



HAL
open science

Dense biased networks with deep priori anatomy and hard region adaptation: Semi-supervised learning for fine renal artery segmentation

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

► To cite this version:

Yuting He, Guanyu Yang, Jian Yang, Yang Chen, Youyong Kong, et al.. Dense biased networks with deep priori anatomy and hard region adaptation: Semi-supervised learning for fine renal artery segmentation. *Medical Image Analysis*, 2020, 63, pp.101722. 10.1016/j.media.2020.101722 . hal-02614243

HAL Id: hal-02614243

<https://hal.science/hal-02614243v1>

Submitted on 20 May 2020

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.

Dense Biased Networks with Deep Prior Anatomy and Hard Region Adaptation: Semi-supervised Learning for Fine Renal Artery Segmentation

Yuting He^a, Guanyu Yang^{a,g,*}, Jian Yang^b, Yang Chen^{a,g}, Youyong Kong^{a,g}, Jiasong Wu^{a,g}, Lijun Tang^d, Xiaomei Zhu^d, Jean-Louis Dillenseger^{c,g}, Pengfei Shao^e, Shaobo Zhang^e, Huazhong Shu^{a,g}, Jean-Louis Coatrieux^{c,g}, Shuo Li^f

^a*LIST, Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing, China*

^b*Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Electronics, Beijing Institute of Technology, Beijing, China*

^c*Univ Rennes, Inserm, LTSI - UMR1099, Rennes, F-35000, France*

^d*Dept. of Radiology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China*

^e*Dept. of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China*

^f*Dept. of Medical Biophysics, University of Western Ontario, London, ON, Canada*

^g*Centre de Recherche en Information Biomédicale Sino-Français (CRIBs)*

Abstract

Fine renal artery segmentation on abdominal CT angiography (CTA) image is one of the most important tasks for kidney disease diagnosis and pre-operative planning. It will help clinicians locate each interlobar artery's blood-feeding region via providing the complete 3D renal artery tree masks. However, it is still a task of great challenges due to the large intra-scale changes, large inter-anatomy variation, thin structures, small volume ratio and small labeled dataset of the fine renal artery. In this paper, we propose the first semi-supervised 3D fine renal artery segmentation framework, DPA-DenseBiasNet, which combines deep prior anatomy (DPA), dense biased network (DenseBiasNet) and hard region adaptation loss (HRA): 1) Based on our proposed dense biased connection, the DenseBiasNet fuses multi-receptive field and multi-resolution feature maps for large intra-scale changes. This dense biased connection also obtains a dense

*Corresponding author

Email address: yang.list@seu.edu.cn (Guanyu Yang)

information flow and dense gradient flow so that the training is accelerated and the accuracy is enhanced. 2) DPA features extracted from an autoencoder (AE) are embedded in DenseBiasNet to cope with the challenge of large inter-anatomy variation and thin structures. The AE is pre-trained (unsupervised) by numerous unlabeled data to achieve the representation ability of anatomy features and these features are embedded in DenseBiasNet. This process will not introduce incorrect labels as optimization targets and thus contributes to a stable semi-supervised training strategy that is suitable for sensitive thin structures. 3) The HRA selects the loss value calculation region dynamically according to the segmentation quality so the network will pay attention to the hard regions in the training process and keep the class balanced.

Experiments demonstrated that DPA-DenseBiasNet had high predictive accuracy and generalization with the Dice coefficient of 0.884 which increased by 0.083 compared with 3D U-Net (Çiçek et al., 2016). This revealed our framework with great potential for the 3D fine renal artery segmentation in clinical practice.

Keywords: renal artery segmentation, semi-supervised learning, dense biased network, dense biased connection, deep priori anatomy, hard region adaptation loss function, 3D fine segmentation, CT angiography image

2010 MSC: 00-01, 99-00

1. Introduction

Fine renal artery segmentation on abdominal CT angiography (CTA) image is one of the most important tasks for pre-operative planning (Ljungberg et al., 2015; Shao et al., 2011, 2012; Porpiglia et al., 2018). It targets on achieving 3D
5 renal artery tree masks that reach the end of interlobar arteries, if successful, clinicians will locate each interlobar artery's blood-feeding region to complete an accurate pre-operative plan (Zhang et al., 2019). As is shown in Fig. 1(a), it will play a key role in the clamping of segmental renal arteries before laparoscopic partial nephrectomy (LPN) (Shao et al., 2011, 2012). The tumor-feeding interlo-

10 bar arteries (red arrows) which attach to the tumor on CTA images are detected to find the segmental renal artery, thus helping the operation go smoothly.

However, obtaining a fine renal artery segmentation automatically is a challenging task due to: 1) Large intra-scale changes. As is shown in Fig. 1(a), there is a large-scale change between different renal artery parts of the same patient. 15 The interlobular arteries are less than 1.5 mm which are one-fifth of the thickest part. Thus, multi-scale feature representations are required to adapt to the artery’s large intra-scale changes in our task. 2) Large inter-anatomy variation. Renal arteries’ branch structures, ostia, accessory arteries and the way entering the kidneys are various among different patients as is shown in Fig. 1(b). For 20 example, Petru et al. (2012) counted 11 different renal artery structures just from 461 kidneys. Therefore, some singular anatomical structures are difficult to be covered in the dataset leading to the model’s poor generalization ability, especially when the labeled dataset is small. 3) Thin structures. The renal artery has an elongated tubular structure where the thinnest part is less than 25 1.5 mm as shown in Fig. 1(a)(b). Such thin objects have more hard regions, such as edges, surfaces, ends, etc., which lead to serious instability of the results and are prone to be over-segmented or under-segmented. 4) Small volume ratio. The renal artery only accounts for 0.27 percent of the kidney’s interest of region which leads to severe class imbalance. Thus, a) the training process will 30 be inefficient because the background region which is the majority of the image will produce low loss, and b) it is easy for the network to judge the pixels as the majority class so that the minority vessels will be ignored. 5) Small labeled dataset. It is challenging to delineate the renal arteries manually which leads to the limitation of fine-labeled data and thus making it difficult for the seg- 35 mentation model to achieve the feature representation of singular anatomical structures.

Semi-supervised learning strategy improves the model generalization ability (Weston et al., 2012) and is suitable for challenges of inter-anatomy variation and data limitation, while previous strategies (Nie et al., 2018; Bai et al., 2017) 40 cannot be applied to our task. These strategies use unlabeled data to get inac-

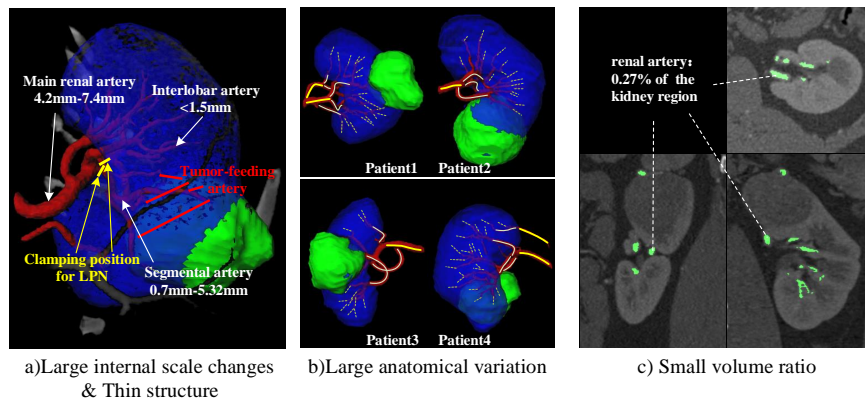


Figure 1: The challenges in fine renal artery segmentation. a) There are large scale gaps between different renal artery parts, with the interlobular arteries being one-fourth of the main arteries. The artery containing large hard regions is much thinner than other organs. b) Different arteries' branch structures, ostia, accessory arteries and the way entering the kidneys are variable between patients. c) Renal arteries account for 0.27 percent of the kidney region leading to serious class imbalance.

curate labels that lose thin structures like vessels easily and take these wrong labels as optimization targets causing the training process interfered and the model performance weakened. We propose the deep priori anatomy strategy which trains an autoencoder (AE) in an unsupervised manner with numerous unlabeled data and embeds its feature representation ability in the supervised model to guide the adaptation of more anatomy structures. Therefore, the unsupervised training in the first stage combines with supervised training in the second stage making it a semi-supervised framework (Hady and Schwenker, 2013). This feature embedding method will not introduce inaccurate labels so that the segmentation quality of thin structures is ensured.

Dense connections (Gao et al., 2017) transmit all feature maps in each layer to every forward layer which fuses multi-receptive field feature maps so that it has powerful representation capability for the objects with different scales. However, the depth and width of the networks based on the dense connection are limited due to the reusing of all feature maps which take up too much

memory. We propose a dense biased connection that compresses feature maps before transmission so that it will keep the advantages of dense connection, simplify the network and save the memory. The whole network’s dense biased connection will also make the fusion of multi-resolution feature maps and guide
60 the adaptation of large-scale changes. Besides, the dense connectivity pattern keeps dense information flow and leads to implicit deep supervision (Lee et al., 2015) which will improve the accuracy and accelerate the training.

Re-dividing the dataset to increase the minority or decrease the majority category images is a general method to cope with the class imbalance (Liu,
65 2009). But in our segmentation task, it is limited due to the class imbalance comes from the intra-image. We propose a hard region adaptation loss function that samples the loss region dynamically according to the segmentation quality intra-image so that the hard-to-segment regions such as edges, surfaces, ends, etc. will be focused and their segmentation quality will be enhanced.

