A Multi-task Contextual Atrous Residual Network for Brain Tumor Detection & Segmentation

Ngan Le, Kashu Yamazaki, Kha Gia Quach, Dat Truong, Marios Savvides

Emails: {thile, kyamazak, tt032}@uark.edu; thihoanl@andrew.cmu.edu, msavvid@ri.cmu.edu

Abstract: We investigate that segmenting a brain tumor is facing to the imbalanced data problem where the number of pixels belonging to the background class (non tumor pixel) is much larger than the number of pixels belonging to the foreground class (tumor pixel). To address this problem, we propose a multitask network which is formed as a cascaded structure. Our model consists of two targets, i.e., (i) effectively differentiate the brain tumor regions and (ii) estimate the brain tumor mask. The first objective is performed by our proposed contextual brain tumor detection network, which plays a role of an attention gate and focuses on the region around brain tumor only while ignoring the far neighbor background which is less correlated to the tumor. Different from other existing object detection networks which process every pixel, our contextual brain tumor detection network only processes contextual regions around ground-truth instances and this strategy aims at producing meaningful regions proposals. The second objective is built upon a 3D atrous residual network and under an encode-decode network in order to effectively segment both large and small objects (brain tumor). Our 3D atrous residual network is designed with a skip connection to enables the gradient from the deep layers to be directly propagated to shallow layers, thus, features of different depths are preserved and used for refining each other. In order to incorporate larger contextual information from volume MRI data, our network utilizes the 3D atrous convolution with various kernel sizes, which enlarges the receptive field of filters.

Object Detection Networks

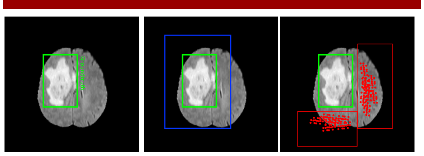


Fig. 1. From left to right: Given an image with groundtruth (left), positive window is defined as contextual around groundtruth (middle) and negative windows which are not covered in the positive windows (right). Each red circle is corresponding to center of each proposal.

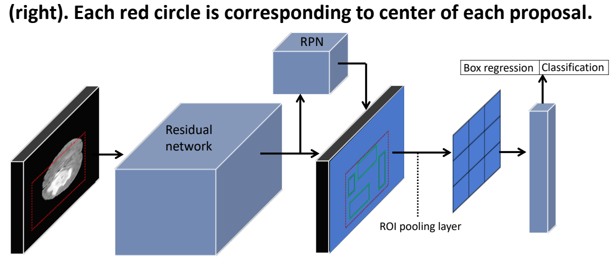


Fig 2. Brain Tumor Detection

Positive windows generation: the positive window is defined as a contextual region around the groundtruth and the size of the positive window is as twice as brain tumor groundtruth. An example of positive windows generation is given in Fig.1(middle) where the groundtruth is presented in green box and the positive window is presented in blue box.

Negative Windows Generation: Although the positive windows cover all the positive proposals, a significant portion of the brain, which is not lesion and maybe considered to be background, is not covered by them. In this step, we eliminate regions that does not contains any object, simply the almost black background will be ignored.

3D Atrous Residual Segmentation Network

The segmentation network takes the detected result from our previous detector component which is first extended on each direction an offset f = 6. The highest resolution and half resolution layers are designed with vanilla convolution whereas the quarter resolution layers are designed with 2x2 and 3x3 atrous convolutions in order to learn long-range representation.

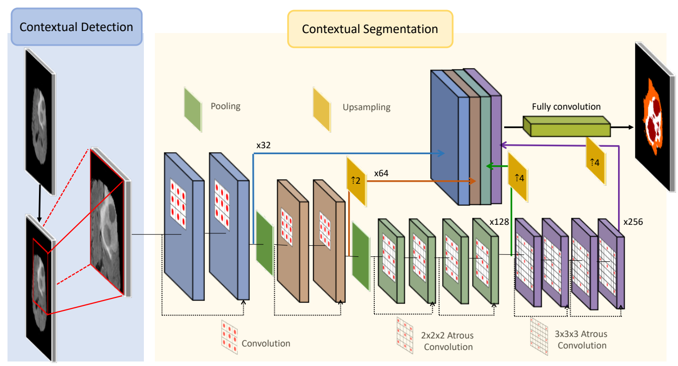
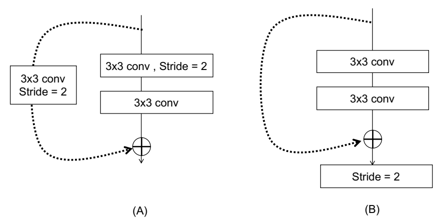


Fig 3. Brain Tumor Segmentation

In order to communicate between different resolution, features from different depth are usually combined by concatenate, residual connection. As shown in Fig.4 (A), the skip connection in Resnet contains a convolutional layer with stride 2 to deal with the inconsistency between the numbers of input channels and output channels. In our proposed network, we concatenate features of very different depths to the final output as shown in Fig.4 (B). By concatenating features from different scales, the semantic meaning of features are also preserved throughout the whole network.



Results

Table 1. Evaluation on online testing set of BRATS 18, BRATS 17, BRATS 15

Metrics	BRATS 2018			BRATS 2017			BRATS 2015		
	WT	TC	ET	WT	TC	ET	WT	TC	ET
Dice score	86.4	82.5	78.2	85.9	82.2	74.2	90.4	82.3	72.4
Sensitivity-TPVF	95.4	88.0	82.8	95.8	85.8	79.3	92.1	83.4	76.2
Specificity-TNVF	98.7	99.6	99.8	98.5	99.6	99.8	91.3	82.2	73.1
Precision	87.8	86.9	84.9	86.7	85.3	78.3	88.4	86.5	74.3
Hauf	4.7	6.3	5.2	8.8	10.1	11.2	5.1	12.8	6.8
ASSD	1.03	1.93	2.48	1.06	2.23	2.94	0.95	2.35	3.46

Table 2. Comparison on BRATS 15 online testing sets

Methods	Dice Score			Sensitivity		
	WT	TC	ET	WT	TC	ET
Pereira et al [20]	78.0	65.0	70.0	-	-	-
Pavel et al [19]	83.0	75.0	77.0	-	-	-
Chang et al [35]	87.0	81.0	72.0	-	-	-
Deep Medic [21]	89.6	75.4	71.8	90.3	73.0	73.0
DMRes [22]	**89.8**	75.0	72.0	89.1	72.1	72 .5
Improved Unet [36]	85.0	74.0	64.0	**91.0**	73.0	72.0
DRLS [7]	88.0	**82.0**	**73.0**	**91.0**	76.0	**78.0**
FSENet [37]	85.0	72.0	61.0	86.0	68.0	63.0
Multi_task [38]	87.0	75.0	65.0	89.0	**85.0**	63.0
Our	**90.4**	**82.3**	**72.4**	**92.0**	**83.4**	**76.7**

Table 3. Ablation study on local BRATS 18 validation set

	Dice			Hauf		
	WT	TC	ET	WT	TC	ET
Without Detection, Only 3D Atrous Residual Network	90.25	85.78	79.59	3.81	6.14	2.98
Contextual Detection & 3D Atrous Residual Network	90.95	88.88	81.40	3.73	5.92	2.70