FCNN-based segmentation of kidney vessels - Towards constraints definition for safe robot-assisted nephrectomy

S. Moccia[1,2], S. Foti[3], S.M. Rossi[3], I. Rota[3], M. Scotti[3], S. Toffoli[3], L.S. Mattos[3], E. De Momi[3], E. Frontoni[2]

[1]*Department of Advanced Robotics, Istituto Italiano di Tecnologia, Genoa (Italy)*

[2]*Department of Information Engineering, Università Politecnica delle Marche, Ancona (Italy)*

[3]*Department of Information, Electronics and Bioengineering, Politecnico di Milano, Milan (Italy)*

sara.moccia@iit.it

INTRODUCTION

Nephrectomy requires careful surgical actions in order to remove tumoral tissue while avoiding the damaging of major vessels such as the renal artery [1]. Robotic nephrectomy is emerging as a powerful solution to avoid vessel damaging through active-constraints (AC) control [1]. To provide AC control, sensitive vascular structures have to be detected in real-time in intra-operative images.

A well established literature on vessel segmentation exists. Vessel-segmentation approaches can be divided in: (i) vessel enhancement, (ii) model-based methods, (iii) tracking, and (iv) machine-learning (ML) approaches [2]. Recently, successful segmentation approaches are mainly based on ML strategies based on fully-convolutional neural networks (FCNNs). FCNNs allows to tackle inter- and intra-patient variability, as well as noise and illumination variation in the images.

The goal of this paper is to investigate the use of FCNN for vessel segmentation in laparoscopic images acquired during nephrectomy procedures.

MATERIALS AND METHODS

The FCNN investigated in this work is the U-Net [3]. U-Net is made of 9 processing steps, which form a contractive and an expansive path. The paths are symmetric to each other. The contractive path acts as feature extractor, while the expansive one performs up-convolution and segmentation. The U-Net used in this paper has only 5 of the original 9 processing steps to lower the segmentation computational cost while preserving segmentation performance [4]. The steps in the contractive path are made of 3x3 convolutional and 2x2 max pooling kernels. The expansive path has 3x3 convolutional and 2x2 up-sampling kernels. The peculiarity of U-Net is the presence of copy layers. The copy layers were introduced in [3] to link the contracting and the expansive path, as to retrieve the information lost in the contracting path and take it into account while performing object localization.

The segmentation of the renal artery – which is recognized as one of the most critical structures in the clinical literature [1] – was studied deploying a deep neural network. However, to the best of authors knowledge, no renal arteries dataset was available to train the FCNN. For this reason, a new dataset was created. It consisted of laparoscopic images manually extracted and labeled from 4 videos of robot-assisted partial nephrectomy, relative to 4 different patients. The images were extracted in such a way that the renal artery was clearly visible in the camera field of view. The number of extracted frame was 65 for the first patient (Pat. 1), 100 for the second one (Pat. 2), 30 for the third one (Pat. 3) and 123 for the last one (Pat. 4). Images size was 640x480 pixels for Pat. 1, 1280x720 for Pat. 2, and 854x480 for Pat. 3 and Pat. 4. Black borders were removed from the images, that were then resized to 296x220 pixels to smooth noise and reduce both processing time and memory usage. The image labeling process involved 4 expert subjects, that manually drew vessel contours using MATLAB 2018A. Challenges in the dataset included: different illumination levels, noise in the images, varying laparoscope pose with respect to the renal artery and high variability in both inter-patient artery architecture and location. Sample images acquired from different patients are shown in Fig. 1.

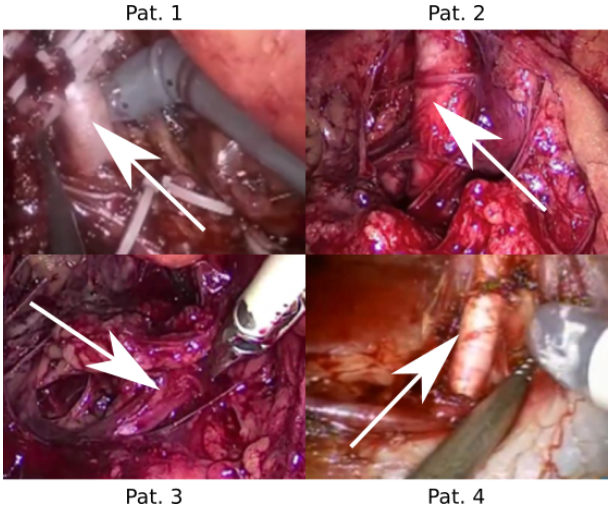

Figure 1: Sample images. Arrows indicate the renal artery.