70 In this paper, we propose a 3D semi-supervised framework (DPA-DenseBiasNet) for the first time to overcome fine renal artery segmentation challenges and achieve this task robustly. It is based on deep priori anatomy (DPA) strategy, dense biased network (DenseBiasNet) and hard region adaptation (HRA) loss to guide the adaptation of inter-anatomy variation steadily, fuse multi-receptive
75 field features for intra-scale changes, and focus on hard regions to keep class balanced dynamically.

1.1. Related work

Vascular segmentation. Vascular segmentation is significant for disease diagnosis and pre-operative planning, many works have made efforts on this task.
80 Some traditional methods (Lesage et al., 2009), such as some centerline-based methods (Yang et al., 2012; Tyrrell et al., 2007), region-growing approaches (Metz et al., 2007; Tschirren et al., 2005), active contours (Toledo et al., 2000; Mille and Cohen, 2009), etc., have achieved success in many vascular segmentation scenarios. These methods require a lot of manual design and are sensitive
85 to parameters. With the development of deep learning, deep learning-based

vascular segmentation methods (Moccia et al., 2018; Wolterink et al., 2019; Bai et al., 2018; Noothout et al., 2018) are widely used for vascular segmentation owing to their fast speed, high accuracy and powerful generalization ability. However, these deep learning-based vascular segmentation methods always rely
90 on large fine labeled datasets whose performance will be limited when datasets are small.

Fine renal artery segmentation has not been reported success apart from the rough renal artery segmentation which are obtained in multi-object segmentation tasks (Li et al., 2018a; Taha et al., 2018). Li et al. (2018a) train a
95 convolutional neural network with 400 arterial-phase CT images to achieve the segmentation of multiple renal structures including the main and segmental renal arteries. Taha et al. (2018) propose a Kid-Net which is trained via 99 cases to segment the kidney vessels, while only the renal arteries that outside the kidney are segmented successfully. These methods cannot meet the requirements
100 of our task because of their inherent flaws: 1) Insufficient fineness. The renal artery part before the interlobular artery corresponds to a large blood-feeding region which impossible to locate the vessel feeding for a specific region, such as the tumor-feeding artery. 2) Dependence on large datasets. These supervised models are trained based on large rough labeled datasets, and if the dataset is
105 small, their generalization ability will be limited.

Semi-supervised learning. Semi-supervised learning has great potential in improving the generalization ability of neural networks via utilizing unlabeled data (Weston et al., 2012; Hady and Schwenker, 2013) which has achieved success in the image processing field (Qiao et al., 2018; Gu and Angelov, 2018; Hung et al., 2018; Tang et al., 2018). Numerous semi-supervised methods such
110 as co-training (Han et al., 2018) and adversarial learning (Hung et al., 2018) have achieved remarkable results. In biomedical image segmentation, some semi-supervised methods (Nie et al., 2018; Bai et al., 2017) have been used to help handle the problem of the limitation of labeled data. Bai et al. (2017) combine
115 conditional random fields with a fully convolutional network for cardiac MR im-

age segmentation and use 240 unlabeled data to improve the model’s accuracy. Nie et al. (2018) take the generative adversarial network to obtain a confidence map so that the segmentation result of the unlabeled data can be used as a label for network training. However, these semi-supervised strategies cannot be
120 applied to vessel segmentation tasks because the unlabeled data was used to get inaccurate labels losing thin structures like vessels easily. Therefore, the wrong labels bringing inaccurate optimization targets will interfere with the training process and weaken the model performance in our task. Our DPA strategy utilizes feature embedding method fusing the knowledge from numerous unlabeled
125 data to the supervised model directly, thus avoiding the inaccurate optimization target and enhancing the accuracy steadily.

Dense connection. Dense connection (Gao et al., 2017) improves the flow of information and gradients which optimize the training process, but it also brings feature redundancy and memory limitation problem which is not suitable for
130 big 3D medical images. Mixed-Scale Dense (MS-D) network (Pelt and Sethian, 2018) take dense connection method and achieve the biomedical image segmentation. Similarly, Li et al. (2018b) design a novel hybrid densely connected U-Net which consists of 2D and 3D DenseUnet and obtain the liver tumor segmentation. DenseASPP (Yang et al., 2018) combines dense connection with
135 dilated convolution to fuse multi-scale information and improve the accuracy of semantic segmentation. Since these networks consume a large quantity of memory with all feature maps reused, the input data size is limited. Especially in the 3D medical image segmentation task, this memory problem will be more serious. Our dense biased connection compresses the feature maps before transmission,
140 thus avoiding feature redundancy, keeping the advantages of dense connection and saving the memory. This method makes it easy to build whole network dense connectivity patterns so that each hidden layer will have additional direct access to the gradients from the loss function through the shorter connections. Therefore, these additional gradients will optimize these layers during the training
145 process making implicit deep supervision (Lee et al., 2015; Gao et al., 2017)

improving the accuracy and accelerating the training.

Autoencoder network (AE). AE (Masci et al., 2011) is an unsupervised neural network that extracts and composes robust features, so it can be used for semi-supervised learning to improve the model’s performance. Three main methods for applying AE to semi-supervised learning can be identified. 1) Gradient embedding. The features from the pre-trained encoder network are used as an auxiliary optimization target for the specific task network, thereby the additional gradients are embedded in the training process (Sedai et al., 2017). 2) Auxiliary task. AE shares a part of parameters with the specific task network and is trained as an additional task at the same time (Rasmus et al., 2015b,a). 3) Feature embedding. The features extracted by the pre-trained autoencoder are embedded in the supervised neural network (Wu et al., 2019; Kingma et al., 2014; Xiao et al., 2018). Different than in the previous two methods, the feature embedding method fuses the knowledge from numerous unlabeled data to the supervised model directly avoiding the impact on the training process which is friendly to our renal arteries (thin structure). Therefore, we use a denoising autoencoder (Vincent et al., 2010) which captures higher-level features from the input images for priori anatomy features extraction and embedding.

Class imbalance. Class imbalance means that there is a large gap between the number of different classes in the dataset, a severe class imbalance will lead to inefficient training process and model degradation (Lin et al., 2017). Wolterink et al. (2016) take a two-stage hard negative mining method in coronary artery calcium scoring task that locates the rare class before its scoring, thus balancing the classes and achieving success. The soft Dice loss (Milletari et al., 2016) calculates the Dice value of different categories as the loss value to avoid the inter-class interaction. The weighted cross-entropy gives higher weight to rare categories making the network pay more attention to them, but the weight is a hyper-parameter and is difficult to determine. The focal loss (Lin et al., 2017) adds to the weighted cross-entropy a modulating factor which calculated from the result so that the network pays more attention to the hard region.

The first two loss functions do not distinguish the difficulty of hard regions in the segmentation task which leads to local defects easily. Although focal loss weakens this problem to some extent, the weighted sum of the main class’s loss values is still much larger than the rare class when the class is seriously imbalanced. In our proposed HRA loss, most of the regions that have been segmented well are not involved in the training process, thus keeping the balance of the class dynamically and making the network focus on the hard-to-segment regions.

1.2. Contributions

We propose a powerful 3D semi-supervised framework (DPA-DenseBiasNet) for the first time to achieve fine renal artery segmentation. The detailed contributions of our work are as follows:

- As far as we know, we achieve fine renal artery segmentation for the first time and review the relevant technical and clinical value of this topic, which will provide a valuable reference for follow-up works.
- We propose a novel 3D semi-supervised framework for fine renal artery segmentation which can play an important role in pre-operative planning and demonstrate its excellent performance via complete experiments.
- We present a steady semi-supervised learning strategy named deep priori anatomy (DPA) to guide the adaptation of more anatomy structures and weaken the data limitation. It extracts priori anatomy features via a pre-trained AE and embeds them in the segmentation network to introduce anatomy features representation ability from numerous unlabeled data.
- Our newly proposed dense biased connection method fuses multi-receptive field and multi-resolution features so that more powerful multi-scale representation ability is achieved to adapt the network to intra-scale changes. A dense biased network (DenseBiasNet) based on this dense biased connection is designed.

- We propose a hard region adaptation (HRA) loss function to keep class
205 balanced. It samples the loss function according to the segmentation quality so that the network will attach significance to the hard-to-segment region dynamically, thus keeping the class balanced and improving the segmentation quality of hard regions.

In this work, we advance our preliminary efforts (He et al., 2019) on fine
210 renal artery segmentation task in the following aspects: 1) conduct more comprehensive review of the techniques related to fine renal artery segmentation, 2) propose a hard region adaptation loss function (HRA) to focus on the hard-to-segment region, thus keeping the class balanced and improving the segmentation quality of hard regions, and 3) demonstrate the performance of our proposed
215 DPA-DenseBiasNet in the fine renal artery segmentation task via carrying out more extensive experiments for performance analysis and comparison.

For the rest of the paper, we demonstrate the details of the proposed DPA-DenseBiasNet in Sec. 2, including deep priori anatomy (DPA), dense biased network (DenseBiasNet) and hard region adaptation (HRA) loss function. Then the
220 dataset, experiment settings, and evaluation measures are described in Sec. 3. Sec. 4 shows the results and analysis of comparison experiments and ablation experiments. Sec. 5 discusses and concludes the paper.

