HAL
open science

BiSC-UNet: A fine segmentation framework for kidney and renal tumor

Chuanxia Wang, Yuting He, Xiaoming Qi, Ziteng Zhao, Guanyu Yang, Xiaomei Zhu, Shaobo Zhang, Jean-Louis Dillenseger, Jean-Louis Coatrieux

To cite this version:

Chuanxia Wang, Yuting He, Xiaoming Qi, Ziteng Zhao, Guanyu Yang, et al.. BiSC-UNet: A fine segmentation framework for kidney and renal tumor. 2019 Kidney Tumor Segmentation Challenge: KiTS19, Miccai 2019, Oct 2019, Shenzen, China. 10.24926/548719.013 . hal-02358682

HAL Id: hal-02358682

https://hal.science/hal-02358682

Submitted on 12 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BiSC-UNet: A fine segmentation framework for kidney and renal tumor

Chuanxia Wang[1], Yuting He[1], Xiaoming Qi[1], Ziteng Zhao[1], Guanyu Yang[1,3](✉), Xiaomei Zhu[4] and Shaobo Zhang[5], Jean-Louis Dillenseger[2,3],, and Jean-Louis Coatrieux[2]

[1] LIST, Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing, China
yang.list@seu.edu.cn
[2] Univ Rennes, Inserm, LTSI - UMR1099, Rennes, F-35000, France
[3] Centre de Recherche en Information Biomédicale Sino-Français (CRIBs)
[4] Dept. of Radiology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China
[5] Dept. of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China

Abstract. Computed tomography (CT) images can provide a view of the patients' internal organs. This property is particular conducive to surgery planning. As one of the most common cancers, renal cancer also can be treated effectively by the laparo-scopic partial nephrectomy (LPN). However, automatic and accurate kidney and renal tumor segmentation in CT images remains a challenge. In this paper, BiSC-UNet framework which combines two different resolution SC-UNets is proposed for kidney and renal tumor fine segmentation. Rough SC-UNet is in charge of locating the kidney and renal tumor roughly to achieve the kidney region of interest (ROI) in original CT images. The other fine SC-UNet utilizes the kidney ROI for final fine kidney and renal tumor segmentation. In the proposed SC-UNet, not only the labeled kidney and renal tumor are used for training, but also the labeled renal tumor edge and kidney edge are utilized for accurate segmentation result. A balanced cross-entropy which selects and dynamic weights different regions and categorys is proposed for category imbalance. Extensive experiments on CT datasets demonstrate the effectiveness of proposed network.

1 Introduction

Renal cancer is one of ten most common cancers in human beings. The established treatments of renal cancer include surgery (radical or partial nephrectomy), chemotherapy, radiation therapy and radiofrequency ablation [2]. The most common treatment is surgery, such as: traditional radical nephrectomy (RN) and minimally invasive laparoscopic partial nephrectomy (LPN) [6]. These treatments are image-based and use computed tomography (CT) or magnetic resonance imaging (MRI). Because of this, major information of kidney and

renal tumor should be extracted from the images before surgery. Manual assessment of lesions characteristics in CT images is currently too tedious and time-consuming. Thus, an automatic or semi-automatic kidney and renal tumor segmentation method would be time-saving and useful in clinical practice.

This segmentation task concerns about two aspects: kidney segmentation and renal tumor segmentation. For kidney segmentation, several approaches have shown good performance in CT or MR images. By utilizing random forest algorithms, Cuingnet et al. [1] proposed a two-step kidney segmentation method. Yang et al. [10] utilized multi-atlas image on a coarse-to-fine segmentation. Torres et al. [9] adapted the B-Spline Explicit Active Surfaces (BEAS) framework for semi-automatic kidney segmentation in CT images. These algorithms only addressed the segmentation of kidney. Some of them may be failed when the shape of kidney changes because of injuries and serious diseases.

