International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Brain Tumor Segmentation Using Inception Modules

Asha K Kumaraswamy[1], Chandrashekar M. Patil[2]

[1]Department of Electronics and Communication, Vidyavardhaka College of Engineering, Mysuru, India
[1]*ashakk06[at]gmial.com*
[2]*patilcm[at]gmail.com*

Abstract: ***Among brain tumors, gliomas are the most common primary brain malignancies and they are very aggressive, thus leading to a very short life expectancy in their highest grade. Therefore, accurate and robust tumor segmentation is key stage for diagnosis, treatment planning and risk factor identification. Magnetic resonance imaging (MRI) is a widely used imaging technique to assess these tumors, but the large amount of MR images generated in clinical routine makes it difficult for manual segmentation. In addition, manual segmentation is time consuming, subjective and depends on the level of individual's experience.Therefore, automatic and reliable segmentation methods are required; however, the large spatial and structural variability among brain tumors make automatic segmentation a challenging problem. In this paper, we propose a novel automatic segmentation method based on Convolutional Neural Networks (CNN). Our method is a combination of U-net and Inception modules. Experiments with BraTS 2020 training set, our proposed method achieved average Dice scores of 0.902,0.797, 0.855 for whole tumor, enhancing tumor core and tumor core respectively. In this work we show that segmentation results can be improved by adding Inception modules to the U-net.***

Keywords: Automatic segmentation, Inception module, Magnetic resonance imaging, Brain tumor

1. Introduction

Cancer can be defined as the uncontrolled, unnatural growth and division of the cells in the body. Brain tumors are the consequence of abnormal growths and uncontrolled cells division in the brain. Depending on their initial origin, brain tumors can be considered as either primary brain tumors or metastatic brain tumors. In primary ones, the origin of the cells are brain tissue cells, where in metastatic, cells become cancerous at any other part of the body and spread into the brain.

Glioblastoma is one of the most aggressive human brain tumors[1] (Bleeker et al., 2012). The term glioma is a general term that is used to describe different types of gliomas ranging from low-grade gliomas (LGG) like astrocytomas and oligodendrogliomas to the high grade (grade IV) glioblastoma multiform (GBM), which is the most aggressive and the most common primary malignant brain tumor[2] . Surgery, chemotherapy and radiotherapy are the techniques used, usually in combination, to treat gliomas [3].

Magnetic resonance imaging (MRI) has the characteristics of significant soft tissue contrast and it can provide abundant physiological tissue information. In the clinical treatment of gliomas, MRI is commonly used method in radiology for analyzing phenotype (appearance and shape) and intrinsic heterogeneity of gliomas, since multimodal MRI scans, as shown in Figure 1 such as T1-weighted, contrast enhanced T1-weighted (T1Gd), T2-weighted, and Fluid Attenuation Inversion Recovery (FLAIR) images, provide complementary profiles for different sub-regions of gliomas. The general FLAIR sequence is suitable for observing edema tissues, T1 images are used for distinguishing healthy tissues and the T1Gd sequence is suitable for observing the active components of the tumor core, whereas T2 images are used to delineate the edema region which produces bright signal on the image.

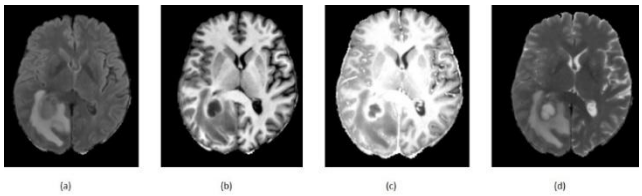

Figure 1: Multimodal MRI scans of glioma (a) FLAIR (b) T1 weighted (c) post contrast T1 weighted (d) T2 weighted.

Glioma segmentation is considered as the first step in the MRI analysis of glioma patients. Manual segmentation of glioma regions requires a lot of time and manpower. In addition, manual segmentation is often based on the brightness of the image, which is easily affected by the quality of the image generated and the individual's experience, thus leads to erroneous segmentation and segmentation of redundant areas. Therefore, in clinical practice, a fully automatic segmentation method with good segmentation accuracy for gliomas is needed. In recent years, deep learning has become the method of choice for complex medical image segmentation due to its high accuracy. Convolutional Neural Networks (CNNs) are artificial neural networks with multiple hidden convolutional layers between the input and output layers. Because of their non-linear properties they are capable of extracting higher level representative features [4](Gu et al., 2018). In this paper, we present a novel deep learning-based framework for segmentation of a brain tumor and its subregions from multimodal MRI scans using 2020 BraTS data set. Our results show that adding inception net increase the segmentation accuracy.

