Computational Longitudinal Modeling of 3D Preoperative Brain Tumor Volume in Glioblastoma Patients

Mihir Garimella
Nashua High School South
Nashua, NH, USA 03062
mgarimella343@gmail.com

Abstract

Glioblastoma (GBM) is a malignant, aggressive brain tumor. The most common form of tumor in the central nervous system, the fast-advancing nature of GBM makes early treatment vital to patient survival, with average survival post-diagnosis only 15 months. The pipeline for GBM diagnosis involves analysis of multi-parametric magnetic resonance imaging (mpMRI), followed by tumor tissue biopsy to identify mutations affecting prognosis. This time-intensive diagnosis with imprecise results (Dice coefficient=0.74-0.85) is a bottleneck in the treatment process, and increases the chance of relapse post-surgery. We introduce a multi-step system for longitudinal modeling of GBM, consisting of a volumetric slice-wise 2D U-Net model for tumor segmentation from low-resolution mpMRI scans, and a convolutional long short-term memory network (ConvLSTM) for prediction of tumor growth. The efficient U-Net designed (<5 seconds, Dice coefficient=0.86) is utilized to generate tumor masks for mpMRI scans from past and current time steps, which are fed into the ConvLSTM to determine the tumor volume at a future time. With an average time step of 26.0±2.86 weeks, our system models the growth of life-threatening tumors 6 months in advance, while maintaining expert-level accuracy and timing, with validation Dice coefficient=0.75, and predictions in under 4 seconds. The novel approach introduced successfully provides an effective method for longitudinal tumor volume forecasting, with radiologist-comparable accuracy and significantly improved diagnosis time.

1 Introduction

Glioblastoma (GBM) is the most common form of cancerous growth in the central nervous system and a fast-advancing tumor developed in the hemispheres of the brain (Hanif et al., 2017). Within the United States, an estimated 12,000 cases of GBM arise every year, making up over 75% of all primary malignant tumors diagnosed. The weak prognosis of patients, despite medical advances, is clear, with median survival between 14 and 15 months (Mesfin et al., 2019; Lee et al., 2018).

The reason for such poor outcomes of GBM cases is due to the time intensive nature of analyzing such tumors. Current protocol for diagnosis involves 3 lengthy steps: a neurological examination, medical imaging, and tissue biopsy (Hanif et al., 2017; Hirahata et al., 2022). As time-to-surgery increases, the rapid-growing nature of GBM decreases survival outcomes, and relapse becomes more likely, requiring intensive chemotherapy to overcome. With an average doubling time of less than 1 month, efficient diagnosis is imperative to GBM patient survival (Müller et al., 2021).

In this study, we introduce a novel system for growth modeling of GBM tumors in the brain. This multi-step approach consists of a 2D U-Net model for real-time tumor segmentation, and a

Preprint. Under review.

convolutional long short-term memory (ConvLSTM) for time series modeling of the growth of GBM tumor volumes (Ronneberger et al, 2015; Shi et al., 2015). While the implementation of a U-Net design for tumor segmentation has been performed, this study aims to explore an alteration while retaining state-of-the-art accuracy (Zheng et al., 2022). GBM volume modeling, on the other hand, has not been attempted prior through means of an LSTM forecasting network (Morris et al., 2006). Past designs of future tumor volume modeling utilize mathematical methods of growth simulation, but these approaches are inadequate when considering the amount of generalizations made to simplify proposed functions, such as a constant diffusion and proliferation rates (Le et al., 2016; Mascheroni et al., 2021; Pei et al., 2020). The introduced system provides a comprehensive, novel approach to evaluating the growth of GBM tumor volumes.

2 Materials and Methods

Two public datasets were used for model development. The following MRI sequences are present for subjects in each (subjects without all sequences were excluded): t1, contrast enhanced t1-weighted (t1ce), t2-weighted, and t2 Fluid Attenuated Inversion Recovery (FLAIR), although only t1ce and FLAIR — selected for effectiveness in determining borders between tumor and tissue — were used, to meet computation constraints (Beheshti et al., 2020). Segmentations performed by radiologists were also provided for each subject. For the U-Net model, the BraTS 2021 dataset was used (Baid et al., 2021; Menze et al., 2015; Bakas et al., 2017). This >1000 subject dataset was split into training (87.875%), validation (7.5%), and testing (4.625%) prior to model development. The LUMIERE dataset was used to design the ConvLSTM model for prediction of tumor growth (Suter et al., 2022). This 90 subject dataset provides a starting point for comprehensive modeling of tumor growth, acknowledging the economical complexity of acquiring long term brain scan data. After removing subjects with less than 3 time steps available, 42 subjects were left to develop the ConvLSTM, split into training (90.25%), validation (5%), and testing (4.75%) batches prior to model development.