2. Methodology

As shown in Fig. 2, the DPA-DenseBiasNet is proposed for fine renal artery
225 segmentation. It has three structures: 1) An AE is pre-trained and used to extract and embed DPA features to guide the adaptation of anatomy variation. This process forms a steady semi-supervised learning strategy and we call it DPA strategy (Sec. 2.1). 2) Based on dense biased connection, a DenseBiasNet (Sec. 2.2) fuses multi-receptive field and multi-resolution feature maps for powerful
230 multi-scale representation ability, thus adapting to the intra-scale changes. 3) Based on an intra-image region sampling strategy, an HRA (Sec. 2.3) loss samples the loss region dynamically according to the segmentation quality so

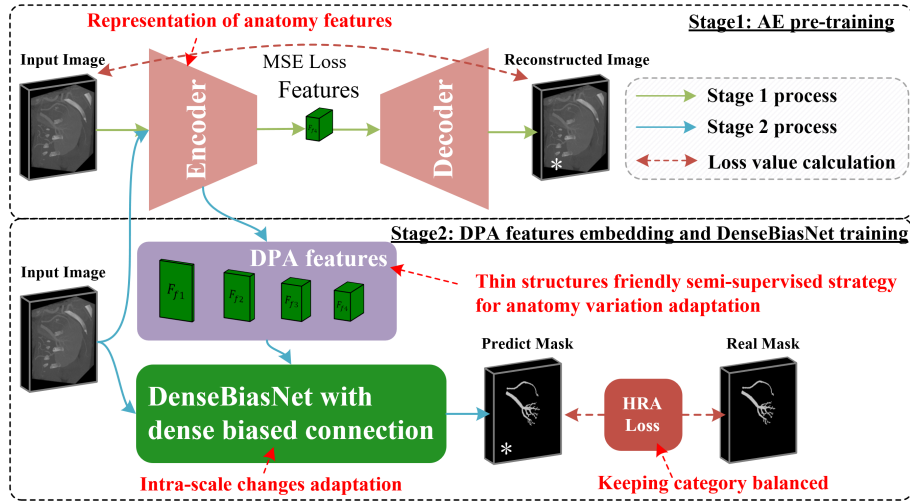


Figure 2: DPA-DenseBiasNet for fine renal artery segmentation: **Stage 1** is the AE pre-training. The AE is trained by unlabeled images, and the encoder network is used to extract DPA features in stage 2. **Stage 2** is DPA features embedding and DenseBiasNet training. The DenseBiasNet with dense biased connection takes DPA features to adapt the anatomical variation and is optimized by the HRA loss function to keep the class balanced.

the network will focus on the hard regions and keep the class balanced. The training process consists of two stages including a) AE pre-training, b) DPA features embedding and DenseBiasNet training. In stage 1, an AE (Masci et al., 2011) is trained with a big unlabeled dataset (unsupervised) to get the representation ability of numerous anatomy features. In stage 2, an encoder network from this AE extracts the DPA features and embeds them in the DenseBiasNet as the priori information to adapt the anatomy variation. To achieve fine renal artery segmentation, the DenseBiasNet is trained by both the labeled data and the DPA features with the HRA loss function.

2.1. Deep priori anatomy for anatomy variation adaptation

DPA strategy is a novel steady semi-supervised learning strategy that utilizes the feature representation ability of pre-trained AE to extracts anatomy features (Sec. 2.1.2) and embeds them in different depth of a segmentation network for anatomy variation adaptation (Sec. 2.1.3) as is illustrated in Fig. 2.

This process pre-trains the encoder, pre-extracts the features which reflect the anatomical information from different depth of the encoder, and fuses these features to different depth of the segmentation network. Thus, we extract three
250 representative keywords and name this strategy 'Deep priori anatomy'.

2.1.1. Advantages of DPA

1) Steady and suitable for thin structures. Other semi-supervised segmentation methods (Nie et al., 2018; Bai et al., 2017) take inaccurate optimization targets and interfere with the training process. Our DPA strategy fuses the
255 knowledge from numerous unlabeled data to the supervised model directly and takes a single accurate optimization target to avoid interference thus bringing a steady training process. This steady process is suitable for the thin structures which are lost easily. 2) Generalization improvement. DPA strategy embeds the feature extraction ability learned from numerous unlabeled images to the
260 segmentation network to guide the adaptation of more anatomy structures so that its generalization ability will be improved.

2.1.2. AE pre-training for DPA features representation

DPA strategy takes an AE to obtain anatomy feature representation capability via numerous unlabeled data. Then the pre-trained encoder of the AE
265 is used to extract DPA features in our framework. As is shown in Fig. 3, we take the denoising autoencoder (DAE) (Vincent et al., 2010) for better higher-level representation and build a convolutional DAE which has an encoder and a decoder. The encoder encodes the image into low-resolution anatomy features via successive downsampling, and then these features are upsampled by the
270 decoder to reconstruct the original image. An unlabeled image x is an input into a dropout (Srivastava et al., 2014) layer to get the image x' corrupted by noise before entering the network. The encoder extracts the feature of x' to obtain the code $h = f(x')$, and then restores the image $\hat{x} = g(h)$ through the decoder. Finally, the mean square error (MSE) loss (Masci et al., 2011) values between
275 the reconstructed image \hat{x} and the original image x are calculated for network

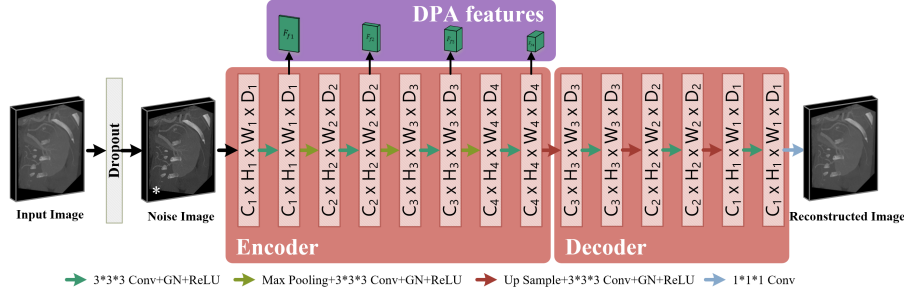


Figure 3: AE for anatomy features representation. The convolutional denoising autoencoder (DAE) is trained with numerous unlabeled data for anatomy feature representation ability and the encoder outputs four sets of DPA features with different resolutions from different depths. The C_n, H_n, W_n, D_n is the channels, height, width and depth of the feature maps for each layer.

optimization. A large amount of unlabeled data is used to train this DAE to equip it with the DPA features extraction ability in stage 1. Its encoder network is used to extract DPA features of different semantic levels from encoder’s different resolution stages for DenseBiasNet in stage 2.

280 *2.1.3. DPA features for anatomy variation adaptation*

DPA strategy extracts the DPA features from input images via the pre-trained encoder and embeds them in the DenseBiasNet for anatomy variation adaptation as is shown in Fig. 2. The image x is put into the frozen encoder network which is trained via numerous unlabeled data during stage 1 to get the DPA features of different semantic levels from different depths 285 $\{F_{f1}, F_{f2}, F_{f3}, F_{f4}\} = f(x)$. These features are concatenated to the input feature maps of the second convolution layer of their corresponding resolution stage in DenseBiasNet’s analysis path as shown in Fig. 5. And the DenseBiasNet utilizes the original image x and these DPA features to obtain the segmentation 290 result $\hat{y} = d(x, F_{f1}, F_{f2}, F_{f3}, F_{f4})$. The loss value, calculated from the result \hat{y} and the label y , is used to optimize the DenseBiasNet $\mathcal{L}(\hat{y}, y)$.

2.2. Dense biased network for multi-scale adaptation

DenseBiasNet (Sec. 2.2.3) adopts the dense biased connection (Sec. 2.2.2) compressing and transmitting all feature maps in each layer to every forward
295 layer, so that multi-receptive field and multi-resolution features are fused for powerful multi-scale representation ability to adapt to intra-scale changes. Due to this dense connectivity pattern, the information flow and gradient flow are completely transmitted in the network, thus the training process will be accelerated and the accuracy will be improved.

300 2.2.1. Advantages of dense biased connection

1) Intra-scale changes adaptation. It fuses the feature maps with different receptive fields and resolutions in each layer so that the network will have powerful multi-scale representation ability, thus adapting to different scale vessels.
2) Training acceleration. The gradient from the loss function optimizes each
305 layer directly along with the connections leading to implicit deep supervision (Lee et al., 2015), so the training process will be accelerated. 3) Complete information flow. The feature maps in each layer are compressed and transmitted to every forward layer so that the complete information is obtained in deep layers.
4) Simplified network structure. Compared to dense connection (Gao et al.,
310 2017), it transmits much fewer feature maps which reduces feature redundancy and makes it easier to build full network dense connectivity pattern, so that it is more suitable for large medical image sizes.

