# DPA-DenseBiasNet: Semi-supervised 3D Fine Renal Artery Segmentation with Dense Biased Network and Deep Priori Anatomy

Yuting He, Guanyu Yang, Yang Chen, Youyong Kong, Jiasong Wu, Lijun Tang, Xiaomei Zhu, Jean-Louis Dillenseger, Pengfei Shao, Shaobo Zhang, et al.

# DPA-DenseBiasNet: Semi-supervised 3D Fine Renal Artery Segmentation with Dense Biased Network and Deep Priori Anatomy

Yuting He[1], Guanyu Yang[1,4], Yang Chen[1,4], Youyong Kong[1,4], Jiasong Wu[1,4], Lijun Tang[3], Xiaomei Zhu[3], Jean-Louis Dillenseger[2,4], Pengfei Shao[5], Shaobo Zhang[5], Huazhong Shu1[1,4], Jean-Louis Coatrieux[2], and Shuo Li[6]

[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, 35000 Rennes, France
[3] Department of Radiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China
[4] Centre de Recherche en Information Biomédicale Sino-Français (CRIBs), France
[5] Department of Urology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China
[6] Department of Medical Biophysics, University of Western Ontario, London, ON, Canada

**Abstract.** 3D fine renal artery segmentation on abdominal CTA image targets on the segmentation of the complete renal artery tree which will help clinicians locate the interlobar artery's corresponding blood feeding region easily. However, it is still a challenging task that no one has reported success due to the large intra-scale changes, large inter-anatomy variation, thin structures, small volume ratio and limitation of labeled data. Hence, in this paper, we propose a novel semi-supervised learning framework named DPA-DenseBiasNet for 3D fine renal artery segmentation. The dense biased connection method is presented for multi-receptive field feature maps merging and implicit deep supervision [5] which enable the network to adapt to large intra-scale changes and improve its training process. The dense biased network (DenseBiasNet) is designed based on this method. We develop deep priori anatomy (DPA) for semi-supervised learning of thin structures. Differ from other semi-supervised methods, it embeds priori anatomical features to segmentation network which avoids inaccurate results sensitive to thin structures as optimizing targets, so that the network achieves generalization of different anatomies with the help of unlabeled data. Only 26 labeled and 118 unlabeled images were used to train our framework and it achieves satisfactory results on the testing dataset. The mean centerline voxel distance is 1.976 which reduced by 3.094 compared to 3D U-Net. The results illustrate that our framework has great prospects in the diagnosis and treatment of kidney disease.

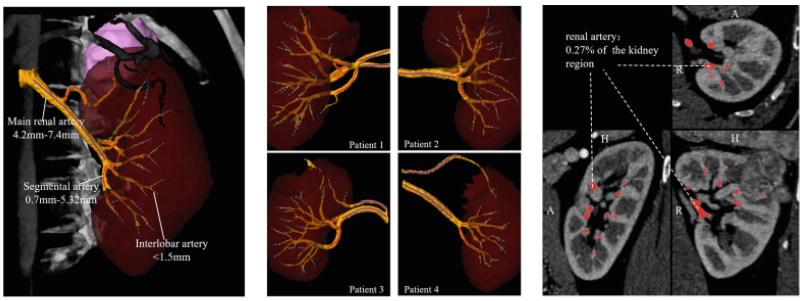
## 1 Introduction

3D fine renal artery segmentation on abdominal CTA image targets on achieving 3D renal artery tree that reaches the end of interlobar arteries. If successful, clinicians will locate the blood feeding region corresponding to each interlobar artery easily which is important for the diagnosis and pre-operative planning of kidney disease [7, 11, 12]. For example, it will show the tumor-feeding artery branches for segmental renal arteries clamping before laparoscopic partial nephrectomy [11, 12]. With the increasing probability of kidney disease [9], 3D fine renal artery segmentation will play an important role in its diagnosis and treatment.

