Fully Automated Brain Tumor Segmentation and Survival Prediction of Gliomas using Deep Learning and MRI

Chandan Ganesh Bangalore Yogananda, M.S.[1], Sahil S. Nalawade, M.S.[1], Gowtham K. Murugesan, M.S.[1], Ben Wagner, B.M.[1], Marco C. Pinho, M.D.[1], Baowei Fei, Ph.D.[2], Ananth J. Madhuranthakam, Ph.D.[1], Joseph A. Maldjian, M.D.[1]

[1] Advanced Neuroscience Imaging Research Lab, Department of Radiology, University of Texas Southwestern Medical Center, Dallas, Texas, USA

[2] Department of Bioengineering, University of Texas at Dallas, Texas, USA

ABSTRACT

Tumor segmentation of magnetic resonance (MR) images is a critical step in providing objective measures of predicting aggressiveness and response to therapy in gliomas. It has valuable applications in diagnosis, monitoring, and treatment planning of brain tumors. The purpose of this work was to develop a fully automated deep learning method for brain tumor segmentation and survival prediction. Well curated brain tumor cases with multi-parametric MR Images from the BraTS2019 dataset were used. A three-group framework was implemented, with each group consisting of three 3D-Dense-UNets to segment whole tumor (WT), tumor core (TC) and enhancing tumor (ET). This method was implemented to decompose the complex multi-class segmentation problem into individual binary segmentation problems for each sub-component. Each group was trained using different approaches and loss functions. The output segmentations of a particular label from their respective networks from the 3 groups were ensembled and post-processed. For survival analysis, a linear regression model based on imaging texture features and wavelet texture features extracted from each of the segmented components was implemented. The networks were tested on the BraTS2019 validation dataset including 125 cases for the brain tumor segmentation task and 29 cases for the survival prediction task. The segmentation networks achieved average dice scores of 0.901, 0.844 and 0.801 for WT, TC and ET respectively. The survival prediction network achieved an accuracy score of 0.55 and mean squared error (MSE) of 119244. This method could be implemented as a robust tool to assist clinicians in primary brain tumor management and follow-up.

Keywords: *Brain tumor segmentation, deep learning, BraTS, Dense-UNet, MRI, Survival prediction, imaging features, Radiomics features, pyradiomics.*

1. INTRODUCTION

Gliomas account for the most common malignant primary brain tumors in both pediatric and adult populations [1]. They arise from glial cells and are divided into low grade and high grade gliomas with significant differences in patient survival. Patients with aggressive high grade gliomas have life expectancies of less than 2 years [2]. Glioblastoma (GBM) are aggressive brain tumors classified by the world health organization (WHO) as stage IV brain cancer [3, 4]. The overall survival for GBM patient is poor and is in the range of 12 to 15 months [5-7]. These tumors are typically treated by surgery, followed by radiotherapy and chemotherapy. Gliomas often

consist of active tumor tissue, necrotic tissue and surrounding edema. Magnetic Resonance Imaging (MRI) is the most commonly used modality to assess brain tumors because of its superior soft tissue contrast. It is routinely used in the clinical work-up of patients for brain tumor diagnosis, monitoring progression and treatment planning [8, 9]. Each MR sequence provides specific information about different tissue sub-components of gliomas. For instance, T1-weighted images with intravenous contrast highlight the most vascular regions of the tumor, called 'enhancing tumor' (ET), along with the 'tumor core' (TC) that does not involve peri-tumoral edema. Conventional T2-weighted (T2W) and T2W-Fluid Attenuation Inversion Recovery (FLAIR) images are used to evaluate the tumor and peri-tumoral edema together defined as the 'whole tumor' (WT) [10].

MRI tumor segmentation is used to identify the subcomponents as enhancing, necrotic or edematous tissue. Due to heterogeneity and tissue relaxation differences in these subcomponents, multi-parametric (or multi-contrast) MRI are often used simultaneously for accurate segmentation [11]. Manual brain tumor segmentation is a challenging and tedious task for human experts due to the variability of tumor appearance, unclear borders of the tumor and the need to evaluate multiple MR images with different contrasts simultaneously [12]. In addition, manual segmentation is often prone to significant intra- and inter-rater variability [12, 13]. Hence machine learning algorithms have been developed for tumor segmentation with high reproducibility and efficiency [12-14]. Following the early success of CNNs [14, 15], they are used as one of the major machine learning methods to achieve great success in clinical applications. [16, 17]. Furthermore, underlying molecular heterogeneity in gliomas makes it difficult to predict the overall survival (OS) of GBM patients based on MR imaging alone [16, 17]. Clinical features [5], along with MR imaging based texture features [18-20] have been used to predict the OS in GBM patients. In this work, we utilized designed a 3D Dense-Unet for segmenting brain tumors into subcomponents and used MRI based texture features from each of these subcomponents for survival prediction in GBM patients. The purpose of this work was to develop a deep learning method with high prediction accuracy for brain tumor segmentation and survival prediction that can be easily incorporated into the clinical workflow.