2.1 Segmentation U-Net

The proposed U-Net is designed to work with easily acquirable imaging modalities, while maintaining relatively low computational intensity. As such, inputs to the model are represented as a 4D vector, accommodating 2 3D MRI sequences (t1ce and FLAIR), allowing for streamlined integration with modern imaging methodology.

The decision to use a 2D U-Net structure was due to its effectiveness shown in past image segmentation challenges (Ronneberger et al, 2015). While a 3D U-Net exists, the computing power required to train such a model is far greater than the minimal resources available in this research (Çiçek et al., 2016). Using a 2D approach over a 3D approach restricts the model from utilizing relevant data from neighboring slices of the MRI scan, potentially leading to issues with multi-slice shape continuity. However, results from the proposed U-Net suggest that while some theoretical limitations may be present, the use of multi-sequence MRI to capture different tumor aspects, and the structure of the model, overcome these challenges to produce highly accurate segmentations.

The U-Net consists of 5 pooling layers followed by 5 upsampling layers, ensuring consistent input-output sizes. All hidden convolutional layers in the model utilize ReLU activation, and batches are normalized before subsequent processing (Agarap, 2018). Between the pooling and upsampling layers, a dropout function is implemented to reduce overfitting. The final output is a 4D array consisting of 4 one-hot encoded 3D masks of different GBM tumor regions provided in the ground truth data: Non-tumor, Core tumor, Edema, and Enhancing tumor. The model was trained on the BraTS 2021 dataset, and utilized the cross-entropy metric as a loss function. Accommodating the probabilistic nature of the cross-entropy metric, the final convolutional layer of the model undergoes softmax activation to determine normalized probability scores for each class and data point (Bridle, 1989). Training took 10 hours on a single NVIDIA RTX 3070, and later evaluation was assessed on the unseen testing split of the data, randomly generated prior to training.

2.2 Longitudinal ConvLSTM

Inputs to the ConvLSTM are represented as a 5D vector, consisting of 2 time steps, each with 2 3D segmentation masks. In order to keep training within computational restraints while minimizing

value loss, the model works with simplified alterations of the tumor segmentation masks produced from the U-Net. Rather than using all 4 classes of tumor regions, inputs are simplified to only contain masks of Non-tumor data points, extracted directly from the predicted segmentation, and Tumor data points, a combination of all remaining 3 classes. The ConvLSTM is designed to input 2 time steps worth of simple tumor segmentations, and output the forecasted segmentation at a future step. After cleaning the dataset, the mean difference in time steps was 26.0±2.86 weeks, determining the future time step to which the model was capable of predicting.

The proposed ConvLSTM consists of 4 convolution steps, with the first 3 steps utilizing a ReLU activation, and the final step using a softmax activation to create probabilistic maps for output compatibility with the cross-entropy metric. A decreasing convolution kernel size is implemented over the steps to most effectively gather information from the time steps available. Layer outputs are normalized and padded to the size of input to retain segmentation size over time. The output of the model is a 4D vector (smaller than the input due to the output containing one time step rather than two steps), consisting of 2 one-hot encoded 3D masks corresponding to data points that are predicted to be Non-tumor and Tumor for the future time step.

Due to the significantly reduced size of training data (<40 subjects) present in the LUMIERE dataset, a custom loss function was used to maximize feature extraction and performance (Suter et al., 2022). The following loss function was defined:

$$Loss(p, q) = CE(p, q) + 2(1 - DCS(p, q))$$

where p are predicted class values and q are ground truth values. The double weighting of the DCS loss was determined through multiple trials of model training. Training took 30 minutes on a single NVIDIA RTX 3070, although trials with different loss functions took from 15 minutes to 2 hours.