	PPV	TPR	DSC	LR
P1-Pat. 1	0.8021	0.8564	0.8284	0.001
P1-Pat. 2	0.8555	0.4693	0.6061	0.001
P1-Pat. 3	0.8195	0.4208	0.5561	0.001
P1-Pat. 4	0.3626	0.7448	0.4878	0.0025
P2	0.7543	0.3814	0.5066	0.003

Table 1: Performance of the two protocols (**P1** and **P2**). For **P1**, results are shown for each of the four patients.

Two segmentation protocols were investigated. The first protocol (**P1**) aimed at evaluating the U-Net performance when training and testing were performed on different images from the same subject. Despite **P1** is not applicable in the actual clinical practice, it allowed investigating the segmentation performance in a trivial scenario. The images from each patient were randomly shuffled and divided in training (90%) and testing (10%).

The second segmentation protocol (**P2**) aimed at investigating the generalization power of U-Net when segmenting images from different patients. Thus, the images from the four patients were merged and randomly shuffled. Also in this case, the 90% of the images were used for training and 10% for testing. Prior to training, data augmentation was performed (9 linear and non-linear transformations were applied).

To train U-Net for the two protocols, the cross-entropy was used as loss function. Adam [5], a method for efficient stochastic optimization, was used as optimization algorithm. The initial learning rate (LR) was set to 0.001, after a trial-and-error procedure. Mini-batch gradient descent technique was used as a trade off between training convergence time and memory usage. Batch size was set equal to 5. TensorFlow (https://www.tensorflow.org/) was used for training and testing purposes.

The U-Net segmentation performance was quantitatively evaluated with respect to manual vessel tracing in terms of positive predictive value (PPV), true positive rate (TPR) and dice similarity coefficient (DSC).

RESULTS

The results relative to P1 and P2 are shown in Table 1. The best segmentation performance was achieved by **P1** for the first patient (Pat. 1). Sample segmentation images are shown in Fig. 2.

CONCLUSION AND DISCUSSION

As can be observed in Table 1, the best result was achieved for **P1**-Pat. 1 with a DSC equal to 82.84%.

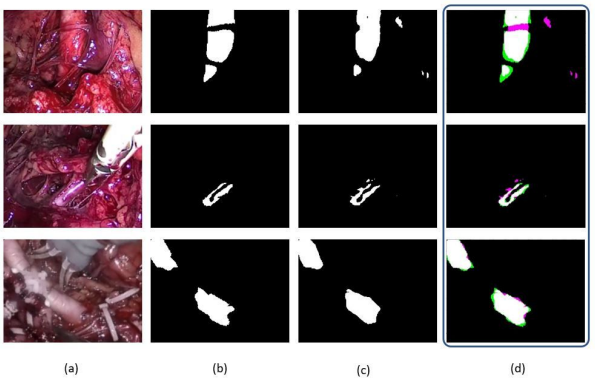

Figure 2: Sample images of U-Net segmentation outcomes. (a) original images, (b) ground-truth manual segmentations , (c) U-Net segmentation outcomes, (d) comparison between ground-truth and segmentation outcomes: false positives and false negatives are represented in magenta and green respectively.

In fact, Pat. 1 images were less challenging with respect to the remaining three patients (Pat. 2, Pat. 3 and Pat. 4), where the presence of cauterization smoke and the movement of the renal artery, caused by the interaction with surgical tools, increased the segmentation challenges. The resulting DSC of 50.66% obtained in P2, reflected the variability between different patients.

Improving the segmentation performance and tackling inter-patient variability by building a larger dataset, could be the natural continuation of the project in view of translation into clinical practice. Finally, instead of performing single-frame segmentation, temporal information could be exploited to track the vascular structures in time and improve segmentation performance [6].

REFERENCES

[1] MacLennan, Steven, et al. "Systematic review of oncological outcomes following surgical management of localised renal cancer." *European Urology* 61.5 (2012): 972-993.

[2] Moccia, Sara, et al. "Blood vessel segmentation algorithms - Review of methods, datasets and evaluation metrics." *Computer Methods and Programs in Biomedicine* 158 (2018): 71-91.

[3] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.

[4] S. Moccia, et al. "Toward Improving Safety in Neurosurgery with an Active Handheld Instrument," *Annals of Biomedical Engineering* (Accepted for publication).

[5] Kingma, Diederik P., and Jimmy Ba. "ADAM: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).

[6] Coskun, Huseyin, et al. "Long short-term memory kalman filters: Recurrent neural estimators for pose regularization." arXiv preprint arXiv:1708.01885 (2017).