For renal tumor segmentation, few research works were devoted to. Linguraru et al. [5] developed a level-set based method to extract the renal tumors. However, the seed point provided by the user inside the lesion is needed. Prevost et al. [8] detected the lesion through a multi-scale radial symmetry transform in 3D contrast-enhanced ultrasound data. In CT images, the kidney and tumors may be similar in intensity distributions and textures. More than this, different tumors between different patients usually show very diverse shape appearance and volumetric size according to the different growth stages. Because of these, a fine automatic segmentation of renal tumors is so challenging.

With the development of convolutional neural network (CNN), it has been applied into various aspects in computer vision, such as: image classification, object recognition, image segmentation, etc. Also, CNN has achieved outcome performance in medical image processing, especially in kidney tumor segmentation. [12, 11] In [12], Yu et al. proposed the crossbar patches, which consists of two orthogonal non-squared patches and trained two sub-models for kidney tumor segmentation. Yang et al. [11] utilized a three-dimensional (3D) fully convolutional network (FCN) which combines a pyramid pooling module (PPM) for kidney and renal tumor segmentation. These two works focused on tumor shape information and contextual information separately. Yu et al. ignored the contextual information in CT images. The network that Yang et al. used is not focused on the shape information of tumor. To improve these shortcomings, we proposed a structural constraint 3D-UNet (SC-UNet) for kidney and renal tumor segmentation task. The proposed method considers not only contextual information but also shape information during training process.

The problem of category imbalance is a common problem in machine learning tasks, and it often has a serious impact on the segmentation results. Lin et al. proposed a focal loss [4] and keeps a category balance in the object detection task. The dice loss [7] balances different categories in the segmentation task by calculating the dice coefficients on each channel. In the kidney and tumor segmentation task, kidney and tumors account for a very small region of the abdominal CT image, leading a serious category imbalance problem. We combines tow SC-UNets for excessive background region on CT images, and propose a

BiSC-UNet framework. The proposed framework first take a rough network to get the kidney region of interest (ROI) for the balance between kidney regions and background and then take a fine network for fine segmentation on kidney ROI. We propose a balanced cross-entropy loss to ensure the balance between small tumors and big kidneys.

The rest of this paper is organized as follows. The details of proposed method are shown in Sec. 2. In Sec. 3, the experimental results are reported. Finally, the conclusion of this paper is drawn in Sec. 4.

2 Methodology

As is shown in Fig. 2, the BiSC-UNet (Sec. 2.2)consists of two cascade-connected network. In the first network, the rough segmentation of kidney and renal tumor will be achieved via a rough SC-UNet (Sec. 2.1) to crop a kidney ROI to get background balance and reduce the data size. Then in the second network, the cropped kidney ROIs are input to a fine SC-UNet for fine segmentation. A balanced cross-entropy loss (Sec. 2.3) is proposed to keep category balanced during the training process dynamically to improve the generalization ability.

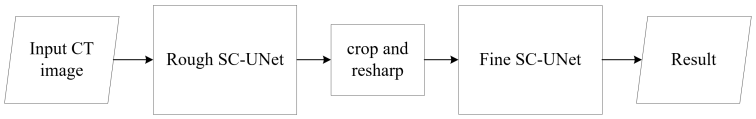

Fig. 1. The structure of our proposed BiSC-UNet framework.

2.1 SC-UNet for shape and contextual information embedding

As is shown in Fig. 2, the SC-UNet consists of down-sampling part and up-sampling part. The down-sampling part is same as 3D-Unet, which consists of three groups of convolutional layers and max-pool layer. The output of each group is represented by $y^i_{down} i \in 1, 2, 3$. In up-sampling part, the outputs of former two groups are represented by y^1_{up} and y^2_{up}. In Fig. 2, it can be found that the y^2_{up} is connected by three up-sampling groups. These three branches are corresponded with kidney and renal tumor label, edge of kidney label and edge of renal tumor label. The up-sampling groups are connected with the down-sampling groups, like 3D-Unet. The edge labels are generated by taking use of Canny edge detection algorithm. The original CT image ($X_{in} \times Y_{in} \times Z_{in}$) is the input of the network and the segmentations of kidney, tumor, the edge of kidney, tumor are the output.

