Tubular and Solid Organ Segmentation Methods and Their Applications to Computer-Aided Kidney Diagnosis

Chenglong Wang

Abstract

Rapid advances in medical image-processing technology have brought the goal of fully automatic computer-aided diagnosis (CAD) systems. With the assists of computers, doctors can obtain precise diagnostic information efficiently. An accurate and reliable CAD system is the basis for realizing this vision. CAD system is composed of a lot of medical image-processing techniques, such as image segmentation, detection, classification and registration. Always, a combination of these processing techniques is required to constitute a CAD system for one specific clinical problem. Among these techniques, segmentation has always been a fundamental part of CAD system. Segmentation is the process of extracting meaningful part from an image. In medical imaging-processing field, the objects of segmentation often correspond to different tissues, organs, or other anatomically relevant structures.

Due to the complexity of human body, segmentation of different anatomical structures may have different factors need to be considered when designing the segmentation methods. Different structures have different appearances in medical imaging. For instance, blood vessels have a bright elongated appearance in contrast-enhanced CT scans, while bronchi are shown in a dark luminal appearance. This thesis mainly focuses on tubular organ and solid organ segmentation. Tubular organ segmentation methods can be used to extract blood vessels and bronchi, solid organ segmentation method can be used to extract abdominal solid organs such as kidneys, livers. To improve the segmentation accuracy, the proposed methods put attention on capturing geometrical and anatomical information for tubular organs and solid organs.

This thesis presents three segmentation methods and two clinical applications to

kidney-related diagnosis. They are categorized into three research topics.

The first research topic presents two segmentation methods designed for tubular organs. These two segmentation methods for tubular structures use different methodologies. One is a traditional image-processing based method, and the other is modern supervised deep-learning based method. Both of these two methods attempt to exploit the geometrical information in their own methodologies. Experimental results showed that the proposed methods outperformed comparative methods in blood vessels and bronchi segmentation tasks.

The second research topic focuses on solid organ segmentation. In this research, deep-learning based method is adopted to segment the organs. To better capture the anatomical information, a spatially aware unit is introduced to explicitly encode the spatial location information. By using this auxiliary module, the proposed network is able to achieve competitive performance compared to other methods with larger training data.

The third research topic presents two clinical applications using the proposed segmentation methods. In this topic, two kidney-related diagnosis applications are demonstrated to evaluate the clinical value of our CAD techniques. The first application presents presurgical simulation method for partial nephrectomy, the second application investigates the renal function variation between pre- and post-operation. Although, large-scale clinical validations have not been conducted yet. From physician's feedback, the proposed CAD systems indeed helped them in real clinical use.

In summary, this thesis presents segmentation methods for both tubular and solid organs. The key point of the proposed methods is capturing the geometrical and anatomical information for specific organs. Clinical applications are also demonstrated to show the clinical value of the proposed CAD systems. Chapter 1 provides an introduction, background and brief overview of each research topic. In Chapter 2, 3, 4 and 5, detailed descriptions of each research topic are given. Finally, in Chapter 6, this thesis is summarized and further research is discussed.

Acknowledgements

This thesis is formally submitted for fulfilling partial requirements for the degree of Doctor of Information Science from Nagoya University. This work would not be accomplished without helps of many people.

Firstly, I would give my thanks to my supervisor, Prof. Dr Kensaku Mori. He gave me the valuable chance to study in this laboratory. Under his supervision, I can finish this work. Not only academically, but also helped throughout my whole time in Japan. His insights in our research area helped me sovle a lot of difficult problems. I cannot attain such excellent academic achievements without his supervision. Prof. Mori is the director of our laboratory. Under his management, our laboratory has an active academic atmosphere, which provides a better academic environment for students.

I would also like to thank Dr Masahiro Oda. Very grateful for his everyday supervision and assistance. Thanks so much for his contribution on revising every paper I have written during my Ph.D course. I greatly appreciate his patience pointing out my mistakes. Under his supervision, my scientific writing skills have significantly improved.

I would like to thank Dr Yuichiro Hayashi for his valuable advises and helps. Academic discussions with him provide new research direction for me. I am extremely grateful for his contributions on my research.

Secondly, I would like to thank people outside of our laboratory. Dr Tokunori Yamamoto and Dr Yasushi Yoshino from Nagoya University Graduate School of Medicine. They are my collaborators throughout my whole research. They pro-

posed a lot of new ideas and helped me from medical aspect. I would also give great thanks to their hard work on preparing medical data.

Prof. Dr Takayuki Kitasaka from Aichi Institute of Technology has given a lot of valuable advises on my research topics. I am really grateful to him for attending the group meeting every week to have a deep discussion on everyone's research progress.

I would also like to express my gratitude to Prof. Dr Hiroshi Murase and Prof. Dr Daisuke Deguchi not only for the advises on my research but also their valuable feedback for this thesis.

I would also like to thank Prof. Dr Kenji Mase and Prof. Dr Hiroaki Kudo for their feedback and insightful comments during the review process.

Last, I would give my special thanks to all members of the Mori Laboratory, who have helped me a lot on both research and daily life. Thanks to all of the students and secretaries for their kind help and support. Their help is just as important as others.

Sincere thanks to my parents

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

With the rapid development of image-processing technology, more and more novel applications have been proposed which significantly improve the social efficiency. For instance, face recognition technique can be used for more convenient security check, scene recognition can be foundation for autonomous driving system, and anomaly detection can be used for video monitoring. Rapid advances in the medical image-processing field also have changed the medical world. Medical image-processing techniques remarkably improve diagnostic efficiency. For example, accurate pulmonary nodule detection helps clinicians to find the lung nodule more easily, and also enabling rapid improvement in early detection of lung cancer. Blood vessel segmentation enables vascular quantitative analysis. Surgery navigation helps physicians to better understand surgical scene, which can improve surgical safety.

This thesis mainly discusses the contribution of medical image-processing techniques in computer-aided kidney diagnosis. In this chapter, I will give a brief introduction of the background this research in Section 1.1 including the aim of this research and overview of computer-aided diagnosis (CAD) in Section 1.1.1 and 1.1.2.

1.1 Background

In this section, I will briefly introduce the aim of our research and literature overview of computer aided diagnosis (CAD). This research is mainly focus on tubular and

solid organ segmentation problems. Specifically, these approaches are applied to computer-aided kidney diagnosis.

1.1.1 Aim of this study

This research aims to proposed tubular (blood vessel) and organ (kidney) segmentation methods with higher accuracy which meet the real clinical demand. By using the proposed methods, it is possible to contribute to realize reliable and quantitative computer-aided kidney diagnosis techniques.

Medical doctors have benefited from the development of image-processing techniques, we called the digital image-processing techniques that assist doctors in the interpretation of medical images, computer-aided diagnosis (CAD). Tracing back to the late 1950s, CAD have over 50 years history used for clinical applications. The basis of CAD system is pattern recognition techniques. Here we list several major categories of CAD system: preprocessing, detection, segmentation, classification, and registration. CAD system has been used in many diagnosis fields, such as breast, lung, prostate and vasculature. We give a more detailed introduction in the next section. The performance of pattern recognition techniques will directly affect the accuracy of CAD system. Improving accuracy of image-processing methods has always been one of our research topics.

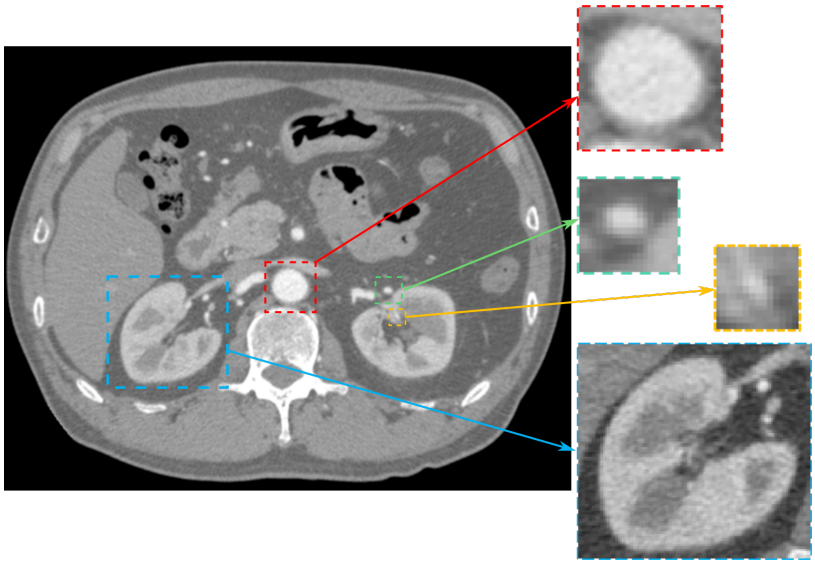
This thesis mainly focuses on computer-aided kidney diagnosis. The kidneys are significantly important solid organs in vertebrates. They are responsible for filtering impurities in the blood, maintaining body water balance and electrolyte balance. According to the statistics from the national kidney foundation, 10% of the world population is suffering from kidney diseases. Every year, millions of patients died because of unaffordable treatments [1]. The development of computer-aided kidney diagnosis become a pressing need. Two main components for analyzing kidney are renal artery and solid kidney tissue. Therefore, the segmentation of renal artery and kidneys become the fundamental tasks for building the computer-aided kidney

diagnosis systems.

1.1.2 Computer-aided diagnosis

Computer-aided diagnosis (CAD) and Computer-aided surgery (CAS) are two major research subjects in medical image-processing field. The development of CAD system can be traced back to early 1960s [3–6]. With the rapid development of image-processing technology, more and more image-processing approaches have incorporated in this filed to realize an accurate and reliable diagnostic and surgical system. The contributions from CAD and CAS have significantly improved people's medical level. The 21st century is the era of big data. Especially in hospital, such as ultrasound scans, computed tomography (CT) scans, and magnetic resonance imaging (MRI) scans have become common tools for baseline diseases diagnosis. One example of 3D contrast-enhanced CT (CE-CT) data is shown in Fig. 1.1. Manual diagnosis of these large amount of data has already become a worldwide problems. With the power of computers, CAD and CAS can provide efficient diagnosis for clinicians. For examples, clinicians and radiologists can immediately obtain the computerized diagnosis results as "second option" and make their final decisions by using CAD systems. One CAD system for diagnosing colorectal lesions with endocytoscopic is shown in Fig. 1.2. Endocytoscopic vascular patterns are automatically classified by machine-learning based method [2]. Physicians can better understand the surgical scene in minimally invasive surgery with the navigation information provided by CAS system. A lot of research suggests that CAD can help clinicians improve the diagnostic performance in real medical applications [7–9].

One major subject of CAD is lesion detection from medical images. Finding small lesions in large 3D medical images is a difficult and time-consuming task. By using the computer-aided detection system, radiologists can quickly localize the lesions and even able to obtain quantitative analysis results. For example, there has been proposed a lot of breast cancer detection research in the literature [10–12]. Detection



(a) One CE-CT slice

(b) 3D volume rendering of CT data

Figure 1.1: One example of 3D contrast-enhanced CT scans

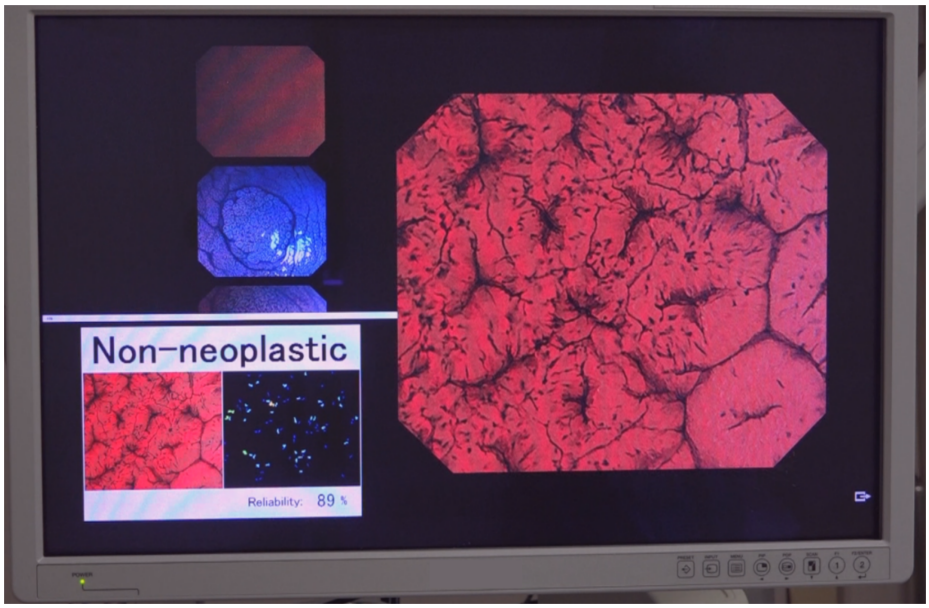


Figure 1.2: CAD system for diagnosing colorectal lesions with endocytoscopic images. Photo is based on material from [2].

of lung nodule is also spotted task in this field [13–16]. Recently, not only these traditional lesion detection tasks but also new detection tasks were introduced with the development of medical imaging technology and new demands. Such as polyp detection in endoscopy videos [17, 18].

Another major subject of CAD is disease diagnosis. With the rapid development of machine learning techniques, computer-aided diagnosis can take full advantage of big-data. Diagnosis often follows after a lesion detection. Traditional machine-learning based methods have been proposed such as Linear Discriminant Analysis (LDA) based methods [19–21], support vector machines (SVMs) based methods [22, 23], and other methods such as genetic algorithm [24], random forest algorithm [25]. Recently, deep learning techniques also have been incorporated for diagnosis tasks. Convolutional neural networks (CNNs) have been used in a lot of related research [26–28].

Segmentation is another important subject in the CAD field, which is also main research topic of this thesis. Generally speaking, segmentation aims to extract target object from images. In medical image-processing field, these target objects can be

organs, tissues, bones, and other body structures. In this thesis, segmentation denotes pixel-/voxel-wise segmentation, which is different with detection tasks. We will give a more detailed introduction of segmentation in Section 1.2.

Besides aforementioned tasks, other research topics also have attracted much attention in CAD field. Such as image registration [29, 30], disease prediction [31–33], and medical reports generation [34, 35]. As we can see from these tasks which cover almost tasks required by clinicians and physicians, CAD techniques aims to make contributions to a better medical environment.

1.2 Segmentation for CAD

In this section, we will give an introduction about segmentation for CAD systems. Segmentation plays a fundamental role in CAD field. Many research subjects are based on segmentation results. Image segmentation has a long history in computer vision field tracing back to 1970s [36, 37]. Image segmentation is a process of dividing pixels/voxels of image into distinct groups. Classical approaches such as region growing [38] and active contour model (snakes) [39] have been early proposed for image segmentation tasks. Based on these fundamental works, more and more excellent approaches have been proposed. Such as geodesic snakes is proposed based on traditional snakes [40], solving the non-intrinsic problem of traditional snakes. Furthermore, inspired by snakes, novel methods such as level-set method (LSM) [41] and gradient vector field (GVF) method [42] were proposed. These methods were not only used for natural images but also medical images. Malladi *et al.* firstly incorporated LSM to segment organs from CT images [43]. Snakes-based approaches also have been used to extract organs [44, 45].

Four major image-related difficulties in medical image segmentation: **1)** *noise*, is mostly introduced at medical imaging processes, such as ring artifact and metal artifact. These noises can alter image intensity and affect segmentation performances.

2) *image non-uniformity*, is mainly caused by different imaging hardware and configurations. Different pixel/voxel sizes, reconstruction functions and scanning time will lead to different image appearances. This non-uniformity will definitely increase the difficulty of segmentation. **3)** *poor imaging quality*, is a fundamental problem in medical image processing. Due to the limitation of imaging techniques, it is impossible to obtain ideal images for specific processing tasks. For example, segmentation for MRI images is difficult due to its low image resolution. Although, CT scans have much higher resolution than MRI, organ segmentation still need to face new problems, such as angiography. Imperfect angiography will directly increase segmentation difficulty such as blood vessel segmentation and pancreas segmentation. **4)** *uncertain boundary*, is mainly caused by complex human anatomical structure. It is difficult to segment adjacent tissues when they have similar intensity. In these cases, other prior information would play an important role in segmentation task, such as shape information and spatial information.

1.2.1 Geometrical and anatomical information

As mentioned in earlier, medical image segmentation remains a difficult topic in image processing field. A lot of efforts have been made in this field to contribute to more accurate and reliable CAD approaches. At early stage, many low-level image process techniques were utilized for segmentation tasks. Most of them are realized by applying filters to images. Those methods are also called heuristic method. Such as thresholding, region growing [46], and region split/merge [47].

Toward further improvement of segmentation accuracy, more abundant features and more better representations are needed. Therefore, high-level approaches have been proposed to achieve better performance. As previously mentioned, deformable models have been successfully applied to medical image-processing field. Such as LSM [43] and GVF [42] have achieved good performance. Considering the particularity of medical image-processing tasks, there have been proposed many specific approaches

designed for medical tasks. Such as statistical shape models (SSMs) [48], atlas-based methods [49] and rule-based methods [50]. SSMs is able to represent a large range of expected and evidence-based variation based on average shape calculated from large training dataset. Since human organs share similar shapes. SSMs is quite effective in organ segmentation problem. Taking advantage of anatomical information, atlas-based methods are able to capture spatial information of organs. Similarly, rule-based methods also use the prior-knowledge of anatomical information. Besides these introduced methods, many other methods also aim to better represent geometrical and anatomical information of specific medical images. Such as model-fitting based methods [51], guided active contour model [52], skeleton representation [53].

The above-mentioned methods mainly focus on capture geometrical and anatomical information from medical images. They are two important features for medical image-processing problems. Our research mainly focuses on tubular and solid organ segmentation tasks which have attracted a lot of attention in CAD field. In the next section, we will give a brief introduction of blood vessel and organ segmentation methods using geometrical and anatomical information.

1.2.2 Blood vessel segmentation

Vascular disease is one of the most urgently need to be resolved health issues around world. A large amount of research has been done in vascular related research topics. Among these topics, blood vessel segmentation is indeed a fundamental task. Based on the segmentation of vascular structures, it will be possible for further vascular analysis such as diagnosis assistance and surgical planning. One example of annotated blood vessel is rendered in Fig. 1.3.

However, automated and complete segmentation of intricate tubular structures remains a major challenge in medical image-processing field. As we mentioned in Section 1.2, three general difficulties of medical image processing. The tiny structures of blood vessel may even amplify these negative effects. Noises, uniformity, and image

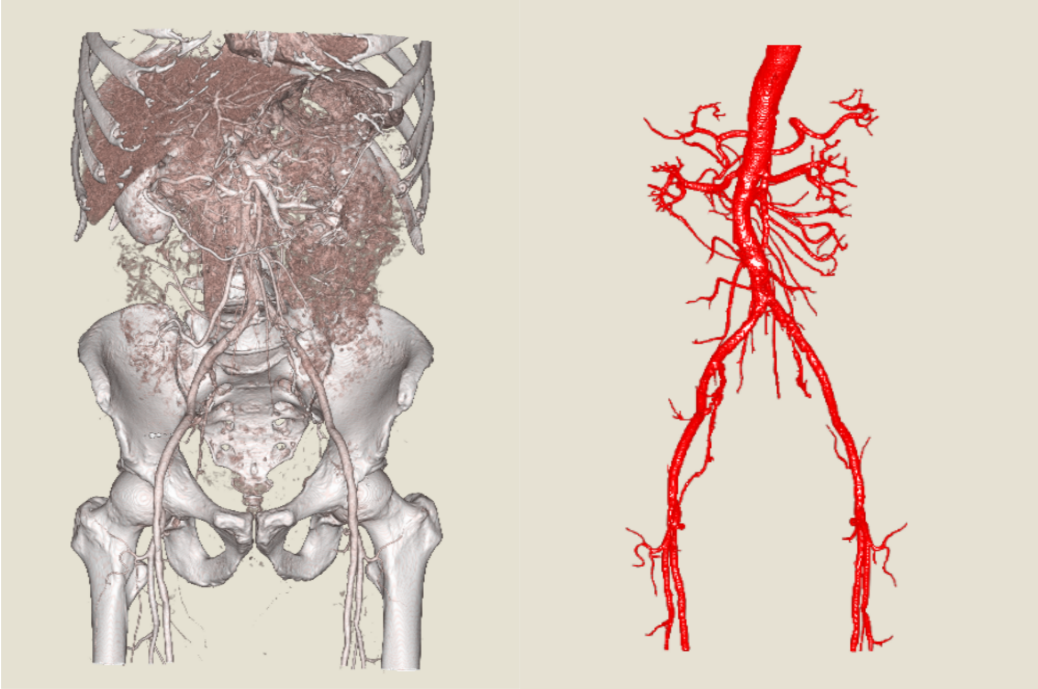


Figure 1.3: One example of abdominal blood vessels. Left is volume rendering of CE-CT data. Right is volume rendering of annotated label of blood vessel.

quality problems may have larger effects on blood vessels than normal solid organs.

A lot of works have been done in blood vessel segmentation problem using various approaches [54]. In this section, we mainly focus on methods using geometrical information. Because we think in blood vessel segmentation tasks, geometrical information is more important than anatomical information. Bright tubular structure is the key characteristic of blood vessel observed in medical images. This feature can be effectively captured by geometric-based methods.

In this thesis, we decide to category blood vessel segmentation methods based on the feature representation manner: explicit model and implicit model. Explicit model denotes explicitly generation of vascular models for searching blood vessels. Implicit model means encoding the geometrical features from images for further analysis.

1.2.2.1 Explicit vascular models

Explicitly generation of vascular models is one major methodology of geometric-based methods. de Bruijne *et al.* utilized active shape models (ASMs) to learn tubular

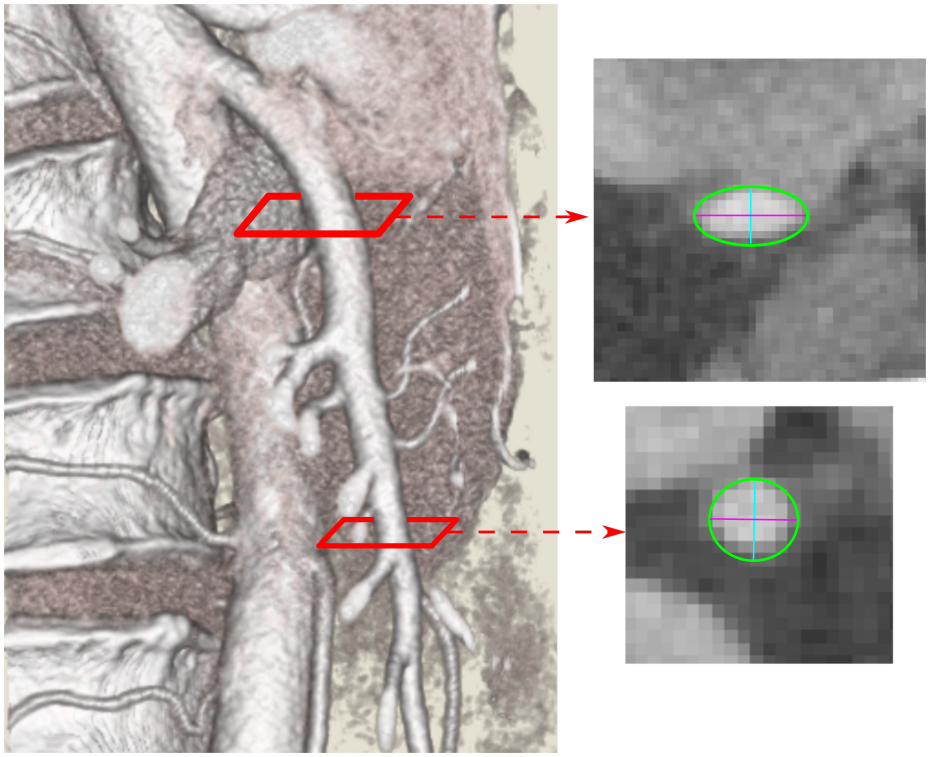


Figure 1.4: Illustration of 2D cross-section based blood vessel segmentation. Given the centerline, blood vessel segmentation is performed by using 2D snakes. Illustration is based on material from [56].

shape from vascular structures of abdominal aortic aneurysms [55]. Both shape and appearance models of ASM are modified for a better shape approximation of aortic aneurysms. The ASM is considered as a variation of active contour model. The ASM incorporates model-based regularity into active contour model.

Active contour models are also used to address tubular structure segmentation problem. In contrast to the original active contour model [39] which was designed for general purpose, in tubular structure segmentation problem, tubular geometrical information is taken into consideration [57–60]. Derived from active contour model, level-set method (LSM) has also been adopted to vessel segmentation problem [61, 62].

Many parametric methods have been investigated on vascular centerline. Frangi *et al.* created B-spline tensor surfaces to model blood vessel segments with a central vessel axis curve coupled to a vessel wall surface [63]. B-spline algorithm was also

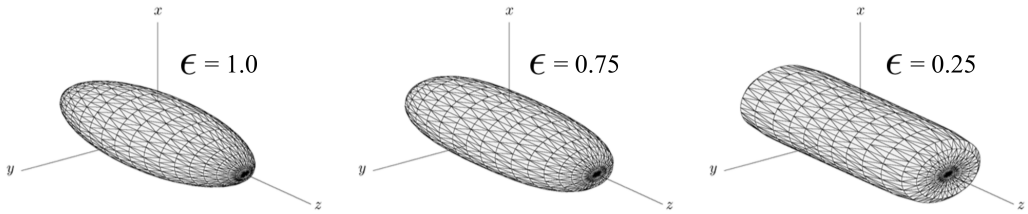


Figure 1.5: Parametric superellipsoids model proposed by Tyrrell *et al.*. ϵ is a shape parameter controlling the model shape. Figure is based on material from their work [69].

widely used for approximate vascular centerline in other works [64, 65].

Along the vascular centerline, 2D cross-section models are generated to segment the blood vessel step-by-step [56, 66]. Example of 2D cross-section method [56] is illustrated in Fig. 1.4. Combination of centerline and cross-section can be considered as a discrete cylinder type model. Friman *et al.* proposed a parametric template cylinder model to track the blood vessels [67]. The generated cylinder model is designed based on a Gaussian line appearance model, consisted with contrast term, template model term, background term and noise term. Similarly, Lacoste *et al.* utilized small cylinder model as particles and performed marked point process algorithm [68]. Tyrrell *et al.* presented a more flexible parametric superellipsoids model [69]. Examples of superellipsoids are illustrated in Fig. 1.5

However, modelling anomalies including stenoses and aneurysms is difficult for explicit vascular models. Next, we will give a brief introduction of implicit modeling of vascular structures.

1.2.2.2 Implicit vascular model

In this thesis, implicit vascular model mainly denotes segmentation methods representing vascular characteristic by extracting vessel-specific features. The features can be extracted from basic intensity information to first- and second-order derivative based detectors.

First-order derivative based methods mainly use image gradient information to

present vascular features. Weickert *et al.* proposed "structure tensor" [70], computing the covariance matrix of gradient vectors. Wiemker *et al.* [71] utilized radial structure tensor, a variation of original structure tensor, to encode tubular and nodular structures. Nergiz *et al.* [72] enhanced the anisotropy between the principal eigenvalues of structure tensors, and segmented retinal vessels by using Otsu thresholding [73] on tensors. Agam *et al.* [74] presented a vesselness enhancement filter based on the eigenvalues of structure tensors. This filter makes it possible to distinguish between vascular structures, bifurcations and nodules. Another well-known first-order derivative based method is earlier mentioned "gradient vector field" (GVF). GVF-based methods have also been used in vessel segmentation problem [75, 76]. Another first-order derivative based method is the optimally oriented flux (OOF)[77] which aims to exploit the distribution of gradient vectors. By measuring the gradient flux through the boundary of local spherical region, OOF is able to suppress the effect of other nearby structures compared to Hessian-based measures.

Second-order derivative information is a very effective feature to characterize the vascular geometry. The assumption of second-order derivative based methods is that tubular structures can be discriminated by the direction and cross-sectional plane, and these two features can be well captured by second-order derivative information. Among the proposed second-order derivative based methods, Hessian matrix is the most widely used tool to encode the vascular geometrical information. It is first used by Koller *et al.* [79] to estimate the vascular orientations. The most commonly used Hessian-based vesselness filters are proposed by Sato *et al.* [80] and Frangi *et al.* [81]. The main contributions of their works are the introduction of a multi-scale Hessian eigenvalues criterion for measuring the relations between the Hessian eigenvalues to distinguish between tubular-, plane- and blob-like structures. Based on these vesselness filters, a lot of blood vessel segmentation approaches have been proposed [78, 82–84]. Illustration of the concept of Hessian matrix is shown in Fig. 1.6. Enhancement results on clinical CT volume is shown in Fig. 1.7. Both single

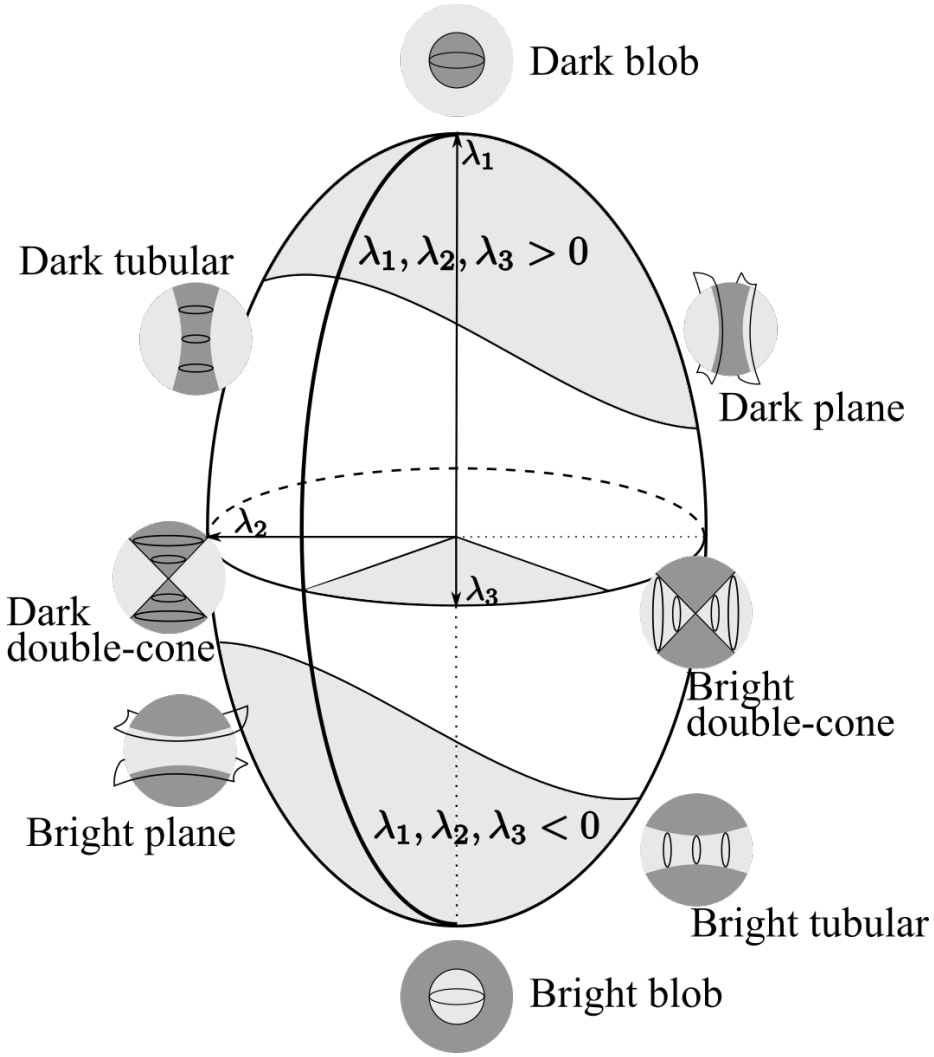


Figure 1.6: Illustration of the concept of Hessian matrix. The ellipsoid shape is defined based on the eigenvalues of Hessian matrix $(\lambda_1, \lambda_2, \lambda_3)$. Typically, blood vessels match the bright tubular prototype, and bronchi match the dark tubular prototype. More detailed please refer to [78]. Illustration is based on material from [78].

scale and multi-scale enhancement results are shown.

Another second-order derivative based method is Weingarten-based filters [85, 86]. Compared to Hessian matrix, Weingarten matrix is less used in the literature, considering possible reason is that Weingarten-based filters was reported to suffer from sensitivity problem to local deformations such as stenoses and aneurysms [54].

As conclusions, in this section, we mainly introduced different blood vessel methods focusing on capture vascular geometrical information. Vascular geometrical informa-

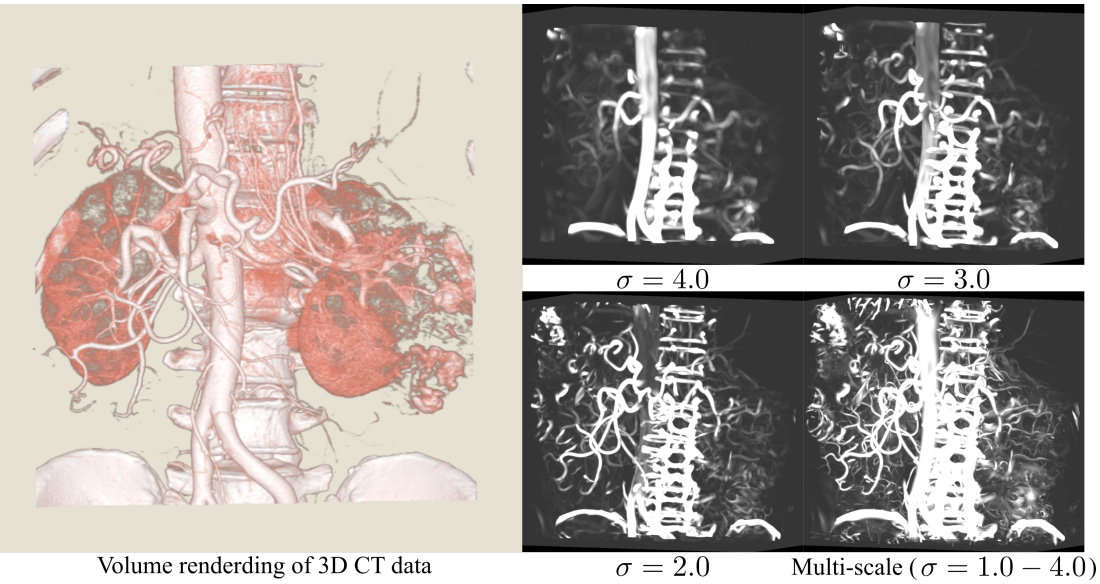


Figure 1.7: Enhancement results of Hessian-based vesselness enhancement filter [81]. Results of both single scale and multi-scale are shown with different scale(σ) settings.

tion is the main characteristic of blood vessels compared to other human tissues. We believe that capturing the geometrical information is one of the most effective and robust approaches for blood vessel segmentation problems. The introduced methods have been summarized in Table 1.1. The reader may refer to [54] for more detailed overview of blood vessel segmentation methods.