3 Results

The combined system utilizes both models together to generate predictions of the volume of GBM tumors based on 2 previous time steps worth of data. Given t1ce and FLAIR scans from past and present time steps with a difference of approximately 6 weeks, the system predicts each time step tumor segmentation mask by U-Net inference, and, after simplifying the classes to only Non-tumor or Tumor, uses the ConvLSTM to model the growth of the tumor approximately 26.0 weeks in the future.

In evaluation of the system, we assess the U-Net and ConvLSTM by analyzing metrics and predictions by the models. All testing results are based on subjects randomly split from training data prior to model development. MRI scan analyses are performed on a specific 2D slice of the predicted and truth scans to ensure compared images are from the same region of the subject brain.

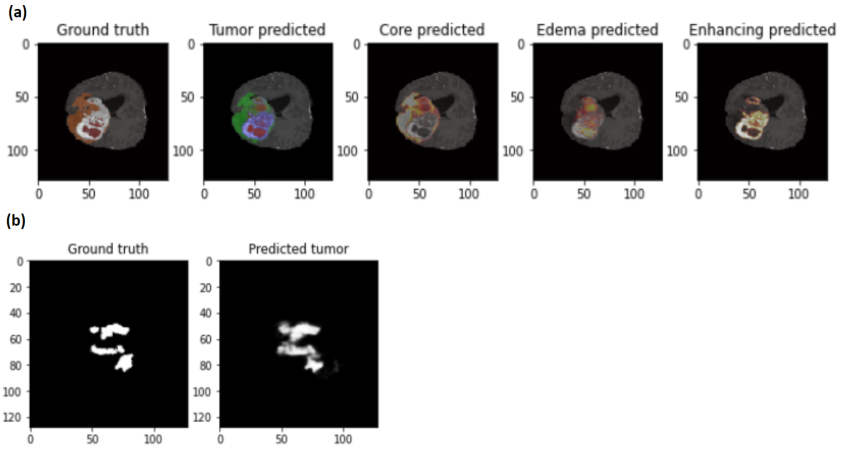

Figure 1: U-Net results. Visualizations include region-specific (a), and total volume predictions (b).

Figure 1 displays predictions made by the U-Net. With a test Dice Coefficient (DSC)=0.86, the U-Net designed performs at a level comparable to radiologists, with experts performing manual

segmentations at DSC=0.74-0.85 (Ben naceur et al., 2020). Figure 1a displays this result, as sections identified are highly similar to what was defined within the ground truth. Furthermore, the predicted total tumor volume displayed in Figure 1b is highly similar to the ground truth segmentation (DSC=0.90), showing the ability of the U-Net to capture the full volume of the tumor. Similar improvements can be seen in the time efficiency of the model. In contrast to the time-intensive GBM segmentation in current diagnosis procedures, the model prediction time is within 5 seconds, with predicted segmentations for both subjects determined in 110 and 152 milliseconds respectively. These improvements on current in-situ volumetric accuracy and time efficiency show the relevance of the slice-focused approach to tumor segmentation.

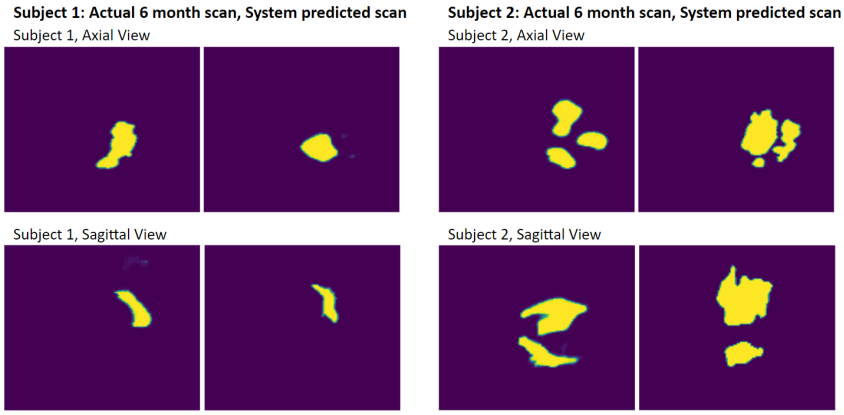

Figure 2: ConvLSTM results. Visualizations include views of true and predicted tumors for 2 subjects.