From the structure of SC-UNet, it can be found that the edge information of kidney and tumor are treated as important as the entire kidney and tumor. By

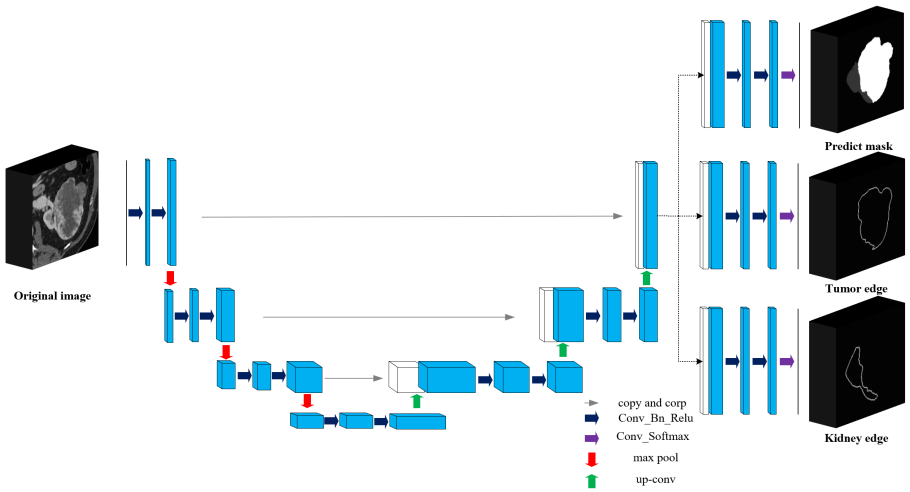

Fig. 2. The structure of SC-UNet.

the structure of three branches, the SC-UNet will focus on related shape information learning and combines the contextual information from 3D convolution, the SC-UNet can ensure both contextual and shape information be learned on training process.

2.2 BiSC-UNet for fine segmentation

As shown in Fig 1, the proposed BiSC-UNet framework consists of a rough SC-UNet and a fine SC-UNet which share same structure as is shown in Fig 2 to obtain fine segmentation gradually. The rough SC-UNet takes low resolution images as input and gets rough kidney and tumor segmentation. A maximum connected component analysis is used to remove the noise region and get the kidney ROI in original resolution CT image completely. The fine SC-UNet takes the original resolution kidney ROI image to obtain the fine kidney and kidney tumor ROI segmentation. Finally, two kidney ROI segmentation masks for each image are restored to the original size and generated a complete segmentation mask.

2.3 Balanced cross-entropy for category imbalance

The proposed balanced cross-entropy loss function utilize the $L1$ distance to measure the segmentation quality of each region, then the hard regions are selected to calculate the pixel-wise cross-entropy value. The pixel-wise loss value is averaged in each channel to get channel-wise loss value, so that each category will be divided by their size to weight dynamically. Finally, the channel-wise loss

values are averaged in the channel direction for image-wise loss value:

$$\mathcal{L}_{BCE} = \frac{1}{C}\sum_{c}^{C}\frac{1}{\sum_{n}^{N} y_{c,n}}\sum_{n}^{N}(Sig(|y_{c,n} - \hat{y}_{c,n}| > T)y_{c,n}log\hat{y}_{c,n}) \quad (1)$$

where C is the number or channels, N is number of pixels in each channel. The function $Sig(|y_{c,n} - \hat{y}_{c,n}| > T)$ means if $|y_{c,n} - \hat{y}_{c,n}| > T$, $Sig() = 1$, otherwise $Sig() = 0$. $y_{c,n}$ is the pixel label on the ground truth and $\hat{y}_{c,n}$ is the predicted pixel label probability.

3 Experiments and Results

3.1 Dataset

KiTS19 challenge data [3] is used to evaluate the performance of our proposed framework. This dataset contain 210 training cases and segmentation masks. Each sample contains at least one kidney with a tumor and is resampled to the resolution of 0.781625, 0.781625 and 3 on the X, Y, and Z directions. In our experiment, 168 (4/5) cases are used for training and 42 (1/5) cases are used for testing. Random cropping and random mirroring in three directions are used to augment the data.