1.2.3 Organ segmentation

Organ segmentation is a big segmentation category because human has numerous body organs. In this thesis, our organ segmentation mainly focuses on segmenting solid organs, such as liver, spleen, kidneys, adrenals, pancreas, and lung. Unlike vascular system shares similar geometrical feature, organs have various shapes and appearances on CT images.

Therefore, besides geometrical information, anatomical information is also important for organ segmentation problems. One commonly used anatomical prior-knowledge is the organ position. Atlas-based models are widely used for organ segmentation problem. Atlases are reference anatomical templates which are learned

Table 1.1: Summary of the introduced related blood vessel segmentation methods in this thesis.

		Methodology	Features	Characteristics
Explicit models		Active shape models [55]	- Landmark-based linear shape model - Appearance: non-parametric multi-class model - Shape: the axis and cross-sectional shape models	A combination of appearance and shape models. The axis and cross-section are learned separately.
		Active contour models [57–60]	- Minimal paths + 3D active contour model - Image gradients - Structure tensor - Gaussian derivative filter	Various features can be used for optimizing external energy for specific tasks.
		B-spline [63, 64]	- Flux features - Hessian-based features	The idea is to use spline to model the vascular centerline. Various features can be used to extract the centerline.
		2D cross-section [56, 66]	- Circular medialness features - Ray-casting features	Naive 2D sectional tracking methods. Considering as a 2D version of model-fitting methods.
		Model-fitting [67–69]	- Local fit of the rigid vascular model	Model fitting-based methods translate the segmentation problem into regression problem.
Implicit models	first-order	Structure tensor [71, 72]	- Local distribution of gradient vectors	Cluster the gradient vectors by analyzing their covariance matrix.
		Gradient vector field [75, 76]	- Image gradient + Hessian-based features	A combination of first-order and second-order derivative information.
		Optimally oriented flux [77]	- Gradient flux	Exploit the direction of vessels by estimating the optimal gradient projection axis.
	second-order	Hessian-based filter [80, 81]	- Hessian-based second-order derivative feature	Analysis the Hessian eigenvalues to discriminate between plane-, blob- and tubular-like structures.
		Weingarten-based filter [85, 86]	- Weingarten-based second-order derivative feature	The principal curvatures and directions correspond to the eigenvalues and eigenvectors of the Weingarten matrix.

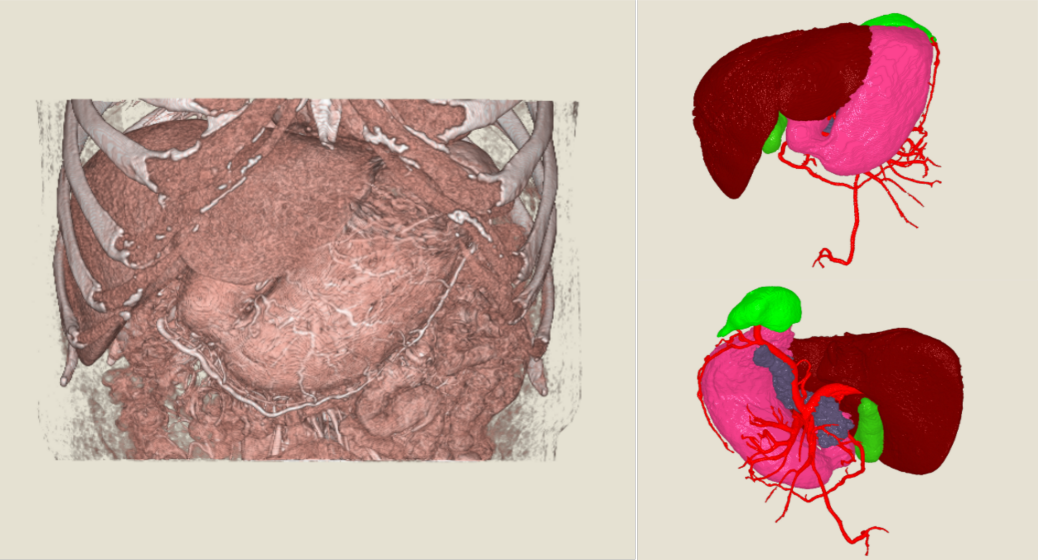


Figure 1.8: One example of multiple abdominal organs. Left is volume rendering of abdominal CE-CT data. Right is volume rendering of annotated labels of multiple organs

from a mount of annotated data. Atlas-based models, including probabilistic atlas-based models and multi-atlas-based models, can capture both organ shape and spatial locations. There have been proposed a fair amount of research using atlas for organ segmentation. Pancreas [49], liver [87] bone [88], and multi-organ [89–91] segmentation have been realized by using atlas-based methods.

Although atlas-based models can effectively capture spatial location information, they do not explicitly capture the internal relations between organs. Atlas-based models and other shape-based models can be regarded as a kind of individual models, which only capture individual features for each target organ. Unlike individual models, sequential models have been introduced to tackle the inter-organ relation issue.

Sequential models perform step-wise processes for multiple organs. The key assumption is that the analysis of stable and related organs can benefit other related and challenging organs. For instance, to improve the segmentation accuracy of pancreas and gallbladder which are challenging tasks, a lot of works has been done to use surrounding organs such as spleen and liver which are much easier to segment

[92–95]. Benefit from human anatomical prior-information, sequential models have quite good performances compared to common individual models.

Lots of research has been done to exploited the human anatomical information. For instance, Camara *et al.* utilized the symmetry of left and right kidney for kidney segmentation [96], the inclusion relation between lung and thoracic cavity [96, 97], and the intersection relation of hepatic vasculature inside liver [98].

Except for above introduced traditional methods, machine-learning and deep-learning based techniques also have been incorporated in organ segmentation problems. Criminisi *et al.* utilized random forest as regressor to localization multiple organs [99]. They captured the intra-organ spatial relations by using context-rich features. Graphical models directly model the inter-organ relations as their edge connections, and organ features as nodes. Nimura *et al.* and Bhole *et al.* utilized conditional random field to extract multiple organs from CT scans [100, 101].

With the development of deep-learning techniques, deep-learning based methods have achieved significant progress. Deep-learning techniques tackle the shortage of traditional machine-learning based methods which need handcraft feature design. Deep-learning is a kind of data-driven learning technique. Features can be learned from prepared training data automatically. In organ segmentation problem, many related deep-learning based methods have been done. Thanks to the power of deep-learning techniques, effective features can be learned without handcraft design process. At early stage, many works utilized general 2D convolutional neural networks (CNNs) to segment organs [102–104]. Segmentation problem is performed by classifying pixels or patches.

Recently, 3D end-to-end fully convolutional networks (FCN) have been widely used in organ segmentation problems. Milletari *et al.* have validated that using 3D data has better performances than 2D and 2.5D data for brain segmentation from MRI and ultrasound volumes [105]. One of the most popular FCN for organ segmentation is U-Net architecture [106, 107]. It has been used for many segmentation works

as backbone network [108–112], and achieved several state-of-the-art performances of organ segmentation challenges [108, 113]. With the progressive understanding of deep-learning techniques, more and more novel architectures and modules have been proposed in medical image-processing field. The potential of multiple networks have been explored for organ segmentation [114, 115]. New loss function, *Dice loss*, has been presented for organ segmentation instead of traditional cross-entropy loss [116].

In this section, we give a brief introduction of several organ segmentation methods including traditional non-machine-learning methods, machine-learning based methods and deep-learning based methods. From the introduction, we can find that besides geometrical information, anatomical information also plays an important role in organ segmentation problems. Compared to deep-learning based methods, non-machine-learning based methods are more explicit to adopt anatomical prior-knowledge. We believe the anatomical information can also contribute to deep-learning based method to achieve better performances.

1.3 Research overview

Segmentation is a fundamental task in medical image-processing field. Especially in CAD, segmentation plays an indispensable role. Many follow-up diagnosis processes are based on segmentation results. This thesis mainly focuses on tubular and solid organ segmentation, and demonstrate clinical applications about computer-aided kidney diagnosis using our approaches. In Table 1.2, we briefly categorize the segmentation targets of tubular and solid organs.

All of segmentation approaches presented in this thesis are designed for 3D CT scan imagery. The high resolution and contrast make it possible to achieve high performance available for real clinical practice. Like many other segmentation studies, the first priority of this research is to achieve high accuracy, compared to computing speed and operational complexity.

Table 1.2: Simple categories of segmentation targets in this thesis.

	In the literature	In this thesis
Tubular organ segmentation	Blood vessel [55–69], bronchus [117–120], nervous system [121, 122], *etc.*	Renal artery, bronchus
Solid organ segmentation	Liver [87], spleen [123], pancreas [49, 111, 124], lung [125],*etc.*	Kidney

The first research topic focuses on tubular organ segmentation, including blood vessel and bronchus. Two main methods are represented in Chapter 2 and 3. We tried to tackle tubular structure segmentation problem from both traditional unsupervised image-processing method and modern supervised deep-learning technique routes. The second topic is organ segmentation and its application to kidney segmentation. By incorporating anatomical information into neural network, segmentation performance has been significantly improved. The third research topic is clinical applications. Based on our developed approaches, this research demonstrates the significance of CAD system in contributing to clinical kidney-related applications.

1.3.1 Topic 1: Tubular organ segmentation

In this research topic, two methods for tubular organ segmentation problems are presented. As we introduced in Section 1.2.2, tubular structure segmentation remains a challenging task such as blood vessel segmentation. In this thesis, we present two segmentation methods for tubular structure using different methodologies. The first one is a traditional image-processing based method. The second is a modern supervised deep-learning based method. To take advantage of human prior-knowledge, the proposed traditional image-processing based method is performed in unsupervised fashion. One advantage of this type of methods is that they do not need annotated training data. Generating annotated labels for 3D tubular structures is time-consuming and labor-intensive. In contrast, the second proposed method is a supervised deep-learning based method. Thanks to the power of deep-learning tech-

niques, deep-learning based methods can automatically learn useful features from manually labelled data, making it possible to further improve segmentation accuracy compared to classic methods.

We have introduced related research working on blood vessel segmentation problem in Section 1.2.2. We mainly focus on the methods using geometrical information. We believe the geometrical information plays a very important role tubular structure segmentation problem. Our first method, named "tensor-cut", captures the geometrical feature by second-order tensor. First-order Markov random field (MRF) is constructed by incorporating both intensity and tensor information. By this strategy, both appearance (bright intensity) and geometry (tubular structure) can be modelling into one unifying model. We successfully applied tensor-cut to kidney blood vessel segmentation problem and achieved better performance than other classical blood vessel segmentation.

Deep-learning technique is adopted in the second tubular structure segmentation method. We aimed to exploit the potential of 3D FCN in tubular structure segmentation problem. Conventional neural networks still suffer under-segmentation problem for tiny tubular structures. To address this problem, we presented two new modules designed for tubular structures: 3D recurrent convolutional layer and radial distance loss. These two modules can be easily Incorporated into general FCN architecture. Experimental results showed that our modules significantly improved the segmentation accuracy.

1.3.2 Topic 2: Solid organ segmentation

Organ segmentation is also an important research topic in CAD field. Like blood vessel segmentation, organ segmentation is a fundamental step for many other medical processes. In this research topic, we present one deep-learning based organ segmentation method.

Conventionally, we utilized semi-automatic approach to segment organs from CT

scans. Graph-cut algorithm is used to semi-automatically extract specific organs. Although graph-cut algorithm is a high computational efficiency and easy-to-use segmentation tool. The semi-automatic procedure is not suitable for large dataset. Fully automated organ segmentation methods are the expected approaches for real clinical practices.

We present a spatially aware fully convolutional network for organ segmentation. By explicitly introduce spatially aware unit, we can directly capture the anatomical information (relative spatial location). Contrast to atlas-based methods which encode the spatial information as probabilistic model, we explicitly encode the organ's relative spatial location information as latent representation, and feed into FCNs. Experimental results demonstrated that our spatially aware unit can effectively capture the spatial information and improved the segmentation accuracy.

1.3.3 Topic 3: Applications to computer-aided kidney diagnosis

Although our segmentation methods are generally designed. They can be adopted into other applications. In this thesis, we focus on computer-aided kidney diagnosis. In this topic, we will demonstrate how to fully support computer-aided kidney diagnosis using our approaches, and how clinical applications can benefit from CAD system.

In Chapter 5, we will give two clinical applications about kidney diagnosis. Different organs may have different analysis processes for various clinical needs. A combination with other processes may need to meet specific clinical needs. In kidney-related CAD systems, one major research subject is nephrectomy-related diagnosis. Nephrectomy surgery is one common treatment for kidney cancer. Consequently, a lot of effort has been put into the development of CAD system for nephrectomy.

The first application aims to to provide computer-aided diagnostics for the presurgical simulation of partial nephrectomy surgery. This research can directly contribute to help physicians design a better surgical plan for partial nephrectomy. The medical

contribution of this research is to provide a quantitative standard for partial nephrectomy. To accomplish this task, we need pixel-wise labels for kidney region, renal arteries, and renal vascular dominant regions. In our research, by using our proposed methods, this three information can be extracted automatically from CT scans, which significantly reduced manual labor. The experimental results demonstrate that our CAD framework can achieve reasonable accuracy for presurgical simulation.

The second application aims to quantitative analyze the renal function variation between pre- and post-surgery. The quantitative analysis can give a better understanding of renal function changes influenced by nephrectomy. The renal cortex and renal medulla are two main regions of the kidney. The volume change is one major index to reflect renal function variation. Therefore, segmentation of kidney, renal cortex and medulla become a necessary step for follow-up analysis. In this application, we utilized LSM-based methods to extract renal cortex and medulla regions from kidney regions which is pre-extracted by our proposed FCN described in Chapter 4.

1.4 Thesis structure

This thesis consists of six chapters one appendix. An overview of the relationship between each chapter is illustrated in Fig. 1.9.

Chapter 1 has discussed the background, related works of this research, and also has given a brief overview of each research topic. Through this chapter, we hope readers can have a basic understanding of motivation and contributions of our research.

Chapter 2 and 3 give a detailed descriptions of our proposed tubular organ segmentation methods. Two methods have been validated in different applications. Traditional non-machine-leaning method, tensor-cut, was validated in renal artery segmentation problem. Deep-learning based method described in Chapter 3 was evaluated in bronchus segmentation problem.

Chapter 4 describes our proposed spatially aware FCN method for organ segmen-

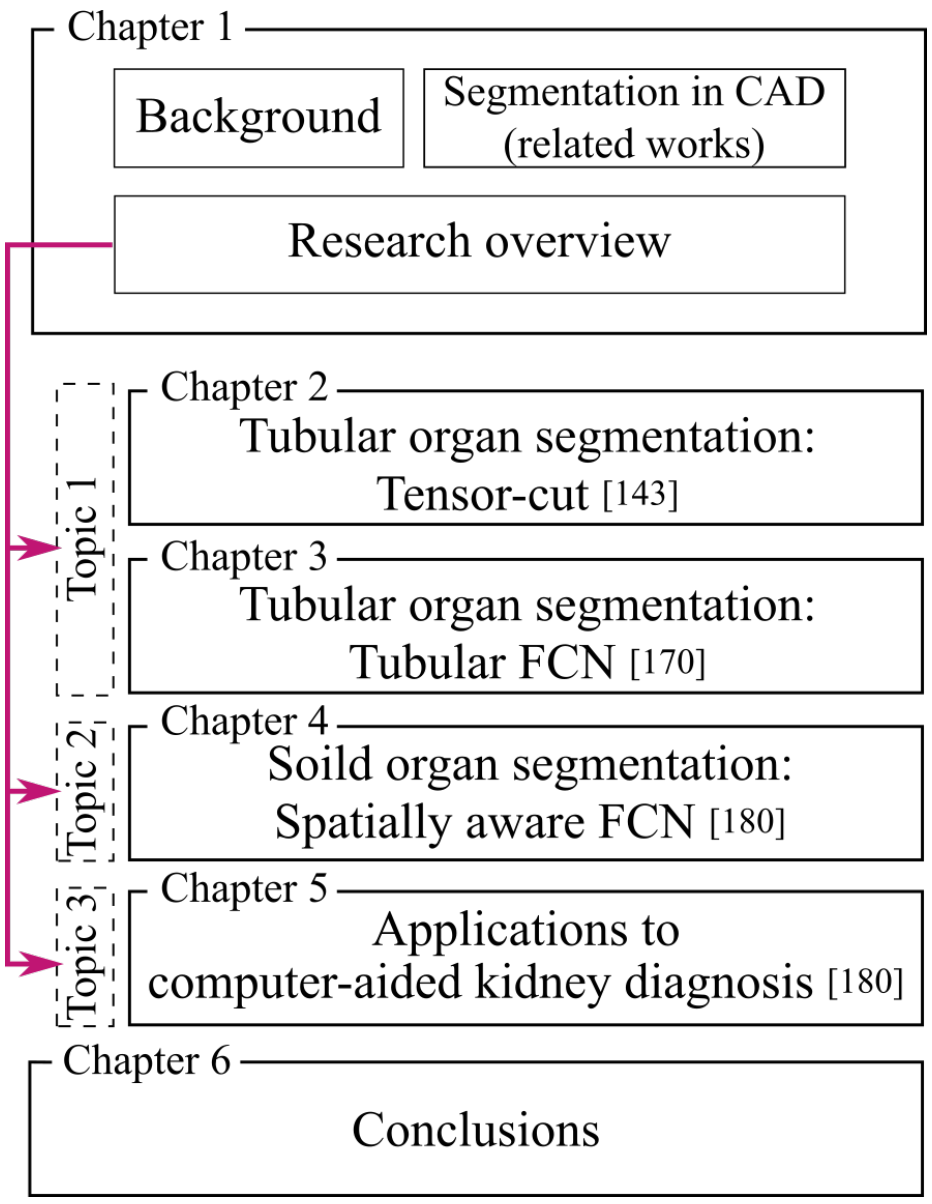


Figure 1.9: Overview of the chapters of this thesis.

tation, accuracy evaluation was performed on kidney segmentation problem.

Chapter 5 presents the third research topic: applications to computer-aided kidney diagnosis. In this chapter, several kidney-related clinical applications are briefly introduced. We want to demonstrate the contribution of CAD system in real clinical practices.

Chapter 6 contains the summary and future research of research area.

Chapter 2

Tubular organ segmentation: Tensor-cut

2.1 Background

In Chapter 1, we have introduced the importance of segmentation methods in CAD system. Segmentation objects can be any anatomical structure in medical images, such as heart, lung, liver, and prostate. We can roughly classify these structures into several categories: solid organs, hollow organs, tubular organs, and lesions. Solid organs have firm internal tissue consistency, such as liver, spleen, pancreas, kidney. In contrast, hollow organs have a cavity which serves vital functions, such as stomach, prostate, and bladder. Lesions indicate abnormal human tissues such as nodules, cancers. As for tubular organs, typical tubular organs contain blood vessel, bronchus, and nerve. Because the processing methodology for tubular organs may large differ from other organs, thus we separate these objects as a single category.

As we introduced in Section 1.2.2, numerous blood vessel segmentation methods have already been proposed in the literature for various blood vessels. Different image modalities, such as digital subtraction angiography, fundus photography, CE-CT, and MRI angiography, have different characteristics of vessel features. Even in the same modality, blood vessels in different tissues may still have slight differences in feature representations due to signal intensity variations, inhomogeneous contrast filling, or

imaging artifacts. Therefore, a lot of existing techniques exploit such vessel prior knowledge as tubular structures.

The method we present in this chapter is initially designed for kidney blood vessels segmentation problem. However, due to the generally designed methodology, we found this method also can be applied to other vasculature, theoretically. The motivation of this work is to propose a blood vessel segmentation which is able to extract tiny vessels in an unsupervised fashion. In Chapter 3, we present another tubular structure segmentation method which aims to extract tubular organs in a supervised fashion. Deep-learning techniques have been utilized to achieve high segmentation accuracy. Supervised deep-learning based methods are highly data-demanded. Manual annotation process is very time-consuming and labor-intensive, especially in 3D medical image processing field. The proposed method described in this chapter is developed for unsupervised processing, which means no additional annotated labels are needed for segmentation task.

2.2 Related works

A general introduction of blood vessel segmentation methods is described in Section 1.2.2. In this section, we will introduce several previous research more closely related to our method. In section 2.2.1, we will introduce direct related works which inspired our work. In section 2.2.2, we will describe our previous attempts to propose fine blood vessel segmentation methods. We hope readers can have a brief understanding of the history of our research.

2.2.1 Related segmentation methods

We attempted to capture vascular geometrical information by using second-order derivative of intensity information. Second-order derivatives information, such as the widely used Hessian matrix, models the tubular geometric structure of blood

vessels [80, 81]. For 3D data, Hessian matrix is a 3×3 symmetric matrix. The magnitudes of its three eigenvalues indicate the local geometric structure of the local image. A fair amount of approaches use Hessian-based vesselness features for tubular structure representation [78, 83, 126, 127]. We also used Hessian-matrix to capture the geometrical information in our works.

The second closely related work is graph-cut algorithm. The graph-cut algorithm, which has been widely used for many such low-level computer vision problems as segmentation and registration [128–130], can find the exact solution in polynomial time for submodular energy. However, the local neighboring Markov random field (MRF) model, where the voxels are only connected to their neighboring voxels, has a side effect, short-boundary bias, which results from the influence of the pairwise term, which penalizes the elongated boundaries. A high-order MRF model has been presented to tackle this problem by coupling similar edges [131–133]. Many approaches have also adopted graph-cuts into blood vessel segmentation problems using additional information, such as shape prior [134] and geometrical moments [135].

The third related work is multi-scale nonlinear structure tensor (MSNST) proposed by Han *et al.* [136]. Han *et al.* constructed a MSNST space to extract textures from natural images, and calculate dissimilarity measure and probability distribution of features in the Riemannian space instead of Euclidean space. The authors succeed in evaluating structure tensors in Riemannian space. Furthermore, they also incorporated graph-cut algorithm in solving the optimization problem. Our method, tensor-cut, is also directly inspired by Han *et al.*'s work.

2.2.2 Previously proposed methods

Before introducing our proposed "tensor-cut" approach, we first give a brief introduction of our previously proposed methods putting efforts on capturing tubular geometrical information using Hessian-matrix. Our previous approaches will be introduced in next sections. Section 2.2.2.1 briefly describes a hybrid method combine

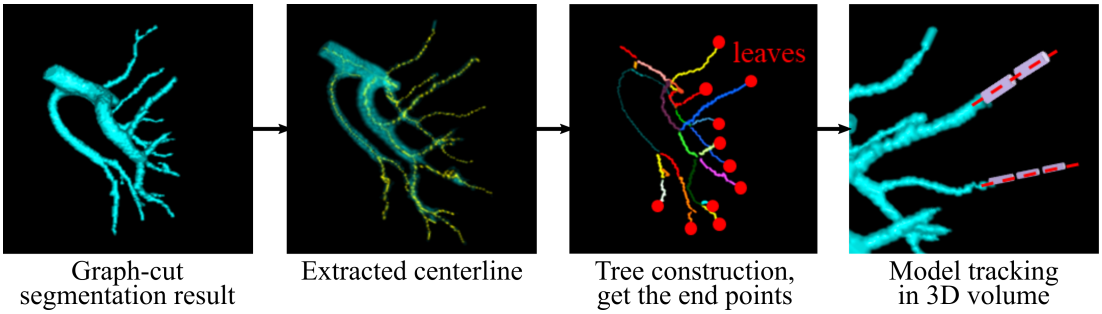


Figure 2.1: A simple workflow of hybrid method.

graph-cut with model-fitting methods. In Section 2.2.2.2, another proposed concept is briefly introduced.

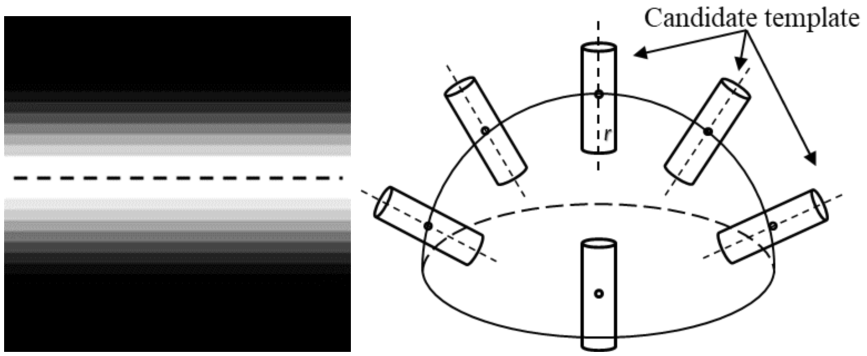
2.2.2.1 Hybrid method

For renal artery segmentation, hybrid method utilizes the graph-cut method for a rough segmentation to extract the thick blood vessels, and utilize template model tracking to focus on the tiny blood vessels. After rough segmentation of thick blood vessels, template model tracking will continue to extract the tiny blood vessels from the end of each thick blood vessel branch. A simple workflow of hybrid method is illustrated in Fig. 2.1. To utilize the graph-cut method for segmentation, foreground and background information should be specified (i.e. the voxels belong to source S or sink T). Therefore, we utilize a Hessian-based vesselness enhancement filter to extract the foreground region. Detailed process is described in Section 2.4.

After thick blood vessel is extracted, multiple hypothesis template model tracking (MHT) [67] will start at the end point of each branch and continue to extract the tiny blood vessels. A cylinder model is generated based on a Gaussian line appearance model, consisted with contrast term, template model term, background term and noise term. The image model $M(\mathbf{x})$ is described as

$$M(\mathbf{x}) = kT(\mathbf{x}; r, \mathbf{x}_0, \hat{\mathbf{v}}) + m + \epsilon(\mathbf{x}), \tag{2.1}$$

where k is the vessel contrast, m is the image background intensity and $\epsilon(\mathbf{x})$ represents



(a) (b)

Figure 2.2: Illustration of template model. (a) Illustration of 2D template model. (b) Illustration of generation of candidate models in 3D space.

noise. $T(\mathbf{x}; r, x_0, \hat{\mathbf{v}})$ is the template function which is given by:

$$T(\mathbf{x}; r, \mathbf{x}_0, \hat{\mathbf{v}}) = \frac{r^\gamma}{(d^2(\mathbf{x}; \mathbf{x}_0, \hat{\mathbf{v}}))^{\gamma/2} + r^\gamma}, \tag{2.2}$$

where r is the radius of the template model, $\mathbf{x}_0$ is the center point of the template model, $\hat{\mathbf{v}}$ is the direction, γ is the steepness parameter and d is the distance from voxel $\mathbf{x}$ to the centerline of the model. A 2D template model is illustrated in Fig. 2.2.

Furthermore, the mathematically defined model has a low computation cost. Levenberg-Marquardt algorithm is applied to solve the non-linear problem in model fitting procedure. The following minimization problem

$$\min_{r, \mathbf{x}_0, \hat{\mathbf{v}}, k, m} \|\mathbf{W}(r, \mathbf{x}_0, \hat{\mathbf{v}})[kT(r, \mathbf{x}_0, \hat{\mathbf{v}}) + m\mathbf{1}_n - \mathbf{I}]\| \tag{2.3}$$

is solved to find the best parameters of template model to fit the image, where $\mathbf{W}(r, \mathbf{x}_0, \hat{\mathbf{v}})$ is a diagonal matrix to localize the template model as a weight function. $\mathbf{I}$ is CT image data. $\mathbf{1}_n$ is a vector with a constant 1. Illustration of generation of candidate models is shown in Fig. 2.2. A set of candidate models is generated along the spherical surface in 3D space.

2.2.2.2 Euclidean ellipsoid

In this work, we attempt to exploit the potential of Hessian matrix in graphical model at first time. We involve Hessian matrix analysis approach to graph-cut to improve

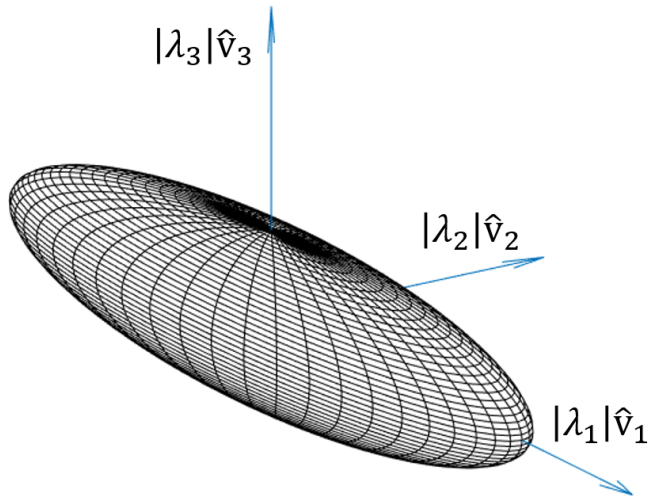


Figure 2.3: Illustration of ellipsoid, three principle axes depend on eigenvalues and eigenvectors of Hessian matrix. Length of principle semi-axis is $|\lambda_i|$, direction is normalized eigenvector $\hat{\mathbf{v}}_i$.

the tubular structure segmentation performance.

Conventionally Euclidean distance of intensity is utilized which is given by

$$V_{m,n}(\mathbf{x}_m, \mathbf{x}_n) = \frac{\exp^{-\eta|I(\mathbf{x}_m)-I(\mathbf{x}_n)|}}{dist(\mathbf{x}_m, \mathbf{x}_n)}, \tag{2.4}$$

where $dist(\cdot)$ is the Euclidean distance operator, parameter η makes the smoothness term applicable to different contrast images. $I(\mathbf{x})$ denotes the intensity of voxels $\mathbf{x}$ of 3D data.

We use GMM to model the foreground and background intensity distributions, and utilizing Hessian matrix analysis method to obtain a novel dissimilarity measure between neighboring voxels. The new smoothness term takes advantage of local geometrical information calculated by Hessian matrix. This scheme will make graph-cut more suitable for tubular structure segmentation problem.

In the Hessian-based vesselness filter, we obtained eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\}$ and eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of every voxel. To calculate the dissimilarity of these structure features, we construct an ellipsoid to unite these features. This idea is inspired by DTI imaging techniques. We utilize Hessian matrix analysis method to simulate

the DTI signals. Assuming that a 3-D ellipsoid EP is centered at origin $(0, 0, 0)$, the columns of 3×3 orthogonal matrix $\mathbf{U}$ are unit vectors in the directions of ellipsoid EP's principal axes. Let Σ be the diagonal matrix with elements $\Sigma_{ii} = \sigma_i$, where $1/\sigma_i$ represents half of the length of EP's ith principal axis. Thus ellipsoid EP is given by

$$EP \equiv \{\mathbf{x} | \mathbf{x}^T \mathbf{U} \Sigma^2 \mathbf{U}^T \mathbf{x} \leq 1\}. \tag{2.5}$$

Therefore, $\sigma_i = 1/\lambda_i$, $\mathbf{U} = [\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \hat{\mathbf{v}}_3]$, where $\hat{\mathbf{v}}_i$ is unit eigenvector, $\mathbf{x}$ are voxels. More details about representation of ellipsoids can be found in Stephen's work [137]. An illustration of the ellipsoid is shown in Fig. 2.3.

To simplify the dissimilarity measure, Euclidean distance is utilized to measure the dissimilarity between the ellipsoids [138]. The dissimilarity between ellipsoid pair (EP_i, EP_j) is given by

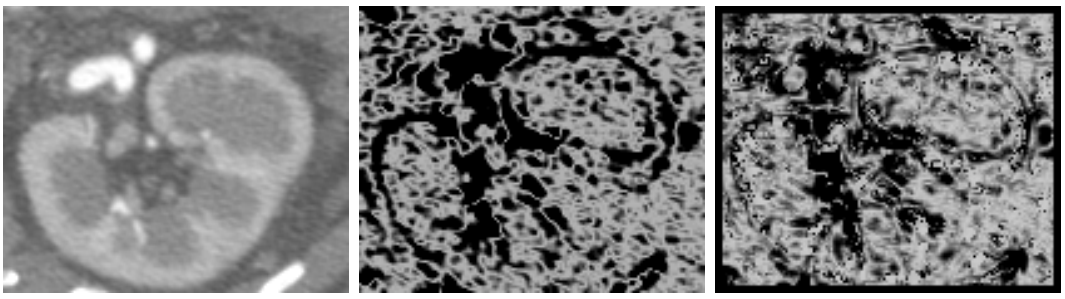
$$s(EP_i, EP_j) = e^{-\|\mathbf{c}_i - \mathbf{c}_j\|} \cdot e^{-\|sin\theta\|} \cdot e^{-\|\xi_i - \xi_j\|}, \tag{2.6}$$

where $\mathbf{c}$ is the center of ellipsoid. θ is the angle between EP_i and EP_j. In this work, only the direction of major principle axis $\mathbf{A}_1 = \|\lambda_1\|\hat{\mathbf{v}}_1$ is calculated, viz., $\theta = arccos(\mathbf{A}_{i1} \cdot \mathbf{A}_{j1}/|\mathbf{A}_{i1}||\mathbf{A}_{j1}|)$. ξ represents shape features that $\xi = (1/\sigma_1, 1/\sigma_2, 1/\sigma_3)^T$. An example of dissimilarity map is shown in Fig. 2.4. Finally, we can obtain the segmented blood vessels assigned with label L_F utilizing the proposed dissimilarity map.