2. Related Works

Automated segmentation of brain tumor from MR images is a necessary step in medical image analysis. This section highlights some important works on brain tumor segmentation from multimodal MR images based on CNN.
Li Sun et.al [5] used ensembles of three different 3D CNN architectures for tumor segmentation. They used this method to reduce model bias and to boost the performance. For survival prediction, they extract 4,524 radiomic features from segmented tumor regions, then, a decision tree and cross validation are used to select potent features. Finally, a random forest model is trained to predict the overall survival of patients.

Daniel E Cahall et.al [6] propose an end-to-end brain tumor segmentation framework which utilizes a modified U-Net architecture with Inception modules to accomplish multi-scale feature extraction. They used Dice Similarity coefficient as loss function.

Guotia wang et. al.[7]show the use of cascade of CNNs for sequential segmentation of brain tumor and the subregions from multi-modal MRI, which decomposes the complex task of multi-class segmentation into three simpler binary segmentation tasks they also proposed 2.5D network structures with anisotropic convolution for the segmentation task.

Wentao Wu et. al [8], trained a deep convolutional neural network fusion support vector machine algorithm (DCNN-F-SVM). Their segmentation model was mainly divided into three stages. In the first stage, a deep convolutional neural network is trained to learn the mapping from image space to tumor marker space. In the second stage, the predicted labels obtained from the deep convolutional neural network training are input into the integrated support vector machine classifier together with the test images. In the third stage, a deep convolutional neural network and an integrated support vector machine are connected in series to train a deep classifier.

Havaei et al. (2016) [9] combined local and global 2D features extracted by a CNN for brain tumor segmentation. Although it outperformed the conventional discriminative methods, the 2D CNN only uses 2D features without considering the volumetric context. To incorporate 3D features, applying the 2D networks in axial, sagittal and coronal views and fusing their results has been proposed (McKinley et al., 2016; Li and Shen, 2017; Hu et al., 2018) [10]. However, the features employed by such a method are from cross-planes rather than entire 3D space.

DeepMedic (Kamnitsas et al., 2017b) [11] used a 3D CNN to exploit multi-scale volumetric features and further encoded spatial information with a fully connected Conditional Random Field (CRF). It achieved better segmentation performance than using 2D CNNs but has a relatively low inference efficiency due to the multi-scale image patch-based analysis

3. Materials and Methods

3.1 Data Gathering

We have used the BraTS 2020 dataset [12-16] to train our methods. The training set contained images from 369 patients, including 293 High-Grade Glioma (HGG) and 76 Low-Grade Glioma (LGG). Each patient's MRI data contained four MRI sequences: T2-weighted (T2), T1-weighted, T1 with gadolinium enhancing contrast (T1C), and Fluid-Attenuated Inversion Recovery (FLAIR) images. All the images were skull-striped and re-sampled to an isotropic $1mm^3$ resolution, and the four sequences of the same patient had been co-registered. The ground truth of segmentation mask was obtained by manual segmentation results given by experts. Annotations comprise the enhancing tumor (ET — label 4), the peritumoral edema (ED — label2), and the necrotic and non-enhancing tumor core (NCR/NET — label 1).

3.2 Data Preprocessing

We applied data augmentation by randomly rotating ± 15° with respect to X, Y, Z axes and translating along X and Y directions. All the original images were resized to a 128x128x128 matrix, while maintaining the aspect ratio. Intensity values are then normalized to zero mean and unit standard deviation.

