Two stages kidney and tumor segmentation(kits2019)

Zhang Yiwen[1]

[1] School of Biomedical Engineering, Southern Medical University, Guangzhou Guangdong 510515, China

Abstract. In order to accurately segment kidney tumors, this paper uses a two-stage segmentation/ Stage one performs a coarse step positioning of the kidney. Stage two performs an accurate segmentation of the kidneys and tumors.

Keywords: Kidney tumor, Segmentation.

1 Stage one(coarse kidney segmentation)

1.1 Data preprocessing

The body mask is obtained by otsu and morphological algorithm. Using CT value -700 and body mask, I can get the air mask inside the human body. Firstly, I multiply the original data with body mask to Remove irrelevant information such as beds. Secondly, minimum circumscribed cuboid of the body mask is cropped followed by 50 pixels are cut inward in each of the upper, lower, left and right directions in transverse plane. According to statistics, this will not cut out the kidney area. Thirdly, I can remove the lungs through the air mask. Specifically, I traverse each slice from the bottom up until $S_a^i / S_b^i >= 0.1$(S_a^i, S_b^i are the air and body area of the slice i respectively) establishs. The slices above i will be removed. In order to finally restore the kidney to its original position, the coordinates of the clipping region are recorded in an excel table. Fourthly, the sizes in transverse plane of all individuals are resized to 288*416. If the data is used in a 3d model, the slice thickness is interpolated to 1 mm. This is not required for 2d model. The nearest neighbor interpolation is used for all interpolation. Lastly, the CT values are clipped to -400 to 600 and then normalize to 0-1.

1.2 Train 2D Model

I only trained a model on transverse plane. Only the slices containing kidney and the 20 slices above and below them respectively are used as training sets. I used BiSeNet[1] model with dice+bce loss function to only segment kidney. Data augmentations include random scaling, elastic deformation, add noise. Details are as follows: the input size is 288*416, batchsize is 8, Optimizer is Adam with max learning rate 0.0001, warming up and decay strategy were used, trained 30 epochs.

[1] https://github.com/ooooverflow/BiSeNet

1.3 Train 3D Model

Like the training set of 2D model, only the slices containing kidney and the 20 slices above and below them respectively are used. The difference is that the slice thickness is interpolated to 1 mm. I used 3D attention-unet[2] model with dice+bce loss function to only segment kidney. Data augmentations include random crop, random scaling, add noise. Details are as follows: the input size is 96*128*256(96*288*416 when test phase), batchsize is 2, Optimizer is Adam with max learning rate 0.0001, warming up and decay strategy were used, trained 200 epochs.

1.4 Post-processing and analysis

Because the model was only trained on data containing kidneys, there were some false positive when it tested on the whole human body. I found 2D model was not as precise as 3D models in Segmentation of renal region but false positive were significantly less. In fact, the accuracy of 2D model is completely satisfied with the coarse segmentation of stage one ,besides its false positives are less and more convenient for post- processing. So I only used the results of the 2D model for post-processing at last. The Post-processing mainly includes two steps: One is removing the targets with connected area less than 20000 pixels and cropping a bounding box of top two targets in Z-axis direction. The bounding box is the minimum circumscribed cuboid of the top two targets. Because the lung has been removed by the pretreatment process, the kidney is always predicted to be in the upper position. Through statistics, I found that false positives of 2D models usually appear near thighs; Two is removing the targets with connected area less than 20000 pixels again and taking the first 2 targets with the largest connection area which is the two kidneys we want. If there is only one target, the patient probably has only one kidney.

2 Stage two(Fine segmentation of kidney and tumor)

2.1 Data preprocessing

In the training phase, Kidney blocks can be obtained directly from labels. I cropped the each kidney with its bounding boxes which is got by expanded the 6 sides of minimum circumscribed cuboid of the kidney by 10 pixels respectively. In the testing phase , Coarse kidney mask can be obtained by stage one. The bounding boxes of each kidney can also be obtained. The size in transverse plane of each kidney block resized to 128*128. Just like stage one, if the data is used in a 3d model, the slice thickness is interpolated to 1 mm. Lastly, the CT values are clipped to -400 to 600 and then normalize to 0-1.

[2] https://github.com/ozan-oktay/Attention-Gated-Networks

2.2 Train 2D Model

I only trained a model on transverse plane. Only the slices containing tumor and the 7 slices above and below them respectively are used as training sets. I used BiSeNet model with dice+bce loss function to segment good kidney, tumor and background. Data augmentations include random scaling, elastic deformation, add noise, flip horizontally and vertically. Details are as follows: the input size is 128*128, batchsize is 5, Optimizer is Adam with max learning rate 0.0001, warming up and decay strategy were used, trained 300 epochs.

2.3 Train 3D Model

All preprocessed data are used as training sets. And the slice thickness is interpolated to 1 mm. I used 3D attention-unet model with dice+bce loss function to segment good kidney, tumor and background. Data augmentations include random scaling, add noise, flip horizontally and vertically. Details are as follows: the input size is 64*128*128, batchsize is 3, Optimizer is Adam with max learning rate 0.0001, warming up and decay strategy were used, trained 400 epochs.

2.4 Post-processing and analysis

The prediction results of the 3D model were significantly finer in morphology than those of the 2D model, however 3D model was easy to miss small tumors. Some 3D model predictions showed no tumor but I found that every patient in the train set had a tumor. The prediction results of 2D model had more false positives and tumor morphology is discontinuous between slices but every patient can be detected at least one tumor. So I decided that the prediction result of the 3D model is taken as the main body, if the prediction results show that there is no tumor, the tumor region predicted by the 2D network is added to the 3D prediction result.