2.3 Contributions

1. Capturing geometrical information using second-order tensor in Riemannian space.

   Instead of using Euclidean metric to reduce high dimensional geometrical information to 1 dimensional space, we second-order tensors to represent geometrical



(a) (b) (c)

Figure 2.4: Dissimilarity maps. (a) One slice of original CT volume. (b) Conventional intensity-based dissimilarity map. (c) The proposed Hessian-based dissimilarity map.

features, and directly measure tensors in high dimensional space using Riemannian metric.

2. Use tensor and intensity info to create 1st-order Markov random field, and use graph-cuts optimization algorithm to get the best solution.

   This method incorporates both tensor and intensity information into one unified model. Tensor and intensity are used to represent geometrical and appearance features.

3. A fine blood vessel segmentation approach designed for 3D artery. High segmentation accuracy on both simulated and clinical dataset.

   To validate the method, we prepared two datasets: simulated vascular dataset and clinical renal artery dataset. Experimental results confirmed the improvements on both simulated and clinical dataset. Especially on clinical dataset, significant improvements were achieved compared to other methods. Renal artery segmentation remains a difficult problem in this field. High segmentation accuracy on renal arteries is a remarkable progress.

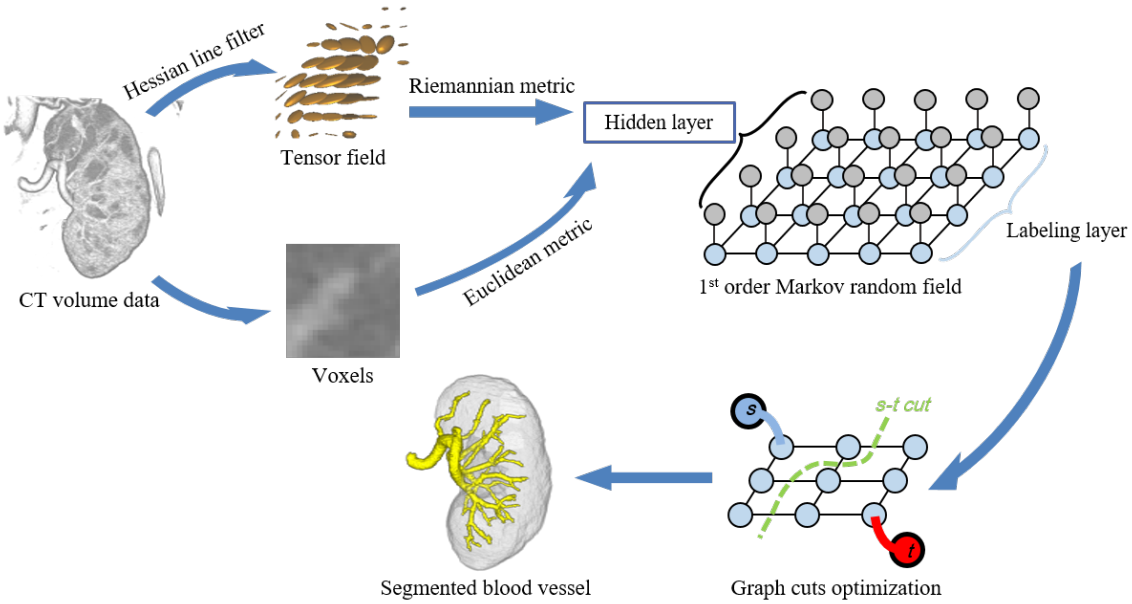


Figure 2.5: Flowchart of proposed algorithm.

2.4 Tensor-cut: a tensor-based graph-cut method

2.4.1 Overview

In this section, we give a detailed description of the proposed method, tensor-cut, a tensor-based graph-cut method designed for vascular structures. By using this method, we attempt to tackle the short-boundary bias issue by introducing tubular geometrical information to the MRF to exploit the advantages of both graph-cut and vesselness-enhancement filtering. We use both intensity and tensor information to construct a first-order MRF. Using tensors for image processing can be traced back to 1980s, Bigün *et al.* first presented structure tensor using gradient information [139, 140]. Knutsson *et al.* presented a local energy-based method to calculated the local structure tensor by analysing the responses of a set of quadrature filters [141]. Lindeberg *et al.* presented Gaussian scale-space for a multi-scale representation for tensors [142]. Our proposed method exploits the second-order information of the geometry structure of blood vessels. Unlike many existing blood vessel segmentation methods that only use the high-dimensional geometrical information in low-dimensional Eu-

clidean space, the proposed method directly utilizes the geometrical information in manifold space that can contribute to a more accurate model. The detailed description of this method is based on material from our published work [143]. The flowchart of this work is shown in Fig. 2.5.

2.4.2 Tensor field

As we described in section 2.2.1, Han *et al.* constructed a multi-scale nonlinear structure tensor (MSNST) space to extract textures from natural images [136]. First, a multi-scale structure tensor (MSST) is constructed by calculating gradient $D^s(x, y)$ (omitted as D^s) of image I at scale s:

$$\tau^s = \alpha^{-2s} \begin{bmatrix} (D_x^s)^2 & (D_x^s D_y^s) \\ (D_x^s D_y^s) & (D_y^s)^2 \end{bmatrix}, \tag{2.7}$$

where $s \in (0, S)$, S is the total of the scales. D_x^s and D_y^s represent two gradient components along the x and y directions. $D_x^s(x, y) = \partial(I * \theta^s)(x, y)/\partial x$, $D_y^s(x, y) = \partial(I * \theta^s)(x, y)/\partial y$. $\theta(x, y)$ is a 2D differential function, and α is a constant parameter.

Then, a nonlinear anisotropic filter [144] is applied to each MSST to obtain the MSNST space. Each pixel in an image is represented as a MSNST $\mathbf{T}$:

$$\mathbf{T} = \{\hat{\tau}_0, \hat{\tau}_1, ..., \hat{\tau}_{s-1}\}. \tag{2.8}$$

In this work, instead of using first-order structure tensor to capture the texture features, we utilize a Hessian matrix to describe the tubular geometric features. Unlike the previous work [136], we do not construct a multi-scale tensor space. Using a multi-scale vesselness-enhancement filter, we get the best scale to represent the blood vessels. This strategy effectively reduces the computing time.

The 3D Hessian matrix at scale s is given:

$$\nabla^2 I^s(\mathbf{x}) = \begin{bmatrix} I_{xx}^s(\mathbf{x}) & I_{xy}^s(\mathbf{x}) & I_{xz}^s(\mathbf{x}) \\ I_{yx}^s(\mathbf{x}) & I_{yy}^s(\mathbf{x}) & I_{yz}^s(\mathbf{x}) \\ I_{zx}^s(\mathbf{x}) & I_{zy}^s(\mathbf{x}) & I_{zz}^s(\mathbf{x}) \end{bmatrix}. \tag{2.9}$$

Here $I^{ij}(\mathbf{x}) = \dfrac{\partial^2}{\partial i \partial j} I(\mathbf{x}), (i, j \in \{x, y, z\})$ represents the second-order partial derivatives of local image I at voxel $\mathbf{x}$. This symmetric matrix resembles a second-order tensor $\mathcal{T}^H$. A multi-scale Hessian-based vesselness-enhancement filter [80, 81] is utilized to find the most appropriate scale to represent the tubular structure. Vesselness-measure $\mathcal{V}^s$ at scale s is given:

$$\mathcal{V}^s = \begin{cases} |\lambda_3^s| \left(\dfrac{\lambda_2^s}{\lambda_3^s}\right)^{\gamma_{23}} \left(1 + \dfrac{\lambda_1^s}{|\lambda_2^s|}\right)^{\gamma_{12}}, & \text{if } \lambda_3^s < \lambda_2^s, \lambda_1^s \leq 0, \\ |\lambda_3^s| \left(\dfrac{\lambda_2^s}{\lambda_3^s}\right)^{\gamma_{23}} \left(1 - \alpha\dfrac{\lambda_1^s}{|\lambda_2^s|}\right)^{\gamma_{12}}, & \text{if } \lambda_3^s < \lambda_2^s, 0 < \lambda_1^s < \dfrac{|\lambda_2^s|}{\alpha}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.10}$$

where $\lambda_1^s, \lambda_2^s,$ and λ_3^s are the eigenvalues of the Hessian matrix at scale s, $(\lambda_1^s > \lambda_2^s > \lambda_3^s)$. Scale $s \in (0, S)$, S is the total number of scales. γ_{12} and γ_{23} are constant parameters. α is introduced to provide an asymmetrical characteristic [80]. The scale with highest vesselness-response $\mathcal{V}_{max}$ is the most appropriate scale to represent the blood vessels, and the most appropriate scale $s' = \arg\max_s \mathcal{V}^s, s \in (0, S)$.

However, scale s' can be obtained for every voxel of image $I(\mathbf{x})$ even if its structure is not tubular. Therefore, we add a constraint to reduce these noisy tensors. In this paper, we introduce *diagonal tensor* $\mathcal{T}^D$ to replace these non-vessel tensors that can be expressed:

$$\mathcal{T} = \begin{cases} \mathcal{T}^H, & \mathcal{V}_{max} > 0 \\ \mathcal{T}^D, & \mathcal{V}_{max} \leq 0 \end{cases}, \tag{2.11}$$

where $\mathcal{T}^D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$ and $\lambda_3 \gg \lambda_2 \approx \lambda_1 > 0$. Here $\mathcal{T}^D$ represents a plate-like structure. This diagonal tensor provides a non-tubular tensor for those voxels that do not belong to the tubular structures. Tensor field $\mathbb{T}$ can be constructed. $\mathbb{T} = \{\mathcal{T}_0, \mathcal{T}_1, ..., \mathcal{T}_n\}, n \in N. \mathcal{T}_i \in \{\mathcal{T}^H, \mathcal{T}^D\}, \forall i \in N$. N denotes voxel number. A visualized example of a tensor field is shown in Fig. 2.6 where only tubular tensors are illustrated. The tensor visualization technique was previously presented [145]. Ellipsoids are constructed with the eigenvalues and the eigenvectors of the Hessian

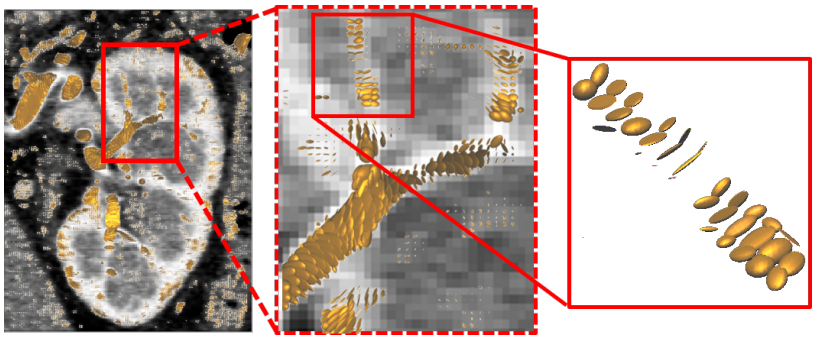


Figure 2.6: Constructed tensor field of one CT slice. Only tensors with $\mathcal{V}_{max} > 0$ are shown.

matrix.

2.4.3 Tensor metric in Riemannian space

In this section, we briefly review the Riemannian metrics for computing tensor statistics. Since second-order tensors cannot be handled by traditional linear statistical methods, non-linear metrics are required for computing statistics on tensors. In our previous work described in Section 2.2.2, we modeled a tensor as an ellipsoid and used Euclidean metrics to calculate the dissimilarity between ellipsoids [146].

Many studies have made efforts on computing the statistics of tensors [147–149]. Tensors are not in traditional vector space so that traditional vector operations cannot applied to tensor computing. Actually, the tensor space forms a type of manifold, viz., a *Riemannian manifold*. Riemannian manifolds $(\mathcal{M}, g)$ are smooth manifolds $\mathcal{M}$ with Riemannian metric g. A manifold consists of a series of linear Euclidean subspaces. Similar to any other curved surfaces, the geodesic distances between two vectors on a manifold is a continuous collection of projection distances in a tangent space at each point. Riemannian metric g makes it possible to calculate the statistics of the manifolds, such as mean value, geodesic distances, and geodesic interpolation.

The Riemannian metrics handle symmetric positive definite (SPD) matrices, which form a convex half-cone in the vector space of the matrices. Many standard operations are stable in this space [148]. However, the Hessian matrix may have negative

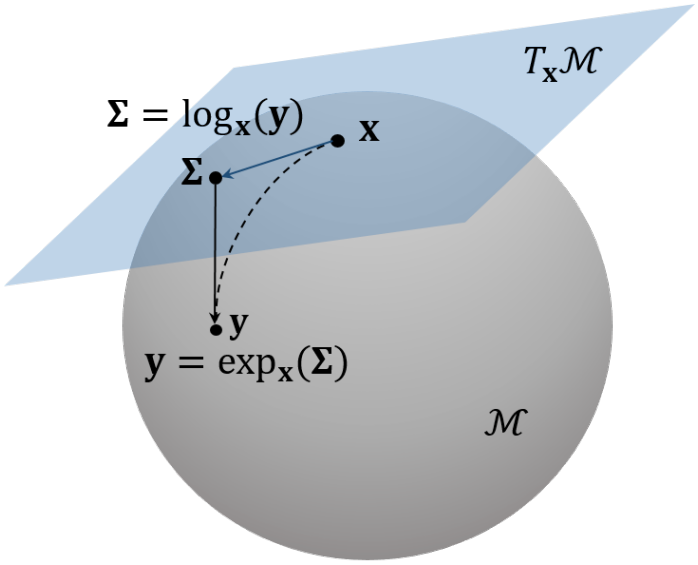


Figure 2.7: Illustration of geodesic distance computing on manifold. $\mathbf{x}$ and $\mathbf{y}$ denote two tensors on manifold $\mathcal{M}$. $\boldsymbol{\Sigma}$ is a vector on tangent space $T_{\mathbf{x}}\mathcal{M}$ at $\mathbf{x}$. $\boldsymbol{\Sigma}$ is projected by a logarithmic mapping $\log_{\mathbf{x}}(\mathbf{y})$ from tensor $\mathbf{x}$ to tangent space, and exponential mapping $\exp_{\mathbf{x}}(\boldsymbol{\Sigma})$ projects $\boldsymbol{\Sigma}$ back to manifold.

eigenvalues, i.e., since it is an indefinite matrix.

In our previous work [150], we used a transformation trick to convert the indefinite Hessian matrices to SPD matrices. The transformation helps map a Hessian matrix to a Riemannian manifold. Let $\mathcal{T}^-$ denotes an indefinite Hessian matrix with negative eigenvalues, and let $\mathbf{U}$ denotes an invertible orthogonal matrix with columns that correspond to eigenvectors. We have $\mathcal{T}^- = \mathbf{U}\mathbf{D}\mathbf{U}^{-1} = \mathbf{U}\mathbf{D}\mathbf{U}^T$, where $\mathbf{D}$ is a diagonal matrix. $\mathbf{D} = \text{diag}(d_i)$, where d_i is the i-th eigenvalue. Positive definite tensor $\mathcal{T}^+$ can be obtained:

$$\mathcal{T}^+ = abs(\mathcal{T}^-) = \mathbf{U}\text{diag}(abs(d_i))\mathbf{U}^T, \mathcal{T}^+ \in Sym^+. \tag{2.12}$$

However, this transformation assumes that tensors are in ideal tubular shape: $\lambda_2 \approx \lambda_3$, $|\lambda_2| \gg |\lambda_1| \approx 0$. In this work, we directly use the absolute value of generalized eigenvalues of the tensors to transform the Hessian tensors to a semi-Riemannian space. The revised transformation is given in Eq. 2.13, which is more mathematically reasonable than previous transformation.

Pennec *et al.* and Moakher *et al.* described an affine invariant Riemannian metric

for statistics to compute SPD matrices (tensors) [148, 151]. As shown in Fig. 2.7, the geodesic path between tensors $\mathbf{x}$ and $\mathbf{y}$ on manifold $\mathcal{M}$ is the shortest curve shown in the dotted line. The geodesic distance is defined as $d(\mathbf{x}, \mathbf{y}) = \|\overrightarrow{\mathbf{xy}}\|_{\mathbf{x}}$. $\overrightarrow{\mathbf{xy}}$ is a vector in tangent space $T_{\mathbf{x}}\mathcal{M}$, which is projected by a logarithmic mapping: $\overrightarrow{\mathbf{xy}} = \log_{\mathbf{x}}(\mathbf{y})$. In our case, the geodesic distance between two tensors can be written:

$$d(\mathcal{T}_1, \mathcal{T}_2) = \|\log_{\mathcal{T}_1}(\mathcal{T}_2)\|_{\mathcal{T}_1} = \left(\sum_{i=1}^{d} \log^2 abs\big(\lambda_i(\mathcal{T}_1, \mathcal{T}_2)\big)\right)^{\frac{1}{2}}, \tag{2.13}$$

where $\lambda_i(\mathcal{T}_1, \mathcal{T}_2)$ is the generalized eigenvalues of tensors $\mathcal{T}_1$ and $\mathcal{T}_2$. d denotes the order of the SPD matrix, which is $d = 3$. Actually, Eq. 2.13 is a kind of semi-Riemannian metric (a.k.a pseudo-Riemannian metric) to handle our Hessian matrices. Semi-Riemannian manifold, is a differentiable manifold, in which the requirement of SPD constraint is relaxed. In semi-Riemannian space, tensors need not be positive definite at every point which allows us to calculate semi-distance between Hessian tensors.

We calculated the mean and the variance of the tensors using the same method given in a previous work [148]. A simple introduction is described in the appendix. A more detailed description is also available [148].

2.4.4 Graph-cut using tensors

A graph-cut is a powerful optimization tool for solving low-level vision problems [128, 152]. Based on the max-flow/min-cut theorem, the minimum cut that corresponds to the best solution of a sub-modular energy function can be found in polynomial time. A traditional graph-cut is based on a first-order Markov random field (MRF), which means that only adjacent nodes are considered. This graphical model easily leads to a short-boundary bias. For blood vessel segmentation problems, this short-boundary bias causes serious under-segmentation of tiny blood vessels. This problem can be addressed with a high-order MRF instead of a first-order MRF considering high-order cliques can also effectively avoid local minimum cuts [132, 133]. Another

direct method to avoid short-boundary bias is to make a more accurate objective distribution model to prevent local minimal cuts. In this paper, we use the local geometric information presented as tensors to improve the accuracy of the GMM model.

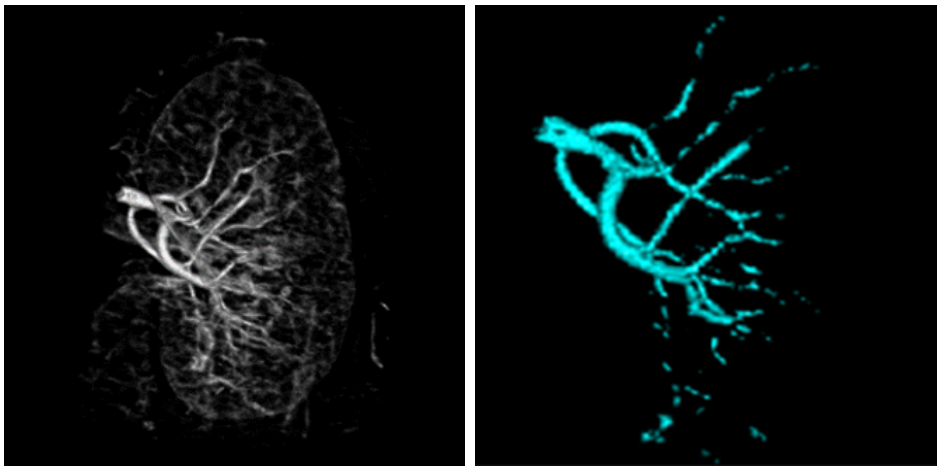
The minimum cut problem, also known as s-t cut, needs pre-specified source nodes (s nodes) and sink nodes (t nodes). However, for 3D blood vessel segmentation, manually creating these *terminal* nodes is time-consuming (including s and t nodes). In this work, we automatically generate terminal nodes using K-means algorithm. Let $I_{\mathcal{V}}(\mathbf{x})$ be an image filtered by a vesselness-enhancement filter, and then we extract the most probable blood vessel region as s nodes with K-means clustering. The most probable cluster $\mathbf{c}'_{max}$ is given:

$$\mathbf{C}' = \arg\min_{\mathbf{C}} \sum_{i=1}^{K} \sum_{\mathbf{x}\in\mathbf{c}_i} ||I_{\mathcal{V}}(\mathbf{x}) - \bar{I}_{\mathbf{c}_i}||^2, \mathbf{c}_i \in \mathbf{C} \tag{2.14}$$

$$\mathbf{c}'_{max} = \arg\max_{\mathbf{c}'_i} \frac{1}{|\mathbf{c}'_i|} \sum_{\mathbf{x}\in\mathbf{c}'_i} |I_{\mathcal{V}}(\mathbf{x})|, \mathbf{c}'_i \in \mathbf{C}', \tag{2.15}$$

where $\mathbf{c}_i$ is the i-th cluster of total K clusters $\mathbf{C}$. $\bar{I}_{\mathbf{c}_i}$ denotes the mean intensity of the voxels in cluster $\mathbf{c}_i$. $\mathbf{C}'$ is the final clusters. In Eq. 2.15, cluster $\mathbf{c}'_{max}$ with the highest mean intensity is selected as the foreground region: the s nodes. In practice, cluster number K is set to five in this paper. An example of a generated foreground region is shown in Fig. 2.8. For the possible background region, we first dilate the foreground region with radius $r = 3$ and use the complement of the dilated foreground region as a possible background region (t nodes).

As same as previous work [129], for improved accuracy, we used a Gaussian mixture model (GMM) to estimate the probability distribution of the blood vessels and the background region instead of histogram approaches. Two GMMs with K_g components are required: one for the foreground region and another for the background region. A GMM can generally be written as $\Pr(u) = \sum_k \eta_k \mathcal{N}(u|\mu_k, \sigma_k)$ ($k \in K_g$). $\mathcal{N}(u)$ denotes the Gaussian distribution. η_k is a weight parameter for each Gaussian component.



(a) (b)

Figure 2.8: Generated foreground region using K-means clustering: (a) Vesselness-enhancement result $I_{\mathcal{V}}(\mathbf{x})$. (b) Most probable blood vessel region $\mathbf{c}'_{max}$ extracted using K-means ($K = 5$).

GMMs for intensity and tensors are given:

$$\begin{aligned} \Pr(\mathbf{x}) &= \sum_{k=1}^{K_g} \gamma_k \mathcal{N}(\mathbf{x}|\bar{I}^k, \sigma_I^k) \\ \Pr(\mathcal{T}) &= \sum_{k=1}^{K_g} \theta_k \mathcal{N}(\mathcal{T}|\bar{\mathcal{T}}^k, \sigma_T^k) \end{aligned}, \tag{2.16}$$

where $\bar{I}$ and σ_I denote the mean and the variance of the intensity. $\bar{\mathcal{T}}$ and σ_T denote the mean and the variance of the tensors; detailed descriptions are given in the appendix. γ and θ are weight parameters that can be obtained using the EM algorithm [153].

For a graph-cut, we utilized both the intensity and tensor information. The energy function is described:

$$\begin{aligned} E(L) = &\underbrace{\underbrace{\sum_{\mathbf{x}\in\mathbb{X}} -\log\Pr(L_{\mathbf{x}}|\mathbf{x})}_{\text{data term}} + \lambda_I \underbrace{\sum_{\{\mathbf{x}_m,\mathbf{x}_n\}\in\mathcal{N}} V_{m,n}(\mathbf{x}_m,\mathbf{x}_n)}_{\text{smoothness term}}}_{\text{intensity term}} \\ &+ \underbrace{\omega\left(\sum_{\mathcal{T}\in\mathbb{T}} -\log\Pr(L_{\mathcal{T}}|\mathcal{T}) + \lambda_T \sum_{\{\mathcal{T}_m,\mathcal{T}_n\}\in\mathcal{N}'} U_{m,n}(\mathcal{T}_m,\mathcal{T}_n)\right)}_{\text{tensor term}}, \end{aligned} \tag{2.17}$$

where $L = \{L_F, L_B\}$ denotes labels and L_F and L_B are foreground and background labels. ω represents the weight parameter to adjust the weight between the intensity

and tensor term energy. The λ_I and λ_T parameters adjust the weight between the data term and the smoothness term in the graphical model. $\Pr(L_{\mathbf{x}}|\mathbf{x})$ and $\Pr(L_{\mathcal{T}}|\mathcal{T})$ denote conditional probabilities of voxels $\mathbf{x}$ and tensors $\mathcal{T}$derived using Bayes rule. In this paper, we use a first-order MRF, which means that we only consider the edges between neighboring voxels. $\mathcal{N}$ and $\mathcal{N}'$ denote pair sets of neighboring voxels and tensors. $V_{m,n}(\cdot,\cdot)$ and $U_{m,n}(\cdot,\cdot)$ denote a dissimilarity measure for a pair of neighboring voxels and tensors: the energy of the edges. The dissimilarity measure for both voxels and tensors are given:

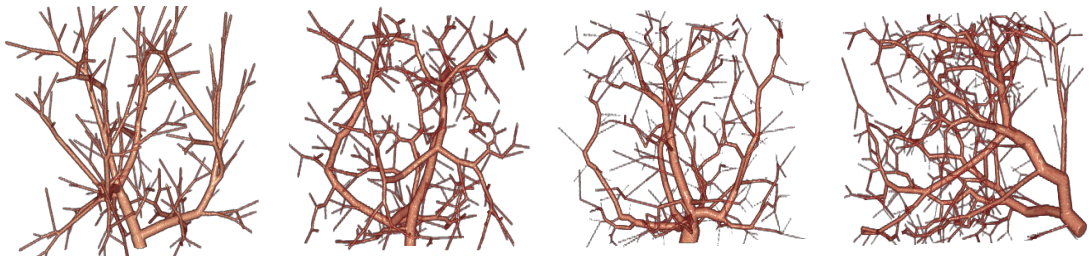
$$\begin{aligned} V_{m,n}(\mathbf{x}_m,\mathbf{x}_n) &= \frac{\exp\left(-\xi d^2(\mathbf{x}_m,\mathbf{x}_n)\right)}{dist(\mathbf{x}_m,\mathbf{x}_n)} \\ U_{m,n}(\mathcal{T}_m,\mathcal{T}_n) &= \frac{\exp\left(-\xi d^2(\mathcal{T}_m,\mathcal{T}_n)\right)}{dist(\mathcal{T}_m,\mathcal{T}_n)} \end{aligned}, \tag{2.18}$$

where ξ is a constant parameter that can be estimated as the noise level and $dist(\cdot,\cdot)$ denotes the Euclidean distance between two voxels or tensors, i.e. the physical distance between two points.

Finally, the minimum cut can be found by the graph-cut algorithm for a given energy function in Eq. 2.17. The terminal nodes belonging to the source are the final segmentation result. A simple pseudocode for our tensor-based graph-cut is shown in Algorithm 1.

Algorithm 1 Tensor-based graph-cut

Input: $\mathbf{x} \in I, \mathcal{T} \in \mathbb{T}$
Output: Segmentation $L = \{L_F, L_B\}$
Initialize: $\overline{\mathcal{T}}_0 = \mathcal{T}_1$, $t = 0, \epsilon = 0.001$
Calculate
$\mathbf{c}'_{max}$ ▷ Terminal nodes (Eq. 2.15)
$\Pr(\mathbf{x}), \Pr(\mathcal{T}),\ (x \in \mathbf{c}'_{max}, \mathcal{T} \in \mathbf{c}'_{max})$ ▷ GMM (Eq. 2.16)
$-\log\Pr(L_{\mathbf{x}}|\mathbf{x}), -\log\Pr(L_{\mathcal{T}}|\mathcal{T})$ ▷ Data terms
$V_{m,n}(\mathbf{x}_m,\mathbf{x}_n), U_{m,n}(\mathcal{T}_m,\mathcal{T}_n)$ ▷ Smoothness terms
$L \leftarrow$ min-cut/max-flow algorithms [152]



(a) (b) (c) (d)

Figure 2.9: Four phantom vascular pieces of data generated by *VascuSynth* with different bifurcation numbers: (a), (b), (c), and (d) respectively have 200, 250, 300, and 400 bifurcations.

2.5 Experimental results

2.5.1 Material

In this work, we did experiments both on simulated data and on clinical contrast-enhanced CT data to validate our proposed algorithm.

We used eleven cases of phantom data, all of which were generated using a public tubular structure generator called vascular synthesizer (*VascuSynth*)[154], which is a simulated vascular synthesizer that generates vascular structure based on such user-defined vascular properties as bifurcation number, oxygen demand map, and other tunable parameters. To the best of our knowledge, *VascuSynth* is the state-of-the-art vascular synthesizer in the literature. However, public vascular datasets[1] have few bifurcation nodes, which can easily be segmented by all methods. Therefore, we generated the phantom datasets with more bifurcation nodes and small bifurcation angles. Four generated pieces of phantom data are shown in Fig. 2.9. In this work, we generated phantom data with different bifurcation numbers that ranged from 200 to 400, and oxygen demand maps and other parameters were randomly decided. Gaussian white noise was added to obtain the final simulated data. We added two noise levels to each phantom data with Gaussian variations $\sigma_n = 0.3$ and 0.8; Thus in the simulated experiments, we have 22 synthetic datasets. To make the simulated

[1]http://vascusynth.cs.sfu.ca

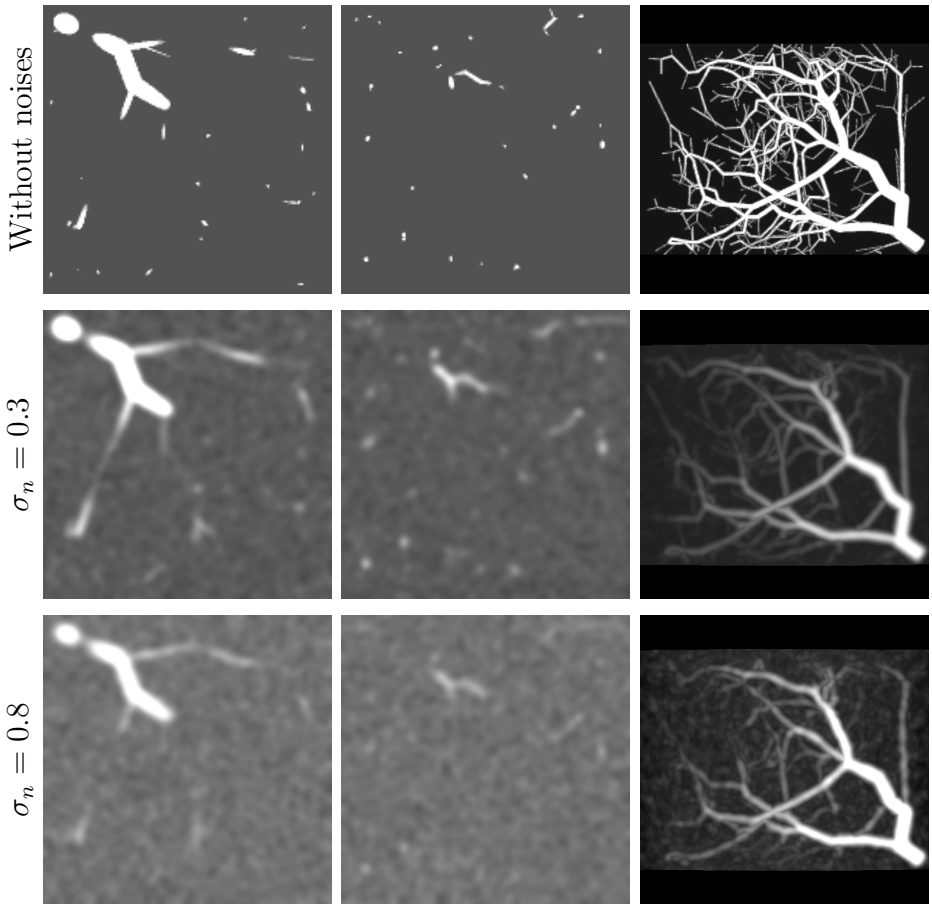
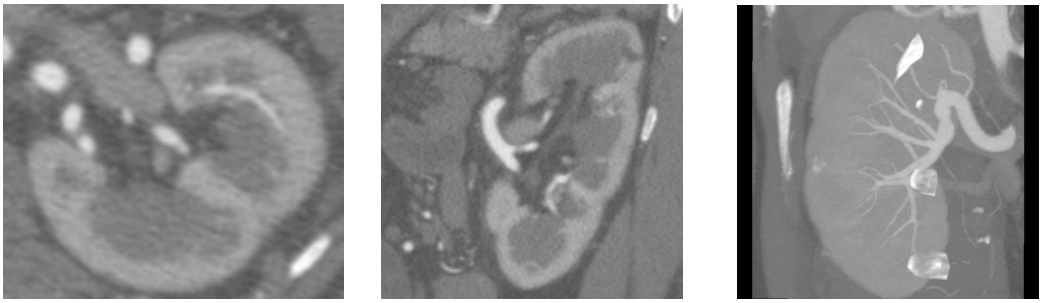


Figure 2.10: Examples of final simulated data: Original phantom data without noise are shown in first row. Simulated data in second row are applied with white noise of $\sigma_n = 0.3$, and third row is applied with $\sigma_n = 0.8$. (a) and (b) two axial slices of the simulated data. (c) maximum intensity projection (MIP) images. Compared to original phantom data, many details are lost by adding white noise and smoothing.

data more closely resemble the clinical CT data, we applied a Gaussian-smoothing filter with a 2.0mm radius to the phantom data. An example of the final generated data is shown in Fig. 2.10. Compared to its original phantom data without noise, many tiny vascular structures are lost by applying with noise and smoothing filters. The sizes of all the 22 simulated bits of data are $160 \times 160 \times 160$ with identical resolution $1.0 \times 1.0 \times 1.0$ mm^3.

We used 19 cases of clinical CT data for the experiments and focused on the renal artery segmentation problem. As pre-processing for the clinical CT data, the volumes



(a) (b) (c)

Figure 2.11: Volume of interest showing kidney in CT data: (a) and (b) one axial and sagittal slice, respectively. (c) MIP of image of VOI data.

of interest (VOIs) of the kidney region were extracted manually. One bit of VOI data contains an entire kidney region. One example of kidney VOI data is shown in Fig. 2.11. Detailed VOI data specification is summarized in Table 2.1.

