architecture that my model implements. This visualization includes the new dimensions of the input image and the number of feature channels as the original input MRI gets downsampled/upsampled through more layers

Model Cross-Comparison

In our study, the performance of the 3D U-Net model was benchmarked against other established segmentation models trained on the same dataset, such as DeepSEG, DeepSCAN, and Ensemble model from a study conducted by Raour A. Zeineldin, Mohamed E. Karar, Oliver Burgert, Franziska Mathis-Ullrich [10]. These other models implement an architecture similar to the U-Net but have some structural difference between them. For example, the DeepSCAN model implements dilated convolutions instead of traditional transition layers and pooling operations, which may not be the most effective for the type of segmentation required in this experiment. The Ensemble model also utilizes multiple models instead of relying on a single model, which in theory would result in each model's unique strengths and biases playing a role in the final segmentation prediction.

Results

The results show that our 3D U-Net model outperforms the other models across all classes: enhancing tumor (ET), tumor core (TC), and whole tumor (WT). These classes are different sub-regions of the larger tumor region and all of the models are designed to segment each sub-region in addition to the tumor as a whole.With DSC scores of 0.9694 for ET, 0.9695 for TC, and 0.9736 for WT, the 3D U-Net model demonstrates a significant improvement in segmentation accuracy compared to DeepSeg, DeepSCAN, and Ensemble models. Notably, the average DSC for the 3D U-Net is 0.9710, which is substantially higher than the other models, reflecting its robustness and precision in segmenting various tumor regions. This superior performance can be attributed to the model's architecture, which may better capture the complexities of tumor shapes and boundaries. Such high accuracy is critical in clinical settings, where precise segmentation directly informs treatment planning and patient outcomes. The results suggest that the 3D U-Net model holds significant potential for clinical application and may set a new standard in medical image segmentation tasks.

Table 1. Comparing our 3D U-Net model's efficacy results with other notable segmentation models. All of these models were ran five times to ensure precision of results and trained on the same data from the same dataset (BraTS 2021).

Model	Metric			
	DSC (ET)	DSC (TC)	DSC (WT)	Average DSC
DeepSeg	0.8356	0.8508	0.9137	0.8667
DeepSCAN	0.8306	0.8683	0.9228	0.8739
Ensemble	0.8438	0.8753	0.9271	0.8821
Our 3D U-Net	**0.9694**	**0.9695**	**0.9736**	**0.9710**

In addition to the Dice Score Coefficient Metric, we included other quantitative machine learning efficacy metrics , including Accuracy and Loss. To calculate Accuracy, you divide the number of correct predictions by the total number of predictions . For the Loss, we used a specific type of loss called Mean Squared Error which is regarded as the simplest type of loss metric to understand and calculate. For some segmentation dataset, our model's mean squared error over the segmentation images is defined as

$$MSE = \frac{1}{HWN}\sum_{i}\sum_{j}(y_j^{(i)} - \hat{y}_j^{(i)})^2$$

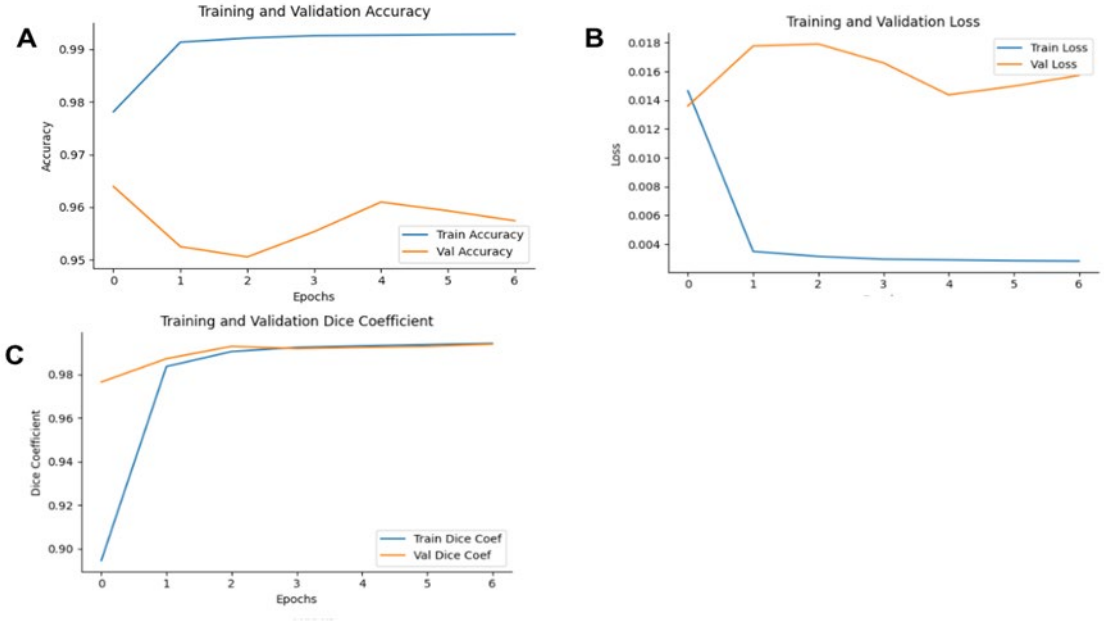

Figure 4. A series of graphs of few machine learning quantitative efficacy metrics, including Accuracy (4A), Loss (4B), and Dice Similarity Coefficient (4C) to quantify our model's impressive results for segmentations on the training data.