3.3 Model Architecture

We propose a new architecture based on the 3D U-Net and convolution Inception module [17-18] in [19] author explored the combination of U-net and Res-Net. An inception module is added after each convolutional layer in the encoder part of the U-Net. Inception module has concatenated connection of three CNN blocks with kernel size of 1x1, 3x3 and 5x5 respectively. One of the important criteria of Inception architecture is their adaption of "Network in Network" approach as shown by Lin et. al [20] which increased the representational power of the neural networks. This had additionally saved them for computational bottlenecks by dimension reduction to 1×1 convolution. The purpose of Inception architecture was to reduce computational resource usage in highly accurate image classification using deep learning [21]. They had focused on finding an optimized position between the traditional way of increasing performance, which is to increase size and depth, and using sparsity in the layers based on the theoretical grounds given by Arora et. al [22]They focused on the approach to generate a correlation statistical analysis to generate groups of higher correlation to feed forward to the next layer. And they took the idea of multiscale analysis of visual information in their 1×1, 3×3 and 5×5 convolution layers. All of these layers then go through dimension reduction to end up in 1×1 convolutions [21].

Each layer on the contracting path of U-net, the height and width of the feature maps are halved and the depth is doubled until reaching the bottleneck i.e., the center of the "U." Conversely, on the expanding path, the height and width of the feature maps are doubled and the depth is

halved at each layer until reaching the output. Furthermore, each set of feature maps generated on the contracting path are concatenated to the corresponding feature maps on the expanding path. In the proposed architecture output convolution layer has 3 filters which are nothing but 3 tumor subtypes. Figure 2 shows the network architecture of our proposed method. A 3D resized volume is used as input. Each conv block performs a 3D convolution followed by batch normalization and leakyReLU activation with negative slope alpha value 0.3. We used Xavier normal initializer for weights. The network is implemented using Keras with a TensorFlow backend on a DGX-1 with 2, 16GB Tesla V1 GPU, batch size = 2, and learning rate = 10-4. The Dice coefficient between the network output and target mask was used as loss function. The network was trained for 200 epochs.

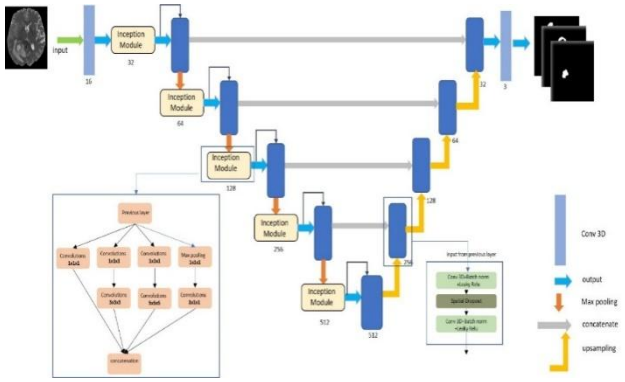

Figure 2: 3D U-net with Inception modules architecture considering a 128x128x128 volume as input. The number shown below each Inception module indicate total number of filters used. Last layer is having 3 filters with sigmoid activation and adam optimizer for back propagation.

3.4 Loss Function

Dice coefficient (DSC) is a F1- oriented statistic used to gauge the similarity of two sets. Given two sets A and B, the dice coefficient between them is given as follows [18]:

$$\text{Dice coefficient} = \frac{2|A \cap B|}{|A|+|B|}$$

In our case, A is the set that contains of all positive examples predicted by a model, and B is the set of all golden positive examples in the dataset.

$$\text{Dice coefficient} = \frac{2TP}{2TP+FN+FP} = \text{F1}$$

Where TP is True Positive, FN is False Negative and FP is False Positive

4. Results

We performed our experiments on BraTS 2020 dataset [12-16] which has images from 369 patients. 80% of the data is used for training, 10% is used for validation and 10% is used for testing. We evaluated the segmentation performance using the mean dice scores which is computed by treating the actual tumor labels as foreground and everything else as background.

For qualitative analysis, we present a sample of segmentation results in Figure 3 and in Figure 4. In both figures, for simplicity of visualization, only the FLAIR image is shown. The green, red and yellow colors show the peritumoral edema, necrotic and non-enhancing tumor core and enhancing tumor respectively. Figure 3 shows glioma subregions in axial, coronal and sagittal view and Figure 4 shows intra-tumoral structures.