2.5.2 Reference approaches

We consider six reference approaches in this paper. As the basic ideas of our proposed method, evaluations of the classic vesselness-enhancement filter (VEF) [80] and the graph-cut algorithm (GC) [152] are critical for proving the effectiveness of our proposed method. Our method is benchmarked against four competing state-of-the-art unsupervised blood vessel segmentation methods: multiscale spherical flux (MSF) [155], multiple hypothesis tracking (MHT) [67], super-ellipsoids (SE) [69], and commercial image processing software AMIRA [156].

VEF: VEF's two main parameters, α_V and γ_V, control the trade-off between branch detection and noise suppression. Based on a previous work [80], $\alpha_V = 0.25$

Table 2.1: Kidney VOI specification

Slice size [pixels]	(98 – 166) × (90 – 148)
Slice number [slices]	157 – 270
Pixel spacing [mm]	0.5 – 0.7
Slice pitch [mm]	0.4 – 0.8

and $\gamma_V = 1.0$ can optimize the trade-off and get robust results. We used this parameter setting for all the experiments in this paper, including both the simulated and clinical CT data. Since VEF is an enhancement filter, we manually threshold the segmentation results to get the best binary segmentation result for a quantitative evaluation. We implemented VEF in C++ based on a previous work [80].

GC: Unlike an interactive segmentation scheme [152], we still use the automatic segmentation scheme cited in Section 2.4.4. The foreground seed region is calculated by K-means, and the possible background region is the complement region of the dilated foreground region. GMM estimates the posterior probability. The energy function is identical as Eq. 2.17 without tensor terms. GC was implemented mainly based on a previous work [152].

MSF: Law *et al.* presented an efficient implement of a multiscale spherical flux computation method [155], which is useful for practical vascular segmentation problems. Multiscale spherical-flux based method was first proposed by Vasilevskiy *et al.* [157]. By measuring the gradient flux through the boundaries of multiscale spheres, divergence of image gradients can be calculated. The maximum response can be obtained when the sphere is located at the centerline of blood vessel with radius same with blood vessel. To accelerate the computation speed, a multiscale spherical flux computation is performed in the Fourier domain. Vesselness is directly formulated to the response of the spherical flux. Similar to other methods [80, 81], MSF is a vascular enhancement filter. An appropriate threshold is needed to obtain the final binary segmentation result. 0.005 is the best threshold value for clinical experiments on MRA data [155]. We used the public Matlab implementation provided by Law *et al.*[2].

SE: SE is a typical model-fitting-based tubular structure segmentation method. A combination of parameterized superellipsoid and intensity appearance model is built for representing tubular structures. Maximum-likelihood method is used to estimate

[2]http://www.cse.ust.hk/∼maxlawwk/

the parameters of superellipsoid models. The corresponding between vessel model and real blood vessel is verified using a generalized likelihood ratio test. Tracking starts from the multiple random seed points for experiments on clinical data and only evaluates the segmentation results inside the kidney mask are evaluated. Model fitting is performed using cylindrical super-ellipsoids to detect the tubular structures. The algorithm was implemented on their Farsight software platform[3]. We used the default parameters in our experiments except for the detection radii.

MHT: MHT obtained the highest score in the coronary artery centerline extraction challenge [158]. Unlike the SE method, a template vessel model is designed based on Gaussian line appearance model instead of super-ellipsoids. Student t-test is used to verify the corresponding instead of likelihood ratio test used in SE. To avoid early stopping caused by a local minimum score, MHT introduces a multiple hypothesis tracking scheme to further improve the segmentation accuracy. The multiple hypothesis approach starts multiple tracking process in parallel, and keep the tracking with the maximum score as the success tracking. The implementation we used is provided on their MeVisLab software platform[4]. The pruning and termination thresholds are the two primary parameters that affect the segmentation performance. In our experiments, the termination thresholds were set twice as high as the pruning thresholds, based on the official suggestion [67]. Since MHT is a semi-automatic method, the initial seed points must be manually specified. For all the experiments, we randomly put ten seed points at each branch level.

AMIRA: AMIRA is commercial biomedical software for visualization and image processing [156]. We used a module named AutoSkeleton for segmentation and implemented a vascular structure segmentation method [159] that can extract the skeleton and the diameter of tubular structures with a distance map. First, distance map is computed using adjusted chamfer distance [160]. Then, centerline and diameter

[3]http://www.farsight-toolkit.org/
[4]https://www.mevislab.de

are calculated using adapt distance ordered homotopic thinning (DOHT) algorithm. In this, work, we directly applied thresholding to distance maps to obtain the final binary segmentation results.

2.5.3 Measurements

We used three types of measurements to validate the segmentation performances in this work: sensitivity (Se), specificity (Spc), centerline overlap (CO), and centerline error (CE). Se, or the true positive rate (TPR), measures the algorithm's ability to correctly extract the target region (blood vessels). Spc, or the true negative rate (TNR), refers to its ability to correctly extract the background region (non-blood vessels). Se and Spc are given:

$$Se = \frac{TP}{TP + FN},\ Spc = \frac{TN}{TN + FP}, \tag{2.19}$$

where TP, FN, TN, and FP respectively denote the voxel numbers of true positives, false negatives, true negatives, and false positives.

The Dice coefficient (DC) is generally used for evaluating the overlap between the ground truth and segmentation results. However, the DC metric is heavily influenced by the volume of the target region [161]. A small volumetric variation of segmented, thick blood vessels may have greater influence than the volumetric variation of tiny blood vessels. Therefore, instead of DC, we introduce the CO index to improve the accuracy evaluation for tiny blood vessel segmentation and use centerline overlap measures [162]. Fig. 2.12 describes how to calculate the centerline overlap ratio. $\Omega_G(\mathbf{x})$ and $\Omega_S(\mathbf{x})$ denote the centerlines of the ground truth and segmentation results, respectively. Overlap region $O_\Omega = \{\mathbf{x} | \mathbf{x} \in \Omega_S(\mathbf{x}) \cap \mathbf{x} \in \mathcal{U}(\mathbf{x})\}$, where $\mathcal{U}(\mathbf{x})$ is dilated $\Omega_G(\mathbf{x})$ with a two-voxels radius. The CO index is given:

$$\text{CO} = \frac{2 * \|O_\Omega\|}{\|\Omega_G\| + \|\Omega_S\|}, \tag{2.20}$$

where $\|\cdot\|$ represents the length.

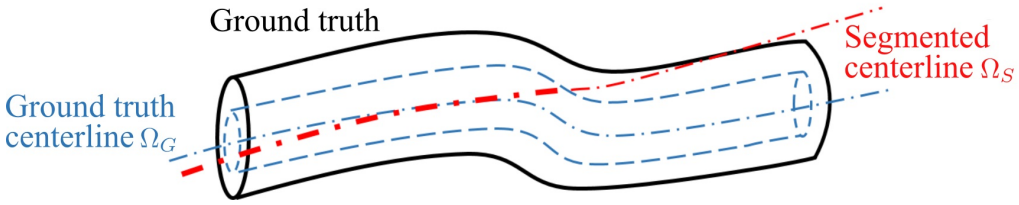


Figure 2.12: Illustration of calculation of centerline overlap CO. Tubular structure in black is gold standard blood vessel, and its centerline $\Omega_G(\mathbf{x})$ is shown as blue dash-dot line. Tube $\mathcal{U}(\mathbf{x})$ is generated by dilating centerline of a two-voxel radius. Red dash line denotes centerline of segmented vessels $\Omega_S(\mathbf{x})$.

Centerline error (CE) has also been used in many blood vessel segmentation works [162, 163] especially for coronary artery which needs quantitative analysis. CE mainly measures the distance between segmentation and ground truth. In this work, our CE index is defined:

$$\mathrm{CE} = \frac{h(\Omega_G(\mathbf{x}), \Omega_S(\mathbf{x})) + h(\Omega_S(\mathbf{x}), \Omega_G(\mathbf{x}))}{2}, \tag{2.21}$$

where $h(\cdot)$ denotes average centerline distance. Definition of centerline distance h is given:

$$h(\Omega_G(\mathbf{x}), \Omega_S(\mathbf{x})) = \frac{1}{\|\Omega_G\|} \sum_{p_G \in \Omega_G} \min_{p_S \in \Omega_S} (|p_G - p_S|) \tag{2.22}$$

For both CO and CE, we only compute the scores on the longest connected vascular trees. Pre-processing is performed to extract the maximum connected voxels. Tensor-cut has three primary parameters that must be specified: λ_I, λ_T, and ω. In both simulated and clinical experiments, we performed a grid search to select these three parameters and to plot the upper bounds of all the results in the ROC curves.

2.5.4 Experimental results on simulated data

We tested the segmentation methods on 22 simulated pieces of data, including two noise levels for each phantom data. The objective of the simulated experiments is to evaluate the general segmentation ability of the tubular structures. Experiments with different noise levels can evaluate the robustness against the noise. The receiver operating characteristic (ROC) curve of three cases is shown in Fig. 2.13. The x- and

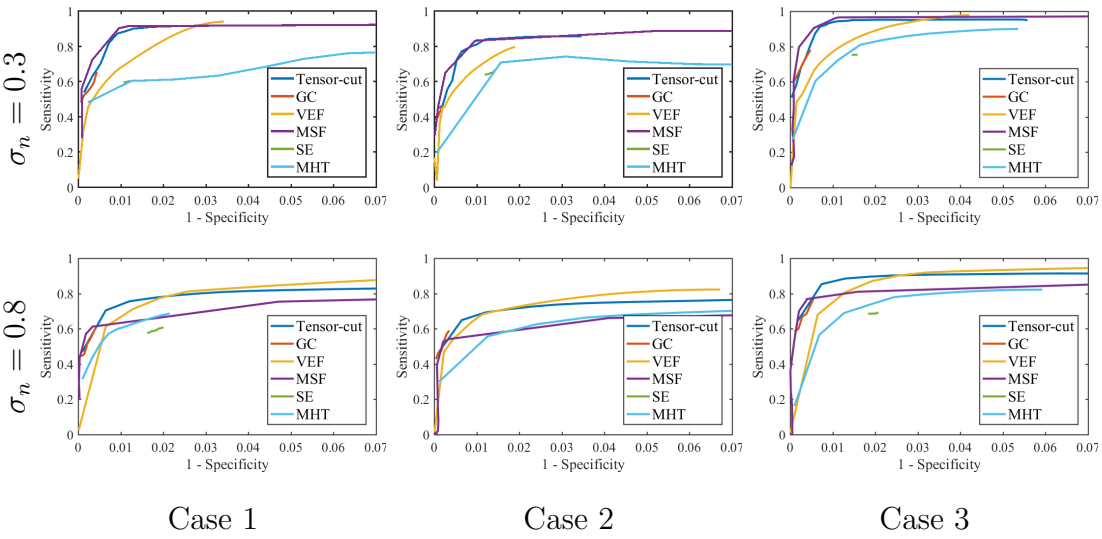


Figure 2.13: ROC curves of three simulated vascular data generated by *VascuSynth* with different bifurcation number. Cases 1, 2, and 3 have 250, 300, and 350 bifurcations. Experimental results of two noise levels ($\sigma_n = 0.3$ and $\sigma_n = 0.8$) are shown in first and second rows.

y-axes denote 1 - *Spc* (FPR) and *Se* (TPR), both of which are measured voxel-wise. No post-processing is performed on the segmentation results. As for methods which have multiple parameters, such as TC and SE, the ROC curves indicate the upper bound of all experimental results using grid-search approach. From the ROC curves, tensor-cut (TC) and MSF have the best performances among the six tested methods on the simulated dataset with a noise level of $\sigma_n = 0.3$. However, TC outperformed MSF on the simulated data with a noise level of $\sigma_n = 0.8$, which proved its robustness against different noise levels.

Experimental results showed that TC had better performances than GC and VEF, which proves the effectiveness of our proposed method. ROC indicates that the TC curves extended the GC curves, which means that TC extracted more structures than GC; i.e., introducing a tensor term effectively suppresses the short-boundary problem.

Two cases are visualized in Fig. 2.14. From the visualized results, TC outperformanced the other approaches. VEF and GC under-segmented the tiny tubular structures. MSF also has good sensitivity, although it is susceptible to noise. MSF

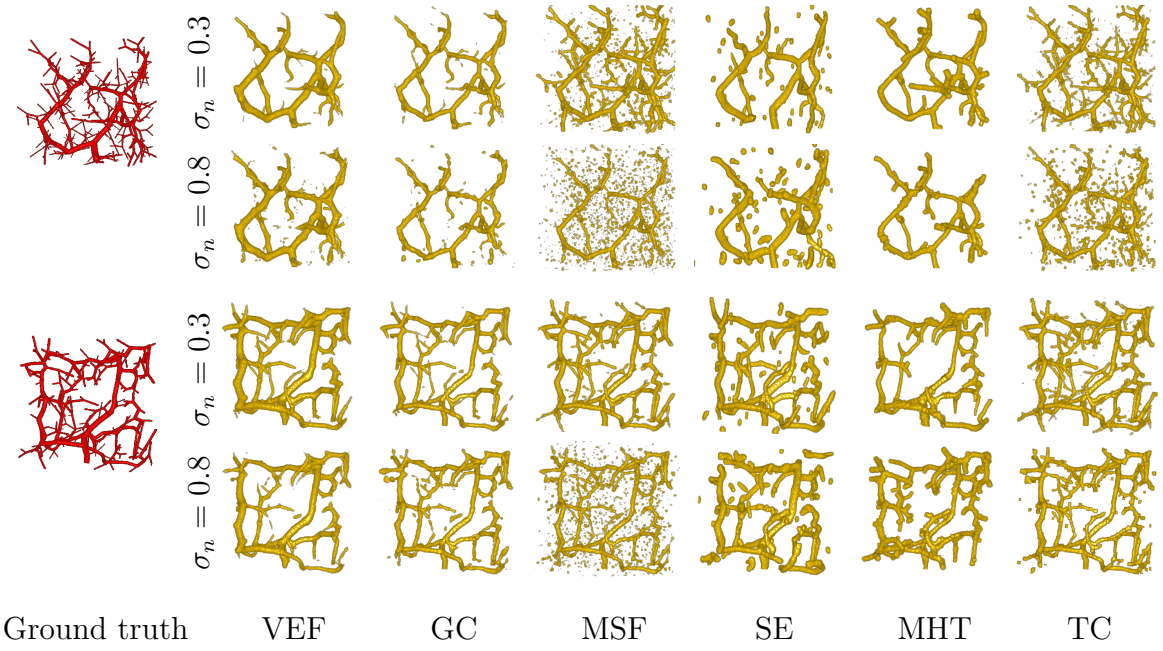


Figure 2.14: Segmentation results of simulated data of Cases 1 and 3. For each, segmentation results of two noise levels ($\sigma_n = 0.3$ and $\sigma_n = 0.8$) are shown in the first and second rows.

achieved similar segmentation results to TC under the $\sigma_n = 0.3$, but underperformed TC under $\sigma_n = 0.8$ condition. The model-fitting-based methods (SE and MHT) are limited to the detection of small tubular structures. ROC also showed that SE and MHT obtained poor results under both two noise levels. The quantitative statistical results of all 22 simulated bits of data are given in Table 2.2 that show the mean *Se* and *Spc* with their respective standard deviations.

2.5.5 Experimental results on clinical CT data

We performed experiments on the 19 clinical CT data (VOI data). To increase the algorithm's focus on the renal arteries, we manually created a kidney mask using the standard graph-cut algorithm [152].

In clinical experiments, the SE and MHT methods did not achieve comparable results on the clinical dataset. The latter only extracted one or two main thick branches, and the former failed to extract renal arteries. Therefore, we omitted

Table 2.2: Quantitative statistical results of simulated dataset. Two noise levels with $\sigma_n = 0.3$ and $\sigma_n = 0.8$ are shown in two columns for each approach. Mean sensitivity, specificity, and centerline errors with their standard deviations are also shown.

	Se (%)		*Spc* (%)		CE (mm)	
	$\sigma_n = 0.3$	$\sigma_n = 0.8$	$\sigma_n = 0.3$	$\sigma_n = 0.8$	$\sigma_n = 0.3$	$\sigma_n = 0.8$
TC	85.5±3.8	75.3±5.4	99.2±0.1	99.3±0.1	1.65±0.23	2.82±0.91
GC	62.2±7.6	66.0±6.4	99.8±0.1	99.6±0.2	5.57±1.98	5.16±1.67
VEF	72.9±4.2	71.2±4.2	98.9±0.2	99.0±0.1	1.57±0.57	3.07±1.09
MSF	87.2±4.2	64.5±6.2	99.0±0.2	99.6±0.1	1.96±0.50	3.84±1.13
SE	67.0±5.1	64.2±5.3	98.7±0.1	98.3±0.1	8.81±3.46	8.09±1.78
MHT	68.0±7.2	60.7±5.6	98.9±0.2	98.8±0.2	3.54±1.22	5.40±1.22

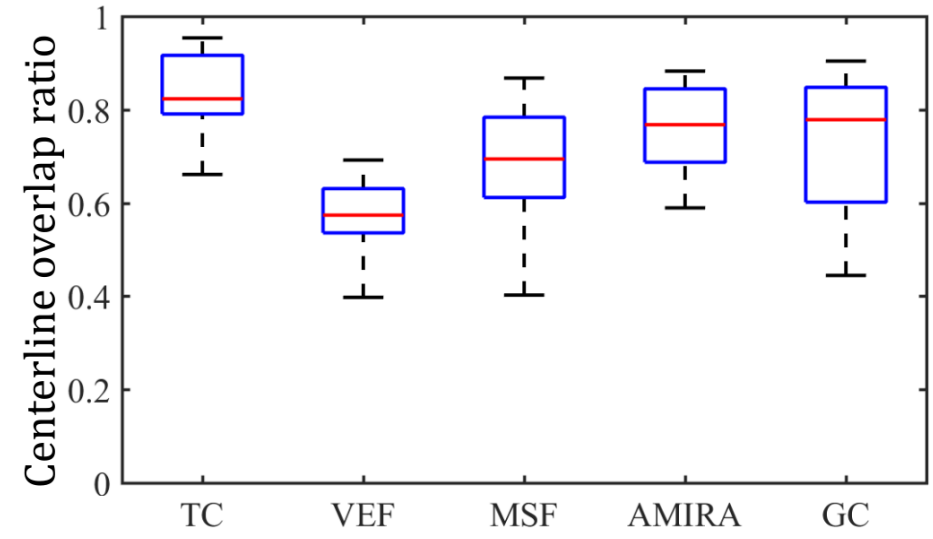
these two methods from our comparisons in the clinical experiments. Although the performances on the simulated and clinical data are different, our proposed method (TC) almost got the best segmentation performances. The quantitative CO and CE evaluation of 19 cases is demonstrated in Fig. 2.15. The TC method outperformed the other methods, and its performances are more robust.

2.6 Discussion and conclusions

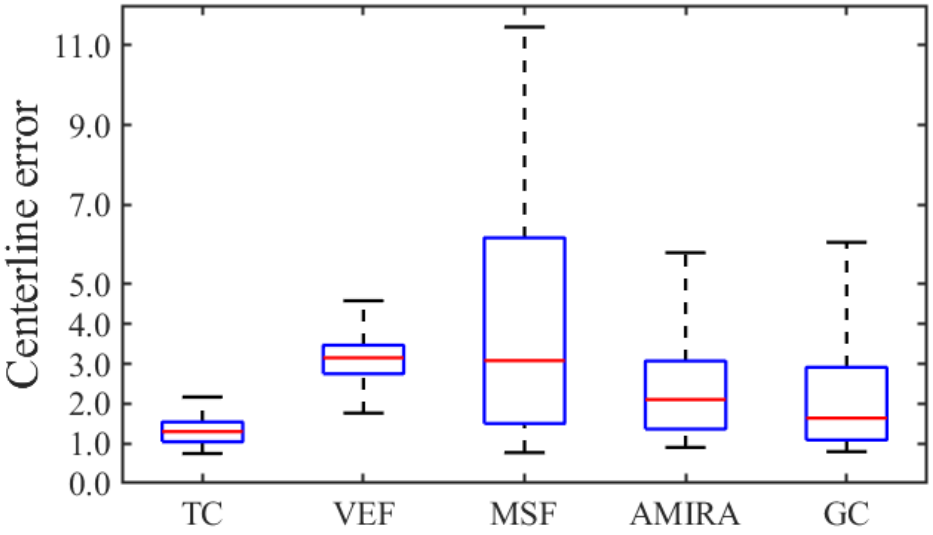
2.6.1 Discussion

Our experiment results on both the simulated and clinical CT data show a high performance of the tensor-cut method for tubular structure segmentation. In this study, we used both the ROC curves (TPR against FPR) and the CO curves (CO against FP) to evaluate the performances of the methods. The ROC curves measure the general segmentation ability, and the CO curves measure the segmentation ability for tiny tubular structures.

Our experiments on simulated data show the segmentation ability for general tubular structures. In Table 2.2, the experimental results with a noise level of $\sigma_n = 0.3$



(a) CO



(b) CE

Figure 2.15: Quantitative evaluation of 19 clinical cases. For each case, we selected the best segmentation results with the highest CO index and the lowest CE index. Red lines indicate median values, and tops and bottoms of blue boxes indicate 25^{th} and 75^{th} percentiles. Upper and lower bounds indicate maximum and minimal CO and CE values.

show that TC and MSF outperformed all the reference approaches. However, on the noisier simulated data with a noise level of $\sigma_n = 0.8$, TC and VEF outperformed the other reference approaches. The experimental results with different noise levels

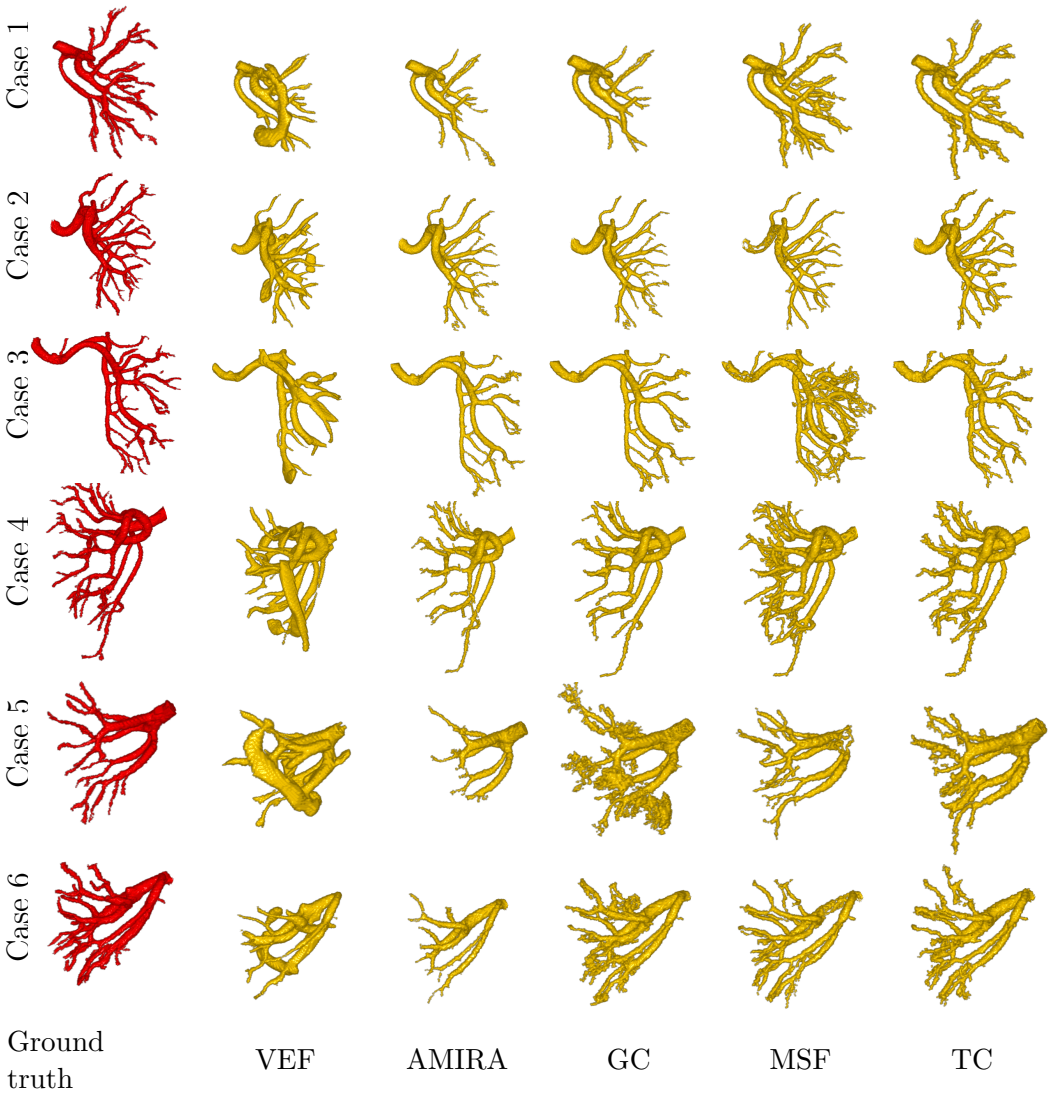


Figure 2.16: Volume rendering of six clinical cases.

indicate that TC is robust to noise and has a stable, higher performance than the other reference approaches. From Fig. 2.14, we identified the limitations of the ability of the model-fitting-based methods (SE and MHT) to segment tiny structures. This weakness results from the size limitation of a single vascular model. Other voxel-wise methods (including TC) do not suffer from this weakness.

Our experiments on clinical CT data aim to show the segmentation ability of blood vessels in actual clinical scenes. We show the overall quantitative CO and CE evaluation of 19 clinical cases (Fig. 2.15). In Fig. 2.15, we plot the best CO and CE

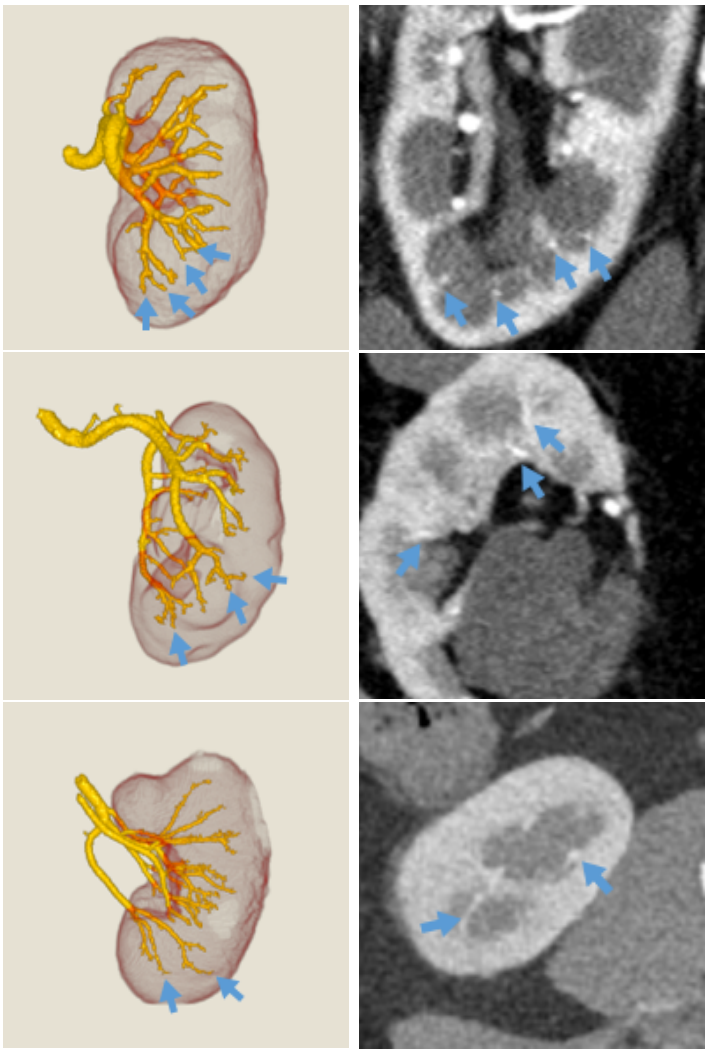
index of each case. TC achieved the best and most robust segmentation performance of all the reference approaches. TC's mean CO index is 84.0% and outperformed the second-best GC of 73.2%. TC's mean CE is 1.37mm outperforming the second-best GC of 2.11mm.

We show several detailed segmentation results on actual clinical CT slices in Fig. 2.17. The arrows indicate tiny renal arteries with small radii and low contrast against the surrounding renal cortex. Although these tiny renal arteries are easily overlooked by humans, our proposed method can segment them.

Similar to the simulated experimental results, TC outperformed the GC and VEF methods in the clinical experiments. GC achieved good segmentation results in Cases 2, 3, and 4 because high-quality CT angiographic data simplify the generation of good initial terminal nodes that contain tiny tubular structures. We show detailed comparison results of Cases 1 and 3 in Fig. 2.18 to evaluate the influence of the initial seed voxels on the segmentation results. GC's performance heavily depends on the initial seed voxels. The poorly extracted initial seeds directly lead to poor segmentation results. TC is more robust to the initial seeds. Owing to the tensor term adopted in TC, the short-boundary bias was suppressed. More tiny blood vessels were extracted by TC than GC for Case 1 (Fig. 2.18).

2.6.2 Conclusions

We presented a novel unsupervised blood vessel segmentation approach using a tensor-based graph-cut method. Our simulated and clinical experiments showed that our proposed method has higher segmentation performances than other reference approaches. In this work, we focused on renal artery segmentation. Our clinical experimental results show that our tensor-cut method is robust and outperformed the other reference methods, especially on tiny blood vessels. Accurate vascular topology is the foundation of follow-up processing such as estimation of vascular dominant regions.



3D volume rendering 2D CT slice

Figure 2.17: Segmentation performances on CT slices: Segmented tiny blood vessels are indicated by arrows on CT slices. These tiny renal arteries can be extracted by our proposed method.

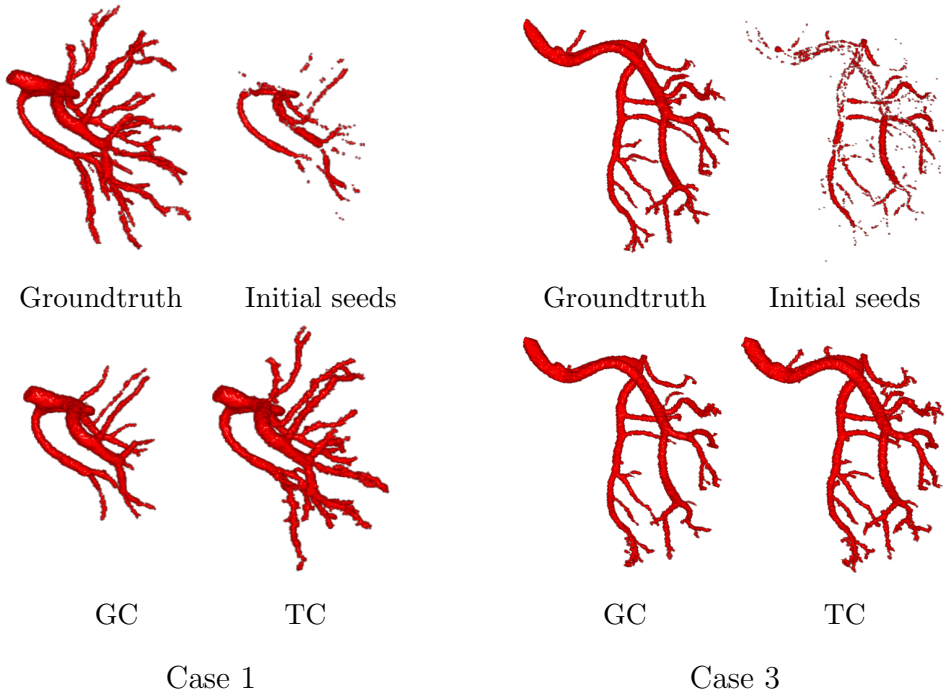


Figure 2.18: Comparison between GC and TC on clinical Cases 1 and 3: Initial foreground seed voxels (s nodes) are shown in upper-right. Segmentation results of GC and TC are shown in lower rows.

Chapter 3

Tubular organ segmentation: Tubular-oriented FCN

3.1 Background

We have introduced the unsupervised blood vessel segmentation method in Chapter 2. In this chapter, we introduce our second tubular structure segmentation method based on supervised deep-learning techniques. Different from traditional unsupervised segmentation approaches, supervised deep-learning based approaches is able to learn the essential features from manual annotated labels (*i.e.* ground truth) automatically. No handcrafted feature-design is needed. As previously mentioned, one major advantage of traditional unsupervised approaches is that they do not need any additional annotated labels. One major disadvantage is the manually-designed features. As contrast, supervised deep-learning based approaches automatically learn the features from the annotated data. They can derive benefits from the big data. However, it is also one disadvantage of supervised deep learning techniques. Generating annotated labels for 3D medical images is a time-consuming and labor-intensive task.

With the rapid development of deep-learning techniques, the medical image-processing community also has adopted the advances of deep-learning techniques, naturally. Compared to traditional machine learning-based methods, deep-learning based meth-

ods do not depend on handcrafted features. One natural advance of deep-learning architecture is that it is able to learn the useful from the training data automatically. Over the past few years, deep learning methods have become the dominant approach in many data analysis fields and also have achieved remarkable advances in the medical image analysis field [164]. There already exists lots of research using deep neural networks for their medical image-processing tasks, including organ segmentation, lesion detection and disease classification. Deep-learning based methods have already showed their significant advances over other traditional methods.

Airways and vasculature pose high demands on image analysis as they are elongated fine structures with calibers ranging from several tens of voxels to voxel-level resolution, branching in deeply multi-scale fashion, and with complex topological and spatial relationships. Recently, airways and vasculature image analysis have also been approached using deep convolutional neural networks (ConvNets) instead of the traditional filter banks and machine learning methods.

In this work, we made efforts to exploit the potential of 3D neural network in tubular structure segmentation problem. With the power of deep-learning techniques, we attempt to build a general segmentation method for tubular structures including vasculature and bronchi. Conventional neural networks still suffer under-segmentation problem for tiny tubular structures. To address this problem, we presented two new modules designed for tubular structures: 3D recurrent convolutional layer and radial distance loss. These two modules can be easily Incorporated into general FCN architecture. Experimental results showed that our modules significantly improved the segmentation accuracy.