2. MATERIAL & METHODS

2.1 Data & Pre-Processing

2.1.1 Brain Tumor Segmentation

335 well curated multi-parametric brain MR images including T2w, T2w-FLAIR, T1w and T1C (post contrast) from the BraTS2019 dataset were used [10, 21-24]. The dataset consisted of 259 high grade glioma (HGG) cases and 76 low grade glioma (LGG) cases. The dataset also included three ground truth labels for a) enhancing tumor, b) non-enhancing tumor including necrosis and c) edema. 125 cases from the BraTS2019 validation dataset was used for evaluating the network's performance. Pre-processing steps included N4BiasCorrection to remove the RF inhomogeneity [25] and intensity normalization to zero-mean and unit variance.

2.1.2 Survival Analysis

259 HGG subjects were provided for the BraTS2019 Survival challenge. Information regarding age, survival days and resection status (Gross Total Resection (GTR), Subtotal Resection (STR) or not available (NA)) were also provided. 210 subjects out of 259 subjects were selected, as the other 49 subjects were either alive or survival days were not available (NA). The results were evaluated on 29 GTR subjects from the validation dataset. Pre-processing steps were similar to the segmentation task including N4BiasCorrection and intensity normalization.

2.2 Network Architecture

Brain tumors contain a complex structure of sub-components and are challenging for automated tumor segmentation. Specifically, the appearance of enhancing tumor and non-enhancing tumor are often different between HGG and LGG. In order to simplify the complex segmentation problem, first a simple convolutional neural network (CNN) was developed to classify HGG and LGG cases. Next, to maintain consistency with the output results provided by the BraTS challenge, separate networks were trained to recognize and predict whole tumor (consisting of enhancing + non-enhancing + necrosis + edema; Whole-net), tumor core (consisting of enhancing + non-enhancing + necrosis; Core-net), and enhancing tumor (enhance-net) as binary features using 3D Dense UNets. Each of these three networks were designed separately for HGG and LGG cases.