Table 1 presents quantitative evaluations with the BraTS 2020 local test set. This method achieves average Dice scores of 0.902,0.797 and 0.856 for whole tumor, enhancing tumor and tumor core respectively. In Figure 5 we provide box plot of dice, sensitivity and specificity for each tumor type.

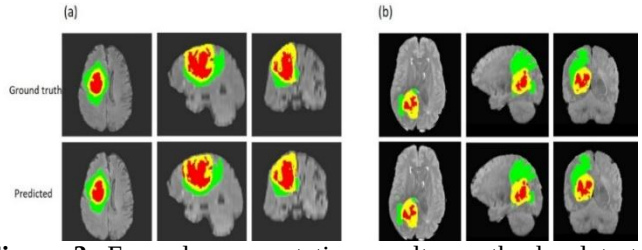

Figure 3: Example segmentation results on the local test subset of BraTS 2020 with ground truth and predicted labels overlaid over FLAIR MRI image in axial, sagittal and coronal slices. The whole tumor (WT) class includes all visible labels (a union of green, yellow and red labels), the tumor core (TC) class is a union of red and yellow, and the enhancing tumor core (ET) class is shown in yellow.

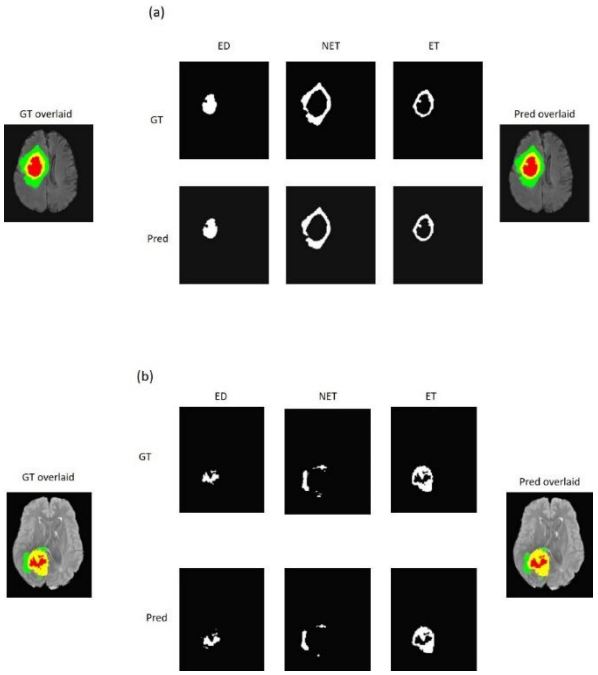

Figure 4: Intra-tumoral structure of same patients shown Figure 3. FLAIR overlaid with ground truth (GT) intra-tumoral structure is shown on the left in both sub-figures for easy visual analysis. Top row shows the GT segments for each intra-tumoral structure (abbreviations used are: ED, peritumoral edema; NET, necrotic and non-enhancing tumor core; ET, enhancing tumor). On the bottom row, the predicted (Pred) segments for each intra-tumoral structure are shown. The last image in each row is the combined predicted segments (i.e., ED, NET and ET) overlaid on FLAIR image.

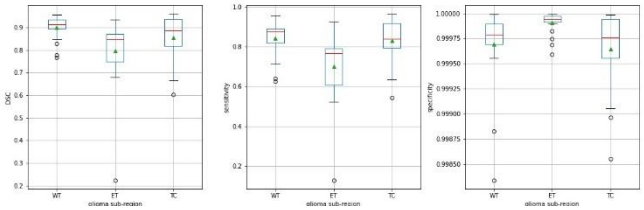

Figure 5: Box plot displaying the results for DSC, sensitivity and specificity. The x-axis is the glioma sub-region. The median value is denoted by the horizontal red line, and the mean is denoted by the green triangle. Abbreviations used are WT, Whole Tumor; TC, Tumor Core; and ET, Enhancing Tumor.

Table 1: Mean dice, sensitivity and specificity values of proposed model

	Whole Tumor	Enhancing tumor core	Tumor core
DSC	0.902	0.797	0.855
Sensitivity	0.841	0.701	0.831
Specificity	0.999	0.999	0.999

5. Discussion

In this work we introduced a novel method using the combination of U-Net architecture and Inception modules. Our results show that adding Inception modules in the encoder part of U-Net helps to achieve better score in segmenting tumor sub-region. We observed that the improvement in the segmentation accuracy is linked to multiple convolutional filters of different sizes employed in each Inception module. These filters with different kernel size can capture the contextual information at multiple scales during the learning process and can retain them.