3.2 Related works

3.2.1 Fully convolutional network

This work is constructed based on fully convolutional network (FCN) architecture. FCN architecture has been more and more popular in segmentation or self-learning tasks. It was first proposed by Long *et al.* [165]. The main feature of FCN is that it enables a end-to-end and pixel-to-pixel segmentation process, instead of a patch-based classification manner. Recently, a widely used FCN architecture U-Net [106] incorporates the concept of skip-connection [166] into FCN, and exceed the state-of-the-art performance in medical segmentation tasks. In this work, 3D U-Net architecture [107] is selected to be the backbone network.

3.2.2 Deep-learning in tubular organ segmentation

Many studies have been made using 2D/2.5D image patches to train their networks [117, 167]. Oda *et al.* [168] used multiple fully connected networks (FCNs) to tackle the problem of segmenting tiny abdominal arteries. Three 2D U-Net networks were employed to respectively segment 2D patches on three anatomical planes, and three 2D segmentation results were merged as a final volumetric segmentation. Similarly, Yun *et al.* [117] used 2D patches on three anatomical planes to train three ConvNets as classifiers to classify whether a voxel belongs to the bronchus. To explore the inherent relationship between such orthogonal 2D patches, Tetteh *et al.* [167] presented 2D orthogonal cross-hair filters to make use of 3D context information but with a reduced computational burden. The above approaches made use of local contexts to address the tubular-structure segmentation. Although, they achieved good performance, the problem remains of taking full advantages of large-range information in the field of view (FOV), which can provide global queues unavailable from small local patches. Meng *et al.* [118] attempted to use an original 3D U-Net combined with a traditional bronchus tracking method to improve bronchus segmentation accuracy. Huang *et al.*

[169] presented a liver-vessel segmentation method using a 3D U-Net with variant Dice loss function. Both works used a large 3D sub-volume size of around $100 \times 100 \times 100$ voxels.

Compared to 2D local patch-based methods, 3D FCNs for tubular-structure segmentation has some limitations that need to be addressed. Foremost is the severe imbalance in the sizes of the large background area when compared to the small foreground (tubular structure). Second, it's unclear whether FCNs can learn useful features from the abundant contexts of 3D data that while capturing coarse-scale long-range interrelationships are also accurate to detect tubular details and topological correctness at finer scales. The motivation of this work is to improve the accuracy of tubular-structure segmentation in respect to these two concerns.

3.3 Contributions

1. We introduce 3D recurrent convolutional layer in FCN architecture for tubular-structure segmentation.

   3D recurrent convolutional layer enables a more effective message-passing mechanism. This will significantly help networks learn the tiny elongate structure more effectively.

2. We propose a novel radial distance loss for 3D tubular-structure segmentation that helps the networks to recover tiny tubular structures.

   Weighted loss functions have been used to solve the imbalance problem between the background and foreground. Tubular organs have a large variation of scales. However, no loss function has been proposed to address the intra-imbalance problem inside the foreground. Our proposed radial distance loss is designed for tubular structures. It aims to solve the imbalance problem inside the tubular structures.

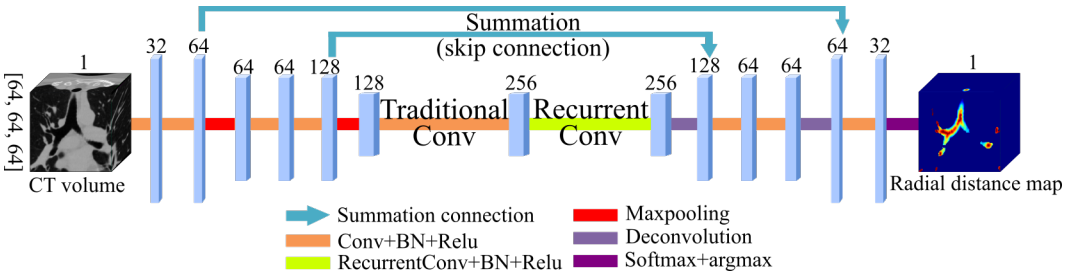


Figure 3.1: Overview architecture

3. Both 3D recurrent convolution layer and radial distance loss are generic and flexible, so they can be easily incorporated in other state-of-the-art networks and used in other applications.

   The proposed two modules are generic designed. They can be easily incorporated in other networks to share the advances of new network architectures.

3.4 Tubular-oriented fully convolutional network

3.4.1 Overview

Our main contribution is twofold from the viewpoint of tackling both network architecture and its loss function. The motivation of this work is to design a new end-to-end FCN for 3D-tubular structure segmentation. Compared with traditional layer-by-layer convolutional layers, the proposed slice-by-slice convolutional layer permits messages to efficiently pass slices. Our proposed loss function places more weight on the centerline of a tubular structure than on the outer border which can help the network pay more attention to tiny structures. Description of the proposed method is based on our published work [170]. A simple illustration of our architecture is shown in Fig. 3.1.

3.4.2 3D spatial FCN

Spatial-CNN was first proposed to address the traffic lane detection problem [171]. The main contribution of spatial-CNN is the introduction of a slice-by-slice convolu-

tional layer. Different from the traditional layer-by-layer CNN, slice-by-slice convolutions perform like recurrent neural networks. The in-layer recurrent convolutions provide more efficient message-passing between neurons in the same layer. This can help networks reinforce structures with strong spatial constraints [171].

In this work, we incorporated Recurrent Convolutional Layers (RCLs) in a 3D FCN architecture for 3D medical imaging applications. Detailed structure of RCLs is illustrated in Fig. 3.2. Convolutions are performed in both forward and backward directions along the width, height, and depth dimensions. As shown in Fig. 3.2, RCL_w, RCL_h, and RCL_d denote RCLs with convolutions in front-to-back direction along the three dimensions. $\text{RCL}_w^{'}$, $\text{RCL}_h^{'}$, and $\text{RCL}_d^{'}$ denote RCLs with convolutions in the reverse direction. Forward computation of RCL_w can be defined as

$$\boldsymbol{\mathcal{Z}}_{c,i,j,k} = \begin{cases} \boldsymbol{\mathcal{X}}_{c,i,j,k}, & \text{if } i = 0 \\ \boldsymbol{\mathcal{X}}_{c,i,j,k} + f(\boldsymbol{\mathcal{Z}}_{c,i-1,j,k} * K), & \text{if } 0 < i \leq W, \end{cases} \tag{3.1}$$

where $\boldsymbol{\mathcal{X}}$ denote an input 4D feature tensor of size $C \times W \times H \times D$, and $\boldsymbol{\mathcal{Z}}$ denotes the output of RCL_w. Let c, i, j, k denote the index of channel, width, height and depth dimensions. K is a 3D convolution kernel of size $w \times w \times w$. $f(\cdot)$ denotes a nonlinear activation function. RELU is used in this work. Similarly, forwarding for RCL_h and RCL_d can be easily derived based on Eq. 3.1.

By introducing skip-connection between encoding and decoding paths, U-net like architecture achieved state-of-the-art performance in many computer vision tasks.

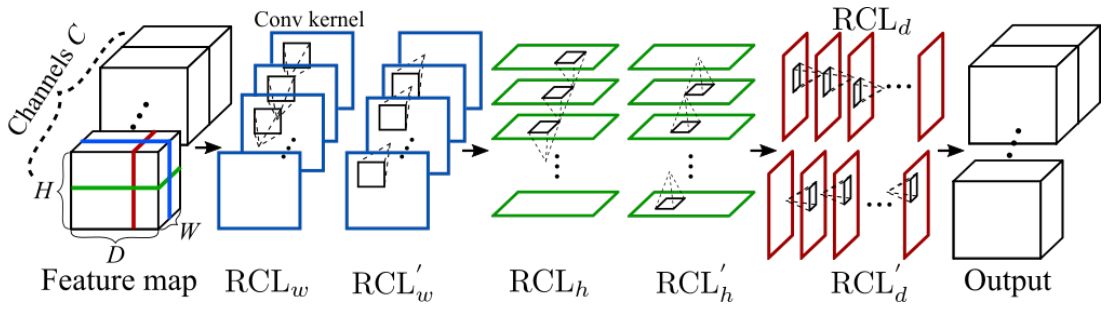


Figure 3.2: RCL architecture: Feature map represents the output of the last convolutional layer with a shape of $N \times C \times W \times H \times D$. Slice-by-slice recurrent convolutions are performed in each RCL.

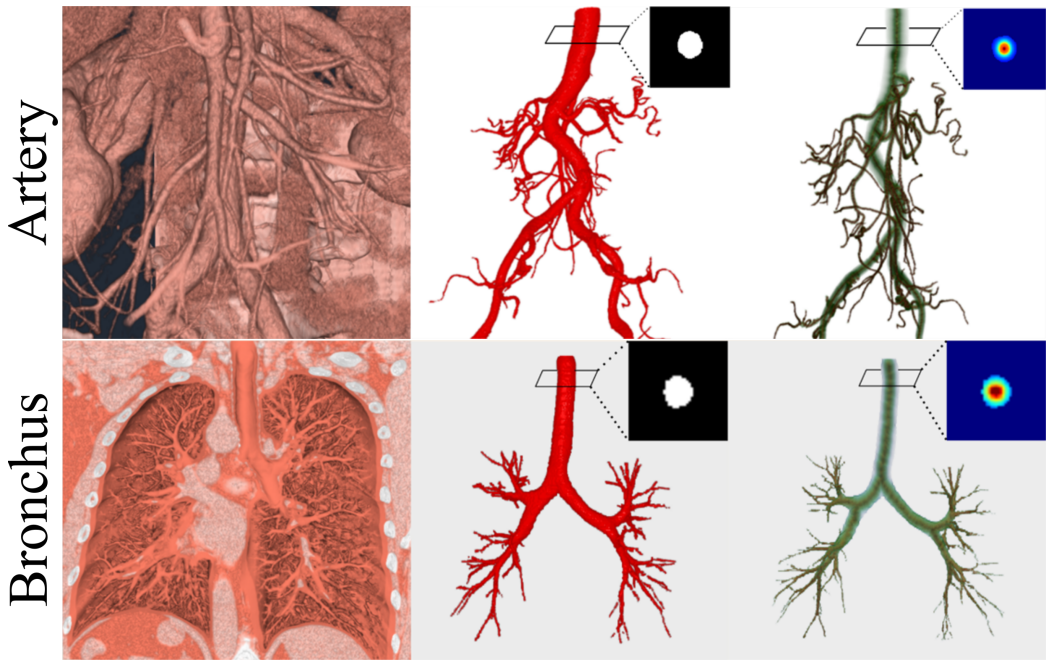


Figure 3.3: Volume rendering of binary ground truth and corresponding radial distance map. Radial distance maps of both artery and bronchus are shown. Radial distance map shown on the right is rendered with pseudo color mapping 0 to blue, and 1 to red.

In this work, we chose U-Net architecture [107, 124] with three multi-scale levels as our backbone FCN. As shown in Fig. 3.1, we placed the RCLs directly after the deepest feature map, i.e. the compressed low-dimensional representation. In previous spatial-CNN [171], the authors placed the RCLs after the output of CNN, under the assumption that the top hidden layer with rich semantic information is an ideal place to apply RCLs. However, in our FCN architecture, applying RCLs to the deepest representation provides better performance than using top hidden layer.

3.4.3 Radial distance loss

The use of Dice loss for medical segmentation tasks was first proposed by Milletari *et al.* [116]. They showed that Dice loss outperformed other losses, especially in severe data-imbalance situations [116]. More recently, Hausdorff distance loss and contour loss were proposed for localization and segmentation tasks [172, 173]. Dice loss measures volumetric variation, and Hausdorff distance loss and contour loss measure the

distance between boundaries. However, no losses were specifically designed for tubular structures. In this work, we propose a novel radial distance loss (RD loss) for tubular structures. The motivation of our proposed loss is to capture the geometric topology loss believed to be more important than volume loss in the tubular-structure segmentation.

The proposal of RD loss is inspired by the centerline overlap (CO) metric used for evaluating blood vessel segmentation [150]. Wang *et al.* used the CO metric to give a more accurate description of tiny blood vessel segmentation than conventional Dice similarity coefficient (DSC). To take advantage of the CO metric while keeping a volumetric measurement, our proposed RD loss is defined as

$$L = -\frac{1}{2}\sum_{k=0}^{1}\mathcal{W}_k\left(\frac{2\sum_i^N p_{i,k}d_{i,k}}{\sum_i^N p_{i,k}^2+\sum_i^N d_{i,k}^2}\right), \tag{3.2}$$

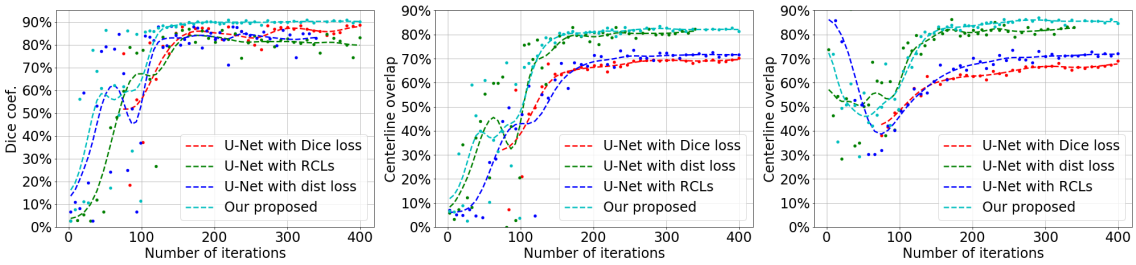
where $p_i \in \mathbf{P}$ and $d_i \in \mathbf{D}$ denote the i-th voxel predicted binary result and the radial distance map. Voxel index $i \in [1, N]$. $k \in [0, 1]$ denote class. Class weights $\mathcal{W}$ is defined as reciprocal volume ratio of each class. Notice that Eq. 3.2 is similar to traditional Dice loss [116] except that we use radial distance map $\mathbf{D}$ instead of binary ground truth $\mathbf{G}$. $\mathbf{D}$ is defined as:

$$\mathbf{D} = -\frac{1}{\max(\mathbf{F})}\mathbf{F} + 1, \tag{3.3}$$

where $\mathbf{F} = \{f_i, i \in N\}$ denotes a distance map created by a Euclidean distance transformation from the centerline of ground truth. $\mathbf{F}$ is defined as:

$$f_i = \begin{cases} \min\{d(f_i, s_j); s_j = 1\}, & \text{if } s_j = 0, \\ 0, & \text{if } s_j = 1 \text{ or } g_i = 0. \end{cases} \tag{3.4}$$

Here, $s_j \in \mathbf{S}$ denotes the j-th voxel of centerline data $\mathbf{S}$ extracted from $\mathbf{G}$. Finally, we obtained radial distance map $\mathbf{D}$ by normalizing $\mathbf{F}$ from 0 to 1 using a simple monomial form. An example of a radial distance map is shown in Fig. 3.3. RD loss can be weighted combination of centerline loss and Dice loss.



(a) DSC' (b) CO' (c) CO''

Figure 3.4: Validation results. Dot markers represent actual validation scores. For a clearer visualization, we plot smoothed results in dashed line using Gaussian filter.

3.5 Experimental results

3.5.1 Materials

In this work, to evaluate our proposed method, we performed bronchus segmentation experiments on 3D clinical CT scans acquired with a standard dose. The sizes of CT slices were 512×512 pixels with a resolution of $0.63 - 0.97$ mm. The number of CT slices ranged from 238 to 851 with varying thickness of $0.63 - 1.00$ mm. Monte Carlo cross-validation (MCCV) was conducted three times. All 38 CT images were randomly divided into the training and validation subsets containing 35 and 3 cases. The model with the best validation accuracy is tested on three unseen datasets acquired in a different hospital than those used for training.

3.5.2 Experimental results on bronchus segmentation

Since we chose U-Net as our backbone FCN, we performed comparison experiments using U-Net with Dice loss as a baseline. U-Net with only RCL structures and U-Net with only radial distance loss were also validated to prove the effectiveness of each proposal. We also compared our proposed method with two other methods, one is a variant U-Net architecture, V-Net [116] (Dice loss and RD loss), 3D voxelwise residual networks, VoxResNet [174] and the other is traditional bronchi tracking method [119].

In training phase, for each epoch, 4 sub-volumes with a size of $64 \times 64 \times 64$ voxels were randomly cropped from each CT image for all 35 training cases. No data aug-

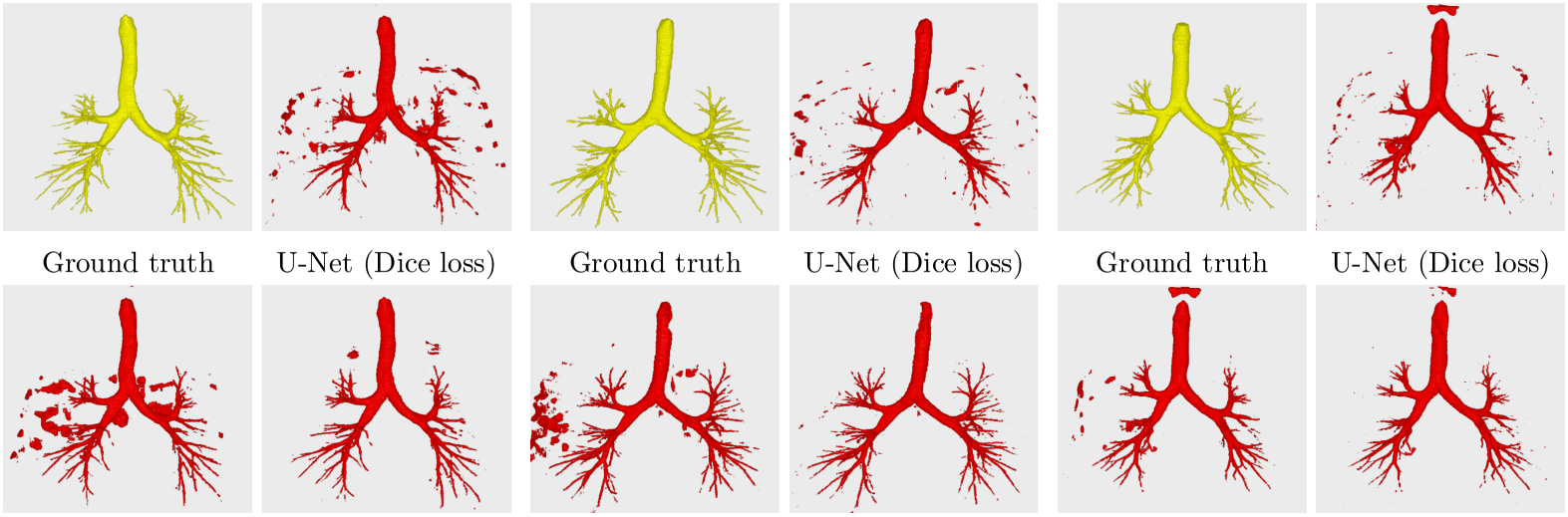


Figure 3.5: Volume rendering of segmentation results of three validation cases. No post-processing was performed. Trained models at 300^{th} iteration were used for prediction.

Table 3.1: Quantitative comparison results. All measurement was shown with mean ± standard deviation values.

Method	Dataset	DSC' (%)	DSC''(%)	Se (%)	CO' (%)	CO'' (%)
(1) In-house dataset						
Yun *et al.* [117]	Train: 59 Test:8	89.9 ± 8.9	-	-	-	-
Meng *et al.* [118]	Train: 30 Test:20	86.6	-	79.6	-	-
(2) Our bronchus dataset						
Kitasaka *et al.* [119]	MCCV Test:3	87.4 ± 2.3	82.8 ± 3.9	58.2 ± 7.8	49.1 ± 10.7	94.8 ± 1.1
VoxResNet (Dice loss) [174]		79.6 ± 3.7	90.0 ± 3.4	72.3 ± 5.0	39.2 ± 2.7	31.0 ± 2.1
V-Net (Dice loss) [116]		65.4 ± 9.9	91.0 ± 2.0	69.0 ± 2.0	28.3 ± 3.9	19.8 ± 1.1
V-Net (RD loss)		83.3 ± 2.0	88.4 ± 0.7	76.3 ± 4.6	53.8 ± 1.0	66.6 ± 4.9
U-Net (Dice loss) [124]		64.0 ± 19.5	92.4 ± 1.6	82.9 ± 5.7	47.2 ± 18.1	54.3 ± 9.0
Our proposed		88.7 ± 1.2	94.5 ± 0.8	86.5 ± 1.0	76.6 ± 6.0	80.6 ± 5.6

mentation was performed in our experiments. Random cropping was performed in this work. The initial learning rate was set to 0.01, and it decayed by 0.2 every 150 epochs. The optimization was realized via stochastic gradient descent (SGD).

DSC and CO were used for quantitative validation. To validate the general segmentation ability of each method, DSC' was measured on the segmentation results with no post-processing, and DSC'' was only measured on thick branches (before 2^{nd} generation of dichotomous branching). We computed CO scores on the results with two post-processing strategies. One is measured on the largest connected component extracted from the segmentation results. The other one is measured on tiny bronchi after the 2^{nd} generation of dichotomous branching and masked by a dilated ground truth with 5 voxels to remove false positives. These two measures are denoted CO' and CO''. CO' measures the ability to segment tubular structures, while CO'' measures the ability of tiny structures. Figure 3.4 shows validation results. Volume rendering of two validation cases are shown in Fig. 3.5. Quantitative comparison results are shown in Table 3.1. Other than DSC and CO, we also compute sensitivity (Se). More detailed comparison results are shown in Fig. 3.6 and 3.7. In Fig.

3.7, we also show a failed case. As we can see from the figure, even in failed case, our proposed method has better performance than other methods. Furthermore, To demonstrate robustness of method, segmentation results of three unseen test data are shown in Fig. 3.8.

3.6 Discussion and conclusions

In this work, we proposed 3D recurrent convolutional layer and radial distance loss, and demonstrated the implementation of these proposals in a widely used U-Net architecture. As experimental results showed, our proposed approaches achieved significant improvement over our baseline architecture, and obtained competitive results with state-of-the-art methods. Our proposed extensions, viz. 3D RCL and radial distance loss, are generic and flexible component that they can be easily incorporated in other deep learning architectures. We demonstrated an application on V-Net architecture using our RD loss, remarkable improvement was obtained comparing with the one using Dice loss.

Figure 3.5 and 3.6 show that thick bronchi are slightly under-segmented. This is a side effect of radial distance loss, since we decreased the weight of the most peripheral voxels. The normalization strategy in Eq. 3.3 can be improved to use more complex functions beyond a simple monomial. However, from Table. 1, DSC'' shows our segmentation accuracy of thick branches is still better than baseline U-Net. Segmentation results of three unseen datasets acquired in different hospital illustrated the robustness of our method. Visually, good segmentation accuracy was achieved.

In conclusion, we choose the challenging bronchus segmentation task to prove the effectiveness of our proposed method regarding general tubular structure segmentation. Experimental results showed that our proposed approaches are proven to be effective in bronchus segmentation task. However, we only evaluate the segmentation performance on bronchus segmentation. Our approaches should theoretically work

Figure 3.6: Volume rendering of more detailed comparison results (1).

Figure 3.7: Volume rendering of more detailed comparison results (2).

(a) Test case 1 (b) Test case 2 (c) Test case 3

Figure 3.8: Segmentation results of three unseen test data. For a clearer visualization, we removed outliers under volume size of 10 voxels.

on blood vessel segmentation problem. Application on blood vessel segmentation will be one of our future works. Additionally, more state-of-the-art networks incorporated with our approaches need to be investigated.

Chapter 4

Solid organ segmentation: Spatially aware FCN

4.1 Background

In this chapter, we introduce another main topic of our research: organ segmentation. Organ segmentation is another important research topic in medical image-processing field. In this work, we present a general designed fully convolutional neural network for organ segmentation and applied it to kidney segmentation. As we introduced in Chapter 1, we have introduced lot of research working on exploiting anatomical information in organ segmentation problem including both traditional methods and machine-learning based method. Although significant improvement has been achieved in organ segmentation with the power of deep-learning techniques, we believe better integrating with anatomical information can contribute to better performance for deep-learning based methods.

4.2 Contributions

1. 3D spatially aware FCN architecture is presented for organ segmentation task.

   In this work, a spatially aware unit is proposed to explicitly capture the anatomical information. This auxiliary module can effectively capture the relative

spatial location information. By integrating this information to the neural network, it enables to explicitly learn the organ's spatial information effectively. This help to suppress the false positives.

4.3 Spatially aware fully convolutional network

4.3.1 Overview

In this section, we describe our proposed methods in detail. We segment the kidney regions with a 3D U-Net-like fully convolutional network (FCN) architecture. U-Net architecture [106, 107], which is an extended version of FCN architecture, consists of a contracting path and a symmetric-expanding path. U-Net can achieve high segmentation accuracy with sparse annotated data [107]. Recently, many U-Net-like architectures have been proposed for segmentation tasks [109, 116, 124, 175–179]. Detailed description of our proposed spatially aware fully convolutional network is based on material from our published work [180].

4.3.2 Spatially aware unit

Our network is based on previous 3D U-Net-like architecture [124, 175]. Roth *et al.* presented a U-Net-like architecture for organ segmentation on 3D medical images and achieved state-of-the-art segmentation results [124]. To tackle the GPU memory limitation problem, they used a sliding-windows strategy for large medical data. However, these cropped sub-volumes were trained independently, i.e., the spatial position information of the sub-volumes was ignored during training. Spatial information is a critical feature for organ segmentation because the relative spatial position of the human organs is generally unchanged between patients. Exploiting spatial information should improve organ segmentation accuracy. Many works have involved spatial information into networks. Brust *et al.* directly incorporated absolute position information into fully connected layers [181]. Akoury *et al.* presented

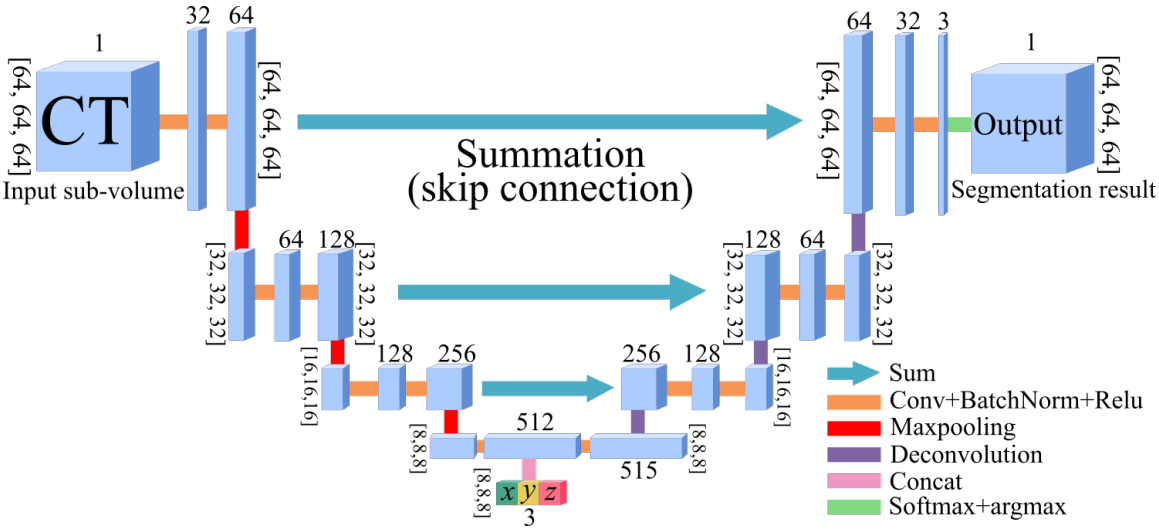


Figure 4.1: Architecture of our presented FCN.

a "Spatial PixelCNN" to impose spatial prior information to maintain the coherence of generated synthetic images [182]. Chen *et al.* incorporated spatial information into the end of an encoder (a bottom feature map) [183, 184]. Zhu *et al.* incorporated spatially structured learning in an adversarial FCN for mammographic mass segmentation [179].

Inspired by the work of Chen *et al.* [183], we introduce spatial position information into 3D U-Net-like architecture to impose the spatial information of each cropped sub-volume into our FCN architecture. Our proposed network is illustrated in Fig. 4.1. The backbone U-Net-like structure, which is based on a previous work [185], consists of four resolution levels. At each level, a skip connection links the contracting path and the corresponding symmetric-expanding path to provide higher resolution features to the symmetric-expanding path. Unlike original U-Net architecture, the skip connections in this network are summation instead of concatenation. Summation connections were first incorporated in U-Net by Roth *et al.* [109]. Their experimental results show that summation connections are slightly better than the original concatenation connections in the pancreas segmentation task. Each resolution level contains two series of convolutional layers, batch normalization and ReLU activation,

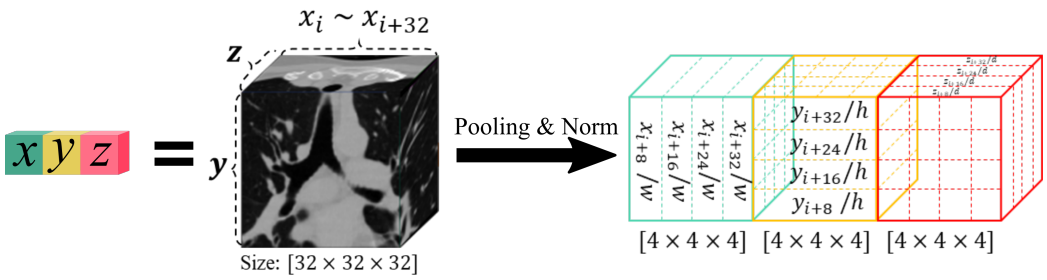
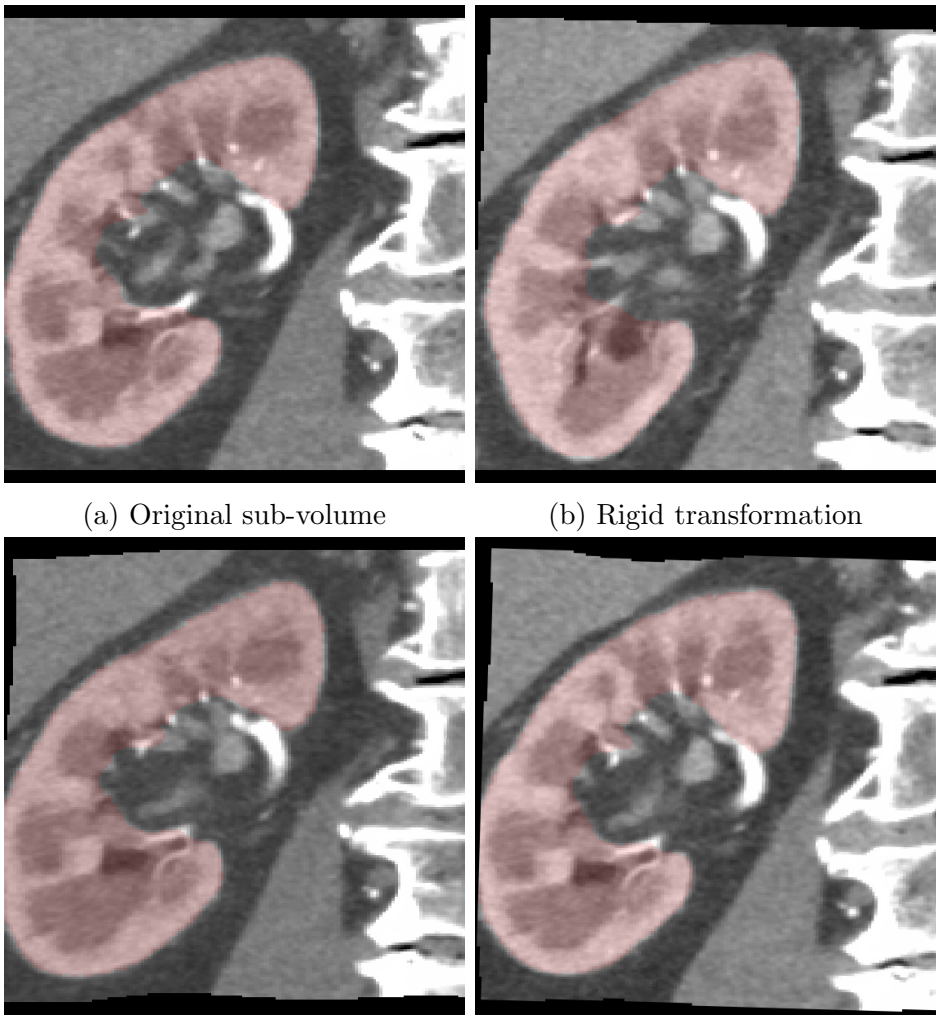


Figure 4.2: Illustration of spatially aware unit. Figure shows an example of spatially aware unit in 3 level FCN with input CT subvolume size of $32 \times 32 \times 32$ voxels.

in both the contracting and symmetric-expanding paths. The kernel sizes of all the convolutional and deconvolutional layers in our network are fixed to $3 \times 3 \times 3$. The kernel size of the max pooling layers is fixed to $2 \times 2 \times 2$.

We concatenated a three-channel feature map, including $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ coordinates, to the bottom feature map to introduce the position information to FCN. The input coordinate information is the relative coordinates of the input sub-volume in the entire CT normalized to $[0, 1]$. Let $\boldsymbol{\mathcal{P}}$ be the three-channel spatial feature map, thus $\boldsymbol{\mathcal{P}}$ is defined as $\boldsymbol{\mathcal{P}} = [\frac{\mathbf{x}}{W}, \frac{\mathbf{y}}{H}, \frac{\mathbf{z}}{D}]$, where $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ denote coordinates of voxels in a sub-volume. W, H, and D denote width, height and depth of a CT image. Unlike a previous study [183] that only considered the center coordinates of the input sub-volume, we used all of the position information and resized the position volume (containing coordinate information) to a suitable input size. An simple illustration of spatially aware unit is shown in Fig. 4.2

Training In this work, the input volume size of our network is fixed to $N_x \times N_y \times N_z$. At each epoch, n sub-volumes are cropped from the original CT volume and fed to the neural network. Here n denotes the batch size. To achieve the best segmentation performance, we exploit the transfer learning technique and pre-train our model on a multi-organ segmentation dataset [109, 175], which doesn't contain any kidney annotation, and fine-tune the model on our kidney dataset. This multi-organ segmentation dataset contains 377 cases, with 340 cases used for pre-training



(a) Original sub-volume (b) Rigid transformation

(c) Elastic transformation (d) Hybrid transformation

Figure 4.3: Examples of data augmentation: (a) original sub-volume extracted from CT data. Red region indicates ground-truth label of kidney. (b) and (c) rigid and elastic transformation results. (d) transformation result containing both rigid and elastic transformations

and 37 cases used for validation. The model with the best validation performance is used for fine-tuning. We used all pre-trained layers except the last classification layer. Fine-tuning was done on all layers.