2.2.2. Dense biased connection as connectivity pattern

DenseBiasNet takes dense biased connection as the connectivity pattern that compresses and transmits the feature maps in each layer to every forward layer, so the multi-scale features are integrated, the integrities of information flow and gradient flow are kept as shown in Fig. 4. Equ. 1 illustrates the calculation process of the l_{th} layer in the network with the dense biased connection. Pooling and upsampling are used to unify the feature maps at different scales.

$$F_l = H_l(F_{l-1} \circ Conv(F_{l-2}; m_{l-2}) \circ \dots \circ Conv(F_0; m_0)) \quad (1)$$

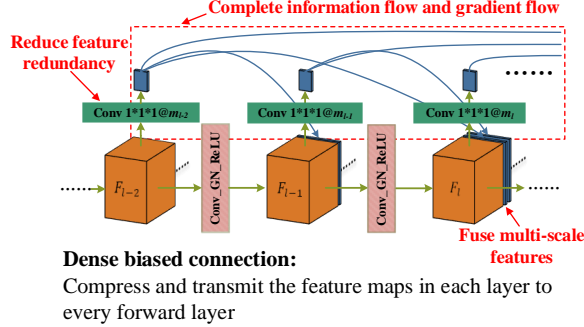


Figure 4: The dense biased connection compresses and transmits the feature maps in each layer to every forward layer to reduce feature redundancy, keep the integrity of information flow and gradient flow and fuse multi-scale features.

where the F_l is the feature maps output from the l_{th} layer and $\{F_0, \dots, F_{l-2}\}$ are the feature maps from previous layers. $Conv(\cdot)$ is a $1 \times 1 \times 1$ convolution layer which compresses the feature maps into a small number of maps, and the $\{m_0, \dots, m_{l-2}\}$ are the bias quantities which represent the number of the output feature maps of these convolution layers. $H_l(\cdot)$ is a combination function that can be group normalization (GN) (Wu and He, 2018), rectified linear units (ReLU), pooling, upsampling or convolution (Conv). In our experiments, we set it as a Conv followed by a GN and a ReLU. ' \circ ' is a concatenation operation that concatenates all feature maps from the $l - 1_{th}$ layer and the compressed feature maps from $\{0, 1, \dots, l - 2_{th}\}$.

2.2.3. DenseBiasNet design

Based on the dense biased connection, the DenseBiasNet takes standard 3D U-Net's analysis path and synthesis path (Çiçek et al., 2016) as the basic structure. Continuous downsampling in the analysis path saves GPU memory effectively and the synthesis path restores the low-resolution feature maps to the one with the original resolution which achieves an end-to-end output with the same resolution. The dense biased connection fuses the feature maps with different resolutions and different receptive fields so that the network has the

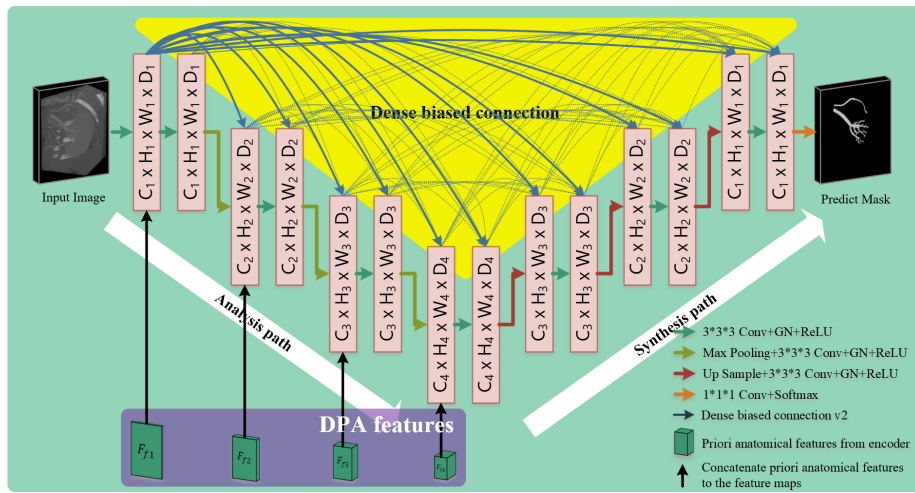


Figure 5: The proposed DenseBiasNet with dense biased connection is embedded with DPA features. The *Dense biased connection* is used throughout the network which builds a dense connectivity pattern for the adaptation of different scale arteries. The *DPA features* from the encoder network are concatenated to different resolution stages of the DenseBiasNet to guide the adaptation of more anatomical structures. The C_n, H_n, W_n, D_n is the channels, height, width, and depth of the feature maps from each layer.

representation ability for features of different semantic levels. As shown in Fig. 5, it has 2 paths and 4 resolution stages with two $3 \times 3 \times 3$ convolution layers which are followed by a GN and a ReLU. Maxpooling is used to reduce the resolution
335 at the end of the stages in analysis path and upsampling is used to improve the resolution at the beginning of the stages in synthesis. A $1 \times 1 \times 1$ convolution layer followed by a softmax is used as the output layer to reduce the number of channels to classes. The dense biased connection is used throughout the network and feature maps' resolution is changed by maxpooling or upsampling to adapt
340 to different resolution stages.

2.3. Hard region adaptation loss for class balance

Hard region adaptation loss function is presented to sample the hard regions intra-image, thus the network will focus on the hard-to-segment region and dynamically keep the class balanced during the training process. Based on our
345 designed intra-image sampling method, it only trains hard region to ensure the intra-class and inter-class balance of the dataset and improving the network accuracy. We will introduce the hard region adaptation loss based on the cross-entropy (CE).

2.3.1. Advantages of HRA

350 1) Training process adaptation. Our HRA selects the regions to calculate the loss function according to the segmentation quality dynamically. Therefore, as the model converges, HRA will gradually exclude the easy-to-segment regions and focus on the hard regions fitting the actual situation of the training. 2) Feature-based class balance. It balances the data with a dynamic selection of
355 hard regions and is more suitable for the real situation of the features than the weighting methods that never focus on regional features.

2.3.2. HRA definition

HRA loss finds hard-to-segment regions according to the segmentation quality of results dynamically and optimizes these regions to keep class balanced

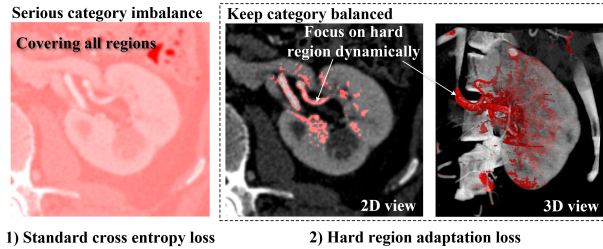


Figure 6: Hard region adaptation loss keeps the class balanced. The red regions are the sampled hard regions. 1) The standard cross-entropy loss is calculated covering all regions which will cause a serious class imbalance problem. 2) The hard region adaptation loss takes an L_1 distance to sample the hard-to-segment regions such as vascular boundary, small ends, etc. and thus keeping class balanced dynamically

dynamically. As is shown in Fig. 6, the red regions of the outputs are calculated for the cross-entropy loss. When training the segmentation network with a standard cross-entropy loss, we calculate the loss value covering all regions and pay the same attention to hard regions and easy regions which causes a serious class imbalance problem. When it comes to HRA loss, the hard regions such as vascular boundary, small ends, etc. are found via calculating the L_1 distance between the results and ground truth, and thresholding for hard-to-segment region mask. Then, the cross-entropy loss is calculated on these regions as is illustrated in Equ. 2:

$$\mathcal{L}_{HRA-CE} = - \sum_n^N I(y_n, \hat{y}_n) y_n \log \hat{y}_n \quad (2)$$

where y_n is the target value and \hat{y}_n indicates the predicted value. $I(y_n, \hat{y}_n)$ is the intra-image sampling strategy and in this paper we set that if $|y_n - \hat{y}_n| > T$,
 360 $I(y_n, \hat{y}_n) = 1$, otherwise $I(y_n, \hat{y}_n) = 0$. The hyper parameter T serves as the threshold to control the sensitivity of hard region, when $T = 0$, the loss value will be calculated on the whole image.

In the training process, the large easy-to-segment regions are fitted well quickly, so the network will calculate the loss value only on hard regions contributing to class balance. If the network is too focused on the hard regions
 365

and starts to reduce the accuracy of easy regions, the L_1 distance of the easy regions will increase, so that the network will pay attention to these regions again. Therefore, the segmentation quality of both the easy and hard regions will be guaranteed dynamically. In this section, we present the definition of HRA with the cross-entropy as an example, while its form is not limited to this. This loss function design method based on the intra-image sampling strategy can be applied to improve other loss functions such as Dice loss and mean square error for class balance.

3. Experiments configurations

3.1. Evaluation dataset

Dataset introduction. Abdominal contrast-enhanced CT images from 196 kidney cancer patients who have a kidney with renal tumor and a healthy kidney were retrospectively selected from the radiology department of Jiangsu Province Hospital in this study. The age of these patients is between 15 to 82 years old. The size of the kidneys is between 85.76ml to 262.78ml and the size of the tumors is between 1.75ml to 144.82ml. The kidney tumor types consist of clear renal cell carcinomas, papillary, chromophobe, angiomyolipoma and eosinophilic adenoma. Each image was acquired on a Siemens dual-source 64-slice CT scanner and the contrast media was injected during the CT image acquisition. The X-ray tube current, the convolution kernel, the exposure time and the voltage are 480mA, B25f, 500ms and 120KV. The pixel size of these CT images is between 0.59mm/pixel to 0.74mm/pixel. The slice’s thickness and the spacing in z-direction were fixed at 0.75mm/pixel and 0.5mm/pixel respectively.

The kidney regions of interests (ROIs) whose sizes were $152 \times 152 \times Z$ were cropped on the abdominal contrast-enhanced CT images automatically via multi-atlas-based approach (Yang et al., 2014) in the pre-processing firstly, thus leading to 392 kidney ROI images. The renal artery boundary that reaches the end of interlobar arteries was semi-automatically delineated by a trained researcher and manually verified by a clinician in 156 images. 52 of them were used

395 as the training set and the other 104 as the test set. The remaining 236 unlabeled images were used for the unsupervised training of the convolutional DAE. In the training processing, these kidney ROIs were cropped into $152 \times 152 \times 64$ online before putted into the network due to the limitation of hardware memory.

Data augmentation. Random mirroring and random rotation methods were used to augment the image during the training process. The flipping method was performed randomly in the x, y and z axis of the kidney ROI and the random rotation angle was $[-10^\circ, 10^\circ]$.

