Segmentation of CT Kidney and kidney tumor by MDD-Net

Ball Chen

Abstract. Accurate segmentation of kidney and kidney tumor is an important step for treatment. Due to the wide variety in kidney and kidney tumor morphology, it's really a challenging task. In this work, we propose the Multi-level double-dimension Network to automatically segmentat kidney and kidney tumor. We select the modified FPN as backbone and aggregate different scale information from multi levels to make the final prediction. In the KiTS 2019, we use 170 CT scans for training and the remaining 40 CT scans are used to evaluate the model. At the time of submission, we obtained the best result by averaging multiple models.
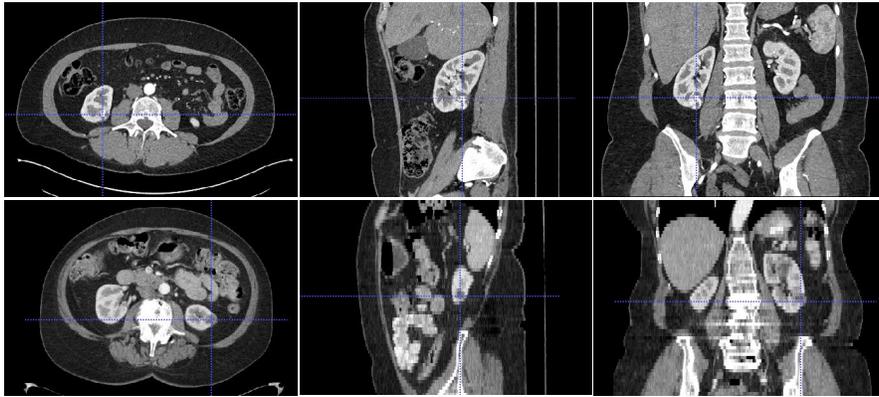
1 Introduction

Fig. 1. Observe the kidney in the CT scans from three axes(xy, yz, xz). In the first row, the spacing rate in x and y axes is 0.91mm and that in z axes is 0.5mm. In the second row, the spacing rate in x and y axes is 0.68mm and that in z axes is 5mm. It's obvious that when the gap between different axes is too large, the kidney and tumor are too ambiguous to identify accurately.

There are more than 400,000 new cases of kidney cancer each year, and surgery is its most common treatment. Due to the wide variety in kidney and

kidney tumor morphology, accurate segmentation of them is still a challenging task. Besides, images from collected from different devices are not consistent in resolution especially the spacing rate in z direction, causing great difficulty. In this task, the spacing rate in x and y direction both range 0.5mm to 1mm, but the spacing rate in z direction ranges 0.5mm to 5mm. As shown in Fig.1, when the gap between different direction about spacing rate is too large, it will cause damage to the structure of kidney, and 3D convolution could't fully extract its complete information. However, just operating convolution on each 2D slice can only extract local regional information, lack of structural information from front and back frames. Based on above, we propose the Multi-level double-dimension Network to accurately segment the kidney and kidney tumor.

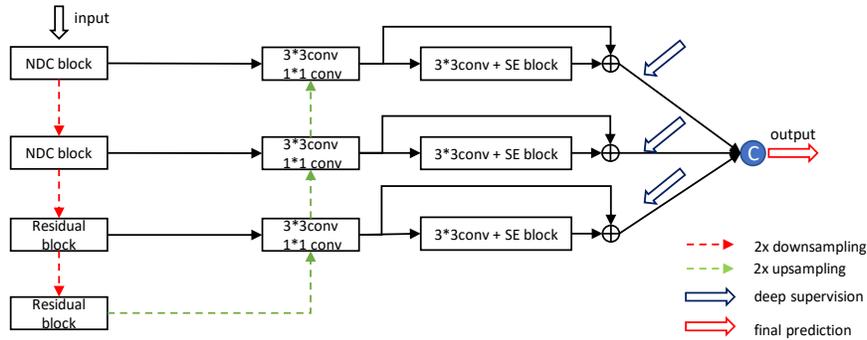## 2 Method

2.1 Network Architecture

Fig. 2. The Network Architecture.

Our network architecture is illustrated in Fig.2, it looks like a FPN structure. In the beginning, each nested dilation convolution block (NDC) contains three convolution layers with dilation rate 1, 2 and 4, which help acquire larger receptive field in shallow layers. In the decoder path, each branch goes through a 3x3 convolution layer and a Sequeeze-and-Excitation block [1] and we extra add deep supervision in each branch. After that, we aggregate the feature maps from top three levels in decoder path to make final prediction.

2.2 loss

In this task, we select the cross-entropy loss to train the kidney and tumor. We extra utilize the dice loss to supervise the kidney tumor because there are lots of small tumors in datasets.

3   Experiments

Our networks were implemented in Keras using the Tensorflow backend. Adaptive moment estimation was used as an optimizer with an initial learning rate of 0.0001, a momentum of 0.9, and a weight decay of 0.00001. Training was implemented on an Tesla P40 for 50 epochs. We did not use a dropout but rather L2 regularization and BN for the whole network structure. We cropped 160 x 160 x 64 patches as inputs accoding to the annotations. The networks were trained from scratch with a batch size of 6.

We randomly select 40 patients as the validation dataset and the others are all used to train the model. It's obvious that our model can segment the kidney and kidney tumor accurately.
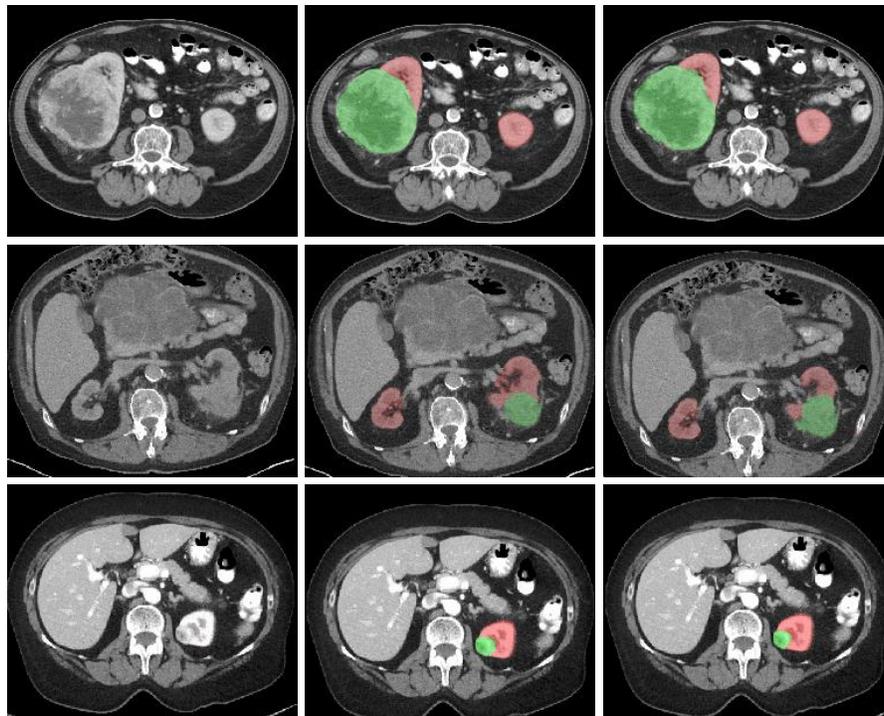

Fig. 3. Examples of segmentation results.

References

1. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132–7141 (2018)