Similar to previous works [106, 107, 109, 124, 175], we used a data-augmentation technique to increase the data variety and robustness. We performed both rigid and elastic transformations to each cropped sub-volume. Rigid transformation includes a translation with a range of $[-10, +10]$ pixels at each dimension and a rotation with

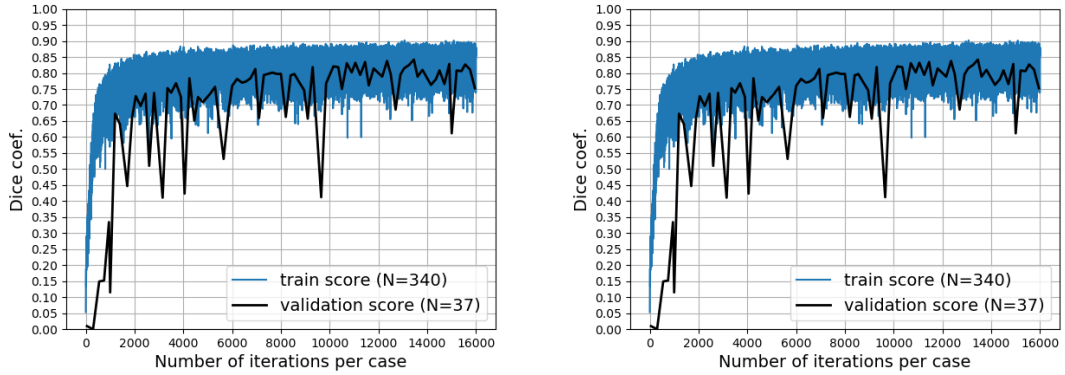
range of $[-15°, +15°]$. A B-spline deformation is employed as an elastic transformation to each sub-volume, like in previous works [107, 109, 124]. For each sub-volume, the deformation fields are randomly sampled from a uniform distribution with a maximum displacement of 3, and the number of B-spline control points in each dimension is set to 3 for all data-augmented experiments. In this work, we performed an in-place deformation operation. In a single iteration, we fed one original sub-volume and $n-1$ hybrid deformed sub-volumes to the network. A data-augmentation example is shown in Fig. 4.3. We individually show the rigid transformation, the elastic transformation, and the hybrid transformation, which combines the rigid and elastic transformations. Deformation is computed on-the-fly at each epoch during training.

In this work, we use pseudo Dice loss instead of conventional cross-entropy loss. Dice loss was first introduced in 2016 [116]. The definition of multi-class Dice coefficient loss D used in this work can be written:

$$D = -\frac{1}{K}\sum_{k=1}^{K}\left(\frac{2\sum_{i}^{N} p_{i,k}g_{i,k}}{\sum_{i}^{N} p_{i,k}^{2} + \sum_{i}^{N} g_{i,k}^{2}}\right). \tag{4.1}$$

where $g_{i,k} \in G$ and $p_{i,k} \in P$ denote voxels from the ground-truth volume and segmentation results for class k of total K classes. N is the total voxel amount of the volume.

Testing The input volume size for testing is the same as in the training. We used a sliding-window strategy to obtain sub-volumes with size $N_x \times N_y \times N_z$ for testing. After predicting all of the sub-volumes, we restored these predictions to a complete 3D labeling data based on their respective positions. The probabilities of the overlapping regions were computed as average probabilities: $p(x) = \frac{1}{R}\sum_{r}^{R} p_r(x)$, where $p_r(x)$ denotes the probability of voxel x from r-th sub-volume, $r \in R$.



(a) Pre-trained (b) Fine-tuned

Figure 4.4: Pre-training and fine-tuning learning curves: Blue curve denotes training Dice coefficient (DSC), and black curve denotes validation DSC. Maximum validation DSC of pre-training achieved 88.8%, and maximum validation DSC of kidney segmentation achieved 96.7%.

4.4 Experimental results

4.4.1 Materials

In this work, we used 27 pieces of abdominal contrast-enhanced CT volume data to evaluate the performance of our proposed method. The pixel spacing ranged from 0.665 to 0.742 mm, and slice pitch ranged from 0.4 to 2.0 mm. All 27 cases containing 54 kidneys were used to evaluate the accuracy of the kidney segmentation. Eight cases containing 14 kidneys were used for evaluating the dominant-region estimation. The ground truth of the kidneys was created by an engineer with medical knowledge.

4.4.2 Experimental results on kidney segmentation

We used an 8-fold cross-validation scheme to evaluate the accuracy of kidney segmentation. We divided all 27 CT volume data into train/validation/test splits at a ratio of 0.8/0.1/0.1. The model with the best validation performance is to be used for test. For a quantitative evaluation, we used three measures: the Dice Similarity coefficient (DSC), Sensitivity (Se), and the Hausdorff distance (HD). DSC is a commonly used measure in image segmentation. It is able to reflect the general segmentation

ability. Se, also known as recall rate, measures the ability of the method to extract correct kidney regions. HD is introduced to measure the surface distance between segmentation results and ground truth. We use the HD metric to focus on segmentation accuracy of the kidney itself. Therefore, as post-processing, we extracted the top 2 largest connected components as kidney regions. HD is used to measure the post-processed segmentation results, to reflect the actual segmentation ability of end-to-end FCN, DSC and Se are still used to measured the original segmentation results without any post-processing. DSC, Se, and HD are defined as follows:

$$\begin{aligned} DSC = \frac{2 * TP}{2 * TP + FP + FN}, Se = \frac{TP}{TP + FN}, \\ HD = \max(h(S_{gt}, S_{seg}), h(S_{seg}, S_{gt})), \end{aligned} \tag{4.2}$$

where True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) were measured in a voxel-wise way. $h(S_{gt}, S_{seg}) = \max_{p_{gt} \in S_{gt}} \min_{p_{seg} \in S_{seg}} \|p_{gt} - p_{seg}\|$, where p denotes the voxel coordinate that belongs to surface S. S_{gt} and S_{seg} represent the surfaces of the manually annotated ground-truth label and post-processed segmentation results.

Experiments were performed on an NVIDIA Quadro P6000 with 24 GB memory. Training on 21 cases took about $7-8$ hours for 4000 iterations, and the testing phase took about five minutes for a single case. For all experiments in this work, we set the learning rate to 0.01, the batch size n to 6, the sub-volume size $N_x \times N_y \times N_z$ to $64 \times 64 \times 64$, and the epoch number to 4000 for fine-tuning and 16000 for pre-training.

The FCN model was first pre-trained on a multi-organ dataset of seven organ labels, including the liver, the spleen, the stomach, and the pancreas. This multi-organ dataset did not contain any kidney labels. Its latest segmentation DSC is given in previous work [175], which achieved an average accuracy of 87.3% (excluding the background regions). The pre-training curve of our FCN model is plotted in Fig. 4.4(a). The maximum validation DSC score was 88.8%. The learning curve of the first fold's fine-tuning on the kidney dataset is shown in Fig. 4.4(b). The maximum validation DSC score of this fold was 96.7%. To demonstrate the improvement

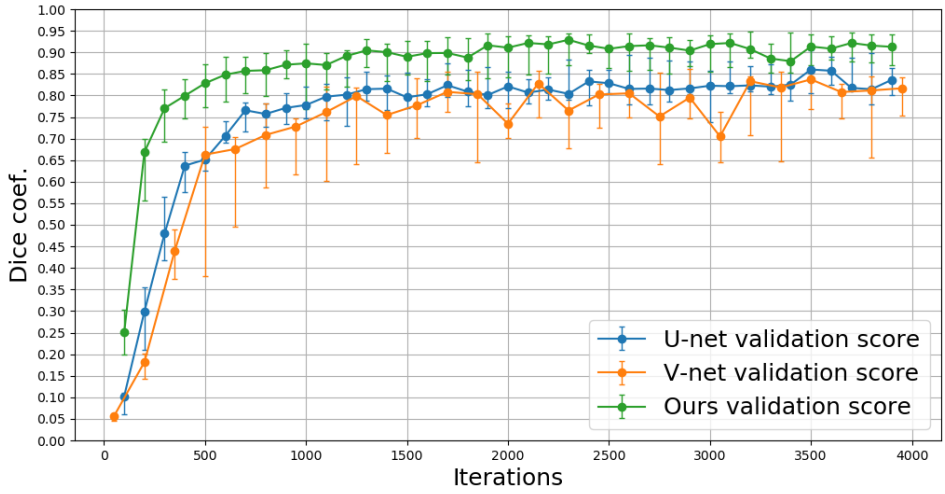


Figure 4.5: Validation *DSC* of U-Net [124], V-Net [116], and our proposed spatially aware FCN. All three of these networks were pre-trained on a multi-organ dataset. Points on curves denote median values of all 8-fold cross validations. Upper and lower bounds of error bars denote the first and third quartiles.

made by introducing the spatial unit, we also evaluated the baseline U-Net [124], i.e. the proposed architecture without the spatial unit. Furthermore, we implemented a variant U-Net architecture, V-Net [116], for comparison on our kidney dataset. All comparison experiments have the same hyperparameter settings. Validation results of these three methods are shown in Fig. 4.5. The quantitative results of unseen test data of all 8-fold cross validations are shown in Table. 4.1. For reference, we also list several similar works on the kidney segmentation task, even though we used a different dataset and annotations. Two detailed segmentation examples using the proposed approach are shown in Fig. 4.6. Comparison segmentation results are shown in Fig. 4.7.

Table 4.1: Comparison of kidney segmentation methods. Three measurements, Dice coef., sensitivity, and Hausdorff distance, are shown with median[1st - 3rd quartile] or mean ± standard deviation values. CE-CT denotes contrast-enhanced CT, LOO denotes leave-one-out cross-validation, and 8-fold denotes 8-fold cross validation. All listed methods used own in-house dataset.

Method	Modality	Case num	*DSC* (%)		*Se* (%)	*HD* (mm)	
			Left	Right		Left	Right
(1) In-house dataset							
Atlas-based random forest [186]	CT	Train: 233 Test: 179	96.0 [93.0-97.0]	96.0 [93.0-97.0]	-	7.0 ± 10.0	7.0 ± 6.0
Atlas-based graph-cut [187]	CT (portal-phase)	100 (LOO)	90.0 ± 5.0		-	-	
CNN+MSL [188]	CE-CT (multi-phase)	Train: 370 Test: 78	90.5		-	-	
Shape-constrained Level-set [189]	CE-CT	10	86.2		-	19.6	
2D patch-based CNN [190]	CE-CT	Train: 79 Test: 20	93.6 [92.0-95.0]	92.5 [88.8-94.5]	93.78 [90.7–95.6]	4.8 [2.7-9.5]	7.0 [4.1-17.5]
(2) Our kidney dataset							
Baseline U-Net [124]	CE-CT	27 (8-fold)	84.6 [79.6-90.3]	91.0 [84.5-92.7]	91.1 [83.1-94.2]	12.0 [7.3-15.4]	5.2 [4.8-8.0]
V-Net [116]	CE-CT	27 (8-fold)	84.1 [74.3-95.1]	86.7 [74.6-92.9]	85.6 [81.5-94.5]	13.9 [7.5-20.8]	8.1 [5.2-28.1]
Ours	CE-CT	27 (8-fold)	87.3 [83.0-90.8]	94.7 [86.4-95.5]	90.5 [86.3-95.5]	7.1 [5.0-9.8]	4.6 [3.8-6.9]

4.5 Discussion and conclusions

4.5.1 Discussion

As shown in Table 4.1, compared with other U-Net-like architectures, our proposed spatially aware FCN achieved better segmentation results. Figure 4.7 shows that the introduced spatial information effectively suppressed the FPs. Furthermore, our approach demonstrated competitive segmentation accuracy with related kidney segmentation methods. In this work, our *DSC* and *Se* are directly measured on the segmentation results without any post-processing. A 2D patch-based CNN method [190] performed post-processing, including opening, closing and extraction of two largest connected components. Therefore, a comparison of *HD* is more useful for

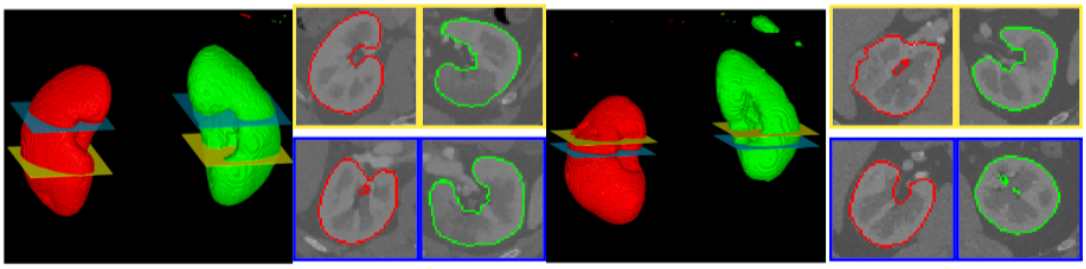


Figure 4.6: Two kidney segmentation examples using proposed FCN. Both 3D volume rendering and 2D segmented kidneys ROIs are shown. Red and green lines indicate contour lines of segmented kidneys. Yellow and blue sections shown in volume rendering correspond to 2D ROIs with same colors.

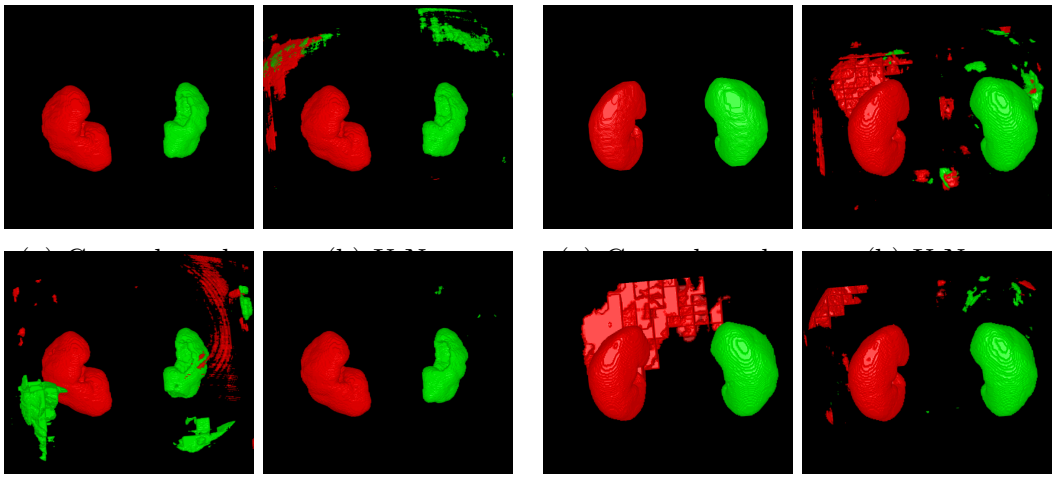


Figure 4.7: Two comparison examples. (a), (b), (c), and (d) denote ground truth and segmentation results of U-Net [124], V-net [116], our proposed network.

demonstrating our segmentation ability. Considering our limited dataset, we believe our proposed spatially aware FCN architecture has potential to achieve competitive results with state-of-the-art kidney-segmentation methods. Furthermore, the spatially aware unit can be easily incorporated in other architectures.

Although our network achieved good kidney segmentation results, its performance remains limited on such pathology patterns as kidney cysts and some late stage cancer. Several segmentation results are shown in Fig. 4.8. Figs. 4.8(a) and (b) show the under-segmented results of kidney cysts. Fig. 4.8(c) shows a case with late-stage

(a) (b) (c)

Figure 4.8: Limited segmentation performance: Yellow arrows indicate under-segmented regions. (a) and (b): limited segmentation performance on kidney cysts. (c): a case with kidney cancer at late stage.

cancer, and the proposed method failed to segment the whole kidney. One reason is that we have only one case that contains this pattern in our dataset; increasing the data with this pattern may improve the segmentation performance. Among all 27 tested cases, two cases failed in segmentation ($DSC < 20\%$). This has also happened in other experiments using U-Net and V-Net. The major reason for this is that the slice thickness of these two failed cases is 2.0 mm, while that of all other cases ranges from 0.4 to 0.8 mm.

4.5.2 Conclusions

Our proposed automatic organ segmentation method achieved competitive results with state-of-the-art methods for kidney segmentation problem in our in-house data. Since our network is generally designed, it can be also adopted to other organ segmentation problem, theoretically. In our research, we mainly focus on kidney-related CAD system. The high segmentation accuracy of kidney helped us build a reliable CAD system in the future.

Chapter 5

Computer-aided kidney diagnosis

5.1 Overview

All the methods we proposed in this thesis can be utilized for real clinical applications. CAD system will become meaningless without clinical applications. In this chapter, we introduce three clinical kidney-related applications based on our proposed method: optimizing surgical planning, quantitative analysis of renal function variation, and quantitative evaluation of kidney tumors.

Kidneys are very important solid organs in vertebrates. They are responsible for filtering impurities in the blood, maintaining body water balance and electrolyte balance. They also have endocrine functions to regulate blood pressure. An illustration of human urinary system is shown in Fig. 5.1. Kidneys are important part of human urinary system. As the figure shows, the kidney has three main anatomical regions: cortex region at outside, medulla region in the middle, and the renal pelvis. Nephrons, which is the minimal functional unit of kidney, located in renal cortex and medulla regions. In other words, renal cortex and medulla are the main functional regions of kidney.

Nephrology and urology are two medical specialties address kidney diseases. Urology mainly addresses kidney diseases such as renal cancer, renal cysts and kidney stones. Urology prefers physical treatments to address these diseases. In our research, we also focus on urology. According to the guidelines of renal cancer treatment from

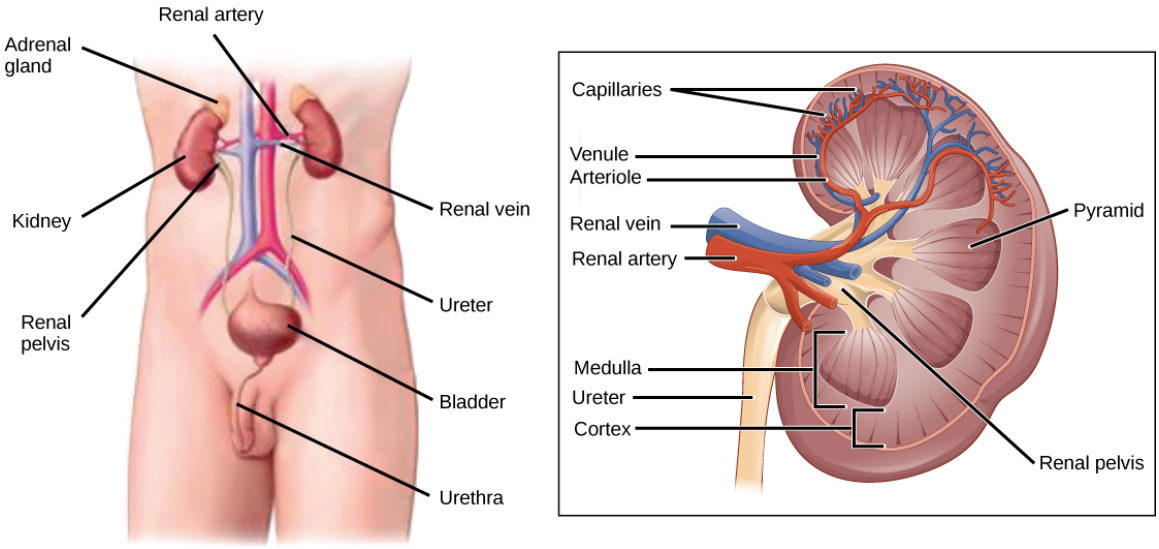


Figure 5.1: Illustration of human urinary system. On the right, the internal anatomical structures of kidney is shown. Photo is based on the materials from [191].

both the U.S. [192] and Japan [193], nephrectomy is the most recommended treatment of renal cancer. Our three applications are also related with nephrectomy. Although we cannot give a quantitative medical validation of our CAD methods, we mainly focus on the concept of our CAD methods. We hope our application can contribute to realize a better computer-aided kidney diagnosis system.

5.2 Applications

5.2.1 Presurgical simulation of partial nephrectomy

5.2.1.1 Background

Partial nephrectomy (PN), which has recently become one of the most common treatments for kidney cancer, can maintain a high residual renal function during surgery [194–196]. A critical problem during PN is that blood vessel clamping directly influences the quality of the surgery. However, PN remains unstandardized. Due to the trade-off between residual renal function and surgical difficulty, it is difficult to design a criterion for PN surgery. In this work, we provide a better and more accurate

computer-assisted diagnosis for PN by estimating the dominant region of each renal artery that facilitates identifying the blood vessels which feed the tumor. Physicians can easily make a surgical plan using such diagnosis information to determine the blood vessel clamping.

The feasibility of computer-aided diagnosis (CAD) for PN has been proven [197–199]. Ukimura *et al.* [197] performed PN on four patients who underwent 3D reconstruction for surgical navigation. The kidney surface was extracted by thresholding, the renal arteries were segmented by a simple region-growing method, and the tumor was manually segmented. Komai *et al.* [198] and Isotani *et al.* [199] used commercial software called "Vincent" to perform the computer analysis for PN clamping. The kidney was extracted by a semi-automatic region-growing method, and the renal arteries were segmented by applying facial detection technology using multi-phase information. Then the vascular dominant regions were estimated by applying a Voronoi diagram. All of the above research focused on the clinical study of CAD's feasibility and its accuracy for PN surgery rather than engineering studies.

The main contribution of this work is the computer-aided kidney diagnosis framework using our proposed automated and precise renal vascular dominant estimation approach. As a preliminary study on CAD system for PN, this work shows the potential of using medical image-processing techniques in improving precise PN.

5.2.1.2 Methodological Overview

The workflow of this CAD framework is shown in Fig. 5.2. First, the kidney regions are extracted using a deep learning approach. Second, a fine renal artery segmentation method is performed to segment the renal arteries inside the bounding-box of the kidney regions. After extracting the kidneys and the renal arteries, we estimate the vascular dominant regions with a Voronoi diagram. The relative statistics of the dominant regions are calculated for further surgical planning.

The kidney segmentation and renal artery segmentation methods are described

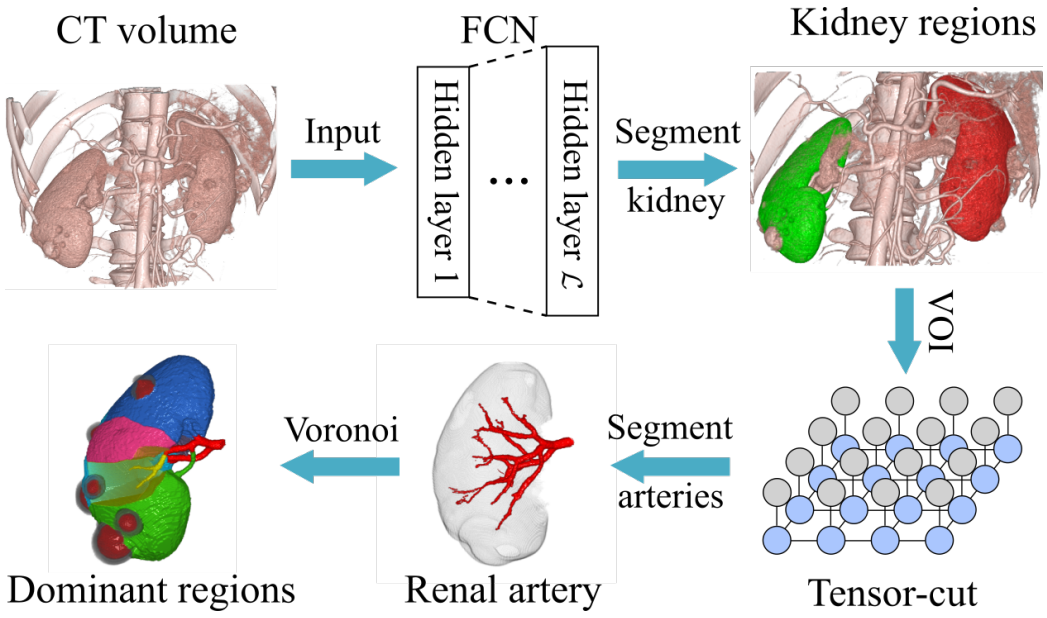


Figure 5.2: Workflow: Our precise estimation approach can be divided into three parts: kidney segmentation, renal artery segmentation and estimation of vascular dominant regions.

in Chapter 4 and 2. These two methods constitute the basis of this CAD system. Combining with other general image-processing methods, this CAD system can be constructed successfully.

5.2.1.3 Estimation of vascular dominant region

We estimated the renal vascular dominant regions with a Voronoi diagram that is widely used for calculating vascular dominant regions [197–199]. Considering the capillaries along the arteries, each branch of the renal arteries is treated as a set of seed points of a Voronoi diagram instead of using the end points of arteries. Let $B_i \in \mathbf{B}, (i \in \mathbb{R})$ be a branch of renal arteries, and define Voronoi cell $\mathcal{C}_i$ function:

$$\mathcal{C}_i = \{\mathbf{x} \in \mathbf{X_v} | d(\mathbf{x}, B_i) \leq d(\mathbf{x}, B_j)\}, \text{for all } i \neq j, \tag{5.1}$$

where image voxels $\mathbf{x}$ are inside of kidney region $\mathbf{X_v}$ extracted by the FCN approach. $d(\cdot)$ is the Euclidean distance between two points, and $d(\mathbf{x}, B_i)$ denotes the minimal Euclidean distance from point $\mathbf{x}$ to vessel branch B_i. This is a simple simulation of a real cell getting nutrition from blood vessels.

Table 5.1: Quantitative evaluation results of renal dominant regions. Eight cases were tested. Second and third rows present number of kidneys and dominant regions of each case. Fourth-row shows Dice score, and median values with 1st and 3rd quartiles values are demonstrated for each case.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Mean
Kidney #	2	2	2	2	1	1	2	2	
Regions #	12	22	16	10	12	7	10	14	
Dice coef. (%)	81.3 [70.2 - 85.1]	77.2 [66.9 - 87.8]	79.7 [73.3 - 86.0]	82.6 [71.9 - 87.0]	82.2 [78.0 - 87.9]	77.3 [73.0 - 81.2]	80.0 [70.6 - 82.4]	77.5 [70.4 - 87.2]	79.9 [70.4 - 86.2]

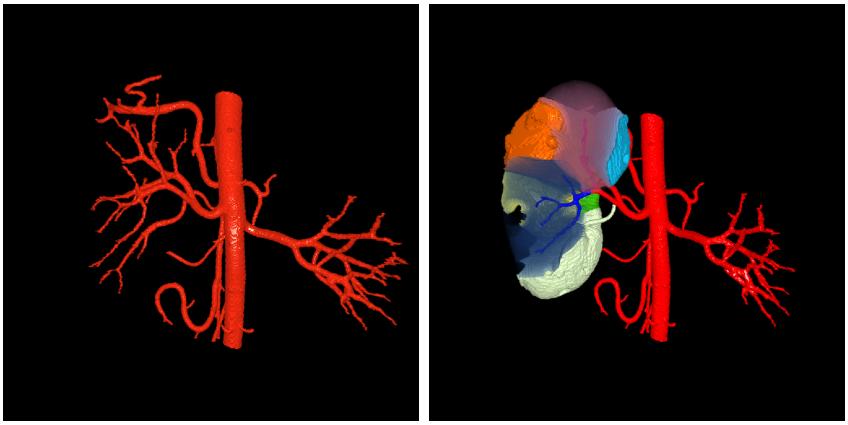


Figure 5.3: Experimental result of Voronoi diagram performed on original 3-D CT volume. Left: Abdominal blood vessels segmentation result. Thick blood vessels are segmented by region-grow semi-automatically, and renal arteries are extracted by proposed method. Right: Estimation result of renal vascular dominant regions that utilized Voronoi diagram.

As demonstrated in Eq. 5.1, the main factors affecting the accuracy of a Voronoi diagram are kidney region $\mathbf{X_v}$ and renal artery branches $\mathbf{B}$. To obtain a quantitative measure for the estimation of the vascular dominant regions, we calculated the volume and the volume ratio of each dominant region denoted by Vol and R. If dominant regions are adjacent to a tumor, we also calculated the contact area of each adjacent region.

5.2.1.4 Experimental results

We conducted a quantitative evaluation of the estimation of dominant regions in eight cases involving 14 kidneys. We measured each estimated dominant region's Dice coef. with the ground truth. Since we cannot obtain the anatomical ground truth of the

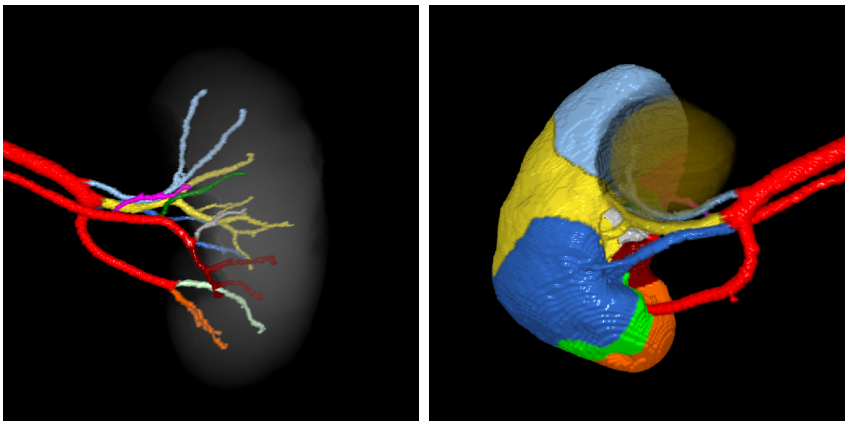


Figure 5.4: Left: Segmented renal arteries labelled with different colors for each branch. Right: Estimation result of vascular dominant regions. Colors correspond to their dominated blood vessels. 5-mm margin was taken outside of tumor. Blue, yellow, green, and fuchsia regions are directly adjacent to a tumor.

renal dominant regions, we used the ground truth of both kidney and renal artery to calculate a simulated ground truth of the dominant regions. The quantitative results are shown in Table 5.1. One partition example on the original CT volume is shown in Fig. 5.3.

Furthermore, one clinical validation was performed. The quantitative analysis results are shown in Table 5.2. Two measures were investigated: the volume of the vascular dominant region (Vol) and the area of the dominant region adjacent to tumor ($Area$). Nine regions were partitioned utilizing Voronoi diagram. Visualization of vascular dominant regions is shown in Fig. 5.4. Among these regions, regions 1, 5, 6, and 7 are directly adjacent to tumors. Nephrectomy surgery was performed using a selective artery clamping scheme shown in Fig. 5.5.

5.2.1.5 Discussion and conclusions

In this work, we present a CAD system for assisting partial nephrectomy (PN). We attempt to propose a reliable and standardized PN. Our proposed kidney and renal artery segmentation methods (described in Chapter 2 and 4) can achieve high accurate segmentation results. They are the basis of this CAD system.

From the quantitative analysis results of the clinical case, we confirmed that four

Table 5.2: Estimation result of renal vascular dominant regions. Regions 1, 5, 6, and 7 are directly adjacent to tumors. *Vol* denotes volume of dominant region, and *Area* denotes adjacent area of dominant region and tumor. Ratio of both *Vol* and *Area* are given in next row.

	1 Yellow	2 Blue	3 Light-green	4 Orange	5 Light-blue	6 Fuchsia	7 Green	8 Gray	9 Brown	all
Vol (mm^3)	49596.7	22567.1	22311.6	14275.1	20234	1603.6	8986.3	460101	16719.3	160895
Vol Ratio (%)	**30.8**	**14.0**	**13.9**	**8.9**	**12.6**	**1.0**	**5.6**	**2.9**	**10.4**	**100**
Area (mm^3)	17.6	–	–	–	761.5	174.9	22.5	–	–	976.5
Area Ratio (%)	**1.8**	–	–	–	**78.0**	**17.9**	**2.3**	–	–	**100**

regions are adjacent to the tumor, and thus at most four blood vessel branches should be clamped based on our simulated results. Actually, surgeons clamped two blood vessel branches that dominate regions 5 and 6. An operation report shows that slight bleeding remains in regions 1 and 7. However, this is considered as a valuable trade-off is found between surgical quality and residual renal function. From surgeon's feedback, our proposed approach helped surgeons build preoperative surgical plans and focus on critical arteries during operations. This is the aim of this research. We hope surgeon can benefit from our CAD system for PN.

As a pilot study on the estimation of renal vascular dominant regions, our experimental results in 8 cases demonstrated that our estimation approach achieved reasonable accuracy. It shows the potential of our system in real clinical applications. However, more clinical evaluations using large-scale database are needed to prove the feasibility of our approach for clinical PN surgical planning.

