Kidney and Tumor Segmentation Using Modified 3D Mask RCNN

Cong Chen[1], Longfei Ma[2], Yan Jia[1] and Panli Zuo[1]

[1] Huiying Medical Technology Co., Ltd, Room A206, B2, Dongsheng Science and Technology Park, HaiDian District, Beijing City, 100192, People's Republic of China
[2] Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China
chencong@huiyihuiying.com

Abstract. Detection of kidney tumors and accurate evaluation of their size are crucial for tracking cancer progression. Automating 3D volume detection and segmentation can improve workflow as well as patient care. We adapt the state of the art architecture for 2D object detection and segmentation, Mask RCNN, to handle 3D images and employ it along with U-net to detect and segment kidney and kidney tumor from CT scans. We report on competitive results for the kidney segmentation and kidney tumor segmentation on the 2019 Kidney Tumor Segmentation Challenge data set.

Keywords: Kidney Tumor Segmentation, 2019 Kidney Tumor Segmentation Challenge, Mask RCNN.

1 Method

1.1 Architecture

Mask RCNN is a 2-stage object detector, including Region Proposal Network (RPN) followed by Region based Convolutional Neural Network (RCNN) and a segmentation model (MASK). We modify the 2D implementation of Mask RCNN [1] to handle 3D images and to account for small object detection. The modified 3D Mask RCNN is used to detect and segment kidney and tumor simultaneously from the challenge data. U-Net is further used to refine the segmentation results of kidney and tumor, and output the optimal results.

The 3D Mask RCNN is composed of four parts: backbone, RPN, RCNN for classification and bounding box regression and another CNN for pixel segmentation of objects, which we refer to as MASK.

The 3D U-net consists of a contracting path and an expansive path [2]. The contracting path follows the typical architecture of a convolutional network. One important modification in our architecture is that in the upsampling part we have also a large number of feature channels, which allow the network to propagate context information to higher resolution layers.

1.2 Training

3D Mask RCNN and 3D U-net are fully trainable end to end. Nonetheless, convergence is faster when training the backbone and RPN together first, and then training only the second stage heads. Focal loss [3] and Intersection over Union (IoU) loss improve results in the class and MASK heads respectively. Training both segmentation and detection tasks simultaneously improves detection rate [4]. We use dropout and heavy augmentation during training to avoid over fitting.

1.3 Inference

We scan each image with overlapping sliding windows. Overlapping boxes are filtered using Non Max Suppression (NMS). To reduce False Positive (FP)s, we keep only boxes with a segmentation mask volume > 0. We use our in house kidney mask CAD to remove tumors detected outside of the kidneys.

2 Experiments

We tested our model on the 2019 Kidney Tumor Segmentation Challenge data set, which includes a 210 patients of training data, and a 90 patients of testing data. Segmentation overlap is measured with the Dice Similarity Coefficient (DSC). 190 patients from training data were used to train our model, and the other 20 patients from training data were used to validate the model. We achieve an average DSC of 96% for kidney segmentation and 43% for kidney tumor segmentation.

References

1. Waleed Abdulla. Mask r-cnn for object detection and instance segmentation on keras and tensorflow. https://github.com/matterport/Mask_RCNN, 2017.
2. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. Convolutional Networks for Biomedical Image Segmentation, 2015
3. Tsung-Yi Lin, Priya Goyal, Ross B. Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. CoRR, abs/1708.02002, 2017.
4. Paul F. Jaeger, Simon A. A. Kohl, Sebastian Bickelhaupt, Fabian Isensee, Tristan Anselm Kuder, Heinz-Peter Schlemmer, and Klaus H. Maier-Hein. Retina u-net: Embarrassingly simple exploitation of segmentation supervision for medical object detection. CoRR, abs/1811.08661, 2018.