Kits2019 Challeges:Brief Descriptions of the Algorithm and Process

Songxiong Wu[1] , Bin Huang[1], Zongyou Cai[1],

[1] School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China
Songxiong Wu,503405075@qq.com

Abstract. Chronic kidney disease affects mang people around the world. Computed tomography (CT) is a widely used imaging modality for kidney disease diagnosis and quantification. However, automatic pathological kidney segmentation is still a challenging task due to large variations in contrast phase, scanning range, pathology, and position in the abdomen, etc. In this work, we propose to combine different Window wide window position as a multi-channel input and Unet, for robust kidney or kidney tumors segmentation.

Keywords: Unet, kidney segmentation,multi-channel

1 Brief Introduction

First, we directly used the Unet (Figure 1) to perform kidney segmentation, but the results were not ideal. Based on the unet, we modify the last downsampling module of contracting path (left side) into the dialated convolution architecture (Figure 2). The other parts of unet were not modified, and then the kidney segmentation results were relatively ideal. We decided to use this network to segment the kidney. Based on the results of kidney segmentation, we further used the kidney segmentation network to segment the tumor. The result was very poor. Therefore, we changed the dialated convolution architecture to the ASPP structure (Fig. 3) and adopted different Window wide window position as a multi-channel input for further segmentation of kidney tumors

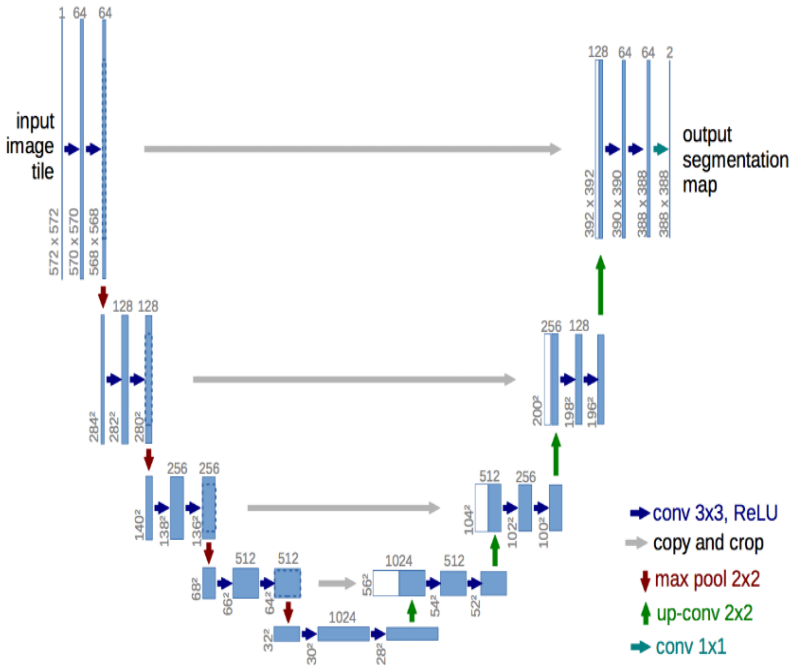

Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

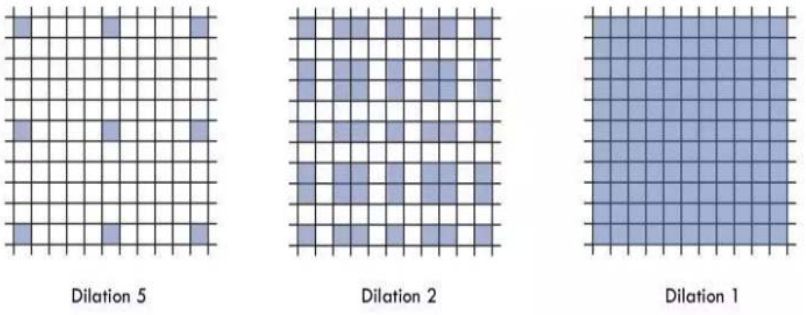

Fig. 2. Dialated convolution architecture

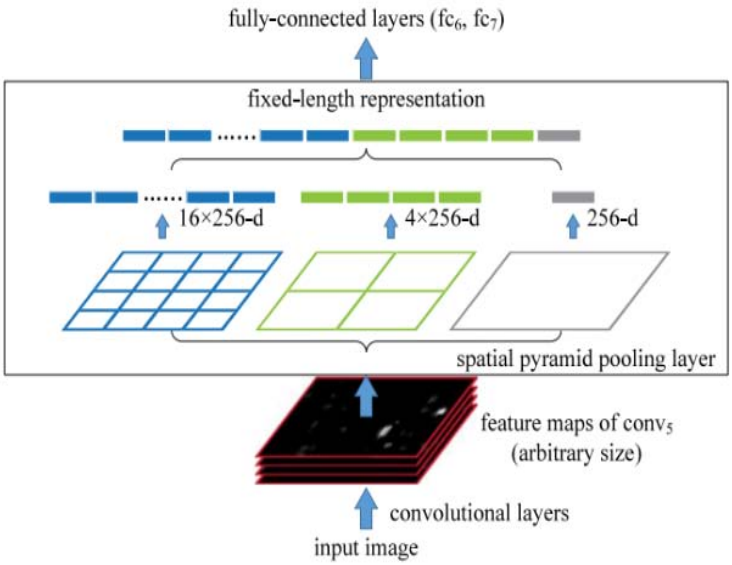

Fig. 3. Spatial Pyramid Pooling architecture

2 Detailed introduction

2.1 Kidney Segmentation

2.1.1 Network Architecture

The network architecture is illustrated in Figure 1. It consists of a contracting path (left side) and an expansive path (right side). The contracting path follows the typical architecture of a convolutional network. It consists of the repeated application of two 3x3 convolutions (unpadded convolutions), each followed by a rectied linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. At each downsampling step we double the number of feature channels. Every step in the expansive path consists of an upsampling of the feature map followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In total the network has 23 convolutional layers. To allow a seamless tiling of the output segmentation map (see Figure 2), it is important to select the input tile

size such that all 2x2 max-pooling operations are applied to a layer with an even x- and y-size. Based on the unet, we modify the last downsampling module of contracting path (left side) into the dialated convolution architecture.

2.1.2 Training:

The input images and corresponding segmentation maps are used to train the network by pytorch. The size of the output image and input image are equal. we split the dataset into three parts, with the last 10 patients as the test set, and split the training set and validation set in 5:1 ratio in the remaining patients. We first divided the 3D images of each patient into 2D images by slices, and marked whether each layer had a kidney region according to the gold standard. Our model used the cross entropy loss function, Adam optimizer.The learning rate was set to 1e-5, and the bacth size was set to 4, for example 3 maps with a kidney and 1 map without a kidney region were dropped for training each time. Dice index is used to evaluate the quality of the model with validation set.

2.2 Kidney tumors Segmentation:

We first selected the highest Dice score(0.96) for kidney segmentation. Based on the results of kidney segmentation, we further performed kidney tumor segmentation. We modified the dialated convolution architecture into a Spatial Pyramid Pooling architecture, and then we selected different window widths and window levels as multi-channel input according to the kidney and kidney tumors.The remaining steps were similar to those of kidney segmentation.

References

1. Ronneberger, Olaf , P. Fischer , and T. Brox . "U-Net: Convolutional Networks for Biomedical Image Segmentation." (2015).
2. He, Kaiming , et al. "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition."IEEE Transactions on Pattern Analysis & Machine Intelligence 37.9(2014) :1904-16.