5.2.2 Quantitative analysis of renal function variation

5.2.2.1 Background

In this work, we aim to quantitatively analyze the renal functional variation between pre- and post-operation. Kidney has a self-healing mechanism when kidney is injured. However, there has been no related work to exploit the detailed healing process. Through this work, we hope to give a more detailed description of the self-healing

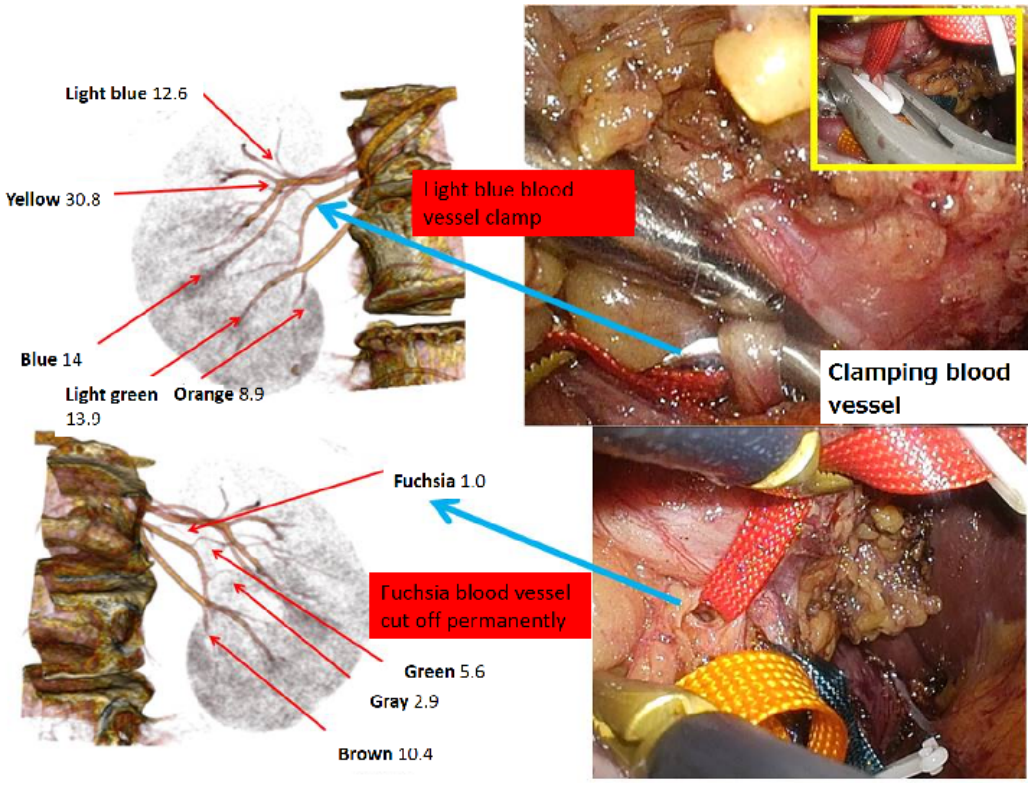


Figure 5.5: Selective artery clamping is performed in nephrectomy surgery. Blood vessel in light blue shown in Fig. 5.4 is clamped to prevent bleeding. Blood vessel in fuchsia is cut off permanently because it is the main blood vessel supply of nutrition to tumors.

mechanism.

The renal cortex and renal medulla are two main functional regions of kidney. The volume change of these two regions is one major index to reflect renal function variation. To have a better understanding of renal functional variation between pre- and post-operation, extracting the cortex and medulla regions from CT volumes is necessary for quantitative analysis.

However, it is a time-consuming task to segment these regions manually. To solve this problem, we present a semi-automatic computerized analysis approach to quantitative evaluate the functional variation between pre- and post-operation. Like the first application described in Section 5.2.1, to realize the quantitative evaluation, estimation of vascular dominant region is essential. Therefore, kidney segmentation and

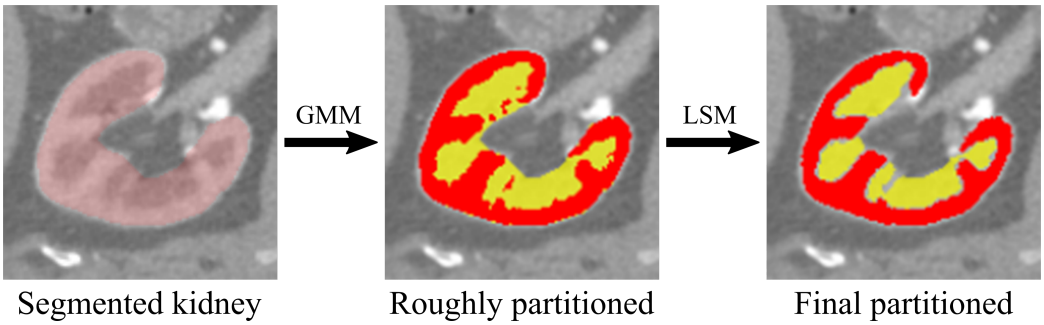


Figure 5.6: Simple workflow of The renal cortex and medulla segmentation.

renal artery segmentation is also necessary in this work. The kidney and renal artery segmentation methods are same with the ones used in the first application. Besides, we used Gaussian mixture model and LSM methods to extract the renal cortex and medulla regions.

5.2.2.2 Methodological overview

As we mentioned above, the methods extracting kidney and renal arteries are same with the first application. Two methods are detailed described in Chapters 2 and 4, respectively. However, this work focuses on quantitative analysis of renal function not methodology. All processing results are manually checked for ensuring the correctness. In this work, after kidney is extracted, the Gaussian mixture model (GMM) and the LSM is used to extract the renal cortex and medulla regions from kidney region. Then, Voronoi diagram is used to estimate the vascular dominant regions of both renal cortex and medulla regions. In the next section, we mainly describe the renal cortex and medulla regions segmentation approaches. A simple workflow in shown in Fig. 5.6

5.2.2.3 The renal cortex and medulla segmentation

In this paper, the Gaussian mixture model (GMM) and the LSM is used to extract the renal cortex and medulla regions. The LSM is one kind of active contour method. By adding an auxiliary dimension to active contour function, it can be solved by perform-

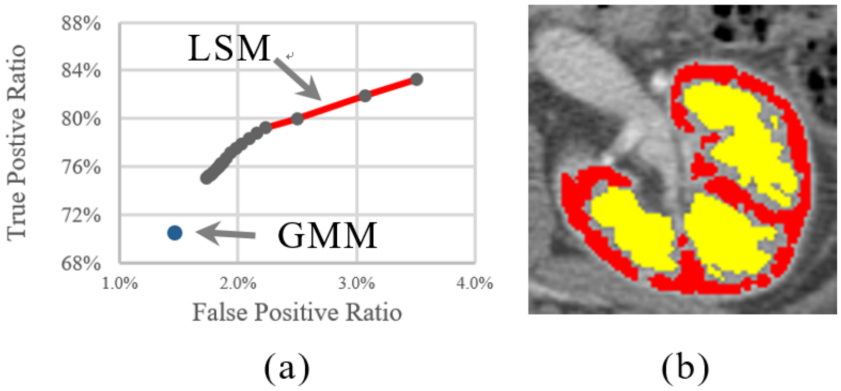


Figure 5.7: (a) ROC curve of the LSM. (b) Segmentation result of the LSM without post-processing. Segmented renal cortex region is shown in red, the renal medulla is shown in yellow.

ing numerical computations rather than parameterizing the surface. We also newly introduce a confidence interval to remove the over-segmentation of kidney regions.

The objective function of the LSM is consisted of three terms, propagation term, advection term and curvature term, parameters α, β and γ can be adjusted to control these three terms, respectively. The speed function $g(I)$ of the LSM used in this work is given as: $g(\mathbf{I}) = \exp(-|(\nabla * G_\sigma)|)$, $I\mathbf{I}$ denotes image, G_σ is a Gaussian kernel with standard deviation σ.

The GMM is used to extract the renal cortex and medulla roughly. The boundary of the extracted regions will be used as initial contour for the LSM. Firstly, the K-Means is used to cluster the voxels inside of kidney region into 2 classes, the renal cortex and medulla. Then, the MAP estimation will be performed on the voxels in each class to estimate their intensity distributions. We use the boundary of roughly extracted renal cortex and medulla regions as the initial boundary of the LSM. The input of the LSM is CT volume and initial boundary data. Segmented renal cortex and medulla regions can be obtained as final output. Closing operator with radius r will be performed as a post-processing.

5.2.2.4 Experimental results

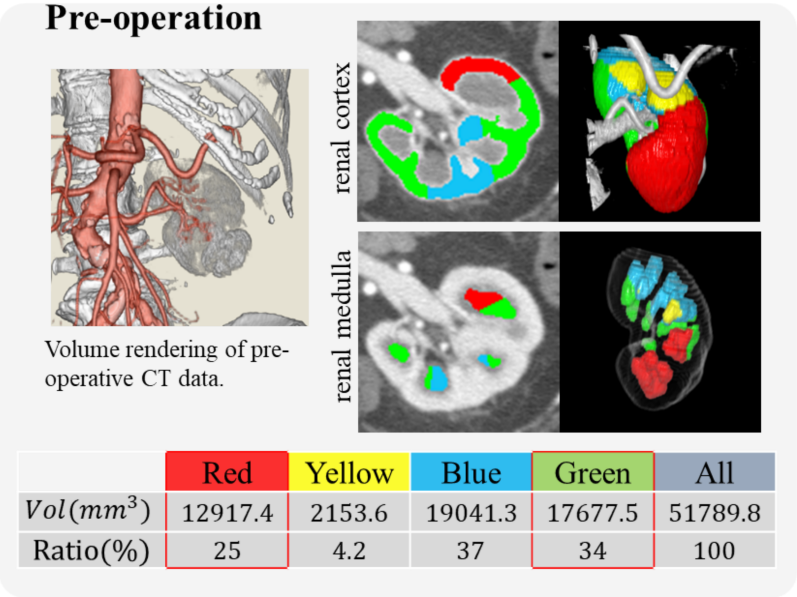
Since all the processing results are finally checked manually. Better segmentation results can decrease human labor on manual annotation process. In this work, eight cases of CT data were tested. We used one case to obtain the optimal parameters using grid search method. Segmentation and evaluation of the rest seven cases were performed using the optimal parameters. ROC curve is shown in Fig. 5.7(a). Figure 5.7 shows one slice of segmented result. Parameters were set as: $\sigma = 2.0, \beta = 3.0, \gamma = 0.2, r = 2$, iteration number of the LSM was set to 30. Mean Dice coefficient of seven cases was 88.2%, mean sensitivity was 89.7%, mean specificity was 99.1%.

One clinical case is demonstrated to show the renal function variation between pre- and post-nephrectomy. As shown in Fig. 5.8, the kidney region in red and green color are partially removed by operation. Volume of these two regions are decreased. However, the volume of kidney region in yellow increases after nephrectomy.

5.2.2.5 Discussion and conclusions

According to the clinical experimental result, we confirmed that kidney has self-healing mechanism when kidney is injured. In partial nephrectomy, kidney increases part of its functional region to cover the lost renal function. Traditionally, physicians only know that when one side kidney is removed in radical nephrectomy, the other one will grow to cover the lost renal function. Though this work, we confirmed that in partial nephrectomy, even in single kidney, it has self-healing mechanism to heal itself. Moreover, we suppose that this self-healing mechanism has some underlying rules. From the clinical result, we can find that only kidney region in yellow increases. In contrast, the kidney region in blue remain unchanged. Such behavior may reveal another research topic in a deeper analysis of how the kidneys work.

In conclusion, this section presented a CAD method for quantitatively analysis of renal functional variation between pre- and post-operation. By using our method, we can easily analyze the renal functional variation from CT data. Our work confirms



	Red	Yellow	Blue	Green	All
$Vol(mm^3)$	12917.4	2153.6	19041.3	17677.5	51789.8
Ratio(%)	25	4.2	37	34	100

	Red	Yellow	Blue	Green	All
$Vol(mm^3)$	9637.8	9763.4	18109.1	17677.5	48806.6
Ratio(%)	20↓	20↑	37	23↓	100

Figure 5.8: One clinical case of pre- and post-nephrectomy.

kidney's self-healing mechanism in partial nephrectomy. A fully-automatic and high accurate method is one of our future works.

5.3 Conclusions

In this chapter, two kidney-related clinical applications are presented. For both applications, the segmentation of kidney and renal arteries are necessary. They constitute the basic processes of kidney-related CAD system. In the first application, we provide necessary precise diagnostic for surgeons. Surgeons can make better surgical plan based on the quantitative results. As a pilot study, this work confirmed that this CAD framework can contribute to realize a standardized partial nephrectomy. However, this complex CAD framework is far from large-scale clinical use. Both precision and reliability need to been proven by more clinical validation. The second application focus on renal function analysis. Unlike the first application, the second one does not directly relate to nephrectomy or treatment. The newly discovered behavior may reveal another research topic in a deeper analysis of how the kidneys work.

The presented two applications demonstrate how our segmentation methods support the computer-aided kidney diagnosis, and also show the importance of precision of segmentation methods. More precise methods help to build a more accurate and reliable CAD system. It is the basis of a CAD system for its practical use in real clinical scene.

Chapter 6

Conclusion and future overview

6.1 Summary

The research described in this thesis aims to present tubular and solid organ segmentation methods with high accuracy. Combination of these segmentation methods constitutes a reliable and accurate CAD system. Recent years, CAD techniques have been rapidly developed benefiting from the fast development of medical image-processing techniques. CAD systems have already played a huge role in many real clinical applications, such as lesion detection, early disease diagnosis, and quantitative pathological analysis. Such applications have already been used in hospitals to support doctors for diagnosis. To realize these applications, segmentation of tubular organ and solid organ has always been the fundamental task for other complex CAD system. We can extract useful diagnostic information from the segmentation results for further analysis. In this thesis, our research mainly focuses on tubular organ and solid organ segmentation, and also demonstrate two clinical applications using our methods to assist kidney-related diagnosis.

The first topic, presented in Chapters 2 and 3, proposed two methods aiming to address the fine tubular organ segmentation problem. Segmentation of tubular organs, such as blood vessel and bronchus, remains a challenge medical image-processing task. In Chapter 2, tensor-based graph-cut method was used to perform a highly precise blood vessel segmentation. There have already been proposed amount of research

working on precise blood vessel segmentation. I believe that capturing the geometrical information (tubular structure) of blood vessel is the key of this problem. Based on this assumption, a tensor-based graph-cut method is proposed, named *tensor-cut*, to capture both geometrical and appearance information. Geometrical information is modeled by a second-order tensor using Hessian matrix. Then, both tensor and intensity are integrated using a graphical model. Tensor (geometrical) and intensity (appearance) are combined into one unified model. The experimental results show that our method outperforms other competitive methods, and have a good performance in renal artery segmentation task.

In Chapter 3, deep-learning technique was adopted in tubular organ segmentation problem. We attempted to exploit the potential of deep-learning techniques in 3D tubular structure segmentation problem. In this work, we presented two deep-learning based modules designed for tubular structure segmentation: 3D recurrent convolutional layer (RCL) and radial distance loss (RD loss). 3D RCL structure enables a more effective message-passing mechanism. It leads to a strong prior-shape constraint. Tubular structures can be then emphasized. Another module, RD loss, makes efforts on proposing a new loss function designed for tubular structures. This new loss function pays more attention on skeleton rather than volume. This strategy can effectively solve the imbalance problem inside of tubular structures which have large variance in scales.

The second topic, presented in Chapter 4, used an spatially aware fully convolutional network to address the solid organ segmentation problem. Many works have incorporated anatomical information in their methods to improve the organ segmentation accuracy. In this work, we explicitly integrated anatomical information, *i.e.* relative spatial location, to deep neural network. This allows the method to more effectively capture the anatomical information from the 3D CT data. Deep-learning is a large data-demand technique. Large annotated training dataset is required to an accurate deep-learning model. By explicitly feeding auxiliary information to neural

network, the proposed method can work with less training data. The experimental results demonstrate that the proposed network successfully captured the anatomical information of organs. False positives can be effectively suppressed.

The third topic, presented in Chapter 5, described two clinical applications constructed based on our tubular and solid organ segmentation methods. Normally, CAD system is developed to serve real clinical applications. The true value of CAD system will be reflected in clinical applications. In this topic, two kidney-related diagnosis applications were given to evaluate the clinical value of our CAD techniques. More specifically, we mainly focus on nephrectomy-related diagnosis. Nephrectomy has been a commonly used and recommended physical treatment for renal cancer. Partial nephrectomy is highly recommended than radical nephrectomy. However, partial nephrectomy still remains unstandardized. Due to the trade-off between residual renal function and surgical difficulty, it is hard to design a criterion for surgical plan of partial nephrectomy.

Our first CAD system aims to provide more accurate computer-assisted diagnostic information by estimating the dominant regions of renal arteries that facilitates identifying the blood vessels feeding the tumor. One clinical validation shows that our CAD system successfully helped surgeons make a better surgical plan. The second application attempt to explore the self-healing mechanism of kidney. There has been no related work to exploit the detailed healing process. In this work, by measure the volume changes of renal cortex and medulla regions, we can have an estimation of renal function variation between pre- and post-nephrectomy. To realize this quantitative analysis, processing methods, such as blood vessel segmentation, organ segmentation, are also necessary. According to the clinical experimental result, we confirm that kidney has self-healing mechanism when kidney is injured. Furthermore, we find renal self-healing also happens in single kidney. Such behavior shows another future work to explore the deeper analysis of how kidney works.

These three topics summarized above constitute my major works during my Ph.D

study. The first and second topic mainly contribute to medical image-processing field from an engineering perspective. These methods have the potential to form a useful basis for a accurate and reliable organ segmentation system. The third topic mainly contribute to clinical medicine field. Our kidney-related CAD systems show the feasibility in supporting physicians in real nephrectomy applications.

6.2 Future research

Although, we have proposed several segmentation problems to address specific issues in medical image segmentation. They are still far from perfect. Nowadays, CAD usually cannot substitute doctors, but rather plays a supporting role. This situation is similar with autonomous driving technology, which also plays a supporting role in current applications. Precision and reliability still remain challenging tasks for most CAD systems. These issues will become major research topics in the medical image-processing field. Here, we show some our opinions on this question.

Improving accuracy is a natural demand for most engineering approaches. More specifically, in our research, improve the segmentation accuracy of tubular and solid organs has always been the first priority of our research topics. Needless to say, deep-learning techniques have taken CAD to another level in both accuracy and robustness. Deep-learning techniques are still fast developing. Many new methods are proposed in this field, such as graph convolutional networks (GCNs), attention mechanism. At early stage, we just adopted deep-learning to medical image-processing field without modification. Now, more and more research have proposed specific neural networks using particular information extracted from medical images. In our solid organ segmentation method, we explicitly use anatomical information extract from 3D CT data to improve the segmentation accuracy. As one future research direction, we believe integrate prior-knowledge of medicine to deep-learning can effectively improve performance. For instance, Qin *et al.* proposed a neural network using the prior-knowledge

that tiny structures of bronchi mostly locate near the lung borders [120]. Therefore, we think taking advantage of prior-knowledge is one potential research direction to improve performance.

In some clinical applications, reliability can be more important than precision. In our consideration, beside of algorithm robustness, another important research topic is uncertainty estimation. This topic has also attracted a lot of attention in this field [200–203]. One future application of uncertainty estimation is used to improve accuracy and robustness by minimizing the uncertainty. Another application is to show the uncertainty to the doctors. Understanding of both certain processed results and uncertain information is vital for doctors to make definite, informed conclusions. We think the auxiliary uncertainty information can contribute to realize a more reliable CAD system.

6.3 Conclusions

Computer-aided diagnosis (CAD) system contributes to help doctors have a precise and efficient computer-assisted diagnosis. The basis of CAD system is the medical image-processing algorithms. Better processing methods can help to realize more accurate and reliable CAD system.

As stated in Chapter 1, this research aims to proposed tubular and organ segmentation methods with higher accuracy which meet the real clinical demand. The proposed methods presented in Chapter 2, 3 and 4 have demonstrated accurate and reliable segmentation results. In Chapter 2, tensor-cut algorithm has address the under-segmentation problem of tiny blood vessels comparing with other unsupervised methods. In Chapter 3, deep-learning technique has been adopted in tubular structure segmentation problem. 3D U-Net architecture is used incorporated with the proposed modules, 3D RCL and RD loss. This method has explored the potential of 3D FCN architecture in tubular structure segmentation problem. In Chapter 4, spatially aware

FCN architecture is used for solid organ segmentation task. The explicitly integrated spatial location helps the network more easily capture the anatomical information.

The segmentation methods presented in this thesis take fully advantage of geometrical and anatomical information. As the technology develops, CAD techniques will be subdivided into more specific research fields. Single unified framework is difficult to handle all these topics. Capturing the anatomical features of each specific research target will become more important.

In the future, CAD system will without doubt play an important role in modern medical field. We can imagine in the near future, such applications like automatic full-body pre-diagnosis, standardized and low-risky surgery will become reality. With the improved diagnosis efficiency, hospital appointments will become more convenient. It is very likely that CAD systems will become essential equipment in hospitals and change everyone's medical experience.

Bibliography

[1] World Kidney Day, *Chronic kidney disease*, https://www.worldkidneyday.org/facts/chronic-kidney-disease, 2015.

[2] Y. Mori, S. Kudo, P. W. Y. Chiu, R. Singh, M. Misawa, K. Wakamura, T. Kudo, T. Hayashi, A. Katagiri, H. Miyachi, *et al.*, "Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: An international web-based study," *Endoscopy*, vol. 48, no. 12, pp. 1110–1118, 2016.

[3] G. S. Lodwick, C. L. Haun, W. E. Smith, R. F. Keller, and E. D. Robertson, "Computer diagnosis of primary bone tumors: A preliminary report," *Radiology*, vol. 80, no. 2, pp. 273–275, 1963.

[4] G. S. Lodwick, A. H. Turner, L. B. Lusted, and A. W. Templeton, "Computer-aided analysis of radiographic images," *Journal of Chronic Diseases*, vol. 19, no. 4, pp. 485–496, 1966.

[5] F. Winsberg, M. Elkin, J. Macy Jr, V. Bordaz, and W. Weymouth, "Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis," *Radiology*, vol. 89, no. 2, pp. 211–215, 1967.

[6] M. L. Giger and K. Doi, "Investigation of basic imaging properties in digital radiography. i. modulation transfer function," *Medical Physics*, vol. 11, no. 3, pp. 287–295, 1984.

[7] H.-P. Chan, B. Sahiner, M. A. Helvie, N. Petrick, M. A. Roubidoux, T. E. Wilson, D. D. Adler, C. Paramagul, J. S. Newman, and S. Sanjay-Gopal, "Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: An ROC study," *Radiology*, vol. 212, no. 3, pp. 817–827, 1999.

[8] F. Li, M. Aoyama, J. Shiraishi, H. Abe, Q. Li, K. Suzuki, R. Engelmann, S. Sone, H. MacMahon, and K. Doi, "Radiologists' performance for differentiating benign from malignant lung nodules on high-resolution CT using computer-estimated likelihood of malignancy," *American Journal of Roentgenology*, vol. 183, no. 5, pp. 1209–1215, 2004.

[9] K. Suzuki, M. Hori, E. Mcfarland, A. Friedman, D. Rockey, and A. Dachman, "Can CAD help improve the performance of radiologists in detection of difficult polyps in CT colonography?," Jan. 2009, p. 872.

[10] H. Qi, P. T. Kuruganti, and Z. Liu, "Early detection of breast cancer using thermal texture maps," in *Proceedings IEEE International Symposium on Biomedical Imaging*, 2002, pp. 309–312.

[11] A. Cruz-Roa, H. Gilmore, A. Basavanhally, M. Feldman, S. Ganesan, N. N. Shih, J. Tomaszewski, F. A. González, and A. Madabhushi, "Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent," *Scientific Reports*, vol. 7, p. 46 450, 2017.

[12] J. Xu, L. Xiang, Q. Liu, H. Gilmore, J. Wu, J. Tang, and A. Madabhushi, "Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 119–130, 2015.

[13] S. G. Armato, F. Li, M. L. Giger, H. MacMahon, S. Sone, and K. Doi, "Lung cancer: Performance of automated lung nodule detection applied to cancers missed in a CT screening program," *Radiology*, vol. 225, no. 3, pp. 685–692, 2002.

[14] M. G. Penedo, M. J. Carreira, A. Mosquera, and D. Cabello, "Computer-aided diagnosis: A neural-network-based approach to lung nodule detection," *IEEE Transactions on Medical Imaging*, vol. 17, no. 6, pp. 872–880, 1998.

[15] N. Tajbakhsh and K. Suzuki, "Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs," *Pattern Recognition*, vol. 63, pp. 476–486, 2017.

[16] H. Jiang, H. Ma, W. Qian, M. Gao, and Y. Li, "An automatic detection system of lung nodule based on multigroup patch-based deep learning network," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 4, pp. 1227–1237, 2017.

[17] Y. Mori, S.-e. Kudo, M. Misawa, and K. Mori, "Simultaneous detection and characterization of diminutive polyps with the use of artificial intelligence during colonoscopy," *VideoGIE*, vol. 4, no. 1, pp. 7–10, 2019.

[18] H. Itoh, H. R. Roth, L. Lu, M. Oda, M. Misawa, Y. Mori, S.-e. Kudo, and K. Mori, "Towards automated colonoscopy diagnosis: Binary polyp size estimation via unsupervised depth learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 11071, 2018, pp. 611–619.

[19] M. F. McNitt-Gray, E. M. Hart, N. Wyckoff, J. W. Sayre, J. G. Goldin, and D. R. Aberle, "A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution CT: Preliminary results," *Medical Physics*, vol. 26, no. 6, pp. 880–888, 1999.

[20] M. Aoyama, Q. Li, S. Katsuragawa, H. MacMahon, and K. Doi, "Automated computerized scheme for distinction between benign and malignant solitary pulmonary nodules on chest images," *Medical Physics*, vol. 29, no. 5, pp. 701–708, 2002.

[21] K. Mori, N. Niki, T. Kondo, Y. Kamiyama, T. Kodama, Y. Kawada, and N. Moriyama, "Development of a novel computer-aided diagnosis system for automatic discrimination of malignant from benign solitary pulmonary nodules on thin-section dynamic computed tomography," *Journal of Computer Assisted Tomography*, vol. 29, no. 2, pp. 215–222, 2005.

[22] Z. Jing, L. Bin, and T. Lianfang, "Lung nodule classification combining rule-based and SVM," in *IEEE International Conference on Bio-Inspired Computing: Theories and Applications*, 2010, pp. 1033–1036.

[23] T. W. Way, B. Sahiner, H.-P. Chan, L. Hadjiiski, P. N. Cascade, A. Chughtai, N. Bogot, and E. Kazerooni, "Computer-aided diagnosis of pulmonary nodules on CT scans: Improvement of classification performance with nodule surface features," *Medical Physics*, vol. 36, no. 7, pp. 3086–3098, 2009.

[24] M. C. Lee, L. Boroczky, K. Sungur-Stasik, A. D. Cann, A. C. Borczuk, S. M. Kawut, and C. A. Powell, "Computer-aided diagnosis of pulmonary nodules using a two-step approach for feature selection and classifier ensemble construction," *Artificial Intelligence in Medicine*, vol. 50, no. 1, pp. 43–53, 2010.

[25] J. Ma, Q. Wang, Y. Ren, H. Hu, and J. Zhao, "Automatic lung nodule classification with radiomics approach," in *Medical Imaging 2016: PACS and Imaging Informatics: Next Generation and Innovations*, vol. 9789, 2016, p. 978 906.

[26] K.-L. Hua, C.-H. Hsu, S. C. Hidayati, W.-H. Cheng, and Y.-J. Chen, "Computer-aided classification of lung nodules on computed tomography images via deep learning technique," *OncoTargets and Therapy*, vol. 8, 2015.

[27] C. Wang, A. Elazab, J. Wu, and Q. Hu, "Lung nodule classification using deep feature fusion in chest radiography," *Computerized Medical Imaging and Graphics*, vol. 57, pp. 10–18, 2017.

[28] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1207–1216, 2016.

[29] A. Varnavas, T. Carrell, and G. Penney, "Fully automated 2D–3D registration and verification," *Medical Image Analysis*, vol. 26, no. 1, pp. 108 –119, 2015.

[30] S. Zhou, Z. Xiong, C. Chen, X. Chen, D. Liu, Y. Zhang, Z.-J. Zha, and F. Wu, "Fast and accurate electron microscopy image registration with 3D convolution," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019, pp. 478–486.

[31] P. W. Wilson, R. B. D'Agostino, D. Levy, A. M. Belanger, H. Silbershatz, and W. B. Kannel, "Prediction of coronary heart disease using risk factor categories," *Circulation*, vol. 97, no. 18, pp. 1837–1847, 1998.

[32] C. Lian, S. Ruan, T. Denœux, F. Jardin, and P. Vera, "Selecting radiomic features from FDG-PET images for cancer treatment outcome prediction," *Medical Image Analysis*, vol. 32, pp. 257–268, 2016.

[33] H. Mi, C. Petitjean, P. Vera, and S. Ruan, "Joint tumor growth prediction and tumor segmentation on therapeutic follow-up pet images," *Medical Image Analysis*, vol. 23, no. 1, pp. 84–91, 2015.

[34] B. Jing, P. Xie, and E. Xing, "On the automatic generation of medical imaging reports," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, vol. 1, Melbourne, Australia, 2018, pp. 2577–2586.

[35] J. Yuan, H. Liao, R. Luo, and J. Luo, "Automatic radiology report generation based on multi-view image fusion and medical concept enrichment," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 11769, 2019, pp. 721–729.

[36] C. R. Brice and C. L. Fennema, "Scene analysis using regions," *Artificial Intelligence*, vol. 1, no. 3-4, pp. 205–226, 1970.

[37] A. Rosenfeld, *Digital picture processing*. Academic press, 1976.

[38] T. Pavlidis, "Waveform segmentation through functional approximation," *IEEE Transactions on Computers*, vol. 100, no. 7, pp. 689–697, 1973.

[39] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[40] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[41] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations," *Journal of Computational Physics*, vol. 79, no. 1, pp. 12–49, 1988.

[42] C. Xu, J. L. Prince, *et al.*, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.

[43] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.

[44] A. Chakraborty, L. H. Staib, and J. S. Duncan, "Deformable boundary finding in medical images by integrating gradient and region information," *IEEE Transactions on Medical Imaging*, vol. 15, no. 6, pp. 859–870, 1996.

[45] C. S. Poon and M. Braun, "Image segmentation by a deformable contour model incorporating region analysis," *Physics in Medicine & Biology*, vol. 42, no. 9, p. 1833, 1997.

[46] S. Hojjatoleslami and F. Kruggel, "Segmentation of large brain lesions," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 666–669, 2001.

[47] N. Nikolaidis and I. Pitas, *3-D Image Processing Algorithms*, 1st. New York, NY, USA, 2000.

[48] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543–563, 2009.

[49] K. Karasawa, M. Oda, T. Kitasaka, K. Misawa, M. Fujiwara, C. Chu, G. Zheng, D. Rueckert, and K. Mori, "Multi-atlas pancreas segmentation: Atlas selection based on vessel structure," *Medical Image Analysis*, vol. 39, pp. 18–28, 2017.

[50] A. Pitiot, H. Delingette, P. M. Thompson, and N. Ayache, "Expert knowledge-guided segmentation system for brain mri," *NeuroImage*, vol. 23, S85–S96, 2004.

[51] S. Wörz and K. Rohr, "A new 3D parametric intensity model for accurate segmentation and quantification of human vessels," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2004, pp. 491–499.

[52] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[53] P. A. Yushkevich, H. Zhang, and J. C. Gee, "Continuous medial representation for anatomical structures," *IEEE Transactions on Medical Imaging*, vol. 25, no. 12, pp. 1547–1564, 2006.

[54] D. Lesage, E. D. Angelini, I. Bloch, and G. Funka-Lea, "A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes," *Medical Image Analysis*, vol. 13, no. 6, pp. 819–845, 2009.

[55] M. de Bruijne, B. van Ginneken, M. A. Viergever, and W. J. Niessen, "Adapting active shape models for 3D segmentation of tubular structures in medical images," in *Biennial International Conference on Information Processing in Medical Imaging*, 2003, pp. 136–147.

[56] M. Hernández-Hoyos, M. Orkisz, P. Puech, C. Mansard-Desbleds, P. Douek, and I. E. Magnin, "Computer-assisted analysis of three-dimensional mr angiograms," *RadioGraphics*, vol. 22, no. 2, pp. 421–436, 2002.

[57] R. Toledo, X. Orriols, P. Radeva, X. Binefa, J. Vitria, C. Canero, and J. Villanuev, "Eigensnakes for vessel segmentation in angiography," in *Proceedings of the International Conference on Pattern Recognition*, vol. 4, 2000, pp. 340–343.

[58] P. J. Yim, J. J. Cebral, R. Mullick, H. B. Marcos, and P. L. Choyke, "Vessel surface reconstruction with a tubular deformable model," *IEEE Transactions on Medical Imaging*, vol. 20, no. 12, pp. 1411–1421, 2001.

[59] H. Li and A. Yezzi, "Vessels as 4-D curves: Global minimal 4-D paths to extract 3-D tubular surfaces and centerlines," *IEEE Transactions on Medical Imaging*, vol. 26, no. 9, pp. 1213–1223, 2007.

[60] M. Chung, J. Lee, J. W. Chung, and Y.-G. Shin, "Accurate liver vessel segmentation via active contour model with dense vessel candidates," *Computer Methods and Programs in Biomedicine*, vol. 166, pp. 61–75, 2018.

[61] L. M. Lorigo, O. D. Faugeras, W. E. L. Grimson, R. Keriven, R. Kikinis, A. Nabavi, and C.-F. Westin, "Curves: Curve evolution for vessel segmentation," *Medical Image Analysis*, vol. 5, no. 3, pp. 195–206, 2001.

[62] T. Lv, G. Yang, Y. Zhang, J. Yang, Y. Chen, H. Shu, and L. Luo, "Vessel segmentation using centerline constrained level set method," *Multimedia Tools and Applications*, vol. 78, no. 12, pp. 17 051–17 075, 2019.

[63] A. F. Frangi, W. J. Niessen, R. M. Hoogeveen, T. Van Walsum, and M. A. Viergever, "Model-based quantitation of 3-D magnetic resonance angiographic images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 946–956, 1999.

[64] S. Bouix, K. Siddiqi, and A. Tannenbaum, "Flux driven automatic centerline extraction," *Medical Image Analysis*, vol. 9, no. 3, pp. 209–221, 2005.

[65] N. Flasque, M. Desvignes, J.-M. Constans, and M. Revenu, "Acquisition, segmentation and tracking of the cerebral vascular tree on 3D magnetic resonance angiography images," *Medical Image Analysis*, vol. 5, no. 3, pp. 173–183, 2001.

[66] T. O'Donnell, M.-P. Dubuisson-Jolly, and A. Gupta, "A cooperative framework for segmentation using 2D active contours and 3D hybrid models as applied to branching cylindrical structures," in *Proceedings of IEEE International Conference on Computer Vision*, 1998, pp. 454–459.

