Coarse-to-fine kidney and tumor segmentation with fully convolutional networks

Chen Shen[1], Chenglong Wang[1], Masahiro Oda[1], and Kensaku Mori[1,2]

[1] Nagoya University, Furo-cho, Chikusa-ku Nagoya, Japan
[2] Research Center for Medical Bigdata, National Institute of Informatics, Tokyo, Japan cshen@mori.m.is.nagoya-u.ac.jp

Abstract. Segmentation is one of the most important tasks in medical image analysis. With the development of deep leaning, fully convolutional networks (FCNs) have become the dominant approach for this task and their extension to 3D achieved considerable improvements for automated organ segmentation in volumetric imaging data. In this paper we demonstrate a coarse-to-fine segmentation method using FCNs for Kits 2019 challenge.

Keywords: kidney · tumor · coarse to fine

1 Introduction

Partial nephrectomy is a common treatment for kidney tumors. Due to the wide variety of tumors anatomical structures, relations between tumors morphology and surgical outcomes have gained a lot of research interests. Quantitative evaluation of kidney tumor can help physicians to have a better understand of tumor status. Manual segmentation of kidney and tumor is a time-consuming and expensive task. In this work, we used a cascade 3D fully convolutional network to automatically segment kidneys and tumors. By using deep-learning based method, kidneys and tumors can be precisely extracted from 3D CT scans. In this work, we used a coarse-to-fine segmentation strategy to precisely extract the kidneys and tumors.

2 Methods

We present the training schemes for this challenge in detail in the following. The baseline of network architecture we utilized is 3D U-Net [2] like FCN. 3D U-Net is one of the most famous network architecture in medical image segmentation which extended from U-Net [1]. Some modifications have made to improve the segmentation performance on the Kits 2019 challenge.

2.1 Pre-processing

Since the intensity value of CT volume has a large difference among the ct scanners, normalization makes a critical role in this work. The linear normalization
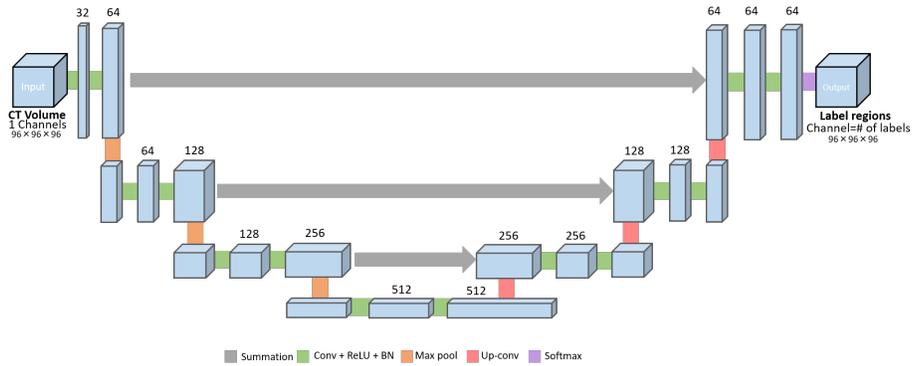
Fig. 1. The FCN architecture which is similar to 3D U-Net.

was utilized to adjust the intensity from -300 to 200. With this process, the tumor regions is more clear in the whole CT volume. In deep learning, it would work better if the resolution of the CT volumes is the same. We interpolate the CT volumes and ground truth into isotropic with is the same resolution as x axis. Cubic interpolation and nearest neighbourhood was adopt for CT volumes and ground truth respectively.

2.2   FCN architecture

With the improvement of convolutional networks architecture and GPU memory, we are able to train a larger number of annotated 3D medical CT volumes to improve the segmentation results. In this work, we utilize a FCN architecture as shown in fig. 1 which is similar to 3D U-Net, which employed an encoder-decoder structure. The input size of the network is constant, which is randomly cropping subvolumes from the training data. Hence, we can obtain a trained model, which can segment the full 3D volume through subvolume prediction in testing. We chose an input size of $96 \times 96 \times 96$ that allows the use of mini-batch size of six subvolumes sampled from different training patients. The skip connection is summation instead of concatenation in original 3D U-Net.