However, no one has reported success in 3D fine renal artery segmentation because it is a challenging task [3]: (1) Large intra-scale changes. The thickest renal artery of a patient can up to 7.4 mm, which can be more than 5 times of the thinnest artery as is shown in Fig.1(a). This makes the network have to sensitive to different scale features, which increases the difficulty in feature extraction. (2) Large inter-anatomy variation. 11 different renal artery structures were found just from 461 patients [3]. The number of ostia, branch and accessory renal artery are variable between patients as Fig.1(b) shows. This makes it difficult for a small dataset to cover all variation and causes the network to overfit easily. (3) Thin structure. The thinnest artery is less than 1.5 mm which is much smaller than other organs in the same region so that the network is prone to lose these structures. (4) Small volume ratio. Renal arteries just account for 0.27% of the kidney region which will cause serious class imbalance problem, so the network is difficult to train. (5) Limitation of labeled data. It is difficult to learn the feature representation of different renal artery anatomies on a small data set, which limits the network's generalization ability. Therefore, how to overcome these challenges and to achieve 3D fine renal artery segmentation is an urgent problem.

There is no success to achieve 3D fine renal artery segmentation automatically, although some rough segmentation methods [6, 13] that only segmented up to segmental arteries have been proposed. Li et al. [6] only segmented the main and the thick segmental renal arteries using 400 images. Taha et al. [6, 13] used 99 cases to achieve the main arteries out of the renal. These methods cannot be applied to our task directly for two reasons: (1) The rough results. Main and segmental arteries usually correspond to multi blood feeding regions so that these rough results cannot be used to locate the specific blood feeding region of each vessel. (2) Large labeled dataset requirements. These methods used supervised learning which relies on a large labeled dataset and made it difficult to achieve satisfactory results when the 3D fine renal artery labeled dataset is small.

Semi-supervised learning gives us a tool to solve the limitation of labeled data because it uses unlabeled data to improve the model's performance [15]. Nie et al. [10] designed an attention-based network trained by 35 labeled and 20 unlabeled data to achieve the segmentation on a pelvic dataset. Bai et al. [1], applied CRF to a full convolutional network, and achieved the cardiac MR image segmentation with 240 unlabeled and 80 labeled images. **However**, these methods cannot be used for our tasks directly because they use unlabeled data




a)Large intra-scale changes b)Large inter-anatomy variation c) Small volume ratio & Thin structure

**Fig. 1.** The challenges of 3D fine renal artery segmentation. (a) The thickest artery of a patient is up to 7.4 mm which can be more than 5 times of the thinnest artery. (b) The number of ostia, branch and accessory renal artery are variable between patients. (c) Renal arteries account for 0.27% of the kidney region and are much thinner than other organs.

to get inaccurate labels which lose thin structures easily so that the model is optimized in wrong targets and its performance is weak in our task.

We propose a novel semi-supervised framework (DPA-DenseBiasNet) for 3D fine renal artery segmentation in this paper. Our work has the following contributions: (1) To the best of our knowledge, this work is the first achievement in 3D fine renal artery segmentation. (2) It presents a dense biased connection method which merges multi-receptive field feature maps to adapt to large intra-scale changes. Further, each layer has direct access to the gradients from the loss function, leading to deep supervision [5], to simplify the training process. Dense biased network (DenseBiasNet) is designed based on this method. (3) Deep priori anatomy (DPA) is proposed for semi-supervised learning of thin structures. It embeds priori anatomical features from the encoder trained by unlabeled data to the DenseBiasNet to avoid inaccurate results sensitive to thin structures as optimizing target like other semi-supervised methods. Therefore, DenseBiasNet trained by a small labeled dataset has a higher generalization of different anatomies and keeps the segmentation ability of thin structures. The experimental results show that our framework has enormous potential for clinical application.

## 2 Methodology

As is illustrated in Fig.2, DPA-DenseBiasNet adopts a dense biased network (DenseBiasNet) and deep priori anatomy (DPA) for 3D fine renal artery segmentation. DenseBiasNet (Sect.2.1) is a segmentation network which uses dense biased connection method for multi-receptive field feature maps merging and implicit deep supervision to achieve the segmentation with large scale changes and simplify the training process. DPA (Sect.2.2) is a semi-supervised method of thin