[67] O. Friman, M. Hindennach, C. Kühnel, and H.-O. Peitgen, "Multiple hypothesis template tracking of small 3D vessel structures," *Medical Image Analysis*, vol. 14, no. 2, pp. 160–171, 2010.

[68] C. Lacoste, G. Finet, and I. E. Magnin, "Coronary tree extraction from x-ray angiograms using marked point processes," in *Proceedings IEEE International Symposium on Biomedical Imaging*, 2006, pp. 157–160.

[69] J. A. Tyrrell, E. di Tomaso, D. Fuja, R. Tong, K. Kozak, R. K. Jain, and B. Roysam, "Robust 3-D modeling of vasculature imagery using superellipsoids," *IEEE Transactions on Medical Imaging*, vol. 26, no. 2, pp. 223–237, 2007.

[70] J. Weickert, "Coherence-enhancing diffusion filtering," *International Journal of Computer Vision*, vol. 31, no. 2-3, pp. 111–127, 1999.

[71] R. Wiemker, T. Klinder, M. Bergtholdt, K. Meetz, I. C. Carlsen, and T. Bülow, "A radial structure tensor and its use for shape-encoding medical visualization of tubular and nodular structures," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 3, pp. 353–366, 2012.

[72] M. Nergiz and M. Akın, "Retinal vessel segmentation via structure tensor coloring and anisotropy enhancement," *Symmetry*, vol. 9, no. 11, p. 276, 2017.

[73] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[74] G. Agam and C. Wu, "Probabilistic modeling based vessel enhancement in thoracic CT scans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, 2005, pp. 649–654.

[75] C. Bauer and H. Bischof, "A novel approach for detection of tubular objects and its application to medical image analysis," in *Joint Pattern Recognition Symposium*, 2008, pp. 163–172.

[76] C. J. Goch, X. Wang, H.-P. Meinzer, and I. Wegner, "Liver vessel segmentation using gradient vector flow," in *Bildverarbeitung für die Medizin 2011*, 2011, pp. 104–108.

[77] M. W. Law and A. C. Chung, "Three dimensional curvilinear structure detection using optimally oriented flux," in *European Conference on Computer Vision*, 2008, pp. 368–382.

[78] Q. Lin, "Enhancement, extraction, and visualization of 3D volume data," PhD thesis, Linköping University Electronic Press, 2003.

[79] T. M. Koller, G. Gerig, G. Szekely, and D. Dettwiler, "Multiscale detection of curvilinear structures in 2-D and 3-D image data," in *Proceedings of IEEE International Conference on Computer Vision*, 1995, pp. 864–869.

[80] Y. Sato, S. Nakajima, N. Shiraga, H. Atsumi, S. Yoshida, T. Koller, G. Gerig, and R. Kikinis, "Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," *Medical Image Analysis*, vol. 2, no. 2, pp. 143–168, 1998.

[81] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 1998, pp. 130–137.

[82] Q. Li, S. Sone, and K. Doi, "Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional CT scans," *Medical Physics*, vol. 30, no. 8, pp. 2040–2051, 2003.

[83] H. Shikata, E. A. Hoffman, and M. Sonka, "Automated segmentation of pulmonary vascular tree from 3D CT images," in *Medical Imaging 2004: Physiology, Function, and Structure from Medical Images*, vol. 5369, 2004, pp. 107–116.

[84] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2538–2548, 2012.

[85] N. Armande, P. Montesinos, and O. Monga, "A 3D thin nets extraction method for medical imaging," in *Proceedings of the International Conference on Pattern Recognition*, vol. 1, 1996, 642–646 vol.1.

[86] V. Prinet, O. Monaga, C. Ge, S. L. Xie, and S. D. Ma, "Thin network extraction in 3D images: Application to medical angiograms," in *Proceedings of the International Conference on Pattern Recognition*, vol. 7276, Washington, DC, USA, 1996, pp. 386–401.

[87] N. Farzaneh, S. Habbo-Gavin, S. M. R. Soroushmehr, H. Patel, D. P. Fessell, K. R. Ward, and K. Najarian, "Atlas based 3D liver segmentation using adaptive thresholding and superpixel approaches," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2017, pp. 1093–1097.

[88] H. Arabi and H. Zaidi, "Comparison of atlas-based techniques for whole-body bone segmentation," *Medical Image Analysis*, vol. 36, pp. 98–112, 2017.

[89] B. Oliveira, S. Queirós, P. Morais, H. R. Torres, J. Gomes-Fonseca, J. C. Fonseca, and J. L. Vilaça, "A novel multi-atlas strategy with dense deformation field reconstruction for abdominal and thoracic multi-organ segmentation from computed tomography," *Medical Image Analysis*, vol. 45, pp. 108–120, 2018.

[90] Z. Xu, R. P. Burke, C. P. Lee, R. B. Baucom, B. K. Poulose, R. G. Abramson, and B. A. Landman, "Efficient multi-atlas abdominal segmentation on clinically acquired CT with SIMPLE context learning," *Medical Image Analysis*, vol. 24, no. 1, pp. 18–27, 2015.

[91] M. Baiker, J. Milles, J. Dijkstra, T. D. Henning, A. W. Weber, I. Que, E. L. Kaijzel, C. W. Löwik, J. H. Reiber, and B. P. Lelieveldt, "Atlas-based whole-body segmentation of mice from low-contrast Micro-CT data," *Medical Image Analysis*, vol. 14, no. 6, pp. 723–737, 2010.

[92] M. Erdt, M. Kirschner, K. Drechsler, S. Wesarg, M. Hammon, and A. Cavallaro, "Automatic pancreas segmentation in contrast enhanced CT data using learned spatial anatomy and texture descriptors," in *Proceedings IEEE International Symposium on Biomedical Imaging*, 2011, pp. 2076–2082.

[93] M. Hammon, A. Cavallaro, M. Erdt, P. Dankerl, M. Kirschner, K. Drechsler, S. Wesarg, M. Uder, and R. Janka, "Model-based pancreas segmentation in portal venous phase contrast-enhanced CT images," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1082–1090, 2013.

[94] A. Shimizu, T. Kimoto, H. Kobatake, S. Nawano, and K. Shinozaki, "Automated pancreas segmentation from three-dimensional contrast-enhanced computed tomography," *International Journal of Computer Assisted Radiology and Surgery*, vol. 5, no. 1, p. 85, 2010.

[95] Y. F. Huang, Z. Z. Xu, and C. R. Yan, "Automatic gallbladder location and segmentation based on anatomical knowledge," *Applied Mechanics and Materials*, vol. 556, pp. 4697–4700, 2014.

[96] O. Camara, O. Colliot, and I. Bloch, "Computational modeling of thoracic and abdominal anatomy using spatial relationships for image segmentation," *Real-Time Imaging*, vol. 10, no. 4, pp. 263–273, Aug. 2004.

[97] J. K. Udupa, D. Odhner, A. X. Falcao, K. C. Ciesielski, P. A. Miranda, P. Vaideeswaran, S. Mishra, G. J. Grevera, B. Saboury, and D. A. Torigian, "Fuzzy object modeling," in *Medical Imaging 2011: Visualization, Image-Guided Procedures, and Modeling*, vol. 7964, 2011, 79640B.

[98] J.-B. Fasquel, V. Agnus, J. Moreau, L. Soler, and J. Marescaux, "An interactive medical image segmentation system based on the optimal management of regions of interest using topological medical knowledge," *Computer methods and programs in biomedicine*, vol. 82, no. 3, pp. 216–230, 2006.

[99] A. Criminisi, D. Robertson, E. Konukoglu, J. Shotton, S. Pathak, S. White, and K. Siddiqui, "Regression forests for efficient anatomy detection and localization in computed tomography scans," *Medical Image Analysis*, vol. 17, no. 8, pp. 1293–1303, 2013.

[100] Y. Nimura, Y. Hayashi, T. Kitasaka, K. Misawa, and K. Mori, "Automated torso organ segmentation from 3D CT images using conditional random field," in *Medical Imaging 2016: Computer-Aided Diagnosis*, vol. 9785, 2016, p. 97853M.

[101] C. Bhole, N. Morsillo, and C. Pal, "3D segmentation in CT imagery with conditional random fields and histograms of oriented gradients," in *International Workshop on Machine Learning in Medical Imaging*, 2011, pp. 326–334.

[102] H.-C. Shin, M. R. Orton, D. J. Collins, S. J. Doran, and M. O. Leach, "Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1930–1943, 2012.

[103] H. R. Roth, C. T. Lee, H. Shin, A. Seff, L. Kim, J. Yao, L. Lu, and R. M. Summers, "Anatomy-specific classification of medical images using deep convolutional nets," in *Proceedings IEEE International Symposium on Biomedical Imaging*, 2015, pp. 101–104.

[104] Z. Yan, Y. Zhan, Z. Peng, S. Liao, Y. Shinagawa, S. Zhang, D. N. Metaxas, and X. S. Zhou, "Multi-instance deep learning: Discover discriminative local anatomies for bodypart recognition," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1332–1343, 2016.

[105] F. Milletari, S.-A. Ahmadi, C. Kroll, A. Plate, V. Rozanski, J. Maiostre, J. Levin, O. Dietrich, B. Ertl-Wagner, K. Bötzel, *et al.*, "Hough-CNN: Deep learning for segmentation of deep brain regions in mri and ultrasound," *Computer Vision and Image Understanding*, vol. 164, pp. 92–102, 2017.

[106] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234–241.

[107] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*, 2016, pp. 424–432.

[108] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain tumor segmentation and radiomics survival prediction: Contribution to the BRATS 2017 challenge," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 2018, pp. 287–297.

[109] H. R. Roth, H. Oda, X. Zhou, N. Shimizu, Y. Yang, Y. Hayashi, M. Oda, M. Fujiwara, K. Misawa, and K. Mori, "An application of cascaded 3D fully convolutional networks for medical image segmentation," *Computerized Medical Imaging and Graphics*, vol. 66, pp. 90–99, 2018.

[110] H. R. Roth, C. Shen, H. Oda, T. Sugino, M. Oda, Y. Hayashi, K. Misawa, and K. Mori, "A multi-scale pyramid of 3D fully convolutional networks for abdominal multi-organ segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 417–425.

[111] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, *et al.*, "Attention U-Net: Learning where to look for the pancreas," in *International Conference on Medical Imaging with Deep Learning*, 2018.

[112] H. Kakeya, T. Okada, and Y. Oshiro, "3D U-JAPA-Net: Mixture of convolutional networks for abdominal multi-organ CT segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 426–433.

[113] F. Isensee and K. H. Maier-Hein, "An attempt at beating the 3D U-Net," *arXiv preprint arXiv:1908.02182*, 2019.

[114] C. Wachinger, M. Reuter, and T. Klein, "DeepNAT: Deep convolutional neural network for segmenting neuroanatomy," *NeuroImage*, vol. 170, pp. 434 –445, 2018, Segmenting the Brain.

[115] V. V. Valindria, N. Pawlowski, M. Rajchl, I. Lavdas, E. O. Aboagye, A. G. Rockall, D. Rueckert, and B. Glocker, "Multi-modal learning from unpaired images: Application to multi-organ segmentation in CT and MRI," in *2018 IEEE Winter Conference on Applications of Computer Vision*, 2018, pp. 547–556.

[116] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision*, 2016, pp. 565–571.

[117] J. Yun, J. Park, D. Yu, J. Yi, M. Lee, H. J. Park, J.-G. Lee, J. B. Seo, and N. Kim, "Improvement of fully automated airway segmentation on volumetric computed tomographic images using a 2.5 dimensional convolutional neural net," *Medical Image Analysis*, vol. 51, pp. 13 –20, 2019.

[118] Q. Meng, H. R. Roth, T. Kitasaka, M. Oda, J. Ueno, and K. Mori, "Tracking and segmentation of the airways in chest CT using a fully convolutional network," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2017, pp. 198–207.

[119] T Kitasaka, K Mori, J Hasegawa, and J Toriwaki, "A method for extraction of bronchus regions from 3D branch tracing and image sharpening for airway tree chest X-ray images by analyzing structural features of the bronchus," *Forma*, vol. 17, pp. 321–338, 2002.

[120] Y. Qin, M. Chen, H. Zheng, Y. Gu, M. Shen, J. Yang, X. Huang, Y.-M. Zhu, and G.-Z. Yang, "Airwaynet: A voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019, pp. 212–220.

[121] P. Lu, L. Barazzetti, V. Chandran, K. Gavaghan, S. Weber, N. Gerber, and M. Reyes, "Highly accurate facial nerve segmentation refinement from CBCT/CT imaging using a super-resolution classification approach," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 1, pp. 178–188, 2017.

[122] H. Bergwerf, M. Bechakra, I. Smal, J. L. Jongen, and E. Meijering, "Nerve fiber segmentation in bright-field microscopy images of skin biopsies using deep learning," in *Proceedings IEEE International Symposium on Biomedical Imaging*, 2019, pp. 232–235.

[123] J. Liu, Y. Huo, Z. Xu, A. Assad, R. G. Abramson, and B. A. Landman, "Multi-atlas spleen segmentation on CT using adaptive context learning," in *Medical Imaging 2017: Image Processing*, vol. 10133, 2017, pp. 65–71.

[124] H. Roth, M. Oda, N. Shimizu, H. Oda, Y. Hayashi, T. Kitasaka, M. Fujiwara, K. Misawa, and K. Mori, "Towards dense volumetric pancreas segmentation in CT using 3D fully convolutional networks," in *Medical Imaging 2018: Image Processing*, vol. 10574, 2018, 105740B.

[125] A. P. Harrison, Z. Xu, K. George, L. Lu, R. M. Summers, and D. J. Mollura, "Progressive and multi-path holistically nested neural networks for pathological lung segmentation from CT images," in *International Conference on Medical Image Computing And Computer-Assisted Intervention*, 2017, pp. 621–629.

[126] O. Wink, W. J. Niessen, and M. A. Viergever, "Multiscale vessel tracking," *IEEE Transactions on Medical Imaging*, vol. 23, no. 1, pp. 130–133, 2004.

[127] H. Skibbe, M. Reisert, S.-i. Maeda, M. Koyama, S. Oba, K. Ito, and S. Ishii, "Efficient monte carlo image analysis for the location of vascular entity," *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 628–643, 2014.

[128] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[129] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," in *ACM Transactions on Graphics*, vol. 23, 2004, pp. 309–314.

[130] T. W. Tang and A. C. Chung, "Non-rigid image registration using graph-cuts," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2007, pp. 916–924.

[131] S. Vicente, V. Kolmogorov, and C. Rother, "Graph cut based image segmentation with connectivity priors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.

[132] S. Jegelka and J. Bilmes, "Submodularity beyond submodular energies: Coupling edges in graph cuts," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2011, pp. 1897–1904.

[133] P. Kohli, A. Osokin, and S. Jegelka, "A principled deep random field model for image segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1971–1978.

[134] C. Bauer, T. Pock, E. Sorantin, H. Bischof, and R. Beichel, "Segmentation of interwoven 3D tubular tree structures utilizing shape priors and graph cuts," *Medical Image Analysis*, vol. 14, no. 2, pp. 172–184, 2010.

[135] S. Esneault, C. Lafon, and J.-L. Dillenseger, "Liver vessels segmentation using a hybrid geometrical moments/graph cuts method," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 2, pp. 276–283, 2010.

[136] S. Han and X. Wang, "Texture segmentation using graph cuts in spectral decomposition based riemannian multi-scale nonlinear structure tensor space," *International Journal of Computer Theory and Engineering*, vol. 7, no. 4, pp. 259–263, 2015.

[137] S. B. Pope, "Algorithms for ellipsoids," Cornell University, FDA, 08-01, 2008.

[138] M. Moshtaghi, T. C. Havens, J. C. Bezdek, L. Park, C. Leckie, S. Rajasegarar, J. M. Keller, and M. Palaniswami, "Clustering ellipses for anomaly detection," *Pattern Recognition*, vol. 44, no. 1, pp. 55–69, 2011.

[139] J. Bigün, "Optimal orientation detection of linear symmetry," in *Proceedings of IEEE International Conference on Computer Vision*, 1987, pp. 433–438.

[140] J. Bigün, G. H. Granlund, and J. Wiklund, "Multidimensional orientation estimation with applications to texture analysis and optical flow," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 775–790, 1991.

[141] H. Knutsson, C.-F. Westin, and M. Andersson, "Representing local structure using tensors II," in *Scandinavian Conference on Image Analysis*, 2011, pp. 545–556.

[142] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *Journal of Applied Statistics*, vol. 21, pp. 224–270, 1994.

[143] C. Wang, M. Oda, Y. Hayashi, Y. Yoshino, T. Yamamoto, A. F. Frangi, and K. Mori, "Tensor-cut: A tensor-based graph-cut blood vessel segmentation method and its application to renal artery segmentation," *Medical Image Analysis*, vol. 60, p. 101 623, 2020.

[144] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of MRI data," *IEEE Transactions on Medical Imaging*, vol. 11, no. 2, pp. 221–232, 1992.

[145] A. Barmpoutis, B. C. Vemuri, T. M. Shepherd, and J. R. Forder, "Tensor splines for interpolation and approximation of DT-MRI with applications to segmentation of isolated rat hippocampi," *IEEE Transactions on Medical Imaging*, vol. 26, no. 11, pp. 1537–1546, 2007.

[146] C. Wang, M. Kagajo, Y. Nakamura, M. Oda, Y. Yoshino, T. Yamamoto, and K. Mori, "Precise renal artery segmentation for estimation of renal vascular dominant regions," in *Medical Imaging 2016 : Image Processing*, 2016, pp. 97842M–97842M.

[147] P. T. Fletcher and S. Joshi, "Riemannian geometry for the statistical analysis of diffusion tensor data," *Signal Processing*, vol. 87, no. 2, pp. 250–262, 2007.

[148] X. Pennec, P. Fillard, and N. Ayache, "A Riemannian framework for tensor computing," *International Journal of Computer Vision*, vol. 66, no. 1, pp. 41–66, 2006.

[149] A. Ghosh, M. Descoteaux, and R. Deriche, "Riemannian framework for estimating symmetric positive definite 4th order diffusion tensors," *Medical Image Computing and Computer-Assisted Intervention*, pp. 858–865, 2008.

[150] C. Wang, M. Oda, Y. Hayashi, Y. Yoshino, T. Yamamoto, A. F. Frangi, and K. Mori, "Tensor-based graph-cut in Riemannian metric space and its application to renal artery segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2016, pp. 353–361.

[151] M. Moakher, "A differential geometric approach to the geometric mean of symmetric positive-definite matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 3, pp. 735–747, 2005.

[152] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in nd images," in *Proceedings of IEEE International Conference on Computer Vision*, vol. 1, 2001, pp. 105–112.

[153] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[154] G. Hamarneh and P. Jassi, "Vascusynth: Simulating vascular trees for generating volumetric image data with ground truth segmentation and tree analysis," *Computerized Medical Imaging and Graphics*, vol. 34, no. 8, pp. 605–616, 2010.

[155] M. W. Law and A. C. Chung, "Efficient implementation for spherical flux computation and its application to vascular segmentation," *IEEE Transactions on Image Processing*, vol. 18, no. 3, pp. 596–612, 2009.

[156] AMIRA3-D, *Analysis software for life sciences [online]*, http://www.vsg3d.com/amira.

[157] A. Vasilevskiy and K. Siddiqi, "Flux maximizing geometric flows," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1565–1578, 2002.

[158] M. Schaap, C. T. Metz, T. van Walsum, A. G. van der Giessen, A. C. Weustink, N. R. Mollet, C. Bauer, H. Bogunović, C. Castro, X. Deng, *et al.*, "Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms," *Medical Image Analysis*, vol. 13, no. 5, pp. 701–714, 2009.

[159] C. Fouard, G. Malandain, S. Prohaska, and M. Westerhoff, "Blockwise processing applied to brain microvascular network study," *IEEE Transactions on Medical Imaging*, vol. 25, no. 10, pp. 1319–1328, 2006.

[160] C. Fouard and G. Malandain, "3-D chamfer distances and norms in anisotropic grids," *Image and Vision Computing*, vol. 23, no. 2, pp. 143–158, 2005.

[161] T. Rohlfing, "Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable," *IEEE Transactions on Medical Imaging*, vol. 31, no. 2, pp. 153–163, 2012.

[162] C. Metz, M. Schaap, T. van Walsum, A. van der Giessen, A. Weustink, N. Mollet, G. Krestin, and W. Niessen, "3D segmentation in the clinic: A grand challenge II-coronary artery tracking," *Insight Journal*, vol. 1, no. 5, pp. 6–18, 2008.

[163] S. Moccia, E. D. Momi, S. E. Hadji, and L. S. Mattos, "Blood vessel segmentation algorithms — review of methods, datasets and evaluation metrics," *Computer Methods and Programs in Biomedicine*, vol. 158, pp. 71 –91, 2018.

[164] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[165] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.

[166] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[167] G. Tetteh, V. Efremov, N. D. Forkert, M. Schneider, J. Kirschke, B. Weber, C. Zimmer, M. Piraud, and B. H. Menze, "Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-D angiographic volumes," *arXiv preprint arXiv:1803.09340*, 2018.

[168] M. Oda, T. Kitasaka, K. Misawa, M. Fujiwara, and K. Mori, "Abdominal artery segmentation from CT volumes using fully convolutional network for small artery segmentation," in *International Journal of Computer Assisted Radiology and Surgery, Sup. 1*, vol. 13, 2018, pp. 20–21.

[169] Q. Huang, J. Sun, H. Ding, X. Wang, and G. Wang, "Robust liver vessel extraction using 3D U-Net with variant dice loss function," *Computers in Biology and Medicine*, vol. 101, pp. 153–162, 2018.

[170] C. Wang, Y. Hayashi, M. Oda, H. Itoh, T. Kitasaka, A. F. Frangi, and K. Mori, "Tubular structure segmentation using spatial fully connected network with radial distance loss for 3d medical images," in *Medical Image Computing and Computer-Assisted Intervention*, 2019, pp. 348–356.

[171] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial as deep: Spatial CNN for traffic scene understanding," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[172] J. Ribera, D. Güera, Y. Chen, and E. Delp, "Weighted hausdorff distance: A loss function for object localization," *arXiv preprint arXiv:1806.07564*, 2018.

[173] S. Jia, A. Despinasse, Z. Wang, H. Delingette, X. Pennec, P. Jaïs, H. Cochet, and M. Sermesant, "Automatically segmenting the left atrium from cardiac images using successive 3D U-Nets and a contour loss," in *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*, vol. 11395, 2019, pp. 221–229.

[174] H. Chen, Q. Dou, L. Yu, J. Qin, and P.-A. Heng, "Voxresnet: Deep voxelwise residual networks for brain segmentation from 3D mr images," *NeuroImage*, vol. 170, pp. 446 –455, 2018.

[175] C. Shen, H. R. Roth, H. Oda, M. Oda, Y. Hayashi, K. Misawa, and K. Mori, "On the influence of dice loss function in multi-class organ segmentation of abdominal CT using 3D fully convolutional networks," *arXiv preprint arXiv:1801.05912*, 2018.

[176] W. Zhu, Y. Huang, L. Zeng, X. Chen, Y. Liu, Z. Qian, N. Du, W. Fan, and X. Xie, "Anatomynet: Deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy," *Medical Physics*, vol. 46, no. 2, pp. 576–589, 2019.

[177] W. Zhu, C. Liu, W. Fan, and X. Xie, "Deeplung: Deep 3D dual path nets for automated pulmonary nodule detection and classification," in *2018 IEEE Winter Conference on Applications of Computer Vision*, 2018, pp. 673–681.

[178] W. Zhu, Y. S. Vang, Y. Huang, and X. Xie, "Deepem: Deep 3D convnets with em for weakly supervised pulmonary nodule detection," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 812–820.

[179] W. Zhu, X. Xiang, T. D. Tran, G. D. Hager, and X. Xie, "Adversarial deep structured nets for mass segmentation from mammograms," in *Proceedings IEEE International Symposium on Biomedical Imaging*, 2018, pp. 847–850.

[180] C. Wang, H. R. Roth, T. Kitasaka, M. Oda, Y. Hayashi, Y. Yoshino, T. Yamamoto, N. Sassa, M. Goto, and K. Mori, "Precise estimation of renal vascular dominant regions using spatially aware fully convolutional networks, tensor-cut and voronoi diagrams," *Computerized Medical Imaging and Graphics*, vol. 77, p. 101 642, 2019.

[181] C.-A. Brust, S. Sickert, M. Simon, E. Rodner, and J. Denzler, "Convolutional patch networks with spatial prior for road detection and urban scene understanding," *arXiv preprint arXiv:1502.06344*, 2015.

[182] N. Akoury and A. Nguyen, "Spatial PixelCNN: Generating images from patches," *arXiv preprint arXiv:1712.00714*, 2017.

[183] L. Chen, Y. Xie, J. Sun, N. Balu, M. Mossa-Basha, K. Pimentel, T. S. Hatsukami, J.-N. Hwang, and C. Yuan, "3D intracranial artery segmentation using a convolutional autoencoder," in *IEEE International Conference on Bioinformatics and Biomedicine*, 2017, pp. 714–717.

[184] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Automatic coronary calcium scoring in cardiac CT angiography using convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 589–596.

[185] H. R. Roth, C. Shen, H. Oda, M. Oda, Y. Hayashi, K. Misawa, and K. Mori, "Deep learning and its application to medical image segmentation," *Medical Imaging Technology*, vol. 36, no. 2, pp. 63–71, 2018.

[186] R. Cuingnet, R. Prevost, D. Lesage, L. D. Cohen, B. Mory, and R. Ardon, "Automatic detection and segmentation of kidneys in 3D CT images using random forests," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2012, pp. 66–74.

[187] C. Chu, M. Oda, T. Kitasaka, K. Misawa, M. Fujiwara, Y. Hayashi, Y. Nimura, D. Rueckert, and K. Mori, "Multi-organ segmentation based on spatially-divided probabilistic atlas from 3D abdominal CT images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2013, pp. 165–172.

[188] Y. Zheng, D. Liu, B. Georgescu, D. Xu, and D. Comaniciu, "Deep learning based automatic segmentation of pathological kidney in CT: Local versus global image context," in *Deep Learning and Convolutional Neural Networks for Medical Image Computing*, 2017, pp. 241–255.

[189] A. Skalski, K. Heryan, J. Jakubowski, and T. Drewniak, "Kidney segmentation in CT data using hybrid Level-Set method with ellipsoidal shape constraints," *Metrology and Measurement Systems*, vol. 24, no. 1, pp. 101–112, 2017.

[190] W. Thong, S. Kadoury, N. Piché, and C. J. Pal, "Convolutional networks for kidney segmentation in contrast-enhanced CT scans," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pp. 1–6, 2016.

[191] C. Molnar and J. Gair, *Concepts of Biology – 1st Canadian Edition*. BCcampus, 2019.

[192] R. J. Motzer, E. Jonasch, N. Agarwal, S. Bhayani, W. P. Bro, S. S. Chang, T. K. Choueiri, B. A. Costello, I. H. Derweesh, M. Fishman, *et al.*, "Kidney cancer, version 2.2017, nccn clinical practice guidelines in oncology," *Journal of the National Comprehensive Cancer Network*, vol. 15, no. 6, pp. 804–834, 2017.

[193] S. Osono, I. Hara, N. Shinohara, H. Kanayama, Y. Tomita, Y. Nagashima, M. Eto, S. Hinotsu, H. Miyake, and S. Takayuki, "Clinical practice guideline for renal cancer," *The Japanese Urological Association*, 2017.

[194] P. Shao, C. Qin, C. Yin, X. Meng, X. Ju, J. Li, Q. Lv, W. Zhang, and Z. Xu, "Laparoscopic partial nephrectomy with segmental renal artery clamping: Technique and clinical outcomes," *European Urology*, vol. 59, no. 5, pp. 849–855, 2011.

[195] P. Shao, L. Tang, P. Li, Y. Xu, C. Qin, Q. Cao, X. Ju, X. Meng, Q. Lv, J. Li, *et al.*, "Precise segmental renal artery clamping under the guidance of dual-source computed tomography angiography during laparoscopic partial nephrectomy," *European Urology*, vol. 62, no. 6, pp. 1001–1008, 2012.

[196] Y. Yoshino, T. Yamamoto, Y. Funahashi, M. Oda, M. Kagajo, C. Wang, K. Mori, and M. Gotoh, "Computational analysis of recovery from ischemic damage to kidney function undergoing robotic partial nephrectomy for renal tumor," in *Supplement of Journal of Endourology*, 2015, pp. 916–924.

[197] O. Ukimura, M. Nakamoto, and I. S. Gill, "Three-dimensional reconstruction of renovascular-tumor anatomy to facilitate zero-ischemia partial nephrectomy," *European Urology*, vol. 61, no. 1, pp. 211–217, 2012.

[198] Y. Komai, Y. Sakai, N. Gotohda, T. Kobayashi, S. Kawakami, and N. Saito, "A novel 3-dimensional image analysis system for case-specific kidney anatomy and surgical simulation to facilitate clampless partial nephrectomy," *Urology*, vol. 83, no. 2, pp. 500–507, 2014.

[199] S. Isotani, H. Shimoyama, I. Yokota, T. China, S.-i. Hisasue, H. Ide, S. Muto, R. Yamaguchi, O. Ukimura, and S. Horie, "Feasibility and accuracy of computational robot-assisted partial nephrectomy planning by virtual partial nephrectomy analysis," *International Journal of Urology*, vol. 22, no. 5, pp. 439–446, 2015.

[200] M. H. Jensen, D. R. Jørgensen, R. Jalaboi, M. E. Hansen, and M. A. Olsen, "Improving uncertainty estimation in convolutional neural networks using inter-rater agreement," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019, pp. 540–548.

[201] R. S. Gill, B. Caldairou, N. Bernasconi, and A. Bernasconi, "Uncertainty-informed detection of epileptogenic brain malformations using bayesian neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019, pp. 225–233.

[202] Z. Eaton-Rosen, T. Varsavsky, S. Ourselin, and M. J. Cardoso, "As easy as 1, 2...4? uncertainty in counting tasks for medical imaging," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019, pp. 356–364.

[203] M. Tardy, B. Scheffer, and D. Mateus, "Uncertainty measurements for the reliable classification of mammograms," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019, pp. 495–503.

[204] H. Karcher, "Riemannian center of mass and mollifier smoothing," *Communications on Pure and Applied Mathematics*, vol. 30, no. 5, pp. 509–541, 1977.

Appendix A: Statistics of Hessian tensors

This section simply describes the mean and variance of Hessian tensors. The Fréchet mean generalizes the centroids to arbitrary matrix space by minimizing the Fréchet variance:

$$\bar{\mathcal{T}} = \arg\min_{\mathcal{T}\in\mathbb{T}} \rho(\mathcal{T}), \tag{A.1}$$

where $\rho(\mathcal{T})$ denotes sum-of-squared distance which is defined:

$$\rho(\mathcal{T}) = \frac{1}{2N}\sum_{i=1}^{N} d^2(\bar{\mathcal{T}}, \mathcal{T}_i). \tag{A.2}$$

Geometric mean tensor $\bar{\mathcal{T}}$ is the center point on a manifold at which the Fréchet variance is minimized for any tensor $\mathcal{T} \in \mathbb{T}$. The existence and uniqueness of geometric mean has been proven by Karcher [204]. Then, Newton gradient descent solves the given minimization problem in Eq. A.1. Also, the gradient of $\rho(\mathcal{T})$ is given by

$$\nabla\rho(\mathcal{T}) = -\frac{1}{N}\sum_{i=1}^{N} \log_{\mathcal{T}}(\mathcal{T}_i). \tag{A.3}$$

Initialize $\bar{\mathcal{T}}_0$ with arbitrary tensor $\mathcal{T} \in \mathbb{T}$, and the $t+1$ step of the Newton gradient descent procedure is given:

$$\bar{\mathcal{T}}_{t+1} = \exp_{\bar{\mathcal{T}}}\left(\frac{1}{N}\sum_{i=1}^{N}\log_{\bar{\mathcal{T}}_t}(\mathcal{T}_i)\right) = \bar{\mathcal{T}}_t^{\frac{1}{2}} \exp\left(\frac{1}{N}\sum_{i=1}^{N}\log(\bar{\mathcal{T}}_t^{-\frac{1}{2}}\mathcal{T}_i\bar{\mathcal{T}}_t^{-\frac{1}{2}})\right)\bar{\mathcal{T}}_t^{\frac{1}{2}}. \tag{A.4}$$

As shown in Fig. 2.7, $\exp_{\mathbf{x}}(\mathbf{\Sigma})$ is called exponential mapping, which is an inverse mapping of logarithmic mapping. Here exponential mapping is to delineates the arithmetic mean in the tangent space back to the manifold space. Gradient descent algorithm workflow is given in Algorithm 2. Detailed deduction information is available [148].

Algorithm 2 Calculate geometric mean tensor $\bar{\mathcal{T}}$

Input: $\mathcal{T}_i \in \mathbb{T}, (i = 1, 2, ..., N)$
Output: Mean tensor $\bar{\mathcal{T}}$
Initialize: $\bar{\mathcal{T}}_0 = \mathcal{T}_1$, $t = 0, \epsilon = 0.001$
Do
$t = t + 1$ ▷ Update step
$X_t = \frac{1}{N} \sum_{k=1}^{N} \log_{\bar{\mathcal{T}}_t}(\mathcal{T}_k)$ ▷ Gradient
$\bar{\mathcal{T}}_{t+1} = \exp_{\bar{\mathcal{T}}_t}(X_t)$ ▷ Update mean tensor
While $\|X_t\| > \epsilon$

Since we know how to calculate the distance and the mean, the variance of tensors σ_T^2 is given:

$$\sigma_T^2 = \mathcal{E}[d^2(\bar{\mathcal{T}}, \mathcal{T})] = \frac{1}{N} \sum_{i=1}^{N} d^2(\bar{\mathcal{T}}, \mathcal{T}_i), \tag{A.5}$$

where $\mathcal{E}[\cdot]$ denotes the expectation, and $\mathcal{E}[d^2(\bar{\mathcal{T}}, \mathcal{T})]$ is the expectation of the squared deviation from mean tensor $\bar{\mathcal{T}}$. The statistics of the tensors given above, geodesic distance, geometric mean and variance, are used to compute a Gaussian mixture model (GMM) to estimate the probability distribution of the foreground and background voxels.