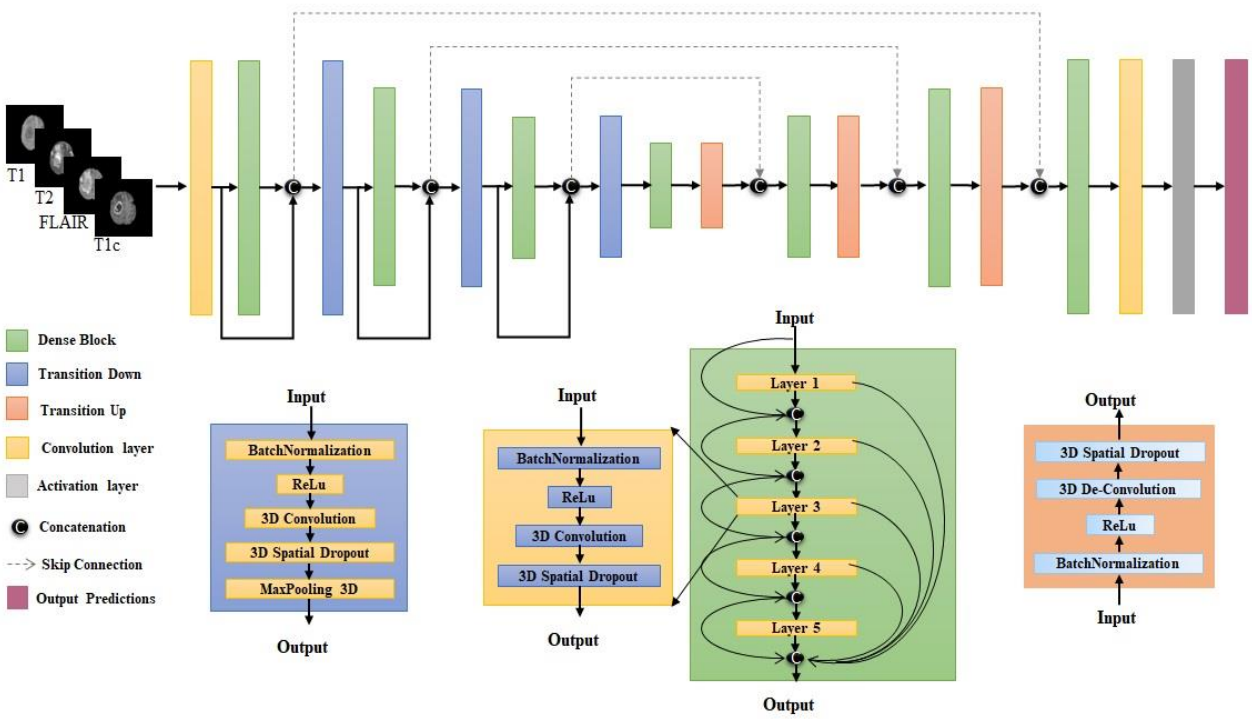

Figure 1: Network Architecture

The networks were designed to predict the local structures for tumors and sub-components in multi-parametric brain MR images. The architecture of each network is shown in Fig 1. On the encoder side of the network, multi-parametric images passed through initial convolution to generate 64 feature maps to be used in

subsequent dense blocks. Each dense block consisted of five layers as shown in Fig 1. Each layer included four sublayers , BatchNormalization, Rectified Linear Unit (ReLu), 3D Convolution and 3D Spatial dropout that were connected sequentially. At each layer, the input was used to generate k feature maps (referred to as growth rate and set to 16) that were subsequently concatenated to the next input layer. The next layer was then applied to create another k feature maps. To generate the final dense block output, inputs from each layer were concatenated with the output of the last layer. At the end of each dense block, the input to the dense block was also concatenated to the output of that dense block. The output of each dense block followed a skip connection to the adjacent decoder . In addition, each dense block output went through a transition down block until the bottle neck block. With this connecting pattern, all feature maps were reused such that every layer in the architecture received a direct supervision signal [26]. On the decoder side, a transition up block preceded each dense block until the final convolution layer followed by a sigmoid activation layer. Two techniques were used to circumvent the problem of maintaining a high number of convolution layers. A) The bottle neck block (Dense block 4 in figure 1 was used such that if the total number of feature maps from a layer exceeded the initial number of convolution maps (i.e. 64) then it was reduced to $1/4^{th}$ of the total generated feature maps in that layer. B) A compression factor of 0.75 was used to reduce the total number of feature maps after every block in the architecture. In addition, due to the large number of high resolution feature maps, a patch based 3D Dense-Unet approach was implemented where higher resolution information was passed through the standard skip connections.

2.3 Training

2.3.1 Brain Tumor Segmentation

Three groups of three Dense-UNets were designed. Each group consisted of networks designed to segment whole tumor, tumor core and enhancing tumor. Group 1 was trained using the dice-coefficient [27] as the loss function while group 2 was trained with binary focal loss [28] as the loss function. Group 3 consisted of 2 sub-groups namely, HGG group and LGG group. As the first step of group 3, a simple convolutional neural network was developed to separate HGG from LGG. The networks in the HGG group were trained using HGG cases only and the networks in the LGG group used the LGG cases only.

All 335 cases from the BraTS2019 dataset were used for training the networks from group 1 and group 2. 259 HGG cases were used to train the networks from the HGG group and 76 LGG cases were used to train the networks from the LGG group. 75% overlapping patches were extracted from the multi-parametric brain MR images that had at least one non-zero pixel on the corresponding ground truth patch. 20% of the extracted patches were used for in-training validation. Data augmentation steps included horizontal flipping, vertical flipping, random rotation, and translational rotation. Down sampled data (128x128x128) was also provided in the training as an additional data augmentation step. To circumvent the problem of data leakage, no patch from the same subject was mixed between training and in-training validation [29, 30].

Dice loss: The Dice co-efficient determines the amount of spatial overlap between the ground truth segmentation (X) and the network segmentation (Y)

$$dice\ loss = \frac{2|X_1 \cap Y_1|}{|X_1| + |Y_1|}$$

Focal loss: The focal loss was designed to address the problem of class imbalance between the foreground and background classes during training.

$$p_t = \begin{cases} p & if\ y = 1 \\ 1 - p & otherwise \end{cases} \quad 'y'\ specifies\ the\ ground\ truth\ and\ 'p' \in [0,1] is\ the\ model's\ predictio$$

$$Focal\ Loss = FL(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log(p_t)$$

HGG & LGG Classifier: 335 cases from the BraTS2019 dataset including 259 HGG cases and 76 LGG cases were used for training the network. Data augmentation steps included horizontal flipping, vertical flipping, random and translational rotation. The dataset was randomly shuffled and split into 60% (155 HGG and 46 LGG) for training, 20% (52 HGG and 15 LGG) for in-training validation and 20% (52 HGG and 15 LGG) for testing.