3.2. Training strategies and implementation

The convolutional DAE and DenseBiasNet were all optimized by Adam whose learning rate, β_1 , β_2 and decay rate was 10^{-4} , 0.9, 0.999 and 10^{-5} respectively. The training batch size was 1 owing to the limitation of memory and was iterated 200 epochs on their corresponding data sets. Our implementation used the Keras library of the Tensorflow backend which was publicly available. We trained the proposed framework on a single 1080 GPU with 8 GB memory.

3.3. Evaluation measures

To demonstrate the advantage of our framework, we compared our proposed DenseBiasNet with three general supervised methods (3D U-Net (Çiçek et al., 2016), V-Net (Milletari et al., 2016), Kid-Net (Taha et al., 2018)) and DPA-DenseBiasNet with two semi-supervised methods (SemiFCN (Baur et al., 2017), ASDNet (Nie et al., 2018)). These models were trained on the same dataset with the same implementation. Mean Dice coefficient, mean centerline distance, mean surface distance and Hausdorff distance were adopted to evaluate our proposed method. All the metrics were calculated for each image and the mean values were got in the end. The standard deviation of each metric was used to measure the dispersion on different images to offer an objective way to evaluate the robustness of our method.

Dice coefficient (Dice). The Dice coefficient was used to evaluate the similarity of the foreground regions in the two images according to Equ. 3.

$$Dice(G, P) = \frac{2 |G \cap P|}{|G| + |P|} \quad (3)$$

where G is the renal artery region in the ground truth, and P is the region in the predicted result.

Mean centerline distance (MCD). In order to evaluate the displacement of the renal artery segmentation, the centerline distance was used as an evaluation metric. The centerline was obtained following paper Boskamp et al. (2005) and the average Hausdorff distance (Taha and Hanbury, 2015) was calculated using the Equ. 4.

Mean surface distance (MSD). It was used to evaluate the lumen thickness and surface coincidence of the renal artery. We first extracted the surface of the blood vessel and then calculated the average Hausdorff distance using Equ. 4. A small number of voxels away from the main segmentation object has a huge impact on the Hausdorff distance, so we used average Hausdorff distance.

$$\begin{aligned}
 HD_{mean}(G, P) &= \max[f(G, P), f(P, G)] \\
 f(G, P) &= \text{mean}(\min_{p \in P} \|g - p\|) \\
 f(P, G) &= \text{mean}(\min_{g \in G} \|p - g\|)
 \end{aligned} \quad (4)$$

Hausdorff distance (HD). It was used to evaluate the outliers which is sensitive in vessel segmentation task. It will give additional information about the segmentation accuracy that cannot be evaluated by surface distance. We follow the Hausdorff distance introduced by Taha and Hanbury (2015).

4. Results and analysis

DPA-DenseBiasNet adapts to the intra-scale changes and inter-anatomy variation, ensures the class balanced, weakens the dataset limitation steady,

optimizes the training process, and achieves the excellent fine renal artery segmentation ultimately. In this part, we thoroughly evaluate and analyze the effectiveness of our proposed DPA-DenseBiasNet in three folds. 1) Firstly, the performances of different existing methods on the 3D fine renal artery segmentation task are quantified on four metrics and their segmentation results were visually analyzed (Sec. 4.1). 2) Next, we analyze the role of HRA in handling class imbalance problems, the help of dense biased connection to accelerate training and improve accuracy, the effectiveness of DPA in smaller datasets, and the enhancement of low-level features in the dense biased connection (Sec. 4.2). 3) Finally, the influence of different hyper-parameters of the framework is analyzed and the most suitable hyper-parameters for our task are selected (Sec. 4.3).

4.1. Excellent fine renal artery segmentation results

The proposed DPA-DenseBiasNet has strong superiority both in visual and evaluation metrics.

Quantitative evaluation supports the advantages. As demonstrated in Tab. 1, the DPA-DenseBiasNet achieves state-of-the-art performance in four metrics. The performance of DenseBiasNet with HRA, whose Dice coefficient, mean centerline distance, mean surface distance and Hausdorff distance are 0.872, 1.440mm, 1.006mm and 27.268mm, is much better than the other three supervised learning methods. The DPA is friendly to small structures and improves the Dice up to 0.884, and will not weaken the network performance like the other two semi-supervised methods which cannot achieve the renal artery mask. The ablation experiments in the second part compare the impact of different strategies and validate that the DPA and HRA improve the model performance obviously.

Qualitative evaluation supports the visual superiority. As shown in Fig. 7, our proposed DPA-DenseBiasNet achieves the fine renal artery segmentation and ensures the segmentation quality of different scales and singular structures. The visual superiority of our framework is demonstrated via four

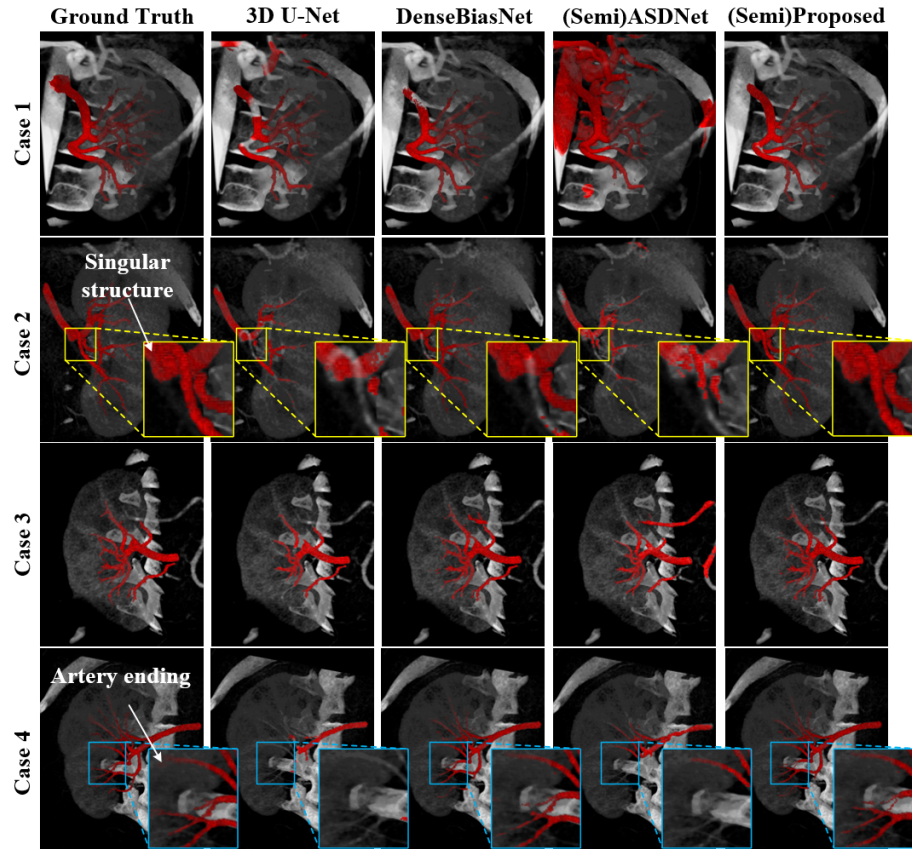


Figure 7: The visual superiority of the proposed framework (DPA-DenseBiasNet). The blue boxes indicate the high segmentation quality of artery endings via our DenseBiasNet and the yellow boxes indicate the high segmentation quality of the singular regions brought by our DAP strategy.

Table 1: The proposed DPA-DenseBiasNet achieves state-of-the-art performance in four metrics. The DenseBiasNet has better performance than three supervised methods (V-Net, 3D U-Net, and Kid-Net), and the DPA strategy improves the generalization ability and will not weaken the network performance like the other two semi-supervised methods (SemiFCN and ASDNet). The HRA loss balances the categories and further improves the performance of standard cross-entropy (CE).

Network	Dice \pm std	MCD \pm std(mm)	MSD \pm std(mm)	HD \pm std(mm)
V-Net (Milletari et al., 2016)	0.715 \pm 0.167	2.856 \pm 3.133	1.893 \pm 2.733	31.976 \pm 15.206
3D U-Net (Çiçek et al., 2016)	0.801 \pm 0.114	2.070 \pm 1.915	1.578 \pm 2.135	32.240 \pm 17.169
Kid-Net (Taha et al., 2018)	0.771 \pm 0.131	2.838 \pm 2.652	1.955 \pm 2.527	33.761 \pm 17.320
(Semi)SemiFCN (Baur et al., 2017)	0.685 \pm 0.281	2.622 \pm 3.054	2.151 \pm 3.387	29.876 \pm 16.960
(Semi)ASDNet (Nie et al., 2018)	0.617 \pm 0.149	4.226 \pm 3.127	3.201 \pm 3.10	38.608 \pm 16.390
DenseBiasNet (CE)	0.869 \pm 0.082	1.428 \pm 1.50	1.038 \pm 1.537	28.263 \pm 16.212
DenseBiasNet (CE+HRA)	0.872 \pm 0.095	1.440 \pm 1.486	1.006 \pm 1.495	27.268 \pm 15.681
(Semi)DPA-DenseBiasNet (CE+HRA)	0.884\pm0.076	1.385\pm1.489	0.940\pm1.529	25.439\pm15.710

465 typical subjects including two left and two right kidney ROIs. The blue boxes
in Case 4 indicate the high segmentation quality of artery endings brought by
our DenseBiasNet, and the other two models lose the thin endings without our
dense connectivity pattern. The yellow boxes in Case 2 show the high segmen-
tation quality of the singular regions brought by our DPA strategy. Without the
470 DPA features, 3D U-Net, ASDNet, and our DenseBiasNet lose these structures
in yellow boxes easily. Our proposed DPA-DenseBiasNet achieves complete and
continuous fine renal artery segmentation reaching the end of the interlobular
artery. As a semi-supervised method, ASDNet segments the spine and other ves-
sels as renal arteries leading to a severe mis-segmentation problem, as a result of
475 the inaccurate optimization target which is sensitive to thin structures. Mean-
while, 3D U-Net has much severe over-segmentation and under-segmentation
problems where part of the bifurcation regions are lost and some other vessels
are segmented as renal arteries.