Conclusion

In our study, we rigorously developed and evaluated various machine learning models, culminating in a 3D U-Net model specifically tailored for the accurate segmentation of glioblastoma multiforme. Our comparative analysis highlighted the 3D U-Net's remarkable performance over other established models, indicating its potential as a superior tool for medical image analysis. This research's real-world importance cannot be overstated, as the precise segmentation of brain tumors is critical for formulating effective treatment plans, including surgical resection and radiation therapy. The improved accuracy of the 3D U-Net model directly translates into potential increased survival rates and better quality of life for patients. By advancing the methodology for glioblastoma segmentation, we contribute to the broader field of medical imaging and computational neuroscience, offering new avenues for clinical application and research. Our findings reinforce the significance of machine learning in healthcare, underscoring its transformative impact on disease diagnosis, treatment personalization, and ultimately, patient care outcomes.

The segmentation model developed in this research, based on the 3D U-Net architecture, stands ready for deployment in clinical settings, where radiologists and neuro-oncologists can use it to swiftly and accurately segment glioblastoma from MRI scans. By integrating this model into diagnostic workflows, medical professionals can enhance the precision of treatment planning, potentially improving patient prognoses. As for next steps, the model could be augmented with Transformer-based architectures, which have shown great promise in capturing long-range dependencies within data. Incorporating these could further improve the model's ability to distinguish between tumor tissue and healthy brain matter, especially in challenging cases with diffuse boundaries . This evolution of the model would represent a significant leap forward, marrying the spatial proficiency

of CNNs with the contextual awareness of Transformers, setting a new benchmark in medical image segmentation.

Acknowledgments

I would like to thank my advisor Erick Siavichay and Inspirit AI for their help and guidance throughout this project. The completion of this project would not have been possible without Erick and Inspirit AI's support.

References

"73 - Image Segmentation Using U-Net - PART1 (What Is U-Net?)." YouTube, 3 Dec. 2019. www.youtube.com/watch?v=azM57JuQpQI&list=PLZsOBAyNTZwbR08R959iCvYT3qzhxvGOE

Aggarwal, M., Tiwari, A. K., Sarathi, M. P., & Bijalwan, A. (2023, April 26). An early detection and segmentation of brain tumor using Deep Neural Network - BMC Medical Informatics and Decision making. BioMed Central. https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-023-02174-8

Chao, R., & Chollet, F. (n.d.). Writing your own callbacks: Tensorflow Core. TensorFlow. https://www.tensorflow.org/guide/keras/writing_your_own_callbacks

Freeman, Claren. "Multi-Parametric Magnetic Resonance Imaging (mpMRI) Scans for de Novo Glioblastoma (GBM) Patients from the University of Pennsylvania Health System (UPENN-GBM)." Cancer Imaging Archive, 5 Dec. 2023. https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=70225642#702256421a2df2bc3e83428fb618c5a85620321a. Accessed 20 Jan. 2024.

Goryawala, Mohammed, et al. "T1-weighted and T2-weighted subtraction MR images for glioma visualization and grading." Journal of Neuroimaging, vol. 31, no. 1, 2020, pp. 124–131. https://doi.org/10.1111/jon.12800

Ronneberger, O., Fischer, P., & Brox, T. (2015, May 18). U-Net: Convolutional Networks for biomedical image segmentation. arXiv. https://arxiv.org/pdf/1505.04597.pdf

S, Premanand. "A Comprehensive Guide to UNET Architecture: Mastering Image Segmentation." Analytics Vidhya, 5 Nov. 2023. www.analyticsvidhya.com/blog/2023/08/unet-architecture-mastering-image-segmentation/

Schettler, Darien. "Brats 2021 Task 1 Dataset." Kaggle, 19 Aug. 2021. www.kaggle.com/datasets/dschettler8845/brats-2021-task1?select=BraTS2021_00621.tar

Singh, Gaurav, and Ashish Phophalia. "Multimodal Brain Tumor segmentation using modified unet architecture." Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, 22 July 2022, pp. 295–305. https://doi.org/10.1007/978-3-031-08999-2_24

Thakkar, Jigisha P., et al. "Glioblastoma Multiforme." American Association of Neurological Surgeons. www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Glioblastoma-Multiforme#. Accessed 20 Jan. 2024.

Zeineldin, R. A., Karar, M. E., Burgert, O., & Mathis-Ullrich, F. (2022). Multimodal CNN Networks for Brain Tumor Segmentation in MRI: A BraTS 2022 Challenge Solution. arXiv. https://arxiv.org/ftp/arxiv/papers/2212/2212.09310.pdf