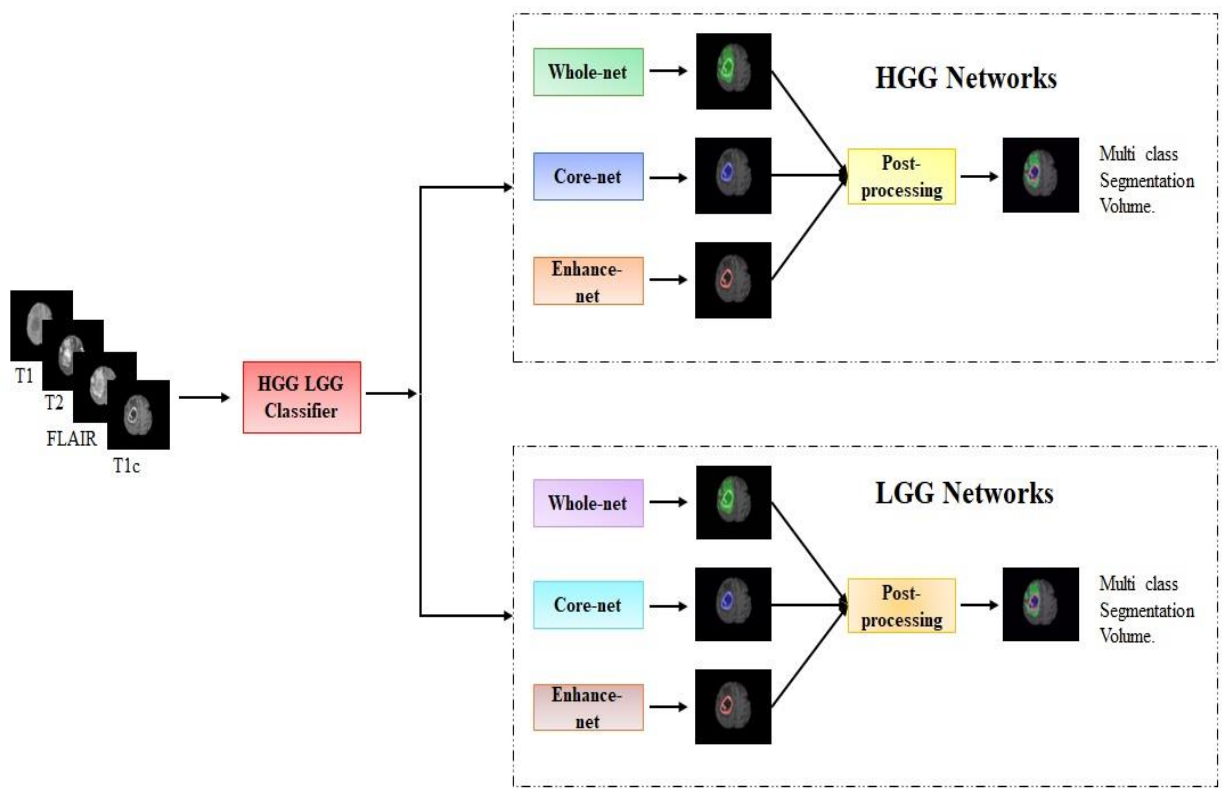

Figure 2: Overview of the 3rd approach using the HGG/LGG classifier.

2.3.3 Survival Analysis

Pyradiomics (a python package) was used to extract imaging (or radiomics) features [31]. Multi contrast MR images along with ground truth labels for each of the subcomponents were used to extract features from the BraTS2019 training dataset. For the BraTS2019 validation dataset, the segmented labels from the three 3D Dense UNets were used to extract features from the multi contrast MR images. 106 features were extracted for each MR imaging sequence from the tumor subcomponents (enhancing, edema and non-enhancing tumor with necrosis). Using these combinations, a total of 1272 features were extracted. Pywavelet, a python toolbox, was also used to extract wavelet based features [32]. 8 wavelet components for each MR image were extracted from level 1 of wavelet transform using coiflets (order=1) [33]. These 8 components (extracted from pywavelet toolbox) for the four MR imaging sequences in combination with 3 tumor subcomponents were used to extract 10,176 features using the pyradiomics package.