3.2 Training strategies and implementation

In the rough SC-UNet, the original image is downsampled to 128×128 in the X and Y directions, and then the patch cropped in the Z direction is input into the network. The rough output is upsampled to the original resolution and is performed the maximum connected component analysis for two kidney masks. After that, the two largest regions are considered to be kidney masks and cropped as the kidney ROIs. The X and Y directions size are 160 or 192 (if bigger than 160), the Z direction size is bigger than the mask size on Z direction and a multiple of 16. In the fine SC-UNet, the kidney ROI in original resolution is input into the network for fine segmentation of kidney and tumor. Finally, the mask in kidney ROIs are restored to the original size for complete segmentation masks.

The optimizer of the rough SC-UNet and fine SC-UNet is Adam whose learning rate was 10^{-4}. The training batch size was 8 and each network was trained 20 epochs who had 4000 iterations. Our framework has been implemented via Pytorch, and trained and tested on a workstation with the CPU of i7-5930K, the RAM of 128GB and one graphics card of TITAN X of 12GB memory.

3.3 Results and analysis

As shown in Fig. 3, our proposed BiSC-UNet achieves fine kidney and kidney tumor segmentation in two steps. The rough segmentation has low resolution

and rough edges, and has the ability to used to identify the location of the two kidneys. The fine segmentation achieves a kidney and tumor mask with smooth edges from the kidney ROI. Cases 1, 2, and 3 show the patients with different sizes renal tumors, and each tumor is segmented accurately. In particular, the tumor that could not be detected by rough segmentation in case 3 is obtained in fine segmentation, demonstrating the validity of our cascade structure.

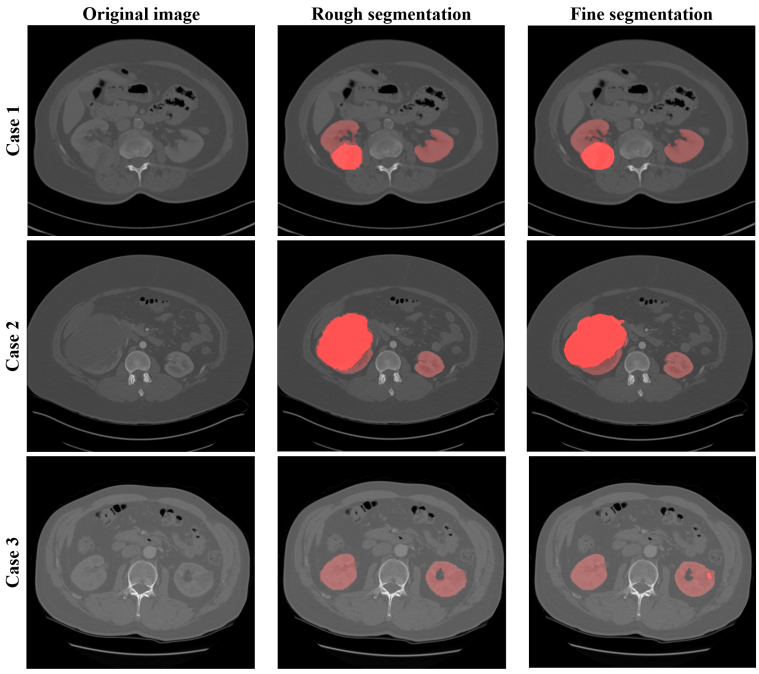

Fig. 3. The visual superiority of our proposed BiSC-UNet. The kidney ROI is cropped according to the rough segmentation and the fine segmentation is achieved from this ROI.

As shown in Tab. 1, our balanced cross-entropy improves the performance of BiSC-UNet and increases the dice coefficient of kidneys and tumors by 0.012 and 0.005 compared with the standard cross entropy loss function.

Table 1. The dice coefficient of kidney and renal tumors.