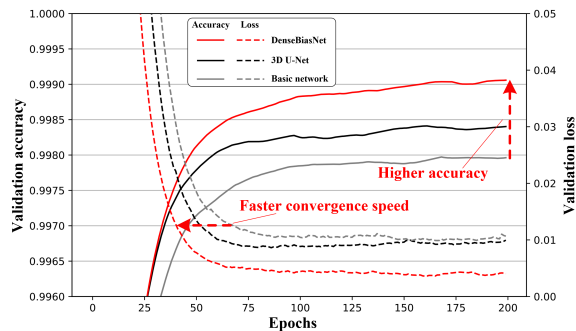


Figure 8: Our dense biased connection accelerates the training process and enhances accuracy. Our DenseBiasNet has faster convergence speed and higher accuracy than the other two networks owing to its dense connectivity pattern that keeps the integrity of the information flow and gradient flow.

4.2. Improvements analysis

480 **Dense biased connection training optimization.** As demonstrated in Fig. 8 and Fig. 10, the dense biased connection optimizes the training process by accelerating the training speed and enhancing the accuracy. The validation accuracy and loss value of DenseBiasNet, 3D U-Net and the basic network structure without dense biased connection are compared. Thanks to the dense connectivity
 485 pattern, the DenseBiasNet keeps the complete information flow and gradient flow so that it has higher accuracy and faster convergence speed than the other two methods. Compared with the basic network structure without dense biased connection, the 3D U-Net gets better training speed and accuracy owing to its partial skip connections. The DenseBiasNet also has higher accuracy than the
 490 other three networks (3D U-Net, V-Net, Kid-Net) according to the ROC and PR curves.

Low-level features enhancement. As is shown in Fig. 9, the low-level features are more sensitive to the arteries than high-level features in red boxes so that transmitting these features from each layer to every forward layer via
 495 our dense biased connection will enhance the sensitivity of the network for thin structures. Thus, the segmentation quality will be ensured not only in large

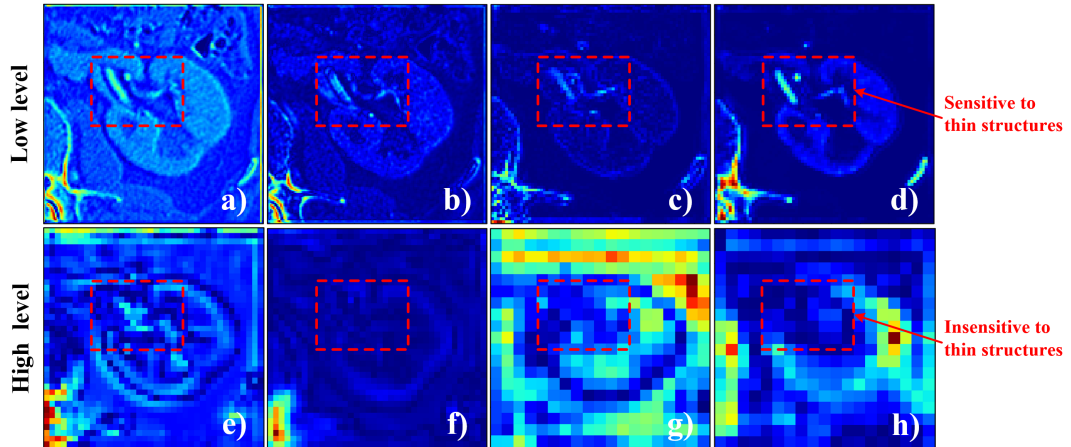


Figure 9: Low-level features are more sensitive to thin structures. a)-h) The feature maps transmitted in dense biased connection from the analysis path of DenseBiasNet with the hyper-parameter m of 1. a)-d) from the first two resolution stages represent the low-level features and e)-h) from the third and fourth stages represent high-level features. The red boxes show that the low-level features in our dense biased connection are more sensitive to arteries than high-level features.

regions but also in thin structures.

HRA for class balance. HRA balances the class dynamically thus preventing the inefficient training process and model degradation. As revealed in Fig. 11, it samples the loss calculation regions according to the segmentation quality to fit the actual situation during the training process. The loss value ratio of background and vessel is calculated and the standard cross-entropy (CE) loss is used as a comparison object. The ratio of the proposed HRA loss declines during the first 50 epochs which means the network focused on hard region gradually. Then the ratio is kept below 10 in the later stages of training indicating that the classes are kept balanced, while the standard CE loss's ratio fluctuates around 100 whose categories are seriously imbalanced which leads to slow convergence and loss of foreground.

As shown in Fig. 10, HRA increases the accuracy of CE loss function owing to the class balance it brings. The ROC and PR curve of the HRA loss cover the

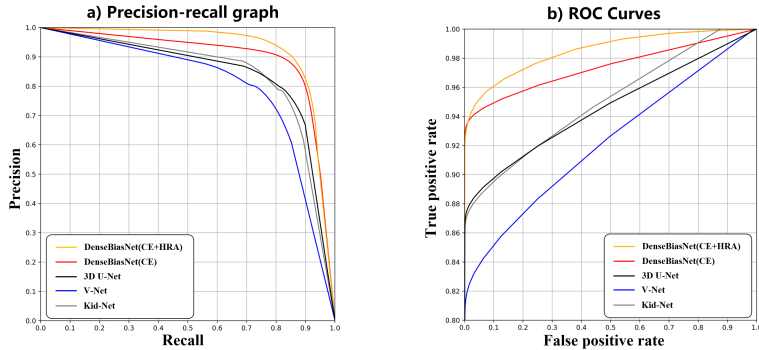


Figure 10: a)-b) The ROCs and PRs show that DenseBiasNet has higher segmentation accuracy than other networks and the HRA loss function improves the accuracy of CE.

standard CE which means the model trained via HRA has powerful performance. Due to the serious class imbalance in our task, the true positive rate will rise rapidly above 0.8 when plotting the ROC curve. For better demonstration, we show the ROC curves with a true positive rate between 0.8 and 1.

515 ***Effectiveness of DPA on smaller datasets.*** As shown in Tab. 2, the DPA for semi-supervised learning greatly improves the model performance especially when the dataset is small. We randomly downsample the training set to 50%, 20% and 10% of the original, and use them to evaluate the effectiveness of DPA on smaller datasets. As the amount of labeled images decreases, the enhance-
 520 ment ability of DPA is gradually reflected that DPA-DenseBiasNet improves the mean Dice coefficient by 7.1% over DenseBiasNet when data size is reduced to 10%.

4.3. Hyper-parameter analysis

As shown in Tab. 3, we performed ablation experiments on the bias quantity
 525 (m) in dense biased connection and the threshold (T) in HRA to analyze the effects of these hyper-parameters. We tried two options for the bias quantity: 1) All layers in the network share the same bias quantity. 2) The bias quantity is a proportion of the number of feature maps that output from each layer (r).

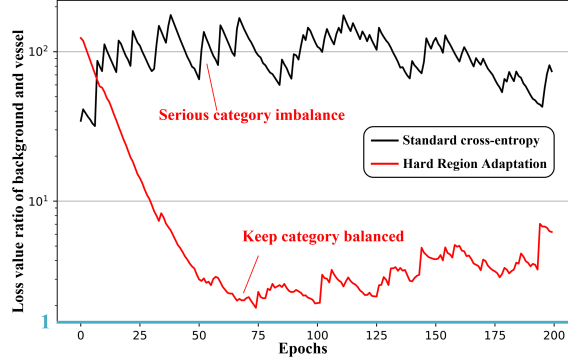


Figure 11: HRA keeps categories balanced thus preventing the inefficient training process and model degradation. The loss value ratio of background and vessel during the training process shows that the HRA loss keeps the ratio below 10 in the later stages of training which keeps class balanced while the cross-entropy loss function’s ratio fluctuates around 100.

Table 2: The DPA for semi-supervised learning keeps the network performance as the dataset decreases. The table shows the Dice coefficient of DenseBiasNet and DPA-DenseBiasNet trained via different labeled images amount.

Amount	DenseBiasNet	DPA-DenseBiasNet
100%	0.872±0.095	0.884±0.076
50%	0.836±0.117	0.853±0.094
20%	0.757±0.146	0.785±0.129
10%	0.630±0.200	0.701±0.183

Table 3: Hyper-parameter analysis. The bias quantity (m) in dense biased connection and the threshold (T) in HRA are studied to analyze their effects.