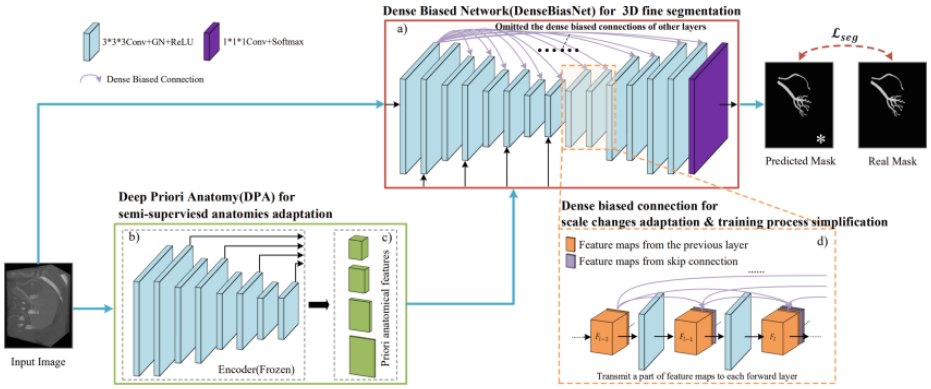



**Fig. 2.** The framework of DPA-DenseBiasNet: (a) and (d) are DenseBiasNet and dense biased connection method which introduced in Sect.2.1. (b) is the encoder network from denoising autoencoder trained by numerous unlabeled data and (c) is the priori anatomical features from the encoder. (b) and (c) constitutes DPA which introduced in Sect.2.2 .

structures which uses a trained encoder to provide priori anatomical features for DenseBiasNet to guide the adaptation of variable anatomical structures.

### 2.1 Dense Biased Network (DenseBiasNet) for Fine Segmentation

DenseBiasNet is a 3D fine renal artery segmentation network which based on dense biased connection method. It connects a part of feature maps in each layer to every other forward layer to build dense connectivity pattern. These feature maps are optimized by richer gradients and play a more important role in the network. Hence, we call the method dense biased connection.

**Advantages of Dense Biased Connection:** (1) It allows the network to adapt to large intra-scale changes via merging multi-receptive field feature maps which have different sensitivities to different scale vessels in each convolutional layer. (2) It simplifies the training process because the gradients from the loss function optimize each layer along the dense biased connection directly.

**Dense Biased Connection:** As is illustrated in Fig.2(d), each layer gets a part of feature maps from all preceding layers as additional inputs and transmits a part of its output feature maps to all forward layers. If the sizes of feature maps do not match, maxpooling or upsample method will be used. We denoted the output of the $l^{th}$ layer as $F_l$. The $l^{th}$ layer receives a part of feature maps of previous layers, $F_0, \ldots, F_{l-2}$, and all the feature maps of $F_{l-1}$ as input: $F_l = H_l(F_{l-1} \circ F_{l-2}[0 : k_{l-2}] \circ \ldots \circ F_0[0 : k_0])$. Where $H_l(\bullet)$ can be a composite function of operations such as group normalization (GN), rectified linear units(ReLU), pooling, upsampling, or convolution (Conv). The symbol $\circ$ refers to the concatenation of the feature maps. $k_0, \ldots, k_{l-2}$ is the number of the feature maps the $l^{th}$ layer can receive from $0,_{\ldots}, l-2$ layers. In our experiments, we set $k_0, ..., k_{l-2} = 1$.

**Structure and Loss Function of DenseBiasNet:** Figure 2(a) illustrates DenseBiasNet's structure. It comprises of 14 3D convolution layers which followed a GN and a ReLU, 3 maxpooling layers, 3 3D deconvolution layers used to change scales and a $1\times1\times1$ convolution layer followed a softmax as the output layer to reduce the number of channels to classes. The dense biased connection is used throughout the network to adapts to different scale arteries.

The network is trained by minimizing the loss function consisting of dice coefficient loss and cross-entropy loss. The dice coefficient loss $\mathcal{L}_{dice}$ helps to establish a balance between artery and background from a global perspective, and the cross-entropy loss $\mathcal{L}_{ce}$ is used for correct classification of each voxel at a local perspective:

$$\mathcal{L}_{seg} = \lambda \underbrace{\left(1 - \frac{1}{C}\sum_{c}^{C}\frac{2\sum_{n}^{N} y_{n,c}\hat{y}_{n,c}}{\sum_{n}^{N} y_{n,c}^2 + \sum_{n}^{N}\hat{y}_{n,c}^2}\right)}_{\mathcal{L}_{dice}} - \underbrace{\frac{1}{N}\sum_{c}^{C}\sum_{n}^{N} y_{n,c}\log\hat{y}_{n,c}}_{\mathcal{L}_{ce}} \quad (1)$$

where $C$ is the number of channels output from the network, and $N$ is the size of each channel. $\hat{y}_{n,c}$ is the predicted result and $y_{n,c}$ indicates the label. $\lambda$ is used to balance these loss functions. In our experiments, we set $\lambda = 0.1$.