Table 1: Selected 17 features for survival analysis

Feature No.	MR Imaging sequence	Tumor Mask	Pyradiomics Feature Name	Wavelet Component	Feature Category
1	T1C	Necrosis	original_glszm_SizeZoneNonUniformity	NA	Imaging
2	T2w-FLAIR	Enhancing	original_firstorder_Skewness	NA	Imaging
3	T2w-FLAIR	Necrosis	original_glszm_LargeAreaLowGrayLevelEmphasis	NA	Imaging
4	T2w-FLAIR	Edema	original_glszm_SmallAreaLowGrayLevelEmphasis	NA	Imaging
5	T2w	Enhancing	original_firstorder_Kurtosis	NA	Imaging
6	T2w	Enhancing	original_firstorder_Maximum	NA	Imaging
7	T2w	Necrosis	original_shape_Maximum2DDiameterRow	NA	Imaging
8	T1w	Enhancing	original_firstorder_Range	NA	Imaging
9	T1w	Enhancing	original_firstorder_Skewness	NA	Imaging
10	T1w	Edema	original_firstorder_Minimum	NA	Imaging
11	T2w-FLAIR	Enhancing	original_firstorder_Skewness	Component 1	Wavelet- Imaging
12	T1C	Necrosis	original_glszm_HighGrayLevelZoneEmphasis	Component 2	Wavelet- Imaging
13	T1C	Edema	original_firstorder_Minimum	Component 2	Wavelet- Imaging
14	T1C	Edema	original_glszm_GrayLevelNonUniformity	Component 4	Wavelet- Imaging
15	T1C	Edema	original_firstorder_Minimum	Component 4	Wavelet- Imaging
16	T1w	Edema	original_glszm_SizeZoneNonUniformity	Component 5	Wavelet- Imaging
17	Age	NA	NA	NA	Non Imaging Feature

These imaging based features were combined with additional features including surface area, volume of tumor core, volume ofwhole tumor, ratio of tumor subcomponents to tumor core, ratio of tumor subcomponents to tumor volume, ratio of tumor core to tumor volume, and variance of enhancing tumor with T1C. The degree of angiogenesis was calculated by subtracting T1w and T1C in the tumor ROI, followed by a threshold of 50%. Finally, age and resection status were added to the feature set [34]. A total of 11,468 features were extracted combining the above features including imaging, texture and wavelet based features. Feature reduction was used to reduce the large number of features based on the feature importance determined by the gradient boost model. This reduced the feature space from 11,468 to 17 features. Table 1 shows a list of the selected 17 features for survival analysis. The 17 features were a combination of 10 Imaging features, 6 wavelet-imaging features and one non-imaging feature. All networks were trained using Tensorflow [35] backend engines, the Keras [36] python package and Pycharm IDEs on Tesla V100s and/or P40 NVIDIA-GPUs.

2.4 Testing

2.4.1 HGG & LGG classifier

The classifier was evaluated on 67 cases including 52 HGGs and 15 LGGs. To generalize the network's performance, a 3 fold cross-validation was also performed. While we used the HGG and LGG classifier as the first step for brain tumor segmentation, the BraTS2019 validation dataset did not include labels for HGG and LGG precluding evaluation of classification accuracy for this initial step on the validation dataset.

2.4.2 Brain Tumor Segmentation

All the networks were tested on 125 cases from the BraTS2019 validation dataset. Patches of size 32x32x32 were provided to the networks for testing. The prediction patches were then used to reconstruct a full segmentation volume. Each group was tested in 3 different ways including (a) non-overlapping patches, (b) 25% overlapping patches and (c) 50% overlapping patches. At the end of testing, each group produced 3 segmentation volumes for a particular label resulting in 9 segmentation volumes across the 3 groups. The 9 segmentation volumes were assigned with equal weights, averaged and thresholded at 0.5 for enhancing tumor, 0.8 for whole tumor and 0.7 for tumor core for every voxel. The same procedure was performed on whole tumor, tumor core and enhancing tumor labels. The ensembled output of whole tumor, tumor core and enhancing tumor were fused in a post-processing step that included the 3D connected components algorithm to improve prediction accuracy by removing false positives.

2.4.3 Survival Analysis

The network was evaluated on 29 cases from the BraTS2019 validation dataset. The Survival analysis was evaluated only for the GTR cases. The predicted overall survival (OS) task classified the subjects as long-survivors (greater than 15 months), mid- survivors (between 10 to 15 months) and short-survivors (less than 10 months).

2.4.3 Uncertainty task

Uncertainty masks were also created for the 125 cases from the BraTS2019 validation dataset. The predicted probability maps for a particular label from each group were assigned with equal weights, averaged and scaled from 0 to 100 such that 0 represents the most certain prediction and 100 represents the most uncertain. In this task, uncertain voxels were removed at multiple predetermined threshold points. The performance of the networks was assessed based on the dice score of the remaining voxels.