We evaluate the performance of our proposed model using DSC and our results also demonstrate that introducing inception blocks in U-net encoder sections improves the intra-tumoral structures of glioma.

6. Conclusion

The model can be improved by adding more training data and with different augmenting methods like applying elastic deformation. In the future, we will explore different network architectures and training strategies to further improve our result.

References

[1] Bleeker, F. E., Molenaar, R. J., and Leenstra, S. (2012). Recent advances in the molecular understanding of glioblastoma. *J. Neuro Oncol.* 108, 11–27. doi: 10.1007/s11060-011-0793-0

[2] Deimling A. Gliomas. Recent Results in Cancer Research vol 171. Berlin: Springer; 2009.

[3] Stupp R. Malignant glioma: ESMO clinical recommendations for diagnosis, treatment and follow-up. Ann Oncol 2007; 18(Suppl 2):69-70.

[4] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., et al. (2018). Recent advances in convolutional neural networks. *Patt. Recogn.* 77, 354–377. doi: 10.1016/j.patcog.2017.10.013

[5] Sun, Li, et al. "Brain tumor segmentation and survival prediction using multimodal MRI scans with deep learning." *Frontiers in neuroscience* 13 (2019): 810.

[6] Cahall, Daniel E., et al. "Inception modules enhance brain tumor segmentation." *Frontiers in computational neuroscience* 13 (2019): 44.

[7] Wang, Guotai, et al. "Automatic brain tumor segmentation based on cascaded convolutional neural networks with uncertainty estimation." *Frontiers in computational neuroscience* 13 (2019): 56.

[8] Wu, Wentao, et al. "An Intelligent Diagnosis Method of Brain MRI Tumor Segmentation Using Deep Convolutional Neural Network and SVM Algorithm." *Computational and Mathematical Methods in Medicine* 2020 (2020).

[9] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., et al. (2017). Brain tumor segmentation with deep neural networks. *Med. Image Anal.* 35, 18–31. doi: 10.1016/j.media.2016.05.004

[10] McKinley, R., Wepfer, R., Gundersen, T., Wagner, F., Chan, A., Wiest, R., et al. (2016). "Nabla-net: A deep dag-like convolutional architecture for biomedical image segmentation," in *Int. MICCAI Brainlesion Work* (Athens), 119–128

[11] Kamnitsas, K., Ledig, C., Newcombe, V. F. J., Simpson, J. P., Kane, A. D., Menon, D. K., et al. (2017b). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* 36, 61–78. doi: 10.1016/j.media.2016.10.004

[12] 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

[13] 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

[14] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge", arXiv preprint arXiv:1811.02629 (2018)

[15] 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

[16] 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

[17] Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Lecture Notes in Computer Science Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015* (Munich), 234–241.

[18] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Las Vegas, NV).
[19] Kumaraswamy, Asha K., and Chandrashekar Patil. "A Cascaded U-net for Kidney and Tumor Segmentation from CT volumes." *Artificial Intelligence in Oncology* 2.1 (2020): 004-008.
[20] M. Lin, Q. Chen, And S. Yan, Network in network, arXiv preprint arXiv:1312.4400, (2013).
[21] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, And A. Rabinovich, Going deeper with convolutions, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
[22] S. Arora, A. Bhaskara, R. Ge, And T. MA, Provable bounds for learning some deep representations., in ICML, 2014, pp. 584–592.

Author Profile

Asha K Kumaraswamy is pursuing Ph.D at Visvesvaraya Technological University, Belagavi under the guidance of Dr. Chandrashekar M Patil, HOD ECE Dept. VidyaVardhaka College of Engineering, Mysuru.

Dr. Chandrashekar M Patil received Ph.D in Signal Processing from Visvesvaraya Technological University, Belagavi. Currently working as Department Head for ECE branch, Vidya Vardhaka College of Engineering, Mysuru.