### 2.2 Deep Priori Anatomy (DPA) Based Semi-supervised Learning

DPA is a novel semi-supervised method that avoids inaccurate optimizing targets which sensitive to thin structures during training. It first trains an autoencoder [14] with numerous unlabeled data, and then uses the encoder part to extract input image's priori anatomical features of different semantic levels in different depth in order to guide the anatomical adaptation of the segmentation network which trained with a small labeled dataset as shown in Fig.2(b)(c).

**Advantages of DPA:** (1) Deep priori anatomical feature representation learned from unlabeled data which adapts to more anatomical structures than manual priori features. (2) It focuses on both local and global anatomical information because different semantic levels features are extracted from different depths. (3) It improves the network's generalization ability utilizing these priori features. (4) It is suitable for thin structures compared with other semi-supervised methods [1, 10] because embedding priori knowledge will not introduce inaccurate labels that are easy to lose thin structures.

**DPA for Semi-supervised Anatomical Adaptation:** Figure 2 shows the process of extracting and embedding deep prior anatomical features into the DenseBiasNet. Prior to this, a convolutional denoising autoencoder was trained with numerous unlabeled data, and then its encoder part was frozen and used to extract anatomical features (Fig.2(b)). The image $\mathbf{x}$ is putted into the encoder to obtain priori anatomical features at different depths $\{F_{f1}, F_{f2}, F_{f3}, F_{f4}\} = f(\mathbf{x};\theta_f)$. These feature maps (Fig.2(c)) together with input image $\mathbf{x}$ are putted into the DenseBiasNet's different convolution layers to predict fine renal artery

segmentation $\hat{\mathbf{y}} = d(\mathbf{x}, F_{f1}, F_{f2}, F_{f3}, F_{f4}; \theta_d)$. The learning process is minimizing the loss function $\mathcal{L}_{seg}(\mathbf{y}, \hat{\mathbf{y}})$ to optimize the DenseBiasNet. $\mathbf{y}$ is the renal artery mask, $f(\cdot)$ is the encoder operation and $d(\cdot)$ is the DenseBiasNet operation.

## 3 Experiments and Results

**Setting:** Abdominal contrast-enhanced CT images of 170 patients who underwent LPN surgery were included in this study. The pixel size of these CT images is between 0.59 $mm^2$ to 0.74 $mm^2$. The slice thickness and the spacing in zdirection were fixed at 0.75 mm and 0.5 mm respectively. 52 of these images have renal artery mask, half of them were used as the training sets and the other half as the test set. The remaining 118 unlabeled images are used to train the denoising autoencoder. On the training set, we mirrored each image on three axes to expand the training data to 104 images. The kidney region of interest which size was $152 \times 152 \times Z$ was extracted firstly. The denoising autoencoder and DenseBiasNet are all trained with Adam where the batch size was 1, the learning rate was $1 \times 10 - 4$ and the decay rate was $1 \times 10 - 5$. They all trained 200 epochs on their corresponding data sets.

To demonstrate the advantage of our framework, we compared our method with 2 supervised methods (3D U-Net and VNet) and 2 semi-supervised methods (SemiFCN and ASDNet). We adopt the following evaluation metrics to evaluate our proposed method: mean dice coefficient (Dice), mean centerline distance (MCD) and mean surface distance (MSD).

**Visual Superiority:** Figure 3 shows the visual superiority of our framework. Case 1 is a left kidney region whose artery has many singular anatomical structures which difficult to segment. Case 2 is a right kidney region whose artery has great scale changes at the branches. Compared with ground truth, DPADenseBiasNet achieves fine segmentation in these cases thanks to the dense biased connection which merges multi-receptive field feature maps and DPA which ensures the segmentation quality of thin structures and generalization of different anatomies. Without the help of DPA, DenseBiasNet loses some parts in case1. SemiFCN cannot realize our task, because CRF removes thin renal arteries so that the inaccurate optimize target weakens the network's performance. ASDNet has serious mis-segmentations due to the more serious class imbalance caused by the confidence map and the instability caused by adversarial learning.