3. RESULTS

3.1 HGG & LGG Classifier

The network achieved a testing accuracy of 89% and an average cross-validation accuracy of 90%.

3.2 Brain Tumor Segmentation

This method achieved average dice scores of 0.95, 0.93 and 0.90 on WT, TC and ET on the training dataset (Table 1). It also achieved average Dice scores of 0.901, 0.844, 0.801 and sensitivity of 0.924, 0.846, 0.796 specificity of 0.991, 0.999, 0.998 for WT, TC and ET respectively on the 125 held-out cases along with Hausdorff distances of 7.60, 8.31 and 5.49 mm respectively.

Table 2: Mean dice scores on BraTS2019 datasets

	Whole Tumor	**Tumor Core**	**Enhancing Tumor**
BraTS2019 Training dataset	0.951	0.930	0.900
BraTS2019 Validation dataset	0.901	0.844	0.801

3.3 Survival task

The linear regression model achieved an accuracy of 55.2% and mean squared error (MSE) of 119244.6 ($\sqrt{109543} = 345\ days$) with median SE of 19881 and standard deviation of 227388.7 and spearman correlation co-efficient of 0.271.

3.4 Uncertainty task

The networks achieved average dice_AUC of 0.894, 0.887, 0.876 and FTP_AUC_ratio of 0.012, 0.061 and 0.046 for WT, TC and ET respectively on the BraTS2019 validation dataset.

Table 3: Uncertainty task results

	Dice AUC Whole tumor	**Dice AUC Tumor core**	**Dice AUC Enhancing tumor**	**FTP Ratio AUC Whole tumor**	**FTP Ratio AUC Tumor core**	**FTP Ratio AUC Enhancing**
BraTS2019 Validation dataset	0.894	0.887	0.876	0. 012	0. 061	0. 046

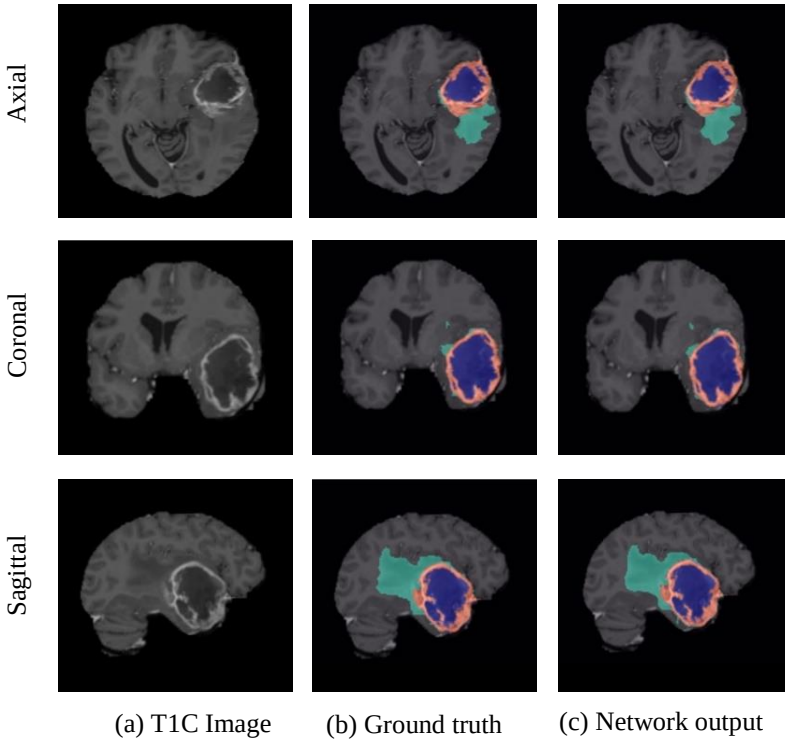

Figure 3: Example segmentation result for a High Grade Glioma (HGG) (a) A post-contrast image. (b) Ground truth (c) Network output. Color Code: Red = Enhancing tumor, Blue = tumor core (Enhancing tumor + non-enhancing tumor + necrosis), Green = Edema, Whole tumor = Green + Blue + Red.

