View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Minnesota Digital Conservancy

Kidney and Kidney Tumor Segmentation Using Two-stage Convolutional Neural Network

Junghyun Lee*, Joonyoung Song*, Serin Yang*, Inhwa Han*, and Jong Chul Ye

Dept. of Bio and Brain Engineering
KAIST, Daejeon, Republic of Korea

{jh200.lee, songjy18, yangsr, inhwahan, jong.ye}@kaist.ac.kr

Abstract. Kidney tumor is typically diagnosed using computed tomography (CT) imaging by investigating geometric features of kidney tumor. For a reliable diagnosis and treatment planning, kidney tumor quantification is necessary. However, manual segmentation by human requires time and expertise. In addition, inter/intra variability of segmentation results can lead to suboptimal decision. In this study, we propose the two-stage segmentation method using 2.5D and 3D convolutional neural network for kidney and kidney tumor delineation. The two-stage model was trained with multi-task loss for pixel-wise cross-entropy loss function for segmentation task and mean square error function for regression task. Experimental results confirm that the proposed method effectively segments kidney and kidney tumor.

Keywords: Kidney Tumor, Two-stage Segmentation Scheme, Convolutional Neural Network.

1 Introduction

Kidney cancer is the seventh most common cancer with a high mortality rate with 175,000 in 2018 [1]. Computed tomography (CT) image is typically used for diagnosing kidney cancer and planning treatment. Specifically, morphological characteristics of a kidney tumor are important criteria, and quantification of geometric features of tumor is a prerequisite for a concrete decision. However, manual delineation by human requires a lot of time and effort. Even in manual segmentation of a kidney tumor, intra/inter-variability can arise. To overcome limitations of manual delineation, automatic segmentation of kidney and kidney tumor is often necessary. Automatic segmentation can save time and effort while segmentation results standardized by the model can be acquired.

Conventionally, hand-crafted features and prior knowledge are required to make automatic segmentation models. However, feature engineering requires expertise and experience. Recently, Convolution Neural Network (CNN) methods have been extensively

* indicates equal contribution

studied as the solutions to traditional approaches, as a CNN is a data-driven method and shows promising results without requiring hand-crafted features.

In this study, we propose the CNN-based two-stage segmentation method for kidney and kidney tumor. We separately design kidney segmentation model and tumor segmentation model to dissect and simplify problems, resulting in an efficient segmentation scheme. To exploit multiple slice information of tumor, we construct 2.5D CNN for kidney segmentation, and 3D CNN for tumor segmentation. Based on dice coefficient of segmentation results, we found that the proposed method effectively segments kidney and kidney tumor.

2 Methods

We use a two-stage approach for kidney and kidney tumor segmentation. The first network is used to segment kidney images from CT abdominal scans, and the second is used to segment kidney cancer images from the segmented kidney map so that the network recognizes the cancer image only from the kidney image. Code would be released at https://github.com/cms4f/kits19

2.1 Kidney segmentation

2D-based segmentation models do not use multi-slice information that is useful for segmenting 3D objects. For efficient segmentation by exploiting multiple slice information, three adjacent slices are stacked as input of the network, and the model is trained to generate a segmentation map corresponding to the center slice of the input. We call this approach 2.5D network. We construct the 2.5D network based on U-Net with attention gates [2]. Attention gate is known to recognize significant features while suppressing unrelated features for a task.

We conduct this process in three directions; axial, coronal, and sagittal. In axial direction, we construct the 2.5D network based on improved version of U-Net [3]. The improved U-Net consists of lossless decomposition which satisfy the *frame condition* in which low and high frequency components of image can be retained [4]. In coronal direction and sagittal direction, we use Max-pooling U-Net. Then the results are combined as union. In this way, we can capture kidney more accurately.

2.2 Tumor segmentation

In tumor segmentation, the model is a 3D network based on max pooling U-Net with attention gates [5]. The images masked by kidney segmentation predictions are used as input of network. At the end of the U-net, we added two headers for tumor segmentation and regression. By adding one additional regression task to the existing segmentation model, we could get more regularized and enhanced results in tumor segmentation.

3 Experiments and results

3.1 Datasets

Our method was trained and evaluated on KiTS19 challenge dataset consisting of CT images from 300 subjects [6]. 210 of these subjects have been released publicly as training set. 90 of 300 subjects are used as test set. The CT images are 3 dimensions, which consist of 512 x 512 x N (N: slice). The label index of segmentation masks is 1 for kidney, 2 for kidney tumor, and 0 except for kidney and tumor. The test set for challenge was released on July 15, 2019. So, we randomly divided training, validation and test set from publicly released training set for developing algorithm. To get rid of the unnecessary information on kidney and tumor segmentation, we clip the data to the range [-1024 to 2125].