**Evaluation Metrics Advantages:** The advantages of our DPA-DenseBiasNet on each metric are demonstrated in Table 1. DPA-DenseBiasNet achieves the best segmentation results compared with other methods. The dice coefficient, mean centerline distance and mean surface distance are 0.861, 1.976 and 1.472, and their corresponding standard deviations are 0.095, 1.394 and 1.738. Ablation experiments in the last two rows validate the importance of DPA which improves the segmentation accuracy obviously.

**Training Process Improvement:** Figure 4 illustrates the improvement of the training process due to dense biased connection. We compared DenseBiasNet, 3D U-Net (3 layers have skip connections) and the network without skip

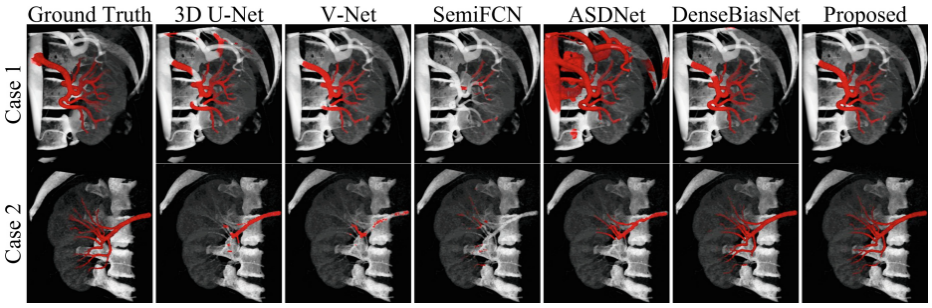



**Fig. 3.** The visual superiority of our framework (DPA-DenseBiasNet).

**Table 1.** The advantages of our method (DPA-DenseBiasNet) on each metrics.

| Network | Dice | MCD | MSD |
|---|---|---|---|
| V-Net [8] | 0.787 (0.113) | 2.872 (2.196) | 2.213 (2.155) |
| 3D U-Net [4] | 0.750 (0.162) | 5.070 (4.949) | 4.385 (4.208) |
| (semi)SemiFCN [2] | 0.388 (0.259) | 8.772 (10.085) | 7.921 (10.593) |
| (semi)ASDNet [10] | 0.555 (0.191) | 8.557 (5.124) | 7.484 (5.132) |
| DenseBiasNet | 0.851 (0.110) | 2.478 (2.090) | 1.920 (2.354) |
| (semi)Proposed | **0.861 (0.095)** | **1.976 (1.394)** | **1.472 (1.738)** |

connections. DenseBiasNet has faster convergence speed and higher test accuracy because the gradients from the loss function optimize each convolution layer directly.

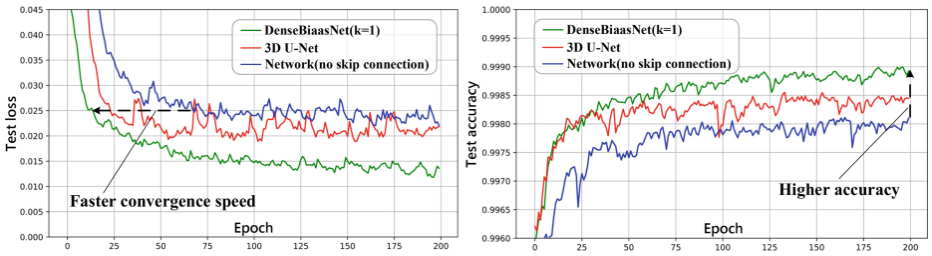



**Fig. 4.** The improvement of the training process by the dense biased connection. DenseBiasNet has faster convergence speed and higher accuracy than the other two networks.

## 4 Conclusion

This paper proposed a novel semi-supervised framework which achieved 3D fine renal artery segmentation. The proposed framework used dense biased connection method to enable DenseBiasNet to adapt to large intra-scale changes and

simplify its training process. Further, our developed semi-supervised method of thin structures, DAP, embedded priori anatomical features from an encoder network to DenseBiasNet to improve its generalization of different anatomies. The performance of our framework was compared with other methods. The results showed that our framework had great prospects in the diagnosis and treatment of kidney disease.