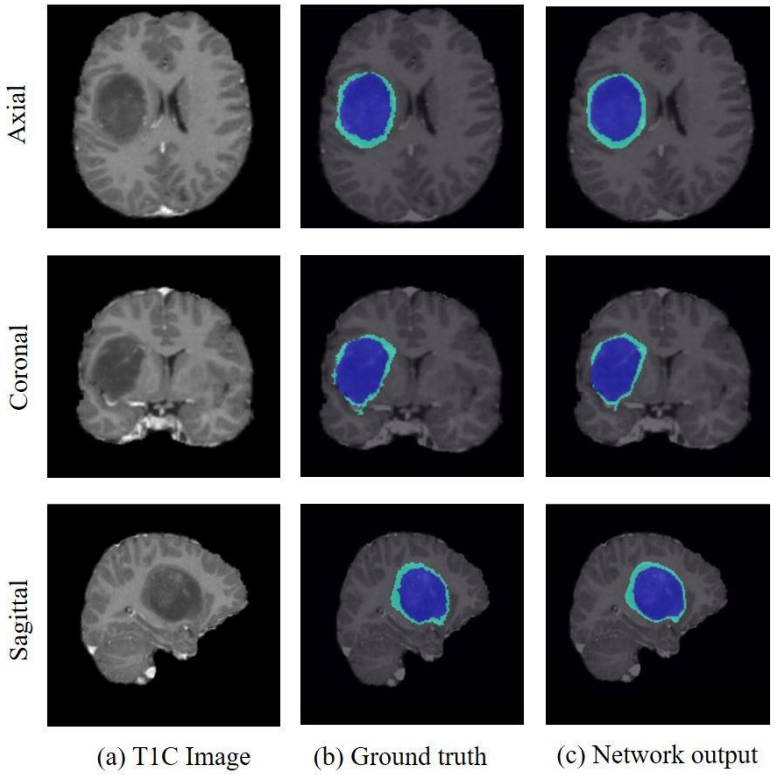

Figure 4: Example segmentation result for a Low Grade Glioma (HGG) (a) A post-contrast image. (b) Ground truth (c) Network output. Color Code: Red = Enhancing tumor, Blue = tumor core (Enhancing tumor + non-enhancing tumor + necrosis), Green = Edema, Whole tumor = Green + Blue + Red.

4. DISCUSSION

Accurate, efficient, and reliable tumor segmentation algorithms have the potential to improve management of GBM patients. Currently the vast majority of clinical and research efforts to evaluate response to therapy utilize gross geometric measurements. MRI-based glioma segmentation algorithms represent a method to reduce subjectivity and provide accurate quantitative analysis to assist in clinical decision making and improve patient outcomes. Although the underlying task can be simply stated as a voxel-level classification, a wide variety of automated and semi-automated tumor segmentation algorithms have been proposed. In this work, we developed a fully automated deep learning method to classify gliomas as high grade and low-grade, segment brain tumors into subcomponents and predict overall survival. This method was tested on the BraTS2019 validation dataset including 125 cases for the tumor segmentation task and 29 cases for the survival analysis task. A three group framework was used, providing several advantages compared to the currently existing methods. Using three binary segmentation networks for segmenting the tumor into its sub-components allowed us to use a simpler network for each task [14]. The networks were easier to train with reduced over-fitting [14]. Furthermore, since all three networks were trained separately as binary segmentation problems, misclassification was greatly reduced, as demonstrated by the results from the uncertainty task.

5. CONCLUSION

Fully automated convolutional neural networks were developed for segmenting brain tumors into their subcomponents. This algorithm reached high performance accuracy on the BraTS2019 validation dataset. High Dice scores, accuracy and speed of this network allows for large scale application in brain tumor segmentation. This method can be implemented in the clinical workflow for reliable tumor segmentation, survival prediction and for providing clinical guidance in diagnosis, surgical planning and follow up assessments.

REFERENCES

1. Ostrom, Q.T., et al., *CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015.* Neuro Oncol, 2018. **20**(suppl_4): p. iv1-iv86.
2. Havaei, M., et al., *Brain tumor segmentation with Deep Neural Networks.* Med Image Anal, 2017. **35**: p. 18-31.
3. Louis, D.N., et al., *The 2007 WHO classification of tumours of the central nervous system.* 2007. **114**(2): p. 97-109.
4. Kleihues, P. and W.K. Cavenee, *Pathology and genetics of tumours of the nervous system*. Vol. 2. 2000: International Agency for Research on Cancer.
5. Lacroix, M., et al., *A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival.* 2001. **95**(2): p. 190-198.
6. Hakin-Smith, V., et al., *Alternative lengthening of telomeres and survival in patients with glioblastoma multiforme.* 2003. **361**(9360): p. 836-838.
7. Johnson, D.R. and B.P.J.J.o.n.-o. O'Neill, *Glioblastoma survival in the United States before and during the temozolomide era.* 2012. **107**(2): p. 359-364.
8. Myronenko, A., *3D MRI brain tumor segmentation using autoencoder regularization.* MICCAI_BraTS_2018_proceedings_shortPapers, 2018.
9. Holland, E.C., *Progenitor cells and glioma formation.* Curr Opin Neurol, 2001. **14**(6): p. 683-8.
10. Menze, B.H., et al., *The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS).* IEEE Trans Med Imaging, 2015. **34**(10): p. 1993-2024.