Figure 2 displays predictions made by the ConvLSTM. The model understands the relative position of different sections of the future GBM tumor, and in the case of less-developed tumors (52 weeks for subject 1), was able to accurately model the size of grown volumes. These results are confirmed when observing subject 2, in which the ConvLSTM captures the quantity and position of the tumor sections. However, when encountering larger tumors which have already grown significantly (72 weeks elapsed for subject 2), the model fails to predict the size of GBM sections. This is expected, as the development of large tumors is heavily dependent on the subject and their personal physical health (Sullivan & Heiden, 2019). A test DSC=0.75, within the range of expert prediction accuracies, indicates the effectiveness of the model in predicting the general structure of the tumor. Additionally, when considering the month-long time-to-surgery considered efficient in current GBM treatment plans, the reduction in tumor prognosis prediction times (<4 seconds for both subjects) while maintaining accuracy is significant (Müller et al., 2021; Kalita et al., 2023).

4 Discussion and Conclusions

In this study, we introduce a novel system for longitudinal GBM tumor growth modeling with improved prediction times and radiologist-comparable accuracy. The significance of the introduced approach extends beyond its predictive capability. Accurate and time-efficient prediction of the prognosis of malignant GBM tumors is crucial to patient outcomes, reducing time-to-surgery and post-surgery relapse possibilities (Müller et al., 2021). While the strength of the proposed system is evident, the method is not without limitations. The use of a small dataset and resource constraints suggest areas for improvement. Future work could involve training and evaluation with larger datasets and computers, and the incorporation of attention-based learning to most effectively use data contained in subject MRI scans. Such efforts would test the robustness of the proposed method, and expand current understanding of tumor growth modeling. The novel longitudinal GBM modeling system introduced in this study signifies an important advancement in medical growth diagnostic techniques. As research is done exploring the intersection of neurology and learning-based algorithms, systems such as the proposed highlight the value of multi-step approaches, paving the way for the delivery of optimally-timed clinical interventions.

Acknowledgments

The author would like to thank teachers at Nashua High School South, namely Statistics teacher Mr. Jeffery Durnford, for their assistance in the ideation of this research.

References

Agarap, Abien Fred. (2018). Deep Learning using Rectified Linear Units (ReLU). ArXiv.org. https://arxiv.org/abs/1803.08375

Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., Colak, E., Farahani, K., Kalpathy-Cramer, J., Kitamura, F. C., Pati, S., Prevedello, L. M., Rudie, J. D., Sako, C., Shinohara, R. T., Bergquist, T., Chai, R., Eddy, J., Elliott, J., Reade, W., & Schaffter, T. (2021). The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification. ArXiv:2107.02314 [Cs]. https://arxiv.org/abs/2107.02314

Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J. S., Freymann, J. B., Farahani, K., & Davatzikos, C. (2017). Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. Scientific Data, 4(1). https://doi.org/10.1038/sdata.2017.117

Beheshti, I., Sone, D., Norihide Maikusa, Kimura, Y., Shigemoto, Y., Sato, N., & Matsuda, H. (2020). FLAIR-Wise Machine-Learning Classification and Lateralization of MRI-Negative 18F-FDG PET-Positive Temporal Lobe Epilepsy. Frontiers in Neurology, 11. https://doi.org/10.3389/fneur.2020.580713

Ben naceur, M., Akil, M., Saouli, R., & Kachouri, R. (2020). Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy. Medical Image Analysis, 63, 101692. https://doi.org/10.1016/j.media.2020.101692

Bridle, J. (1989). Training Stochastic Model Recognition Algorithms 211 Training Stochastic Model Recognition Algorithms as Networks can lead to Maximum Mutual Information Estimation of Parameters. NIPS. https://proceedings.neurips.cc/paper/1989/file/0336dcbab05b9d5ad24f4333c7658a0e-Paper.pdf

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. ArXiv:1606.06650 [Cs]. https://arxiv.org/abs/1606.06650

Hanif, F., Muzaffar, K., Perveen, K., Malhi, S. M., & Simjee, S. U. (2017). Glioblastoma Multiforme: A Review of its Epidemiology and Pathogenesis through Clinical Presentation and Treatment. Asian Pacific Journal of Cancer Prevention : APJCP, 18(1), 3–9. https://doi.org/10.22034/APJCP.2017.18.1.3