**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. Bai, W., Oktay, O., Sinclair, M., Suzuki, H., Rajchl, M., Tarroni, G., Glocker, B., King, A., Matthews, P.M., Rueckert, D.: Semi-supervised learning for network-based cardiac MR image segmentation. In: Descoteaux, M., Maier-Hein, L., Franz, A.M., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 253–260. Springer Cham (2017). https://doi.org/10.1007/978-3-319-66185-8 29
2. Baur, C., Albarqouni, S., Navab, N.: Semi-supervised deep learning for fully convolutional networks. In: Descoteaux, M., Maier-Hein, L., Franz, A.M., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 311–319. Springe Chamr (2017). https://doi.org/10.1007/978-3-319-66179-7_36
3. Bordei, P., Şapte, E., Iliescu, D., Brânzaniuc, K., Baz, R., Matusz, P., Dina, C.: Morphological assessments on the arteries of the superior renal segment. Surgical and radiologic anatomy **34**, 137–144 (2012). https://doi.org/10.1007/s00276-011-0866-y
4. Çiçek, O., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
5. Lee, C.Y., Xie, S., Gallagher, P., Zhang, Z., Tu, Z.: Deeply-supervised nets. In: Artificial intelligence and statistics. pp. 562–570 (2015)
6. Li, J., Lo, P., Taha, A., Wu, H., Zhao, T.: Segmentation of renal structures for image-guided surgery. In: Frangi, A., Schnabel, J., Davatzikos, C., Alberola-Lopez, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11073, pp. 454–462. Springer Cham (2018). https://doi.org/10.1007/978-3-030-00937-3_52
7. Ljungberg, B., Bensalah, K., Canfield, S., Dabestani, S., Hofmann, F., Hora, M., Kuczyk, M.A., Lam, T., Marconi, L., Merseburger, A.S., Mulders, P., Powles, T., Staehler, M., Volpe, A., Bex, A.: EAU guidelines on renal cell carcinoma: 2014 update. European urology **67**, 913–924 (May 2015). https://doi.org/10.1016/j.eururo.2015.01.005
8. Milletari, F., Navab, N., Ahmadi, S.: V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In: Fourth International Conference

on 3D Vision, 3DV 2016, Stanford, CA, USA, October 25-28, 2016. pp. 565–571 (2016). https://doi.org/10.1109/3DV.2016.79
9. National Cancer Institute: Cancer stat facts: kidney and renal pelvis cancer. Tech. rep. (2016), https://seer.cancer.gov/statfacts/html/kidrp.html
10. Nie, D., Gao, Y., Wang, L., Shen, D.: ASDNet: Attention based semi-supervised deep networks for medical image segmentation. In: Frangi, A., Schnabel, J., Davatzikos, C., Alberola-Lopez, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11073, pp. 370–378. Springer Cham (2018). https://doi.org/10.1007/978-3030-00937-3_43
11. Shao, P., Qin, C., Yin, C., Meng, X., Ju, X., Li, J., Lv, Q., Zhang, W., Xu, Z.: Laparoscopic partial nephrectomy with segmental renal artery clamping: technique and clinical outcomes. European urology **59**, 849–855 (May 2011). https://doi.org/10.1016/j.eururo.2010.11.037
12. Shao, P., Tang, L., Li, P., Xu, Y., Qin, C., Cao, Q., Ju, X., Meng, X., Lv, Q., Li, J., Zhang, W., Yin, C.: Precise segmental renal artery clamping under the guidance of dual-source computed tomography angiography during laparoscopic partial nephrectomy. European urology **62**, 1001–1008 (Dec 2012). https://doi.org/10.1016/j.eururo.2012.05.056
13. Taha, A., Lo, P., Li, J., Zhao, T.: Kid-Net: Convolution networks for kidney vessels segmentation from CT-volumes (2018), http://arxiv.org/abs/1806.06769v1
14. Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.A.: Extracting and composing robust features with denoising autoencoders. In: Proceedings of the 25th International Conference on Machine Learning. pp. 1096–1103. ACM (2008). https://doi.org/10.1145/1390156.1390294
15. Zhu, X.: Semi-supervised learning literature survey. Tech. rep., University of Wisconsin, Madison, Computer Science Department (2005)