3.2 Kidney segmentation

We divide full image with size of 512 x 512 to half-sized image with size of 512 x 256 to learn different characteristics of left and right kidney in training and test phase. Final segmentation map is combined with two half-sized maps. We simplify multi-class classification to binary classification problem for kidney segmentation. We set the background index to 0 and otherwise to 1, which means that the network recognizes kidney and kidney tumor as one class.

To train the network, we set the loss function which aims to learn multi-task. We use pixel-wise cross-entropy loss function for segmentation task and mean square error function for autoencoding reconstruction task. We found that exploiting multi-task learning improves the performance of algorithm while preserving fine details of objects. We used ADAM optimizer to optimize parameters of the network for kidney segmentation. The implementation of our algorithm was based Pytorch and Tensorflow library.

3.3 Tumor segmentation

We made a cube which covers whole volume of kidney segmentation and resize the cube into 4 sub-blocks due to memory restriction. We train and inference a network using left and right kidney separately. We use the weighted sum of pixel-wise cross entropy loss function for segmentation task and L1 loss function for regressions task. Also, we weight on loss of small tumor label. The implementation of our algorithm was based Tensorflow library.

3.4 Evaluation

We use dice coefficient as quantitative criterion for image prediction quality. We randomly select 27 validation data set and use it as our criterion for evaluation. As shown in Table 1, the proposed method makes performance with a dice score of 95.22% and 63.12% for kidney and tumor respectively. Figure 1 shows results of kidney and tumor segmentation. As seen in the figure, the model achieves good performance.

Table 1: Kidney and kidney tumor segmentation performance on validation 27 cases

	Kidney Dice	Kidney tumor Dice
Our model	0.9610	0.7080

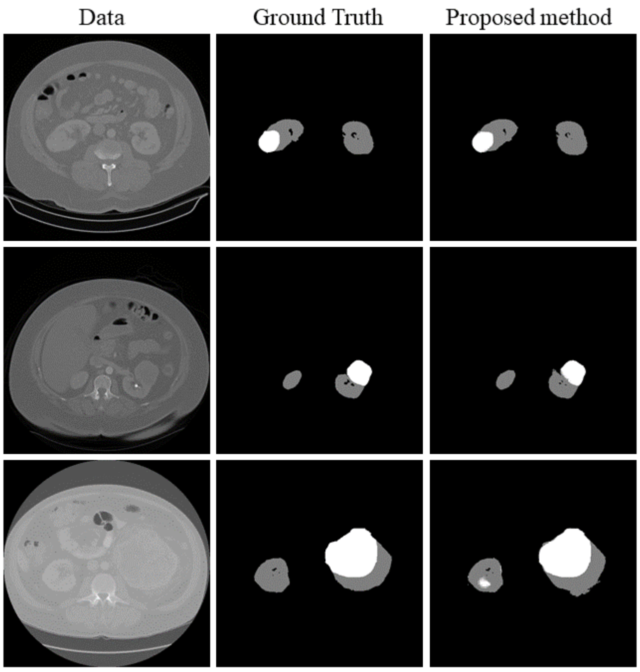

Figure 1: Segmentation results of the validation dataset: grey is kidney and white is tumor

4 Discussion and conclusion

We develop automatic kidney and tumor segmentation method using convolutional neural networks. In the case of kidney segmentation, training with input images of three different axis can capture different features of kidney, thus, combining them improves segmentation performance. Furthermore, adding another task to the segmentation model makes each task regularized and more accurate. In the case of tumor segmentation, 3D model performs well because it can learn and use depth information of kidney and tumor.

References

1. Bray, Freddie, et al. "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." CA: a cancer journal for clinicians 68.6, 394-424 (2018)

2. Oktay, Ozan, et al. "Attention U-Net: learning where to look for the pancreas." In: Proc. 1st Conf. Med. Imaging with Deep Learn. arXiv:1804.03999 (2018)
3. Kim, Boah, and Jong Chul Ye. "Cycle-consistent adversarial network with polyphase U-Nets for liver lesion segmentation." (2018)
4. Ye, Jong Chul, Yoseob Han, and Eunju Cha. "Deep convolutional framelets: A general deep learning framework for inverse problems." SIAM Journal on Imaging Sciences 11.2, 991-1048 (2018)
5. Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." Proceedings of the European Conference on Computer Vision (ECCV). 2018.
6. Heller, Nicholas, et al. "The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes." arXiv preprint arXiv:1904.00445 (2019)