Hirahata, T., ul Quraish, R., Quraish, A. ul, ul Quraish, S., Naz, M., & Razzaq, M. A. (2022). Liquid Biopsy: A Distinctive Approach to the Diagnosis and Prognosis of Cancer. Cancer Informatics, 21(11769351221076062). https://doi.org/10.1177/11769351221076062

Kalita, O., Kazda, T., Reguli, S., Radim Jancalek, Pavel Fadrus, Slachta, M., Pospisil, P., Krska, L., Vrbkova, J., Lumir Hrabalek, Smrcka, M., & Radim Lipina. (2023). Effects of Reoperation Timing on Survival among Recurrent Glioblastoma Patients: A Retrospective Multicentric Descriptive Study. Cancers, 15(9), 2530–2530. https://doi.org/10.3390/cancers15092530

Le, M., Delingette, H., Kalpathy-Cramer, J., Gerstner, E. R., Batchelor, T., Unkelbach, J., & Ayache, N. (2016). MRI Based Bayesian Personalization of a Tumor Growth Model. IEEE Transactions on Medical Imaging, 35(10), 2329–2339. https://doi.org/10.1109/TMI.2016.2561098

Lee, J., Park, S. H., & Kim, Y. Z. (2018). Prognostic Evaluation of Neurological Assessment of the Neuro-Oncology Scale in Glioblastoma Patients. Brain Tumor Research and Treatment, 6(1), 22. https://doi.org/10.14791/btrt.2018.6.e1

Mascheroni, P., Savvopoulos, S., Alfonso, J. C. L., Meyer-Hermann, M., & Hatzikirou, H. (2021). Improving personalized tumor growth predictions using a Bayesian combination of mechanistic modeling and machine learning. Communications Medicine, 1(1), 1–14. https://doi.org/10.1038/s43856-021-00020-4

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M.-A., Arbel, T., Avants, B. B., Ayache, N., Buendia, P., Collins, D. L., Cordier, N., & Corso, J. J. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Transactions on Medical Imaging, 34(10), 1993–2024. https://doi.org/10.1109/tmi.2014.2377694

Mesfin, F. B., & Al-Dhahir, M. A. (2019). Cancer, Brain Gliomas. Nih.gov; StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK441874/ Morris, M., Greiner, R., Sander, J., Murtha, A., & Schmidt, M. (2006). Learning a Classification-based Glioma Growth Model Using MRI Data. Journal of Computers, 1(7). https://doi.org/10.4304/jcp.1.7.21-31

Müller, D. M. J., De Swart, M. E., Ardon, H., Barkhof, F., Bello, L., Berger, M. S., Bouwknegt, W., Van den Brink, W. A., Conti Nibali, M., Eijgelaar, R. S., Furtner, J., Han, S. J., Hervey-Jumper, S., Idema, A. J. S., Kiesel, B., Kloet, A., Mandonnet, E., Robe, P. A. J. T., Rossi, M., & Sciortino, T. (2021). Timing of glioblastoma surgery and patient outcomes: a multicenter cohort study. Neuro-Oncology Advances, 3(1), vdab053. https://doi.org/10.1093/noajnl/vdab053

Pei, L., Bakas, S., Arastoo Vossough, Syed, Christos Davatzikos, & Iftekharuddin, K. M. (2020). Longitudinal brain tumor segmentation prediction in MRI using feature and label fusion. Biomedical Signal Processing and Control, 55, 101648–101648. https://doi.org/10.1016/j.bspc.2019.101648

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv.org. https://arxiv.org/abs/1505.04597 Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., & Woo, W. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. ArXiv:1506.04214 [Cs]. https://arxiv.org/abs/1506.04214

Sullivan, M. R., & Heiden, M. G. V. (2019). Determinants of nutrient limitation in cancer. Critical Reviews in Biochemistry and Molecular Biology, 54(3), 193–207. https://doi.org/10.1080/10409238.2019.1611733

Suter, Y., Knecht, U., Valenzuela, W., Notter, M., Hewer, E., Schucht, P., Wiest, R., & Reyes, M. (2022). The LUMIERE dataset: Longitudinal Glioblastoma MRI with expert RANO evaluation. Scientific Data, 9(1). https://doi.org/10.1038/s41597-022-01881-7

Zheng, P., Zhu, X. & Guo, W. Brain tumour segmentation based on an improved U-Net. BMC Med Imaging 22, 199 (2022). https://doi.org/10.1186/s12880-022-00931-1