11. V, S., Pankajakshan, Vinod *A deep learning architecture for brain tumor segmentation in MRI images.* 2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), Luton, 2017, 2017: p. 1-6.
12. Pei, L., et al., *Improved brain tumor segmentation by utilizing tumor growth model in longitudinal brain MRI.* Proc SPIE Int Soc Opt Eng, 2017. **10134**.
13. K. Kamnitsas , W.B., E. Ferrante, S. McDonagh, M. Sinclair, N. Pawlowski, M. Rajchl, M. Lee, B. Kainz, D. Rueckert, B. Glocker, *Ensembles of Multiple Models and Architectures for Robust Brain Tumour Segmentation.* BrainLes 2017, LNCS 10670, pp. 450–462, 2018. https://doi.org/10.1007/978-3-319-75238-9_38, 2018.
14. Wang, G.L., Wenqi & Ourselin, Sebastien & Vercauteren, Tom., *Automatic Brain Tumor Segmentation using Cascaded Anisotropic Convolutional Neural Networks.* MICCAI Brats Challenge 2017, 2017.
15. Funke, J., et al., *Candidate sampling for neuron reconstruction from anisotropic electron microscopy volumes.* Med Image Comput Comput Assist Interv, 2014. **17**(Pt 1): p. 17-24.
16. Soeda, A., et al., *The evidence of glioblastoma heterogeneity.* 2015. **5**: p. 7979.
17. Shboul, Z.A., et al. *Glioblastoma and survival prediction*. in *International MICCAI Brainlesion Workshop*. 2017. Springer.
18. Yang, D., et al., *Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma.* 2015. **42**(11): p. 6725-6735.
19. Lee, J., et al., *Texture feature ratios from relative CBV maps of perfusion MRI are associated with patient survival in glioblastoma.* 2016. **37**(1): p. 37-43.
20. Sanghani, P., et al., *Overall survival prediction in glioblastoma multiforme patients from volumetric, shape and texture features using machine learning.* 2018. **27**(4): p. 709-714.
21. Bakas, S., et al., *Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features.* Sci Data, 2017. **4**: p. 170117.
22. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R. T., et al. , *Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge*. . https://doi.org/10.17863/CAM.38755.
23. S. Bakas, H.A., 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.
24. S. Bakas, H.A., 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.
25. Tustison, N.J., et al., *Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements.* Neuroimage, 2014. **99**: p. 166-79.
26. Jegou, S., Drozdzal, Michal., Vazquez, David., Romero, Adriana., Bengio, Yoshua, *The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation.* 2017.
27. Taha, A.A. and A. Hanbury, *Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool.* BMC Med Imaging, 2015. **15**: p. 29.
28. Lin, T.Y., et al., *Focal loss for dense object detection.* IEEE Trans Pattern Anal Mach Intell, 2018.
29. Wegmayr V, A.S., Buhmann J, Nicholas Petrick; Kensaku Mori, Editor(s), *Classification of brain MRI with big data and deep 3D convolutional neural networks.* Published in SPIE Proceedings, Medical Imaging 2018: Computer-Aided Diagnosis, 2018. **1057501**.
30. Xinyang Feng, J.Y., Zachary C Lipton, Scott A Small, Frank A Provenzano, *Deep Learning on MRI Affirms the Prominence of the Hippocampal Formation in Alzheimer's Disease Classification.* bioRxiv, 2018. **2018:456277.**
31. Van Griethuysen, J.J., et al., *Computational radiomics system to decode the radiographic phenotype.* 2017. **77**(21): p. e104-e107.
32. Lee, G.R., et al., *PyWavelets: A Python package for wavelet analysis.* 2019. **4**(36): p. 1237.
33. Winger, L.L. and A.N.J.S.P.I.C. Venetsanopoulos, *Biorthogonal nearly coiflet wavelets for image compression.* 2001. **16**(9): p. 859-869.
34. Feng, X., N. Tustison, and C. Meyer. *Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features*. in *International MICCAI Brainlesion Workshop*. 2018. Springer.
35. Abadi, M., et al.,, *Tensorflow: A system for large-scale machine learning.* OSDI, 2016. **pp. 265-284**.
36. Charles, P.W.D. *Keras*. GitHub repository 2013.