Amount	m	T	Dice \pm std	Parameters
100%	1	0.1	0.861 \pm 0.091	1.301M
100%	2	0.1	0.857 \pm 0.089	1.405M
100%	4	0.1	0.872\pm0.095	1.612M
100%	4	0.2	0.848 \pm 0.103	1.612M
100%	0.1r	0.1	0.871 \pm 0.079	1.627M
100%	0.2r	0.1	0.839 \pm 0.119	2.102M

The results show that DenseBiasNet (CE+HRA) achieves the best performance and relatively small number of parameters when $m = 4$ and $T = 0.1$.
530

5. Discussion and conclusion

In this paper, we have proposed and validated the first effective semi-supervised framework, DPA-DenseBiasNet, for 3D fine renal artery segmentation. DPA-DenseBiasNet combines advantages of the DenseBiasNet, deep priori anatomy and hard region adaptation loss so that it achieves powerful generalization ability. Our dense biased connection method which builds the DenseBiasNet is presented to fuse multi-receptive field and multi-resolution features for handling intra-scale changes problem in our task. It also leads the implicit deep supervision which accelerates the training and enhances the accuracy. The DPA strategy is extracted from input images via a trained encoder and embedded in the DenseBiasNet for inter-anatomy variation problem. The HRA makes the network focus on hard regions and keeps class balanced during the training process dynamically.
540

The experiment verified that our DPA-DenseBiasNet with HRA loss has high accuracy compared to the real mask. Tab. 1 shows the high Dice coefficient (0.884), low centerline distance (1.385mm), low surface distance (0.940mm) and low Hausdorff distance (25.439mm) which has verified the effectiveness, robust-
545

ness, and consistency of our proposed framework compared with three advanced supervised methods (Çiçek et al., 2016; Milletari et al., 2016; Taha et al., 2018) and two advanced semi-supervised methods (Baur et al., 2017; Nie et al., 2018). Fig. 7 shows the visual superiority, our framework has the capability to finely segment the thin structures and has high segmentation quality of singular structures such as the case 2 which will help clinicians accurately and directly locate the blood-feeding region.

Clinically, as the first automatic 3D fine renal artery segmentation framework, DPA-DenseBiasNet greatly improves the efficiency of kidney pre-operative plan and reduces the cost of personalized medicine. Our work avoids the time-consuming and challenging renal artery mask delineation process, and provides renal artery masks for the pre-operative plan of LPN with segmental renal artery clamping (Shao et al., 2011, 2012), handling one of the key technical problems for this task. Therefore, with the increasing probability of kidney disease (Bray et al., 2018), more efficient and lower-cost medical services indirectly brought by our work will create greater survival opportunities and higher quality of life for patients.

The success of fine renal artery segmentation task provides a technical reference for similar tasks whose thin structures are important and labeled dataset is small. As shown in Tab. 2, the DPA strategy improves the segmentation performance more significantly with the decrease of labeled data. Therefore, this strategy will improve the performance of the segmentation task (Ge et al., 2019a; Wolterink et al., 2019; Yang et al., 2012) in which the anatomy variation is diverse, the labeling is difficult and the labeled dataset is small. The process of DPA strategy also will not introduce incorrect labels as optimization targets and thus contributes to a stable semi-supervised training strategy that is friendly to thin structures. Besides, the proposed HRA loss keeps class balanced, so that the model will pay attention to the hard-to-segment regions and the segmentation quality of thin structures will be improved.

Based on dense biased connection method, the proposed DenseBiasNet optimizes the training process via the complete information and gradient flow.

Fig. 8 demonstrates the acceleration of the training and the enhancement of
580 the accuracy brought by this novel connection method. As shown in Fig. 9,
more Low-level features are fused and transmitted to the deep layers to supplement the sensitivity of thin structures. The gradient flows directly optimize the shallow layers during back-propagation through the dense connectivity pattern, leading the implicit deep supervision (Lee et al., 2015). Therefore, the training
585 process is greatly optimized and the segmentation quality is also improved. In addition, owing to the advantages of the complete information and gradient flow, this network structure is expected to serve as a new backbone network for other deep learning-based tasks such as quantization (Xue et al., 2017; Afshin et al., 2012; Ge et al., 2019b), super resolution (Ge et al., 2019c), descriptor
590 learning (Zhen et al., 2015), etc.

Our DenseBiasNet greatly enhances the effectiveness in the segmentation of thin structures such as renal arteries, however, the parallelism of the computing process is weakened due to the full network dense connection (Ma et al., 2018). This limitation reduces the speed of our framework resulting in more processing
595 time compared with other frameworks of the same complexity. Fortunately, the fine renal artery segmentation task has low real-time requirements, so it is worth the extra processing time for higher accuracy.

The DPA strategy has enhanced the segmentation generalization performance of our framework, but its encoder brings a lot of parameters as an additional
600 module which will take up more computing resources. Fortunately, there are many lightweight models (Ma et al., 2018; Howard et al., 2017) and network slimming methods (Liu et al., 2017; Zhuang et al., 2018) have been applied to reduce framework parameters successfully. Therefore, further research on the high-performance lightweight fine renal artery segmentation framework for
605 landing is one of our important directions. Besides, using powerful computing equipment in clinical applications is also an effective solution for this limitation.

Acknowledgments

This research was supported by the National Natural Science Foundation under grants (31571001, 61828101,61876037,31800825), the Short-Term Recruitment Program of Foreign Experts (WQ20163200398), Key Research and Development Project of Jiangsu Province (BE2018749), Southeast University-Nanjing Medical University Cooperative Research Project (2242019K3DN08) and Excellence Project Funds of Southeast University.

References

- 615 Afshin, M., Ayed, I.B., Islam, A., Goela, A., Peters, T.M., Li, S., 2012. Global assessment of cardiac function using image statistics in mri, in: Ayache, N., Delingette, H., Golland, P., Mori, K. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 535–543.
- 620 Bai, W., Oktay, O., Sinclair, M., Suzuki, H., Rajchl, M., Tarroni, G., Glocker, B., King, A., Matthews, P.M., Rueckert, D., 2017. Semi-supervised learning for network-based cardiac mr image segmentation, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 253–260.
- 625 Bai, W., Suzuki, H., Qin, C., Tarroni, G., Oktay, O., Matthews, P.M., Rueckert, D., 2018. Recurrent neural networks for aortic image sequence segmentation with sparse annotations, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 586–594.
- Baur, C., Albarqouni, S., Navab, N., 2017. Semi-supervised deep learning for fully convolutional networks, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 311–319.
- 630 Boskamp, T., HAHN, H.K., Hindennach, M., ZIDOWITZ, S., PEITGEN, H.O., OELTZE, S., PREIM, B., 2005. Geometrical and structural analysis of vessel

- systems in 3d medical image datasets, in: *Medical Imaging Systems Technology: Volume 5: Methods in Cardiovascular and Brain Systems*. World Scientific, pp. 1–60.
- 635
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A., 2018. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* 68, 394–424.
- 640
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3d u-net: learning dense volumetric segmentation from sparse annotation, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 424–432.
- 645
- Gao, H., Zhuang, L., Weinberger, K.Q., 2017. Densely connected convolutional networks, in: *IEEE Conference on Computer Vision and Pattern Recognition*.
- Ge, R., Yang, G., Chen, Y., Luo, L., Feng, C., Ma, H., Ren, J., Li, S., 2019a. K-net: Integrate left ventricle segmentation and direct quantification of paired echo sequence. *IEEE transactions on medical imaging* .
- 650
- Ge, R., Yang, G., Chen, Y., Luo, L., Feng, C., Zhang, H., Li, S., 2019b. Pv-lvnet: Direct left ventricle multitype indices estimation from 2d echocardiograms of paired apical views with deep neural networks. *Medical image analysis* 58, 101554.
- 655
- Ge, R., Yang, G., Xu, C., Chen, Y., Luo, L., Li, S., 2019c. Stereo-correlation and noise-distribution aware resvoGAN for dense slices reconstruction and noise reduction in thick low-dose CT, in: Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.T., Khan, A. (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, Springer International Publishing, Cham. pp. 328–338.
- 660
- Gu, X., Angelov, P.P., 2018. Semi-supervised deep rule-based approach for image classification. *Applied Soft Computing* 68, 53–68.

- Hady, M.F.A., Schwenker, F., 2013. Semi-supervised learning, in: Handbook on Neural Information Processing. Springer, pp. 215–239.
- Han, W., Feng, R., Wang, L., Cheng, Y., 2018. A semi-supervised generative framework with deep learning features for high-resolution remote sensing image scene classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 145, 23–43.
- He, Y., Yang, G., Chen, Y., Kong, Y., Wu, J., Tang, L., Zhu, X., Dillenseger, J.L., Shao, P., Zhang, S., et al., 2019. Dpa-densebiasnet: Semi-supervised 3d fine renal artery segmentation with dense biased network and deep priori anatomy, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 139–147.
- Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 .
- Hung, W.C., Tsai, Y.H., Liou, Y.T., Lin, Y.Y., Yang, M.H., 2018. Adversarial learning for semi-supervised semantic segmentation. arXiv preprint arXiv:1802.07934 .
- Kingma, D.P., Mohamed, S., Rezende, D.J., Welling, M., 2014. Semi-supervised learning with deep generative models, in: Advances in neural information processing systems, pp. 3581–3589.
- Lee, C.Y., Xie, S., Gallagher, P., Zhang, Z., Tu, Z., 2015. Deeply-supervised nets, in: Artificial intelligence and statistics, pp. 562–570.
- Lesage, D., Angelini, E.D., Bloch, I., Funka-Lea, G., 2009. A review of 3d vessel lumen segmentation techniques: Models, features and extraction schemes. *Medical image analysis* 13, 819–845.
- Li, J., Lo, P., Taha, A., Wu, H., Zhao, T., 2018a. Segmentation of renal structures for image-guided surgery, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 454–462.