Network	Dice kidney(std)	Dice tumor(std)
BiSC-UNet(BCE)	0.954(0.070)	0.741(0.262)
BiSC-UNet(CE)	0.942(0.078)	0.736(0.241)

4 Conclusion

In this paper, we propose a kidney and renal tumor segmentation framework consists of two cascade-connected SC-UNet. Our proposed SC-UNet combines shape information and contextual information and embeds these information into 3D-Unet. For contextual information, this network takes three-dimensional convolutional layers. For shape information, the edge maps of kidney and tumor are treated as important as the entire labeled map on training process. Then, for category imbalance problem, the framework takes two cascade-connected structures to solve the big background impact and the balanced cross-entropy loss to keep kidney and tumor balanced. The first SC-UNet is taken for kidney ROI, the second SC-UNet is trained for fine segmentation. The balanced cross-entropy loss keeps the weight of each category balanced according to their size dynamically. By the experiments on KiTS19 challenge data, proposed method achieve outcome performance.

Acknowledgements. This research was supported by the National Key Research and Development Program of China (2017YFC0107903), National Natural Science Foundation under grants (31571001, 61828101), the Short-Term Recruitment Program of Foreign Experts (WQ20163200398), Key Research and Development Project of Jiangsu Province (BE2018749) and Southeast University-Nanjing Medical University Cooperative Research Project (2242019K3DN08).

References

1. Cuingnet, R., Prevost, R., Lesage, D., Cohen, L.D., Mory, B., Ardon, R.: Automatic detection and segmentation of kidneys in 3d ct images using random forests. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 66–74. Springer (2012)
2. Grubb III, R.L., Choyke, P.L., Pinto, P.A., Linehan, W.M., Walther, M.M.: Management of von hippel–lindau-associated kidney cancer. Nature Reviews Urology **2**(5), 248 (2005)
3. Heller, N., Sathianathen, N., Kalapara, A., Walczak, E., Moore, K., Kaluzniak, H., Rosenberg, J., Blake, P., Rengel, Z., Oestreich, M., 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)
4. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. pp. 2980–2988 (2017)
5. Linguraru, M.G., Wang, S., Shah, F., Gautam, R., Peterson, J., Linehan, W.M., Summers, R.M.: Automated noninvasive classification of renal cancer on multiphase ct. Medical physics **38**(10), 5738–5746 (2011)
6. Ljungberg, B., Bensalah, K., Canfield, S., Dabestani, S., Hofmann, F., Hora, M., Kuczyk, M.A., Lam, T., Marconi, L., Merseburger, A.S., et al.: Eau guidelines on renal cell carcinoma: 2014 update. European urology **67**(5), 913–924 (2015)
7. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV). pp. 565–571. IEEE (2016)

8. Prevost, R., Cohen, L.D., Corréas, J.M., Ardon, R.: Automatic detection and segmentation of renal lesions in 3d contrast-enhanced ultrasound images. In: Medical Imaging 2012: Image Processing. vol. 8314, p. 83141D. International Society for Optics and Photonics (2012)
9. Torres, H.R., Oliveira, B., Queirós, S., Morais, P., Fonseca, J.C., D'hooge, J., Rodrigues, N.F., Vilaça, J.L.: Kidney segmentation in 3d ct images using b-spline explicit active surfaces. In: 2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH). pp. 1–7. IEEE (2016)
10. Yang, G., Gu, J., Chen, Y., Liu, W., Tang, L., Shu, H., Toumoulin, C.: Automatic kidney segmentation in ct images based on multi-atlas image registration. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 5538–5541. IEEE (2014)
11. Yang, G., Li, G., Pan, T., Kong, Y., Wu, J., Shu, H., Luo, L., Dillenseger, J.L., Coatrieux, J.L., Tang, L., et al.: Automatic segmentation of kidney and renal tumor in ct images based on 3d fully convolutional neural network with pyramid pooling module. In: 2018 24th International Conference on Pattern Recognition (ICPR). pp. 3790–3795. IEEE (2018)
12. Yu, Q., Shi, Y., Sun, J., Gao, Y., Zhu, J., Dai, Y.: Crossbar-net: A novel convolutional neural network for kidney tumor segmentation in ct images. IEEE Transactions on Image Processing (2019)