2.3   Coarse-to-fine segmentation

In the coarse phase, we combine the label of kidney and tumor as the same label. We used the binary segmentation to get the region of interests, which we could utilize as a mask for fine phase. The bounding box of the coarse segmentation result was generalized to focus the region around the kidneys. In the fine phase, the inputs is based on the bounding box we segmented in the coarse phase. Since the tumors volume sizes is much smaller than kidneys for most cases , the tumor regions may not always appear in the subvolumes during the random cropping. We crop on tumor region and randomly in turns during train.

In deep learning, the number of training data make a big difference on the network performance. On-the-fly data augmentation was utilized to artificially increase the dataset which effectively improve the robustness of the network and prevent the overfitting. For augmentation, we apply random 3D translation, rotations, elastic B-spline deformation together with gamma correction.

2.4   Weighted dice loss

The Dice loss function [3] which is inspired by Dice similarity coefficient (DSC) was utilized to calculate the amount of agreement between two overlapping binary regions. This loss function is effective to evaluate the segmentation performance between the prediction result and ground truth.

$$\mathcal{L}_m = -\frac{1}{K} \sum_{k=1}^{K} \left(\frac{2 \sum_i^N s_{i,k} g_{i,k}}{\sum_i^N s_{i,k}^2 + \sum_i^N g_{i,k}^2} \right). \tag{1}$$

Here, $s_{i,k} \in [0, \dots, 1]$ is the continuous values of the *softmax* 3D prediction maps for each class k of K and $g_{i,k} \in [0,1]$ represents the corresponding ground truth value at voxel i. In kidneys and tumors segmentation, the proportions of tumor is relatively small than the kidney. We balancing weight for loss function with the inverse of volume size of each classes.

2.5   Post-processing

For the post-processing, we utilized the results of coarse phase as a mask and reduce the false positive regions in fine segmentation. Nearest neighbourhood interpolation was adapted to all segmentation results in order to fit for their original resolution.

3   Experiments and Results

We divided the training dataset into training and validation with 199 and 20 volumes respectively. We utilized a NVIDIA Quadro P6000 with 24GB memory. We implement all models in the Keras with the TensorFlow backend. For optimization, we trained the network with the initial learning rate 0.0001 using the adaptive moment estimation (Adam) which based on stochastic gradient descent.

Some segmentation results are shown in fig. 2. For most cases, the segmentation results for kidneys and tumor are acceptable.

4   Discussion

In this paper we demonstrated a cascase kidneys and tumors segmentation. We introduced a coarse-to-fine segmentation method which using FCN similar to 3D U-Net [2].

(a) case 1   (b) case 2   (c) case 3
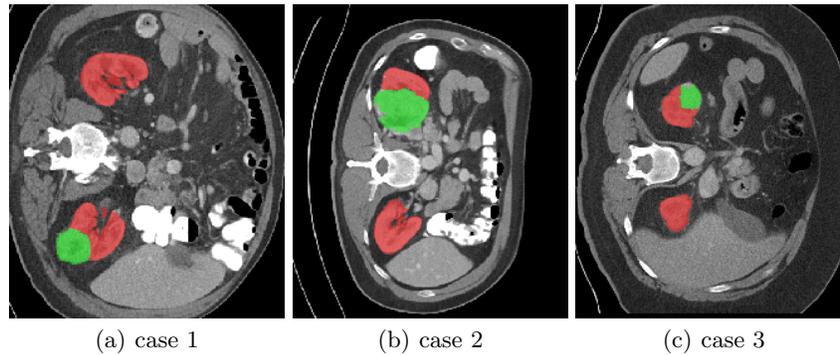
Fig. 2. The segmentation results for some cases.

References

1. O. Ronneberger, P. Fischer, T. Brox, " U-Net: Convolutional Networks for Biomedical Image Segmentation", International Conference on Medical Image Computing and Computer-Assisted Intervention, pp.234–241, 2015.
2. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, " 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation" International Conference on Medical Image Computing and Computer-Assisted Intervention, pp.424–432, 2016.
3. F. Milletari, N. Navab, S. Ahmadi, " V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation", 2016 Fourth International Conference on 3D Vision, pp.565–571, 2016.