- 690 Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A., 2018b. H-denseunet:
hybrid densely connected unet for liver and tumor segmentation from ct vol-
umes. *IEEE transactions on medical imaging* 37, 2663–2674.
- Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal loss for dense
object detection, in: *Proceedings of the IEEE international conference on*
695 *computer vision*, pp. 2980–2988.
- Liu, T.Y., 2009. Easyensemble and feature selection for imbalance data sets, in:
2009 International Joint Conference on Bioinformatics, Systems Biology and
Intelligent Computing, IEEE. pp. 517–520.
- Liu, Z., Li, J., Shen, Z., Huang, G., Yan, S., Zhang, C., 2017. Learning efficient
700 convolutional networks through network slimming, in: *Proceedings of the*
IEEE International Conference on Computer Vision, pp. 2736–2744.
- Ljungberg, B., Bensalah, K., Canfield, S., Dabestani, S., Hofmann, F., Hora,
M., Kuczyk, M.A., Lam, T., Marconi, L., Merseburger, A.S., et al., 2015.
Eau guidelines on renal cell carcinoma: 2014 update. *European urology* 67,
705 913–924.
- Ma, N., Zhang, X., Zheng, H.T., Sun, J., 2018. Shufflenet v2: Practical guide-
lines for efficient cnn architecture design, in: *Proceedings of the European*
Conference on Computer Vision (ECCV), pp. 116–131.
- Masci, J., Meier, U., Cireşan, D., Schmidhuber, J., 2011. Stacked convolutional
710 auto-encoders for hierarchical feature extraction, in: *International Conference*
on Artificial Neural Networks, Springer. pp. 52–59.
- Metz, C., Schaap, M., Van Der Giessen, A., Van Walsum, T., Niessen, W.,
2007. Semi-automatic coronary artery centerline extraction in computed to-
mography angiography data, in: *2007 4th IEEE International Symposium on*
715 *Biomedical Imaging: From Nano to Macro*, IEEE. pp. 856–859.

- Mille, J., Cohen, L.D., 2009. Deformable tree models for 2d and 3d branching structures extraction, in: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, IEEE. pp. 149–156.
- Milletari, F., Navab, N., Ahmadi, S.A., 2016. V-net: Fully convolutional neural networks for volumetric medical image segmentation, in: 2016 Fourth International Conference on 3D Vision (3DV), IEEE. pp. 565–571.
- Moccia, S., De Momi, E., El Hadji, S., Mattos, L.S., 2018. Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics. *Computer methods and programs in biomedicine* 158, 71–91.
- Nie, D., Gao, Y., Wang, L., Shen, D., 2018. Asdnet: Attention based semi-supervised deep networks for medical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 370–378.
- Noothout, J.M., de Vos, B.D., Wolterink, J.M., Išgum, I., 2018. Automatic segmentation of thoracic aorta segments in low-dose chest ct, in: Medical Imaging 2018: Image Processing, International Society for Optics and Photonics. p. 105741S.
- Pelt, D.M., Sethian, J.A., 2018. A mixed-scale dense convolutional neural network for image analysis. *Proceedings of the National Academy of Sciences* 115, 254–259.
- Petru, B., Elena, Ș., Dan, I., Klara, B., Radu, B., Constantin, D., 2012. Morphological assessments on the arteries of the superior renal segment. *Surgical and radiologic anatomy* 34, 137–144.
- Porpiglia, F., Fiori, C., Checcucci, E., Amparore, D., Bertolo, R., 2018. Hyperaccuracy three-dimensional reconstruction is able to maximize the efficacy of selective clamping during robot-assisted partial nephrectomy for complex renal masses. *European urology* 74, 651–660.

- 745 Qiao, S., Shen, W., Zhang, Z., Wang, B., Yuille, A., 2018. Deep co-training for semi-supervised image recognition, in: Proceedings of the European Conference on Computer Vision (ECCV), pp. 135–152.
- Rasmus, A., Berglund, M., Honkala, M., Valpola, H., Raiko, T., 2015a. Semi-supervised learning with ladder networks, in: Advances in neural information processing systems, pp. 3546–3554.
- Rasmus, A., Valpola, H., Raiko, T., 2015b. Lateral connections in denoising
750 autoencoders support supervised learning. arXiv preprint arXiv:1504.08215 .
- Sedai, S., Mahapatra, D., Hewavitharanage, S., Maetschke, S., Garnavi, R., 2017. Semi-supervised segmentation of optic cup in retinal fundus images using variational autoencoder, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 75–82.
- 755 Shao, P., Qin, C., Yin, C., Meng, X., Ju, X., Li, J., Lv, Q., Zhang, W., Xu, Z., 2011. Laparoscopic partial nephrectomy with segmental renal artery clamping: technique and clinical outcomes. *European urology* 59, 849–855.
- Shao, P., Tang, L., Li, P., Xu, Y., Qin, C., Cao, Q., Ju, X., Meng, X., Lv, Q., Li, J., et al., 2012. Precise segmental renal artery clamping under the guidance
760 of dual-source computed tomography angiography during laparoscopic partial nephrectomy. *European urology* 62, 1001–1008.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* 15, 1929–1958.
- 765 Taha, A., Lo, P., Li, J., Zhao, T., 2018. Kid-net: convolution networks for kidney vessels segmentation from ct-volumes, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 463–471.
- Taha, A.A., Hanbury, A., 2015. Metrics for evaluating 3d medical image seg-
770 mentation: analysis, selection, and tool. *BMC medical imaging* 15, 29.

- Tang, Y., Wang, J., Wang, X., Gao, B., Dellandréa, E., Gaizauskas, R., Chen, L., 2018. Visual and semantic knowledge transfer for large scale semi-supervised object detection. *IEEE transactions on pattern analysis and machine intelligence* 40, 3045–3058.
- 775 Toledo, R., Orriols, X., Binefa, X., Radeva, P., Vitria, J., Villanueva, J.J., 2000. Tracking elongated structures using statistical snakes, in: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, IEEE. pp. 157–162.
- Tschirren, J., Hoffman, E.A., McLennan, G., Sonka, M., 2005. Intrathoracic
780 airway trees: segmentation and airway morphology analysis from low-dose ct scans. *IEEE transactions on medical imaging* 24, 1529.
- Tyrrell, J.A., di Tomaso, E., Fuja, D., Tong, R., Kozak, K., Jain, R.K., Roysam, B., 2007. Robust 3-d modeling of vasculature imagery using superellipsoids. *IEEE transactions on medical imaging* 26, 223–237.
- 785 Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.A., 2010. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research* 11, 3371–3408.
- Weston, J., Ratle, F., Mobahi, H., Collobert, R., 2012. Deep learning via semi-
790 supervised embedding, in: *Neural Networks: Tricks of the Trade*. Springer, pp. 639–655.
- Wolterink, J.M., van Hamersvelt, R.W., Viergever, M.A., Leiner, T., Išgum, I., 2019. Coronary artery centerline extraction in cardiac ct angiography using a cnn-based orientation classifier. *Medical image analysis* 51, 46–60.
- 795 Wolterink, J.M., Leiner, T., de Vos, B.D., van Hamersvelt, R.W., Viergever, M.A., Išgum, I., 2016. Automatic coronary artery calcium scoring in cardiac ct angiography using paired convolutional neural networks. *Medical image analysis* 34, 123–136.

- 800 Wu, C., Wu, F., Wu, S., Yuan, Z., Liu, J., Huang, Y., 2019. Semi-supervised dimensional sentiment analysis with variational autoencoder. *Knowledge-Based Systems* 165, 30–39.
- Wu, Y., He, K., 2018. Group normalization, in: *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3–19.
- 805 Xiao, Y., Wu, J., Lin, Z., Zhao, X., 2018. A semi-supervised deep learning method based on stacked sparse auto-encoder for cancer prediction using rna-seq data. *Computer methods and programs in biomedicine* 166, 99–105.
- Xue, W., Nachum, I.B., Pandey, S., Warrington, J., Leung, S., Li, S., 2017. Direct estimation of regional wall thicknesses via residual recurrent neural network, in: Niethammer, M., Styner, M., Aylward, S., Zhu, H., Oguz, I., Yap, P.T., Shen, D. (Eds.), *Information Processing in Medical Imaging*, Springer International Publishing, Cham. pp. 505–516.
- 810 Yang, G., Gu, J., Chen, Y., Liu, W., Tang, L., Shu, H., Toumoulin, C., 2014. 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, IEEE*. pp. 5538–5541.
- 815 Yang, G., Kitslaar, P., Frenay, M., Broersen, A., Boogers, M.J., Bax, J.J., Reiber, J.H.C., Dijkstra, J., 2012. Automatic centerline extraction of coronary arteries in coronary computed tomographic angiography. *International Journal of Cardiovascular Imaging* 28, 921–933.
- 820 Yang, M., Yu, K., Zhang, C., Li, Z., Yang, K., 2018. Denseaspp for semantic segmentation in street scenes, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3684–3692.
- Zhang, S., Yang, G., Tang, L., Lv, Q., Li, J., Xu, Y., Zhu, X., Li, P., Shao, P., Wang, Z., 2019. Application of a functional3-dimensional perfusion model in laparoscopic partial nephrectomy with precise segmental renal artery clamp-
825 ing. *Urology* 125, 98–103.

- Zhen, X., Wang, Z., Yu, M., Li, S., 2015. Supervised descriptor learning for multi-output regression, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1211–1218.
- ⁸³⁰ Zhuang, Z., Tan, M., Zhuang, B., Liu, J., Guo, Y., Wu, Q., Huang, J., Zhu, J., 2018. Discrimination-aware channel pruning for deep neural networks, in: Advances in Neural Information Processing Systems, pp. 875–886.